# UCL

# Minority Report: the impact of predicted grades on university admissions of disadvantaged groups

## Centre for Education Policy and Equalising Opportunities (CEPEO)

**Richard Murphy and Gill Wyness**

**Working Paper No. 20-07**

**March 2020**

Disclaimer

Any opinions expressed here are those of the author(s) and not those of the UCL Institute of Education. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

CEPEO Working Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

# Highlights

- The UK higher education system has the unusual feature that students must apply to degree courses on the basis of predicted exam results (A-levels), rather than their actual grades

- Little is known about how accurate these predictions are, and what the consequences of inaccurate predictions might be for students' enrolment outcomes

- We use aggregate data from on university applications to study the accuracy of predicted grades, and to examine where students with different predictions end up.

- Only 16% of applicants achieve the A-level grade points that they were predicted to achieve by their teachers, based on their best 3 A-levels. And the vast majority of applicants are over-predicted – i.e. their grades are predicted to be higher than they actually go on to achieve.

- But among the highest achieving students, those from disadvantaged backgrounds receive predicted grades that are lower than those from more advantaged backgrounds.

- This may have important consequences for social mobility, since under-predicted students are more likely to be overqualified for the course they eventually enroll in.

# Why does this matter?

Only 16% of A-level grades are accurate; 75% of students receive predictions that are over-inflated.

We also show that among equally high attaining students, disadvantaged students receive predictions that are lower than more advantaged students

Using predicted rather than actual A-level grades could have important consequences for young people's life chances, given that under-predicted students are more likely to be overqualified for the course they eventually enrol in

## 1. Introduction

Despite the possibility of high returns from a university education, many students do not go on to participate in higher education (HE). A great deal of policy interest has focused on the group of non-participants, and the reasons why they choose not to attend university (Carneiro and Heckman, 2002; Chowdry et al, 2013; McNally, 2016). However, even among the elite group of students who do attend university, students from lower socio-economic backgrounds are less likely to attend a high tariff university than their richer counterparts, even when they have similar prior academic attainment (Crawford et al, 2016). Recent research from both the UK and US (Dillon & Smith, 2017; Campbell et al, 2019) finds evidence that a significant proportion of students are 'undermatched' - i.e. their academic credentials would permit them to access a university that is more selective than the one they attend, and that this is more prevalent among disadvantaged students.

This undermatch may arise as a result of a variety of reasons. For example students may apply to universities appropriate to their ability, but fail to gain a place (Boliver, 2013). On the other hand, disadvantaged students may simply fail to apply to selective universities (Hoxby & Avery, 2012). They may prefer to attend their local university, for example, to minimize living costs[1]. Other reasons for undermatching university choice include; wanting to attend the same institution as their peers or siblings (Goodman, Hurwitz & J Smith, 2014); misinformation about the future benefits of high-status universities (McGuigan et al, 2016). In this paper we focus on one particular potential explanation that has not yet been covered in the literature: that disadvantaged students receive more pessimistic predictions about their grades.

---

[1] Note that practically all university degrees charge the maximum tuition fees allowable – meaning there is no variation in fees across degrees. However, disadvantaged students may want to attend their local university to continue living with parents. This would will result in them attending less selective universities on average compared to more advantaged students who are willing to travel further.

In this paper we study the extent, and impact of, misinformation on grades, by examining student application behaviour in the UK. In the UK system, students must make their university choices based on their high-school teachers' predictions of their school-leaving examination grades (A-levels) rather than their actual grades, and commit to their first choice and reserve universities based on these predictions. Thus, teachers' predictions about their students can have a major impact on their future outcomes.

We find this system to generate a high degree of misinformation on students' ability. Only 16% of applicants achieve the A-level grade points that they were predicted to achieve by their teachers, based on their best 3 A-levels. However, the vast majority (75% of applicants) are over-predicted – i.e. their grades are predicted to be higher than they actually achieve. The average student is over-predicted by 1.7 grades. The accuracy of prediction varies dramatically according to the A-level attainment of the student. Lower attaining applicants are far more likely to have their grades over-predicted, whilst high attaining students are more likely to be accurately predicted by their teachers.

In addition to final achievement, we find that the Socio-Economic Status (SES) of the student and the type of school attended are associated with accuracy. We find among students who are equally high achieving, low SES students receive predictions that are lower than those from high SES backgrounds, by around 0.059 grade points (where 1 point is equivalent to a full A-level at the lowest grade). Moreover, high achieving students from state schools also receive lower predictions than those from private schools.

We provide three potential mechanisms for the low SES students being systematically underpredicted; first, teacher bias; second, these students having higher than expected growth in achievement; or third that these students react more positively to predicted grades. Only the latter of these would provide a reason for keeping predicted grades.

Regardless of their origin, we go on to show the consequences of systematic underprediction. High achieving but under-predicted candidates are 10 percentage points less likely to apply to the most selective universities, and 6.9 percentage points more likely to enrol in a university in which they are over-qualified (have grades higher than their fellow peers). Since the university attended has been shown to be linked to a students' eventual earnings (Belfield et al, 2018), such 'undermatch' has potential implications for equity and social mobility.

Our findings support the emerging body of work relating to student undermatch (Campbell et al, 2019; Dillon & Smith, 2017; Hoxby and Avery, 2012), which finds that students from disadvantaged backgrounds are particularly likely to apply to lower ranked universities than their academic credentials would permit. We advance this literature by offering a potential explanation for this: that students (and subsequently universities) may be misinformed about their ability. Our study is also relevant to the emerging body of work on teacher manipulation. Diamond and Persson (2016) show that pupils in Sweden are much less likely to be marked just below a grade threshold, resulting in positive signalling effect in the labour market, whilst Papay et al (2015), find persistent effects of earning a more positive label on college going. We show that students whose teachers over-predict their grades are more likely to end up in more selective institutions.

Finally, our work is relevant to studies of student responses to grades as incentives, as our findings could be driven by differential reactions to predicted grades. Hvidman and Sievertsen (2019) study a reform which led to some students' grades being recoded and ultimately having their GPA being downgraded. They find that students who were downgraded by the recoding performed better on subsequent assessments, indicating the importance of incentives.

This paper proceeds as follows. In section 2 we outline the UK university applications system and background on the use of predicted grades. In section 3 we describe the data used in this analysis, and in Section 4, the methodological strategy. In Section 5 we present the main findings of the analysis. Section 6 concludes.

## 2. The UK university applications system

The UK higher education admissions process is controlled by a centralized body, the University and College Admissions Service (UCAS), which allows young people to apply to multiple courses through a single organisation. Students can apply to both universities and colleges that provide higher education through the UCAS system[2]. Whilst the centralised nature of the admissions system in the UK is clearly of benefit to all students, the process is highly complex, and requires young people to make potentially life-changing decisions far in advance of university entry, based on predicted rather than actual A-level exam grades.

Table 1 describes the current UK admissions process from beginning to end, including the key dates and deadlines to which students must adhere successfully to gain a place. As can be seen from the Table, there are a number of unique elements of the UK system.

As Table 1 illustrates, students must apply to university almost a year in advance of entry, and before they receive the exam results that will determine where they eventually end up. Students intending to go to university sit A-level exams at age 17-18, which, during the period we study, were graded from A* to E, with A* being the top grade. The most selective Russell Group universities typically require three A-levels at grades AAB, though there are many more universities that will accept students with far lower grades. It is also important to

---

note that there is a large degree of heterogeneity within institution, with some courses requiring higher grades than others.

However, instead of applying on the basis of their actual grades, students apply to university based on their predicted grades, which are supplied by their high school teachers for each A-level subject that they take. There is very limited guidance for teachers and careers advisors on the UCAS website. UCAS state that "a predicted grade is the grade of qualification an applicant's school or college believes they're likely to achieve in positive circumstances" (UCAS, 2020). UCAS also suggest that predicted grades should be "aspirational but achievable", but that 'inflated' predictions are "not without risk, and could significantly disadvantage [applicants]". How these predictions are made will very likely vary by school (and teacher), though teachers will have access to the previous exam results of each pupil, most relevantly, their age 16 GCSE results[3], and their age 17 AS-level exam scores[4].

Students may choose to apply to up to five courses (i.e. university-subject) on their UCAS application form. They must also include a personal statement with their application. University admissions tutors, at this point, assess the applications of students on the basis of their predicted grades and personal statement (only a very small number of universities/courses require an interview, mainly Oxbridge, but also medical programmes).

At the end of this selection process, universities must make students offers, conditional on achieving certain grades. Once these offers are received, the student has a short period to choose a first and backup course. This is an important decision: the student is committed to their first choice university course if they gain the required A-level grades specified by the university in their offer, and the university is committed to accepting them. If they do not meet

---

[3] The General Certificate of Secondary Education (GCSE) is a set of compulsory exams taken by students aged 15–16, after two years of study. Most students take between 5 and 12 subjects.
[4] During our period of analysis, the A-level was spread out over two years, with the AS level is taken in the first year, followed by the A2 level in the second year. AS levels could stand as a qualification on their own or could be carried on to A2 the next year to complete the full A level qualification. More recently, this system has been reformed so that results from AS levels no longer count towards the final A level grade.

the grades of their first choice, they are committed to their backup choice, and the backup choice university-course is committed to them, again assuming that they gain the grades required.

Only after this decision process is complete do students then sit their A-level exams, and only once they receive the results will they know which university-course they will attend. Should the student not achieve the required grades, they may enter a process known as 'clearing' in which courses which still have places available are advertised which students are then free to apply to. Around 10% of students gain a university place through clearing (UCAS, 2016).

During the period we study (2013-2015), the fraction of school leavers applying to university was 34-36%, with 89% of applicants eventually accepted at university (UCAS, 2015)5. Among enrolees, 92% are from state (i.e. publicly funded) high schools, whilst 8% are from independent (fee paying) schools – which is roughly representative of the school sector6. However, there are significant discrepancies in terms of the background of students attending university, with those from better-off backgrounds substantially more likely to enrol. Around 18% of students from the poorest quintile (as measured using Polar 3, a definition of socio-economic disadvantage, which is defined according to the university participation rate of the area the applicant is living in) entered university in England, in 2015, versus 45% of those from the richest quintile backgrounds (UCAS, 2015)[7]. However, it is widely understood that the most important reason for the disparity in participation rates among students from different

---

[5] Unfortunately, very little information is available on the precise reasons why students are not accepted at university. To our knowledge there is no available breakdown of applications versus acceptances by school type or Polar3. Even when a student is not accepted at university, they may still reapply in the following year. 8% of those who applied in 2015 were re-applicants (UCAS, 2015), though there is no information on background characteristics of re-applicants.

[6] 93 percent of students attend publicly funded secondary schools in England (Table 2A, DfE, 2010)

[7] The equivalent figures for Scotland, Northern Ireland and Wales are 10% v 36%, 15% v 46%, 16% v 44% respectively.

backgrounds is that those from poorer backgrounds lack the academic credentials to gain a university place (Chowdry et al, 2009).

Table 1 here

## 3. Data and descriptive statistics

### 3.1. Sample

This project uses bespoke data provided by the UK's University Colleges Admissions Service (UCAS), the centralised system used by the vast majority of students to apply to university. The sample being used throughout this paper is of all applicants/acceptances to higher education courses (including those taught in further education (FE) colleges[8] as well as universities) in the UK. We do not have data on the grade predictions of those students who did not apply for HE, such as those who chose to attend a vocational course or to go directly into the labour market. Nor does our study include students applying to UK colleges from the EU or overseas.

The variables included in the dataset are: total A-level points (with points score defined by UCAS as the points score attached to the *highest 3* A level grades achieved by the applicant, with the following points per grade used in the calculation: A*=6, A=5, B=4, C=3, D=2, E=1), the difference in achieved and predicted A-level score (with points score based on the same definition as previous – so that a score of +1 would suggest the students' achieved score was one point higher than their predicted score – the equivalent of going from an E to a D or a D to a C, etc), gender, ethnicity, most recent educational establishment type (which is generally

---

[8] Further education (FE) in the United Kingdom is typically undertaken after age 16, in order to achieve a range of vocational qualifications. In contrast students may choose to continue in secondary school to achieve more academic qualifications, such as A-Levels. FE colleges may also offer HE qualifications such as teaching qualifications, or degrees.

school type but includes other routes such as FE college), Polar3 quintile (a definition of socio-economic disadvantage, which is defined according to the university participation rate of the area the applicant is living in), academic cycle year (for 2013, 2014, 2015) and university tariff band (a measure of university selectivity based on the tariff score of accepted applicants, "higher", "medium" and "lower"[9]). The dataset also contains information on the clearing routes of applicants (firm choice, insurance choice, main clearing, direct clearing or adjustment). A further dataset from the same source contains information on the provider name, A-level points, predicted points again for applicants and acceptances.

The data are aggregate (for reasons of privacy), rather than individual level. We observe the number of unique applicants in each cell, where the cells are defined by the interaction of all the variables described above; A-level points, accuracy of predicted grades, gender, ethnicity, educational establishment type, disadvantage quintile, academic year, and university tariff band. As an example, one cell would contain the number of applicants who: i) have 18 A-level points ii) have a predicted grade score of 18, iii) are female, iv) are ethnic group white, v) who attended a state school, vi) who are from the least disadvantaged quintile, and vii) applied to university for academic cycle 2015[10]. Therefore, while we do not have individual level this data structure does allow for these factors to vary independently.

Starting with a population of around 1,356,000 pupils across our three cohorts, we initially restrict the sample to applicants who have information on gender, school type and disadvantaged status. This restriction reduces our sample to 1,355,130. We then remove all students who have predicted or actual achievement of less than 3 A-levels (because we cannot observe the predicted or actual points for such students). This latter restriction substantially reduces our sample to 858,720 unique applicants, over the three cohorts we study. We present

---

[9] Given that this measure of quality is based on universities' published tariff points, this measure of quality may be prone to error; it is well known that institutions accept students with grades below the advertised requirement (Times Higher Education, 2016).
[10] We apply frequency weights in all calculations

background information for the entire population, and our sample, in Table 2, also splitting our subsample by achieved A-level grades. Although the characteristics of students are fairly similar in our dataset versus our selected sample, restricting to only those with 3 or more A-levels results in our analysis sample being from slightly more advantaged backgrounds, and being more likely to be from an independent or grammar school, versus a state school. Variation in student achievement is also evident from Table 2. In particular, high achieving students disproportionately come from independent (i.e. fee paying) and grammar (i.e. academically selective) schools, as well as the least disadvantaged (richest) quintile.

Table 2 here

## 3.2. Descriptive statistics

We begin by examining the accuracy of predicted grades among the population of university applicants. Figure 1 shows the distribution of the difference between actual and predicted grades for applicants between 2013-2015, where each point on the x-axis represents the achieved point score of the students *minus* the predicted point score of the students. In other words, a *negative* score means that the student achieved a *lower* grade than they were predicted to achieve (the student's grade was over-predicted), whilst a *positive* score means the student achieved a *higher* grade than they were predicted to achieve (the student's grade was under-predicted).

Figure 1 shows that, based on this particular measure[11], the majority of students are over-predicted. In Table 3 and Figure 2, we explore this in more detail. Table 3 shows that

---

[11] It is important to note that this measure of under-over prediction is rather blunt. The measure is based on the total points achieved from the applicants' best 3 A-levels. This means that each teacher in each subject would have to correctly predict the applicants' grade for the total points score to be accurately predicted. There is also room for error in the measure: a student could be predicted to achieve BBB=12 points, and actually achieve AAD=12 points, and appear to be accurately predicted according to this measure.

accuracy is around 16% - i.e. 16% of applicants' grades were correctly predicted. A further 8.5% of applicants were under-predicted, whilst the vast majority (75%) of applicants' grades were over-predicted. This suggests a lower level of accuracy than previous research (Everett & Papageorgiou, 2011) which found that 41% of grades were over-predicted, 51.7% were accurate, and only 6.6% were under-predicted, however the Everett & Papagerogiou research uses a different measure of grade accuracy (based on the actual grades of each individual A-level, rather than based points assigned to the best 3 as in this paper), therefore is not directly comparable with the results presented in this paper. Nevertheless it is notable that in both theirs and this research, under-prediction is uncommon.

As Table 3 also shows, the majority of predicted grades were within 1-2 points of the actual grade achieved (where, as explained, 1 point equates to 1 grade, i.e. the difference between AAA=15 points and AAB=14 points, or DDD and DDE). Nevertheless, 6% of applicants were over-predicted to the tune of 5 points – equivalent to an extra A-level at grade A.

Figure 1 here

Table 3 here

Does grade accuracy vary according to the achievement levels (as measured by actual A-level attainment) of the student? Figure 2 provides the answer to this. Achieved grades are on the x-axis, while predicted grades are on the y-axis. If all students received accurate predictions, every point would be on the 45 degree line. Points above the line show the situation where predicted grades are higher than the grades actually achieved (overpredicted), while points below the line show the situation where predicted grades are lower than the grades students go on to achieve. As the chart clearly shows, the vast majority of points lie above the 45 degree line – i.e. most students are overpredicted. However, it is notable from Figure 2 that over-

prediction particularly occurs at the low end of the A-level attainment distribution, with those students achieving between 3-6 A-level points being over-predicted by up to A-level 4 points (the equivalent of a B grade). It is also notable that as students' A-level attainment increases, so does the accuracy of their predicted grades. Only particularly high achievers tend to be accurately predicted, or under-predicted. Indeed, only students with A-level points at 17 or above (A*A*A) appear to be under-predicted (though note, of course, that there are ceiling effects at play here; straight A* students cant, by definition, be over-predicted. Note also that this chart simply presents the average prediction for each A-level point score – there will still be a distribution of under/over predictions for each point score).

Whilst this research cannot uncover the causes of this inaccuracy, a likely large factor that is driving the results is the presence of ceiling and floor effects for predicted grades. For students that ultimately achieve the lowest points possible it is only feasible for their predicted grades to be correct or over-predicted, meaning that on average these students will be over predicted. Similarly, for the top achieving students, it would be impossible to over-predict their grades, and so on average we would expect under-prediction for high achieving students.

Regardless of the extent of these celling and floor effects, the focus of this paper is the differential predicted grades by student's pre-existing characteristics, which should be unaffected by such mechanical effects. We estimate the gaps in prediction accuracy between students from different backgrounds with the same realised level of achievement; the existence of ceiling and floor effects should not differentially impact students by SES.

Alternatively, there are reasons for the inaccuracies that could be related to the SES of the student. First, student SES directly impacts teachers' expectations of future achievement. Second, any under prediction could be a result of students being more likely to over-perform than expected. Third, students may react to the predicted grades as goals and so act as

incentives. Each of these factors could differ by SES and so are potential mechanisms, we will discuss the implications of each in the conclusions.

Figure 2 here

We can descriptively examine how accurate grades are according to the characteristics of students and schools, examining whether certain types of students, or certain school types are particularly likely to under or over-predict. This is presented in Figures 3-4.

As figure 3 shows, there is a good deal of variation in prediction accuracy according to school type. Independent schools appear to be the most accurate predictors – over 20% of applicants from independent schools' grades were accurately predicted. Meanwhile, academies, state schools and sixth form colleges are more prone to over-predicting their students' grades (as seen by the greater proportions from these school types with a difference between actual and predicted grades below zero). As is evident, there is a limited amount of under-prediction going on; however this will be explored in more detail in Section 3.

*Figure 3 here*

Are applicants from disadvantaged backgrounds more likely to receive inaccurate predictions? The answer can be seen in Figure 4 which shows prediction accuracy according to the quintiles of disadvantage. As Figure 4 shows, that those applicants from the most disadvantaged backgrounds are the least likely applicants to have their grades accurately predicted. They are also more likely to have their grades over-predicted than those from more advantaged groups. Indeed, the gradient in accuracy by advantage level is quite striking in this chart.

For example, applicants from the most disadvantaged backgrounds are more likely to experience moderate to severe over-prediction (i.e. from 2 to 5 points over-predicted) than those the most advantaged backgrounds. 7.5% of the most disadvantaged applicants were over-predicted by 5 points (equivalent to an A grade at A-level), compared with 5% of the most

advantaged students. But the gradient reverses as accuracy improves, so that the most advantaged applicants are more likely than the most disadvantaged to experience relative accuracy (i.e. a difference in achieved and predicted between -1 to +1) in their predicted grades. Interestingly, however, for applicants whose grades are under-predicted by 2 points or more, there appear to be no differences by level of disadvantage.

Figure 4 here

Figure 5 here


Finally, in Figure 5, we examine predicted grade accuracy by ethnicity. Here we can see that, among all ethnic groups, Asian and Black applicants are more likely to be severely over-predicted. White applicants are most likely of all groups to be accurate (though at similar levels to black students) and are also most likely to be slightly over-predicted. Finally, we can also examine whether girls are more likely to have their grades over-predicted than boys. The results (not shown here) reveal no apparent difference in their rates of grade accuracy.

The descriptive statistics in this section showed that those students whose grades are most likely to be accurate are those attending independent or grammar schools, and those from more advantaged backgrounds. They are also more likely to be white students. Meanwhile, ethnic minority applicants, and applicants from state schools and disadvantaged backgrounds are more likely to be over-predicted.

However, it is well documented that disadvantaged students (who are also more likely to attend state schools) tend to be lower attaining than their more advantaged counterparts (e.g. Chowdry et al, 2013), therefore the finding that disadvantaged and state students are more likely to be over-predicted could simply be driven by their increased probability of being low attaining. To account for these confounding factors, we explore these results further in a regression framework.

## 4. Methodology

We are interested in which students are the most likely to have their grades incorrectly predicted and how prediction accuracy correlates with college applications and acceptances.

First we use simple multivariate regression models to examine the likeliness of certain types of students to be under/over predicted, as follows:

$$D_{gt} = \beta X_{gt} + t_t + u_{gt}$$

Where *D* is an integer variable representing the difference between actual and predicted grade points for students in group g (where data are grouped by A-level grade points, predicted grade points, gender, ethnicity, school type, disadvantage quintile of the applicant – see Section 3 for details), at time t (for cohorts of students in 2013, 2014 and 2015) taking values between -5 and +5. A value of +5 would mean that a student scored 5 A-level grades higher than they were predicted (i.e. they were under-predicted), and a value of -5 means that a student achieved 5 grades lower than they were predicted (over-predicted). Thus, the variable represents the degree to which the student is underpredicted. X is a set of characteristics of the student comprising gender, ethnicity, school type and disadvantage quintile, and u is the group error term. All regressions include a set of cohort dummies t. All data is weighted by the number of students in each cell.

To examine the impact of over/under prediction on university applications and acceptances, we run a series of regressions of the type:

$$U_{gt} = \gamma D_{gt} + \delta X_{gt} + t_t + \varepsilon_{gt}$$

Where U represents university type (e.g. high tariff, Russell Group, etc) applied to, or accepted at. Our key explanatory variable is the prediction accuracy of the student *D* as above, and again

we add controls and cohort dummies. In each case we run 3 separate regressions by student achievement, grouping students according to their A-level points score, where low achievement is defined as scoring CCC or below, medium achievement=CCC to ABB, and high achievement=AAB or higher. This is to account for any floor or celling effects that may occur at the extremes of the distribution.

## 5. Results

In this section, we explore students' propensity to be under/over predicted in a regression framework, allowing us to hold constant other factors such as ethnicity, gender or school attended, which may also affect prediction accuracy.

### 5.1. Which students are under or over-predicted?

In Table 4 we examine the difference between students' predictions and their actual grades according to the disadvantage level of the student[12]. Column 1 contains the raw results for the impact of disadvantage level on the difference between actual and predicted grades (as described previously, the outcome variable is a continuous variable representing the difference between actual and predicted grade points), controlling only for academic year. As expected, those from the poorest quintile are less under-predicted (by around 3.3 grade points) than those from the richest quintile (the reference category).

In the second column we add student background characteristics, and in the third column, school type. There most advantaged coefficient is reduced with the addition of schooltype, and there is an increase in the R-squared, indicating that background characteristics and school type do explain some of the variation in prediction accuracy. However, we have

---

[12] In Appendix A, we present regressions with an a binary dependent variable, showing the probability that students from different backgrounds or schools are underpredicted. These results show similar patterns, whereby high achieving students from low SES backgrounds and state schools have a greater probability of being underpredicted than those from high SES backgrounds and independent schools.

seen that the probability of being under or over predicted varies dramatically by student A-level achievement level. Therefore in the next 3 columns we attempt to account for this by sub-dividing the analysis by achieved A-level grades (CCC or less, CCC-AAB, and finally, AAB or more). Looking first at column 3, we can see that, after controlling for school type and student background characteristics, low achieving, disadvantaged students are less under-predicted than low achieving but advantaged students. However, column 4 shows that, for 'medium-attainment' (defined as achieving grades above CCC but lower than AAB) students there is little difference in under-prediction according to disadvantage. Finally, column 5 shows that high achieving (AAB or more)'s grades are significantly more under-predicted than high achieving students from the most advantaged backgrounds, even after controlling for school type, gender, ethnicity and year. The coefficient implies that high achieving students coming from the most disadvantaged compared to the least disadvantaged backgrounds have predicted grades that are 0.059 grade points lower than their more advantaged counterparts (where 1 grade point is the equivalent of a full A-level qualification, at the lowest grade, grade E). This is equivalent to 0.05 of a standard deviation. This is a significant effect and implies that some of our most high achieving students may be misinformed about their likely potential, and that this could prevent them from applying to more selective institutions. In Appendix 1 we also display results where our dependent variable is a binary variable equal to 1 for under-predicted and 0 otherwise. These results show that those high achieving students from the most disadvantaged backgrounds are more likely to be underpredicted than those from the least advantaged backgrounds.

Table 5 now shows similar results for those attending state schools versus independent and selective schools. After controlling for ability and characteristics, those attending state schools receive predictions that are less generous those from independent schools and grammar schools (Columns 3-5). In Appendix 2 we again present results where our dependent variable

is a binary variable equal to 1 for under-predicted and 0 otherwise. These results show that those high achieving students from state schools are more likely to be underpredicted than those from private schools.

If being under-predicted results in students applying to universities to which they are undermatched (i.e. universities where their peers are less qualified than they are), this could have a direct impact on social mobility, potentially skewing social representation across HE institutions. Moreover, it could result in students becoming de-motivated and dropping out of university altogether. On the other hand, students may not pay a great deal of attention to their grade point predictions, therefore the degree of accuracy may not actually affect students' application decisions at all. How grade accuracy impacts students' university choices is the question that we turn to in the next section.

Table 4 here

Table 5 here


**5.2. What is the impact of the predicted grades system on students' university choices?**

We can now turn to the important question of what is the correlation between predicted grades – or rather inaccuracy of predicted grades – on students' university choices. In particular, the question we aim to ask here is whether students whose grades are under-predicted are more likely to apply to universities where they are undermatched, based on their academic credentials.

First, we examine the probability of applying to a high-tariff university based on accuracy of A-level point prediction. In other words, are applicants whose grades are under predicted less likely to apply to a high-tariff university? The definition of a high-tariff university is again somewhat restrictive (see data section, and there are only 3 categories of university, therefore high-tariff universities represent the top third. The results can be found in

Table 6 and again are split by A-level entry points achieved. Note that here we now use a categorical variable for prediction accuracy, to simplify the interpretation.

The results indicate that applicants whose grades are under-predicted are less likely to apply to a high-tariff university, to the tune of around 10 percentage points for the highest achieving students. This is certainly of concern, and is evidence that grade prediction may be reducing the potential of these students.


Table 6 here

This raises the question, do students with inaccurate grade predictions apply to universities that are not appropriate for them, based on their realised A-level points? And importantly, do accepted applicants whose A-level scores were under-predicted, end up at universities in which they are over-qualified (undermatched)?

To answer this question, we turn to a second dataset (again described in Section 3) which contains A-level points and predicted grades at every UCAS university. In order to perform this analysis we first calculate the average A-level point score at every university in our sample (by year, based on the A-level point scores of acceptances). We then define a student as over-qualified for the university they attend to if their own A-level achieved score exceeds the average for that university. Note, that this measure of over-qualification is, of course, subject to error. In particular, whilst some students may have A-level scores well above or below the mean for their university this could simply reflect the particular course that they are studying, since even within universities, courses have different entry requirements. As we do not have information on the course of study, we are unable to take this into account, and therefore these results can be seen as merely suggestive evidence of under/over-qualification. Note also that for this analysis our set of variables is far more restrictive than above – in particular we are not able to control for any student characteristics, or school type. As with the

other results in this paper, we must interpret these findings with caution, since they represent correlations between the explanatory and outcome variables, rather than causal relationships[13].

The results of this analysis can be found in Table 7. Here, we present two sets of results. In columns 1 and 2 we present analysis for unique applicants, in columns 3 and 4 we present results for acceptances at university. In column 1 we can see that under-predicted applicants are 10 percentage points more likely than applicants whose grades were accurate or over-predicted to have applied to a university that they are over-qualified for, based on this definition. In column 2 we can see that the most able applicants who under-predict are also more likely to apply to universities that they are over-qualified for. Column 3 now shows that among those accepted to university, having under-predicted grades increases the likelihood of being over-qualified for university (though perhaps unsurprisingly, the probability drops to just 1.5 percentage points). It is important to note here that being over-predicted actually increases the likelihood of being under-qualified for the university attended (not shown). Finally in column 4 we can see that, among the most able students this is also true, and indeed being under-predicted increases the chances of attending a university which one is over-qualified for by 6.9 percentage points). Finally it is of interest to note that among accepted applicants from the most highly able groups (Column 4), the most disadvantaged students are most likely to be over-qualified for their university. This latter finding implies that the predicted grades system may be skewing the socio-economic mix of students at certain universities; if high-ability in section 3), and these students in turn apply to universities that they are over-qualified for, then this could result in fewer disadvantaged students at top universities.

Table 7 here

---

[13] As a robustness test we replicate the results in Table 6 in the dataset without controls, defining high tariff as Russell Group institutions, and medium and low tariff as all other institutions. We find similar results to those in Table 6, suggesting the lack of control variables may not seriously impact these coefficients.

## 6. Conclusions and discussion

Even among the selective group of students who attend university, students from lower socio-economic backgrounds are less likely to attend a high-status institution than their richer counterparts. In this paper we focus on one potential explanation; that disadvantaged students are given lower predicted grades and therefore apply to less selective courses than advantaged students, conditional on A-level attainment.

Our results show that the predicted grades students in the UK receive – on which students make their application decisions – are quite inaccurate. Only 16% of applicants' predicted grades correctly predicted when the best 3 A-levels are considered. The vast majority of applicants (75%) receive predictions that are too optimistic, to the extent that the average student is overpredicted by 1.7 grades. There are a few potential explanations for this. First, this may be as a result of teachers "target setting", attempting to encourage students to work harder, and is somewhat akin to the evidence on teacher manipulation (Diamond & Persson, 2016; Papay et al, 2015). Second, they may wish to maximise young people's chances of getting into the best possible university, perhaps as a result of parental pressure, or simply to try and improve their pupils' life chances. Third, they may simply be aware that general practise is to overestimate pupil grades, so they are attempting to "level the playing field". Finally, they may have over-confidence in their own abilities as a teacher and may therefore believe their pupils will be higher performing in exams than they actually go on to be.

There are also small socio-economic gaps in predicted grades. High achieving (AAB+) disadvantaged students receive predicted grades that are 0.059 grade points lower than their more advantaged counterparts. State school students also receive less generous predictions than their private school counterparts, to the tune of 0.16 grade points. Grammar school students receive predicted grades that are 0.067 points lower than their private school counterparts, comprehensive students by 0.17 points and FE students by 0.2 points.

While we are unable to explain the mechanisms behind these prediction accuracies, it is useful to consider potential explanations which could lead to the results we observe here, and the resulting policy recommendations. We can consider both players in the education market – teachers and students – and their likely actions.

The first potential explanation of grade inaccuracy could be teacher bias; teachers could simply under-estimate the ability, and hence future attainment, of students from disadvantaged backgrounds. In this case, one potential policy solution would be to better educate teachers about the achievement of low SES students. An alternative, and probably more effective solution would be to simply abolish the use of predicted A-level grades and base university application on actual results (as is the case in all other systems around the world, that the authors are aware of).

A second potential explanation is that those from low SES backgrounds may experience faster growth in ability during the time between receiving their grade predictions and sitting their exams, compared to high SES students. Teachers may provide accurate predictions for a given level of achievement to date, but may not anticipate late growth surges in student attainment among low SES students. This would also result in high attaining, low SES students being under-predicted. Here, useful policy solutions would be to either provide teachers with better information of the likely growth in student achievement between predictions and exams, or again, to simply reform the predicted grade system.

Finally, the patterns we observe in the data could arise as a result of heterogenous incentive effects (e.g. Hvidman and Sievertsen, 2020). High ability low SES children may be more responsive to predicted grades due to incentives/goal setting compared to high ability high SES children. This could be due to high SES students being over-confident. In this case, low-SES students at the top of the achievement distribution would be more likely to be under-

predicted. Abolishing predicted grades in this scenario may actually be damaging, because it might remove the incentive from low SES students to work hard.

Regardless of the specific mechanism it is important to note that predictions themselves are high-stakes; students may be dissuaded from applying to selective courses – or may not be made offers from those they do apply to – on account of less generous predictions, regardless of their eventual attainment. This will have consequences for the socio-economic mix at universities, and for equality and social mobility, given the link between university attended and labour market outcomes (Crawford et al, 2016).

A further concerning implication of the predicted grades system is that it may have repercussions for the efficiency of the university application system. If school teachers continue to inflate the majority of students' grades, universities may adjust their admission policies to reflect this, taking in students whose grades are even higher than advertised. This will only further reduce the value of information on students' potential ability, and may result in universities abandoning the system of predicted grades and introducing their own entry exams.

Figures and tables

*Figure 1: Distribution of the difference between actual and predicted grades*
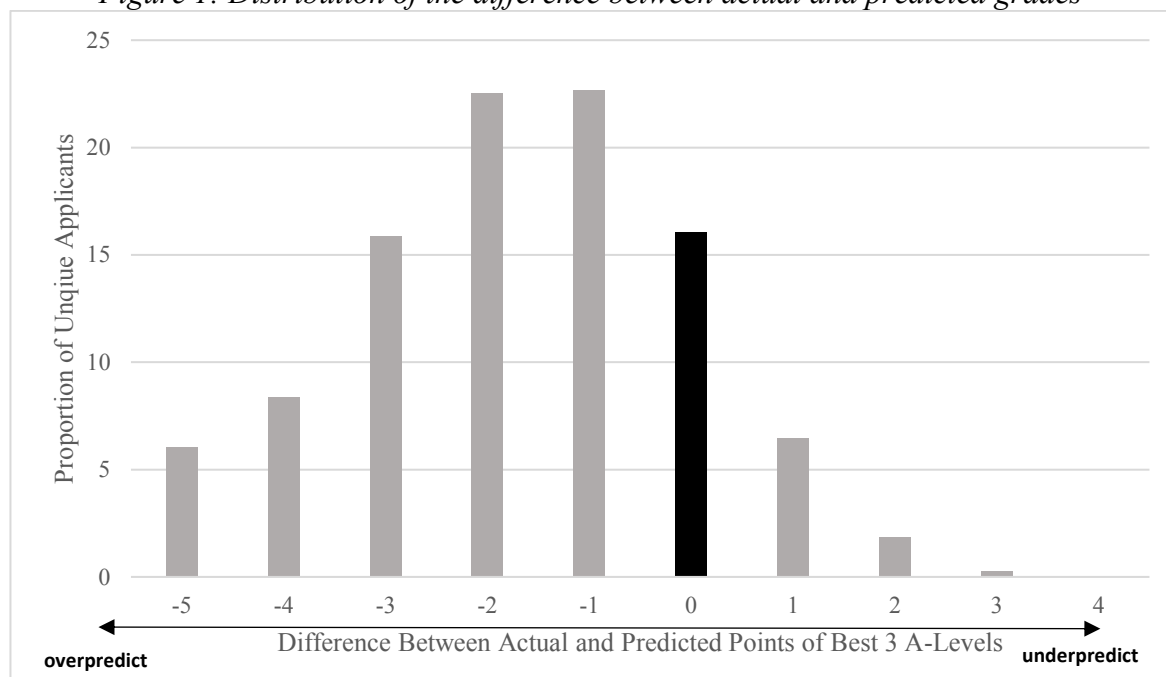


*Notes 1: each point on the x-axis represents the achieved point score of the applicant minus the predicted point score of the applicant*
*Notes 2: Points score is defined by UCAS as the points score attached to the highest 3 A level grades achieved by the applicant, with the following points per grade used in the calculation: A\*=6, A=5, B=4, C=3, D=2, E=1*

*Figure 2: Accuracy of predicted grades, by applicants achieved A-level points score*



*Notes: Mean predicted score is the mean predicted score of all applications for each A-Level points achieved. Points score is defined by UCAS as the points score attached to the highest 3 A level grades achieved by the applicant, with the following points per grade used in the calculation: A\*=6, A=5, B=4, C=3, D=2, E=1. The dashed 45degree line represents students A-Level points being accurately predicted.*
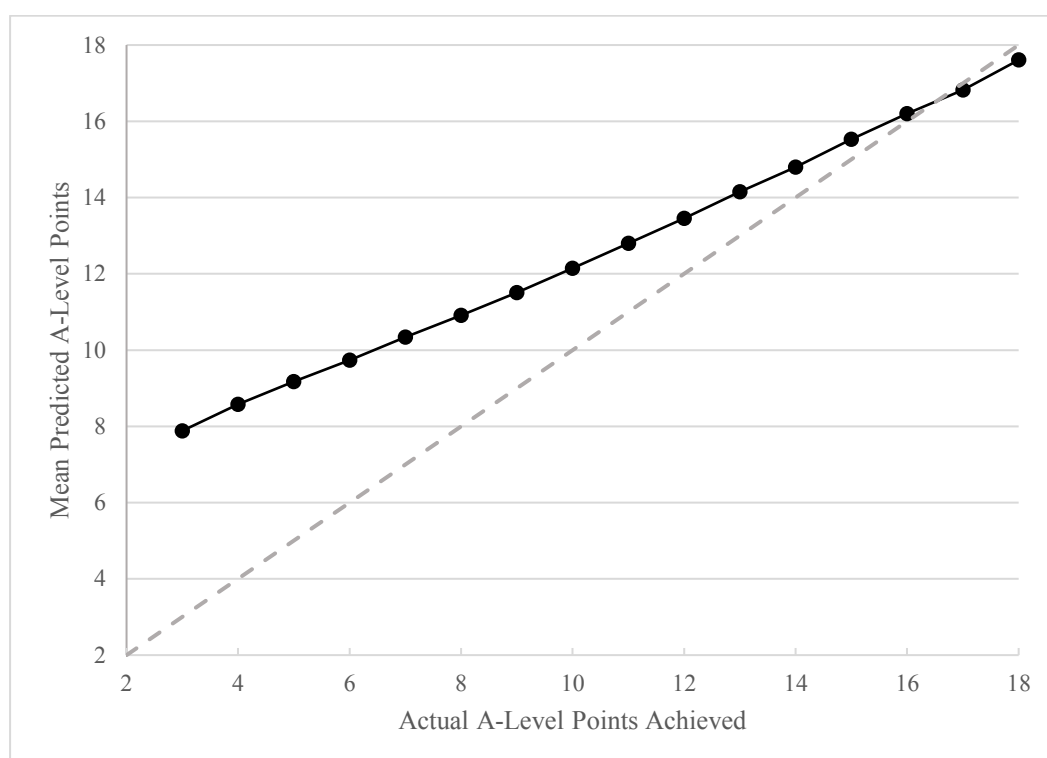
*Figure 3: Difference between actual and predicted grades, by school type*



*Notes 3: each point on the x-axis represents the achieved point score of the applicant minus the predicted point score of the applicant*
*Notes 2: Points score is defined by UCAS as the points score attached to the highest 3 A level grades achieved by the applicant, with the following points per grade used in the calculation: A\*=6, A=5, B=4, C=3, D=2, E=1*

*Figure 4: Difference between actual and predicted grades, by level of disadvantage*
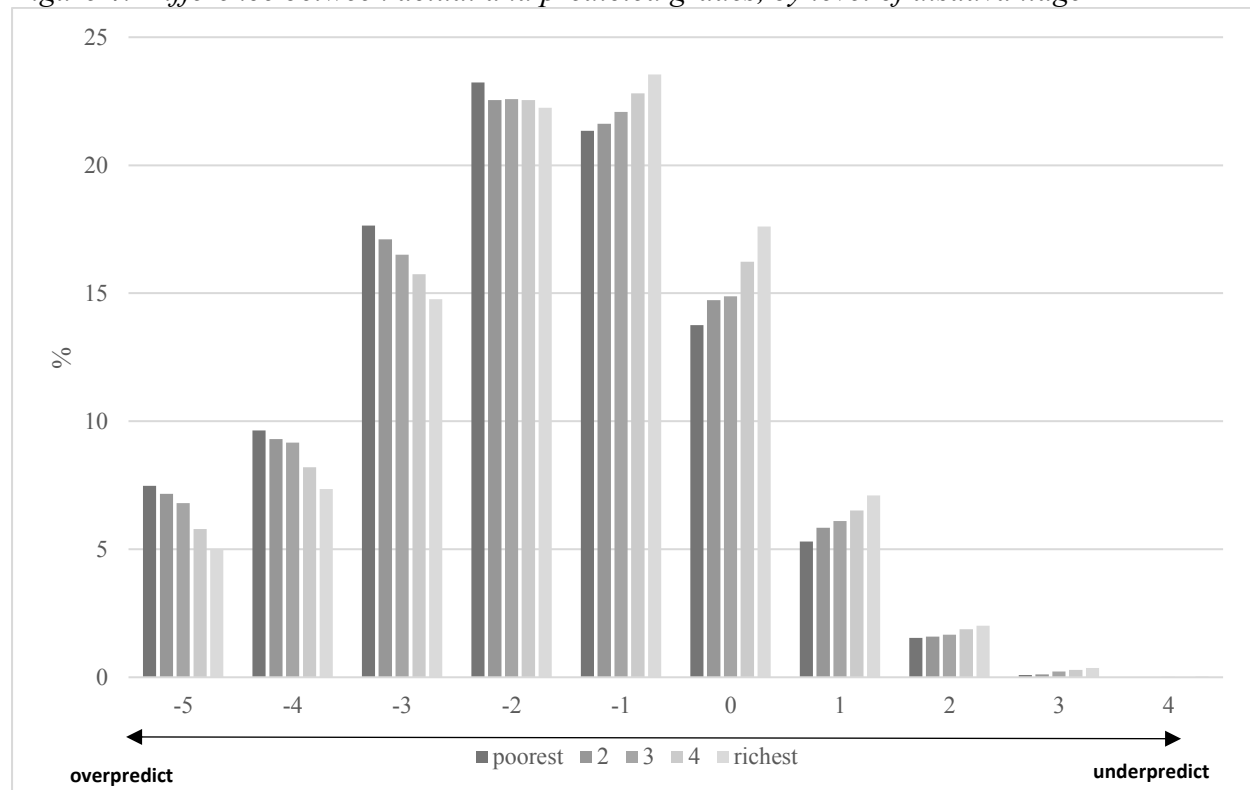


*Notes 4: each point on the x-axis represents the achieved point score of the applicant minus the predicted point score of the applicant*
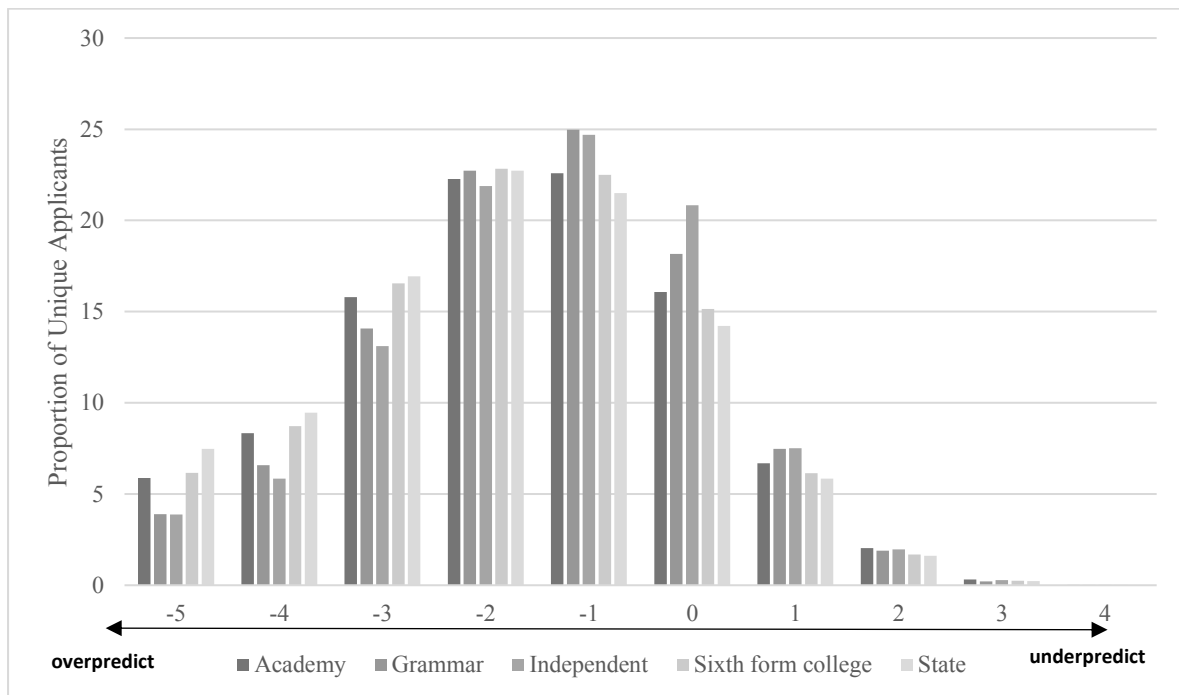*Notes 2: Points score is defined by UCAS as the points score attached to the highest 3 A level grades achieved by the applicant, with the following points per grade used in the  calculation: A\*=6, A=5, B=4, C=3, D=2, E=1*

*Figure 5: Difference between actual and predicted grades, by ethnic group*



*Notes 5: each point on the x-axis represents the achieved point score of the applicant minus the predicted point score of the applicant*
*Notes 2: Points score is defined by UCAS as the points score attached to the highest 3 A level grades achieved by the applicant, with the following points per grade used in the calculation: A\*=6, A=5, B=4, C=3, D=2, E=1*
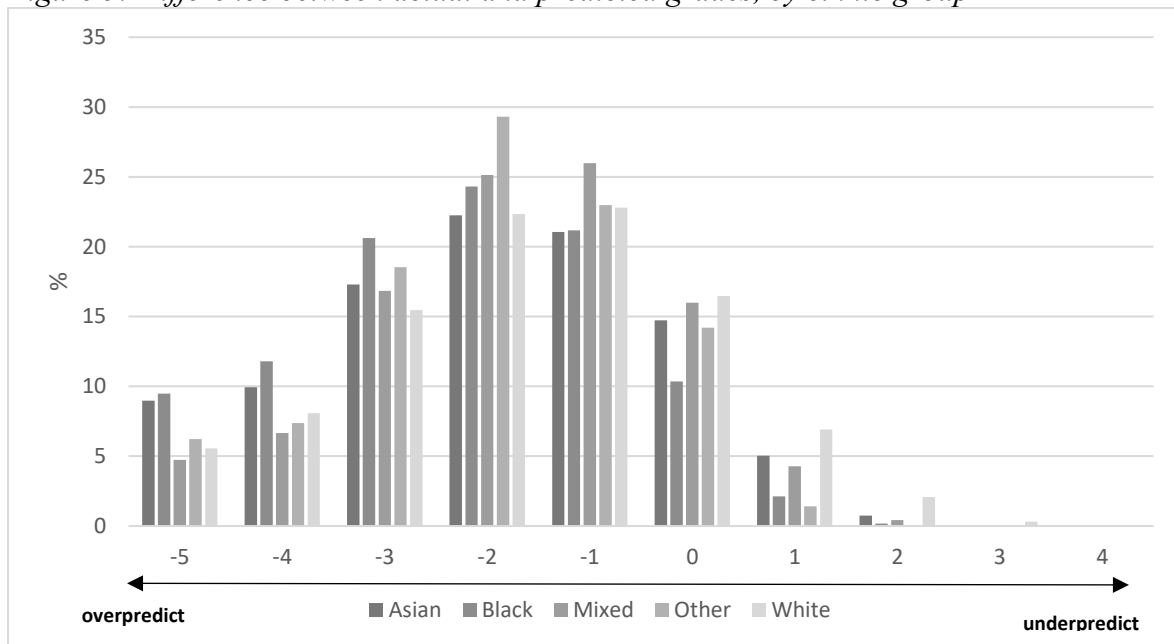*Figure 6: Difference between actual and predicted grades, by ethnic group*

*Table 1: The UK Admissions Process*

| School Year | Term | Age | Event | Deadlines |
|---|---|---|---|---|
| 12 | Summer | 16-17 | All universities will hold Open Days in June and in September for people about to apply through UCAS. Most of those attending will be Year 12 students who are about to complete their UCAS applications. | |
| 13 | Autumn | 17-18 | Students receive predicted A-level grades from schools. | |
| 13 | Autumn | 17-18 | Students apply for up to five courses (no order of preference) based on their predicted grades. They must also include a personal statement with their application. Again, open days happen here - often after students have applied to university. | January application deadline for entry that year.<br><br>September deadline for Oxbridge courses. |
| 13 | Spring | 17-18 | Students receive offers conditional on academic success. They must choose a first choice and insurance course. Students with no offers may enter "Extra", allowing another choice. | Decisions from universities received in May. June deadline to commit to choices.<br><br>"Extra" opens between the 25$^{th}$ of February and 4$^{th}$ of July. |
| 13 | Summer | 17-18 | Students sit A-level exams which will determine entry. | |
| 13 | Summer | 17-18 | A-level results are published. Both student and university are <u>committed</u> to first/insurance choice conditional on results. Students who missed their predicted grades may still be accepted. Students should apply for finance. | August: A-level results day. |
| 13 | Summer | 17-18 | Those without a place go into clearing; those who exceeded their target grades may go into adjustment. | Clearing opens early July. |
| | Autumn | 18-19 | Attend course of choice. | |

*Table 2: Summary statistics*

| | Unique Applicants | Analysis sample | <CCC | CCC-AAB | AAB+ |
|---|---|---|---|---|---|
| *All* | *1,356,055* | 858,720 | 146,890 | 480,610 | 231,220 |
| *Gender* | | | | | |
| male | 0.43 | 0.42 | 0.47 | 0.4 | 0.42 |
| | (0.49) | (0.49) | (0.5) | (0.49) | (0.49) |
| female | 0.57 | 0.58 | 0.53 | 0.6 | 0.58 |
| | (0.49) | (0.49) | (0.5) | (0.49) | (0.49) |
| *Ethnicity* | | | | | |
| white | 0.81 | 0.83 | 0.8 | 0.82 | 0.85 |
| | (0.39) | (0.38) | (0.4) | (0.38) | (0.36) |
| non-white | 0.19 | 0.17 | 0.2 | 0.18 | 0.15 |
| | (0.39) | (0.38) | (0.4) | (0.38) | (0.36) |
| *School Type* | | | | | |
| State | 0.25 | 0.21 | 0.26 | 0.22 | 0.15 |
| | (0.43) | (0.4) | (0.44) | (0.41) | (0.36) |
| Academy | 0.3 | 0.37 | 0.37 | 0.37 | 0.36 |
| | (0.46) | (0.48) | (0.48) | (0.48) | (0.48) |
| Grammar | 0.05 | 0.07 | 0.04 | 0.07 | 0.09 |
| | (0.21) | (0.25) | (0.19) | (0.25) | (0.29) |
| Independent | 0.09 | 0.11 | 0.05 | 0.09 | 0.2 |
| | (0.28) | (0.32) | (0.21) | (0.29) | (0.4) |
| 6th Form College | 0.2 | 0.21 | 0.23 | 0.22 | 0.17 |
| | (0.4) | (0.41) | (0.42) | (0.41) | (0.38) |
| Further Education | 0.09 | 0.04 | 0.05 | 0.04 | 0.02 |
| | (0.29) | (0.19) | (0.22) | (0.2) | (0.15) |
| *Disadvantage quintile* | | | | | |
| Q1 (poorest) | 0.1 | 0.08 | 0.11 | 0.09 | 0.05 |
| | (0.3) | (0.27) | (0.31) | (0.28) | (0.22) |
| Q5 (richest) | 0.31 | 0.35 | 0.27 | 0.33 | 0.44 |
| | (0.46) | (0.48) | (0.44) | (0.47) | (0.5) |
| Diff. actual v predicted | -1.7 | -1.7 | -.3.3 | -1.79 | -0.4 |
| | (1.65) | (1.65) | (1.4) | (1.4) | (1.22) |

*Source: UCAS*

*Table 3: under/over prediction by A-level points*

| Accuracy of Prediction | % |
|---|---|
| -5 | 6.03 |
| -4 | 8.35 |
| -3 | 15.88 |
| -2 | 22.5 |
| -1 | 22.65 |
| 0 | 16.05 |
| 1 | 6.45 |
| 2 | 1.82 |
| 3 | 0.26 |
| 4 | 0.01 |
| Total | 100 |

*Notes 6:Accuracy of prediction is the difference between actual and predicted points. Each point represents the achieved point score of the applicant minus the predicted point score of the applicant*

*Notes 7: Points score is defined by UCAS as the points score attached to the highest 3 A level grades achieved by the applicant, with the following points per grade used in the calculation: A\*=6, A=5, B=4, C=3, D=2, E=1*

*Table 4: Extent of Underprediction (Actual-Predicted grade), by level of disadvantage*

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Sample | All | All | All | CCC | CCC-AAB | AAB+ |
| Quintile of Disadvantage (ref=richest) | | | | | | |
| 1 (Poorest) | -0.334*** | -0.346*** | -0.271*** | -0.034** | 0.016** | 0.059*** |
| | (0.007) | (0.007) | (0.007) | (0.013) | (0.008) | (0.012) |
| 2 | -0.271*** | -0.265*** | -0.201*** | -0.024** | 0.014** | 0.037*** |
| | (0.006) | (0.006) | (0.006) | (0.011) | (0.007) | (0.009) |
| 3 | -0.228*** | -0.202*** | -0.154*** | -0.023** | 0.002 | 0.044*** |
| | (0.005) | (0.005) | (0.005) | (0.011) | (0.006) | (0.008) |
| 4 | -0.113*** | -0.112*** | -0.083*** | 0.004 | 0.011** | 0.008 |
| | (0.005) | (0.005) | (0.005) | (0.010) | (0.005) | (0.006) |
| | | | | | | |
| $R^2$ | 0.007 | 0.015 | 0.021 | 0.015 | 0.008 | 0.014 |
| Observations | 858,720 | 858,720 | 858,720 | 146,890 | 480,610 | 231,220 |
| Year effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Characteristics | | ✓ | ✓ | ✓ | ✓ | ✓ |
| School Type | | | ✓ | ✓ | ✓ | ✓ |

*Notes: source, UCAS*
*Dependent variable takes values between -5 and +5. A value of +5 would mean that a student scored 5 A-level grades higher than they were predicted (i.e. they were under-predicted), and a value of -5 means that a student achieved 5 grades lower than they were predicted (over-predicted).*
*Characteristics in column 2 are gender, ethnicity. Columns 1-3 use the full sample. Columns 4-6 use subsamples according to A-level achievement of students. Standard errors in parentheses. \* p < 0.1. \*\* p < 0.05. \*\*\* p < 0.01*

*Table 5: Extent of Underprediction (Actual-Predicted grade), by school type*

| Sample | (1) All | (2) All | (3) All | (4) CCC | (5) CCC-AAB | (6) AAB+ |
|---|---|---|---|---|---|---|
| School Type (Ref= Private School) | | | | | | |
| Grammar School | -0.084*** | -0.110*** | -0.084*** | 0.059** | 0.110*** | 0.067*** |
| | (0.009) | (0.009) | (0.009) | (0.025) | (0.010) | (0.010) |
| Academy | -0.273*** | -0.244*** | -0.197*** | 0.345*** | 0.199*** | 0.139*** |
| | (0.006) | (0.006) | (0.006) | (0.018) | (0.008) | (0.007) |
| 6th Form College | -0.352*** | -0.320*** | -0.251*** | 0.387*** | 0.200*** | 0.128*** |
| | (0.007) | (0.007) | (0.007) | (0.018) | (0.008) | (0.009) |
| State | -0.449*** | -0.417*** | -0.362*** | 0.283*** | 0.166*** | 0.168*** |
| | (0.007) | (0.007) | (0.007) | (0.018) | (0.008) | (0.009) |
| Further Education College | -0.633*** | -0.666*** | -0.578*** | 0.004 | -0.013 | 0.198*** |
| | (0.010) | (0.010) | (0.011) | (0.023) | (0.012) | (0.017) |
| Other | -0.615*** | -0.670*** | -0.605*** | 0.186*** | 0.072** | 0.033 |
| | (0.031) | (0.031) | (0.031) | (0.060) | (0.033) | (0.068) |
| | | | | | | |
| $R^2$ | 0.007 | 0.018 | 0.021 | 0.015 | 0.008 | 0.014 |
| Observations | 858,720 | 858,835 | 858,720 | 146,890 | 480,610 | 231,220 |
| Year effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Student characteristics | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Polar | | | ✓ | ✓ | ✓ | ✓ |

*Notes: source, UCAS*

*Dependent variable takes values between -5 and +5. A value of +5 would mean that a student scored 5 A-level grades higher than they were predicted (i.e. they were under-predicted), and a value of -5 means that a student achieved 5 grades lower than they were predicted (over-predicted).*

*Characteristics in column 2 are gender, ethnicity. Columns 4-6 refer to A-level achievement of students. Standard errors in parentheses. * p < 0.1. ** p < 0.05. *** p < 0.01*

*Table 6: Probability of applying to a high-tier university*

|  | (1) CCC | (2) CCC-AAB | (3) AAB |
|---|---|---|---|
| Sample |  |  |  |
| Under-predicted | -0.155*** | -0.143*** | -0.109*** |
|  | (0.016) | (0.003) | (0.002) |
|  |  |  |  |
| $R^2$ | 0.006 | 0.017 | 0.045 |
| Observations | 146,890 | 480,610 | 231,220 |
| Year effects | ✓ | ✓ | ✓ |
| Student characteristics | ✓ | ✓ | ✓ |
| School type | ✓ | ✓ | ✓ |
| Polar | ✓ | ✓ | ✓ |

*Notes: source, UCAS*
*Student characteristics are gender, ethnicity*
*Columns 1-3 refer to A-level achievement of students. Standard errors in parentheses. * $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$*

*Table 7: Probability of being over-qualified for university*

| Sample | (1)<br>All | (2)<br>ABB+ | (3)<br>All | (4)<br>ABB+ |
|---|---|---|---|---|
| Under-predicted | 0.102*** | 0.091*** | 0.015*** | 0.069*** |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| Quintile of disadvantage (ref=poorest) | | | | |
| 1 (poorest) | -0.164*** | 0.003* | -0.187*** | 0.014*** |
| | (0.001) | (0.002) | (0.003) | (0.004) |
| 2 | -0.116*** | 0.001 | -0.123*** | 0.003 |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| 3 | -0.089*** | -0.002 | -0.087*** | 0.005** |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| 4 | -0.038*** | -0.001 | -0.036*** | 0.003* |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| | | | | |
| $R^2$ | 0.016 | 0.017 | 0.015 | 0.012 |
| Observations | 2,842,190 | 701,070 | 509,820 | 130,610 |
| Year effects | ✓ | ✓ | ✓ | ✓ |

*Notes: source, UCAS*
*No controls, other than cycle year were available*
*Standard errors in parentheses. * $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$*

# References

AoC (2019). 'College Key Facts 2018-19', Association of Colleges, London

Boliver, V. (2013). How fair is access to more prestigious UK universities?. The British journal of sociology, 64(2), 344-364.

Carneiro, P., & Heckman, J. J. (2002). The evidence on credit constraints in post-secondary schooling. *The Economic Journal*, *112*(482), 705-734.

Campbell, S., MacMillan, L., Murphy, R.J. and Wyness, G (2019) 'Matching in the dark? Inequalities in student to degree match', CEPEO working paper no.20-01, Centre for Education Policy and Equalising Opportunities, UCL, London.

Chowdry, H., Crawford, C., Dearden, L., Goodman, A. and Vignoles, A. (2013) 'Widening Participation in Higher Education: Analysis Using Linked Administrative Data', Journal of the Royal Statistical Society: Series A (Statistics in Society), 176, 431–457.

Dillon, E. W. and J. A. Smith (2017). Determinants of the match between student ability and college quality. Journal of Labor Economics 35(1), 45–66.

Everett & Papageorgiou (2011), 'Investigating the Accuracy of Predicted A Level Grades as part of 2009 UCAS Admission Process'

Havergal, C (2015), 'Elite universities lower entry standards to expand, report reveals', Times Higher Education, online, (accessed October 2016)

HESA (2020), 'Students by HE provider', Higher Education Statistics Agency, online.

Hoxby & Avery (2012), 'The Missing "One-Offs": The Hidden Supply of High-Achieving, Low Income Students', NBER working paper No. 18586, National Bureau of Economic Research

Hvidman, U., & Sievertsen, H. H. (2019). High-Stakes Grades and Student Behavior. Journal of Human Resources, 0718-9620R2.

Lasselle, L., J. McDougall-Bagnall, & Smith. I., (2014), 'School grades, school context and university degree performance: evidence from an elite Scottish institution', Oxford Review of Education, Volume 40, 2014 - Issue 3

McGuigan, M., McNally, S., & Wyness, G (2016), 'Student Awareness of Costs and Benefits of Educational Decisions: Effects of an Information Campaign", Journal of Human Capital

Mcnally, S. (2016). How important is career information and advice?. IZA World of Labor.

Department of Education (DfE). (2010). Schools, pupils and their characteristics. January 2010

Smith, J, Pender,M., & Howell, J., (2013). 'The Full Extent Of Student-College Academic Undermatch'. Economics of Education Review 32: 247-261.

UCAS (2019), '2019 cycle applicant figures – 30 June deadline', University and Colleges Admissions Service, Cheltenham

UCAS (2020), 'Predicted grades: what you need to know", online, University and Colleges Admissions Service, Cheltenham. Available at https://www.ucas.com/advisers/managing-applications/predicted-grades-what-you-need-know (last accessed 4th March, 2020)

UCU (2015), 'Undergraduate application and admission survey', University and College Union

Wilson, J., (2015), 'Students should apply to university once they've received their grades, not before', The Guardian Higher Education Network, blog post, (accessed September 2016)

Appendix 1: Probability of being underpredicted by quintile of disadvantage

| Sample | | (1) All | (2) All | (3) CCC | (4) CCC-AAB | (5) AAB |
|---|---|---|---|---|---|---|
| Quintile of Disadvantage (ref=richest) | | | | | | |
| 1 (poorest) | | -0.026*** | -0.024*** | -0.003*** | -0.001 | 0.011*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.004) |
| | 2 | -0.019*** | -0.016*** | -0.003*** | 0.001 | 0.008** |
| | | (0.001) | (0.001) | (0.000) | (0.001) | (0.003) |
| | 3 | -0.015*** | -0.010*** | -0.002*** | 0.001 | 0.013*** |
| | | (0.001) | (0.001) | (0.000) | (0.001) | (0.003) |
| | 4 | -0.008*** | -0.007*** | -0.001*** | 0.000 | 0.002 |
| | | (0.001) | (0.001) | (0.000) | (0.001) | (0.002) |
| year=2014 | | -0.003*** | -0.003*** | 0.000 | -0.005*** | 0.009*** |
| | | (0.001) | (0.001) | (0.000) | (0.001) | (0.002) |
| year=2015 | | -0.007*** | -0.006*** | 0.000 | -0.007*** | 0.004** |
| | | (0.001) | (0.001) | (0.000) | (0.001) | (0.002) |
| Constant | | 0.098*** | 0.099*** | 0.004*** | 0.053*** | 0.253*** |
| | | (0.001) | (0.001) | (0.000) | (0.001) | (0.003) |
| | | | | | | |
| R-squared | | 0.001 | 0.006 | 0.001 | 0.004 | 0.013 |
| Observations | | 858,720 | 858,720 | 146,890 | 480,610 | 231,220 |
| Student characteristics | | | ✓ | ✓ | ✓ | ✓ |
| School type | | | ✓ | ✓ | ✓ | ✓ |

Notes: source, UCAS

Characteristics in column 2 are gender, ethnicity. Columns 3-5 refer to A-level achievement of students. Standard errors in parentheses. * $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$

Appendix 2: Probability of being underpredicted by school type

| Sample | (1) All | (2) All | (3) CCC | (4) CCC-AAB | (5) AAB |
|---|---|---|---|---|---|
| School Type (Ref= Private School) | | | | | |
| Grammar | -0.002 | -0.004*** | 0.001 | 0.003** | 0.019*** |
| | (0.001) | (0.001) | (0.001) | (0.002) | (0.003) |
| Academy | -0.007*** | 0.000 | 0.005*** | 0.020*** | 0.055*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| 6th form college | -0.017*** | -0.008*** | 0.003*** | 0.018*** | 0.048*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.003) |
| State | -0.021*** | -0.012*** | 0.004*** | 0.015*** | 0.065*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.003) |
| Further Education College | -0.040*** | -0.037*** | 0.000 | -0.007*** | 0.046*** |
| | (0.002) | (0.002) | (0.001) | (0.002) | (0.006) |
| Other | -0.078*** | -0.079*** | 0.000 | -0.036*** | -0.024 |
| | (0.005) | (0.005) | (0.003) | (0.005) | (0.023) |
| year=2014 | -0.003*** | -0.003*** | 0.000 | -0.005*** | 0.009*** |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.002) |
| year=2015 | -0.007*** | -0.006*** | 0.000 | -0.007*** | 0.004** |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.002) |
| Constant | 0.101*** | 0.111*** | 0.001 | 0.038*** | 0.188*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.003) |
| | | | | | |
| R-squared | 0.001 | 0.006 | 0.001 | 0.004 | 0.013 |
| Observations | 858,835 | 858,720 | 146,890 | 480,610 | 231,220 |
| Student characteristics | | ✓ | ✓ | ✓ | ✓ |
| School Type | | ✓ | ✓ | ✓ | ✓ |

*Notes: source, UCAS*
*Characteristics in column 2 are gender, ethnicity. Columns 3-5 refer to A-level achievement of students. Standard errors in parentheses. \* p < 0.1. \*\* p < 0.05. \*\*\* p < 0.01*

@ cepeo_ucl

ucl.ac.uk/ioe/cepeo