

Teacher Grade Predictions for Ethnic Minority Groups: Evidence from England

Hester Burn
Laura Fumagalli
Birgitta Rabe

Institute for Social and Economic Research
University of Essex

No. 2023-3
March 2023



INSTITUTE FOR SOCIAL
& ECONOMIC RESEARCH



Economic
and Social
Research Council



University of Essex

Teacher Grade Predictions for Ethnic Minority Groups: Evidence from England

Non-Technical Summary

In England, students sit national examinations (GCSEs) at the end of secondary school, at age 16. These exams are externally and blindly marked and the final grades are determined by a national regulatory body. However, when schools in England closed due to the Covid-19 pandemic in 2020, GCSEs were cancelled, and instead teachers were asked to predict how they believed their students would have performed had the exams gone ahead. It was widely anticipated that this policy would result in considerably higher grades than those given in previous years. There were also widespread concerns that using predicted grades could result in an increase in attainment gaps between different groups of students. This may have occurred, for example, if teachers' prediction methods differed across schools, or if their predictions differed systematically according to students' characteristics.

In this paper, we use the cancellation of GCSE exams during Covid-19 to examine whether teachers have different predictions for the examination performance of students with different characteristics. We focus on ethnic minority students' outcomes in maths and English. Comparing the GCSE grades of White British and ethnic minority students in 2020 and in the preceding years when GCSE were externally marked, we find that in 2020 ethnic minority students did relatively better than White British students in maths and worse in English. These results do not appear to be driven by trends over time, or differences in the observed characteristics between the cohorts. Our analysis suggests that in 2020 high achieving ethnic minority students did not perform as well as in previous years in English, relative to White British students, partly because they were already performing so well that they had less room for further improvements. Nonetheless a large portion of the differences in the performance of ethnic minorities relative to White British students in 2020 compared to preceding years when the exams were blindly marked is unexplained. We consider alternative reasons but conclude that patterns in differential teacher predictions for the examination performance of different ethnic groups across subjects is a convincing explanation of the results.

The literature comparing teacher assessments with blindly marked exam grades for students in England paints a remarkably consistent story, with students from ethnic minority groups found to receive lower teacher assessed grades than exam grades in English, and either similar or higher teacher assessed grades than exam grades in maths. Our results are not directly comparable with this literature as we examine teachers' predictions of students' examination performance rather than a separate, intentionally teacher assessed measure of students' skills. Nonetheless our results are congruent with this general pattern. From a policy perspective these results are at the very least a useful exploration of some of the likely effects of using teacher predictions in the context of national examinations in England. Overall, the results suggest that the use of teacher predictions may see ethnic minority students receive higher grades relative to White British students in maths but lower grades relative to White British students in English than would otherwise be the case.

Teacher Grade Predictions for Ethnic Minority Groups: Evidence from England ^{*†}

Hester Burn[◇] Laura Fumagalli[§] Birgitta Rabe^{‡ §}

[◇] Institute for Social and Economic Research, University of Essex, UK

[‡] Institute for Social and Economic Research, University of Essex, UK

[§] Institute for Social and Economic Research, University of Essex, UK

Abstract

We explore whether teachers have different predictions for the examination performance of ethnic minority students relative to White British students. We exploit an exogenous change in assessment methods to compare grades based on teacher predictions to grades received through actual blindly marked examinations. Relative to White British students, teachers appear to have higher predictions for ethnic minority students' examination performance in maths and lower predictions for ethnic minority students' examination performance in English. These effects do not disappear when observable differences between groups and cohorts are accounted for, with differential teacher predictions of examination performance across ethnic groups remaining a convincing explanation of the results.

Keywords: educational evaluation schemes, teacher grading, Covid-19, ethnicity, bias

JEL Classification: I21, I24, J15

*Acknowledgment: The authors are grateful for funding from UKRI ESRC under Covid-19 Rapid Response Project ES/V015222/1 and co-funding from the ESRC Research Centre on Micro-Social Change (ES/S012486/1). Burn also acknowledges PhD funding by The Economic and Social Research Council Research Centre on Micro-Social Change (ES/P00072X/1). Corresponding author: Burn, Institute for Social and Economic Research, h.burn@essex.ac.uk.

[†]This work was produced using statistical data accessed via the ONS Secure Research Service. The use of this data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

1 Introduction

The grades that students receive at school determine the choices available to them in their later lives. Systematic gaps in the grades received by different groups of students are concerning. Existing research into the determinants of grade gaps has focused on the roles played by individual, family, community, school, and peer characteristics (Sammons, 2007). However it is also possible that teachers contribute to grade gaps, either through teaching practices that are more effective for some student groups than others, or by differentially assessing their students' skills. Most countries use teacher assessments at least at some stages of children's educational trajectory, and there is a live debate in England about replacing some national examinations with teacher assessments (e.g. Blunkett et al., 2022: p.38). Discerning whether teachers' assessments differ systematically according to students' characteristics is therefore an important contribution to education policy discussions both in England and further afield.

In this paper, we exploit a change in assessment methods during the Covid-19 pandemic to examine whether teachers have different predictions for the examination performance of students with different characteristics. We focus on ethnic minority students' outcomes in two subjects: English and maths. We use administrative data containing a census of all state school students in England and the grades that they receive in a set of national examinations completed at age sixteen (GCSEs). Importantly, these data include the grades that teachers predicted that their students would attain in 2020, which were used in place of examination grades when examinations were cancelled due to the Covid-19 pandemic.

The announcement that GCSE examinations were to be cancelled in England occurred just two months before they were due to begin, with teachers instructed to base their grade predictions only on students' work prior to the announcement rather than on any work completed after school closures. This allows us to study teachers' grade predictions for students with different ethnic backgrounds that are unaffected by any differences that might arise through teaching practices, as any such 'input' effects will have been constant across years. It also means that differences in students' home learning experiences during Covid-19 are not a concern. Therefore our empirical strategy simply compares both raw

and conditional gaps in the grades that students from different ethnic groups received from teachers in 2020 to those that students from different ethnic groups received from blindly marked examinations in the previous year(s). We first describe raw changes in the grade gaps and then use decomposition analyses to ascertain the extent to which these changes can be accounted for by differences in the levels of – or returns to – observed covariates between the groups and years, and the extent to which they might be accounted for by other factors.

Our results indicate that, relative to White British students, teachers in England have higher predictions for the examination performance of ethnic minority students in maths and lower predictions for the examination performance of ethnic minority students in English. We find that these results remain after accounting for differences in the observed characteristics of the ethnic groups between cohorts. They also remain after accounting for differences, between years, in how these observed characteristics are related to grades. The fact that there is an upper limit to GCSE grades appears to contribute to the results by limiting the 'growth potential' of particularly high-attaining students. However these ceiling effects do not sufficiently account for all the changes observed. We conclude that that patterns of differential teacher predictions for the examination performance of different ethnic groups across subjects is a convincing explanation of the results.

This paper is closely related to an economics literature which examines systematic differences in students' attainment in teacher assessments compared to blindly marked examinations. Some studies have found teacher assessed grades to be very similar to those awarded through blindly marked examinations (Dorsey and Colliver, 1995; Reeves et al., 2001; Dhillon, 2005). Other studies have found that teachers' assessments diverge, for example by resulting in higher grades than blindly marked examinations for lower-ability students (Gibbons and Chevalier, 2008; Borcan et al., 2017) or for any students close to a particular cut off (Diamond and Persson, 2016). Crucially, a number of studies have found that teachers' assessments diverge in a way that may indicate conscious or unconscious bias against students based on characteristics such as their socio-economic status, gender or ethnicity (Lindahl, 2007; Terrier, 2016; Lavy and Megalokonomou, 2019; Alesina et al., 2019). For example, low-income students in England tend to receive lower grades from

teacher assessment compared to blindly marked examinations across all subjects and ages (Campbell, 2015; Burgess and Greaves, 2013; Wyness and Murphy, 2020), while boys are often found to receive higher teacher-assessed grades in maths and lower teacher-assessed grades in English relative to girls (Campbell, 2015; Lavy and Sand, 2018; Lindahl, 2007).

Our approach differs slightly from those above. The most common methodology in this literature is a comparison of “blind” (examination) and “nonblind” (teacher) assessments of the same student (e.g. Lindahl, 2007; Burgess and Greaves, 2013; Campbell, 2015). However this methodology is not possible for students at age sixteen in England as there are no contemporaneous teacher and examination assessments at this age. This methodology also rests on the sometimes strong assumption that both assessment methods are attempting to measure the same skills. For example, many studies compare examinations intended to provide a “snapshot” of students’ performance in examination settings with teacher judgements which consider students’ written, practical, and oral classwork over an entire academic year (e.g. Gibbons and Chevalier, 2008). Exploiting an exogenous change in national assessment methods has, to our knowledge, not been done before in the literature, but does offer some unique benefits in this respect. In particular, the teachers providing “nonblind” assessments in 2020 were asked to report precisely the grade that they predicted their students would achieve had they sat blindly marked examinations. This approach therefore provides a level of comparability in outcome which is arguably greater than in some of the examples above. Moreover, teachers’ predictions of students’ performance in national examinations are currently used by both further and higher education providers in the UK (ENgalnd?) to inform admission offers and are therefore a highly relevant outcome.

This paper is also closely related to an existing literature examining teachers’ assessments of students from minority or non-native ethnic groups (e.g. Alesina et al., 2019). In England, the extant evidence paints a remarkably consistent story. Comparing teacher assessments with examination grades at age seven, Campbell (2015) finds that all non-White groups in her sample receive lower teacher assessed grades in reading but similar teacher assessed and examination grades in maths. Comparing teacher assessments with examination grades at age eleven, Burgess and Greaves (2013) find that students from

Pakistani, Black African, and Black Caribbean backgrounds are approximately twice as likely to receive lower teacher assessment grades than White British students in English, whereas the comparable rates for maths are much more similar to White British students for nearly all groups. Comparing teacher assessments with examination grades at age fourteen, Gibbons and Chevalier (2008) find that students from all ethnic minority groups have significantly lower teacher assessments and higher examination grades in English, but higher teacher assessments and lower examination grades in maths.

This literature suggests that ethnic minority students in England receive lower teacher assessed grades than examination grades in subjects relating to English, but not in maths. It also suggests that the likelihood of ethnic minority students receiving higher teacher assessed grades than examination grades in maths may increase as students age. However the evidence regarding students older than age fourteen is scarce. This is notable because ethnic minority students generally attain more highly relative to White British students as they get older (Dustmann et al., 2010), making it likely that teachers' perceptions of their skills and knowledge will also change. Wyness and Murphy (2020) compare teacher predictions with examination grades that students receive at age eighteen but discuss outcomes for ethnic minority students in their naïve descriptive statistics only, not in the multivariate modelling. This paper, by focusing on teachers' predictions of ethnic minority students' examination performance at age sixteen, therefore adds to a growing picture regarding the attainment and experiences of these groups as they progress through schooling in England.

Finally, this paper contributes to a literature considering the impacts of the Covid-19 pandemic in England. Existing analyses find no evidence that the use of teacher predicted grades in 2020 systematically disadvantaged students from typically lower-attaining ethnic minority groups (Wei Lee et al., 2020; Hunt et al., 2022). These analyses, however, did not account for differences in the observed characteristics of the ethnic groups between cohorts, or in how these observed characteristics are related to grades between years. We contribute further evidence about the impact of the change in assessment methods by accounting for these differences. This is important for two reasons. First, as national examinations in England are very high stakes for both students and schools, it is important

to know how comparable the grades awarded in 2020 are to those awarded in previous years. Second, as teacher predictions of student examination performance are used by universities in England to compare applicants, our findings can inform ongoing, national policy debates about current student assessment and university admissions procedures.

Section 2 describes examinations in England both before and during the Covid-19 pandemic, and Section 3 then explains the data used. Section 4 presents descriptive evidence about the grades and cohorts in 2020 in comparison to the previous years. Section 5 explains the decomposition methodology used, and Sections 6, 7, and 8 the results, robustness checks, and alternative explanations considered. Section 9 concludes.

2 Institutional context

2.1 School assessments in England

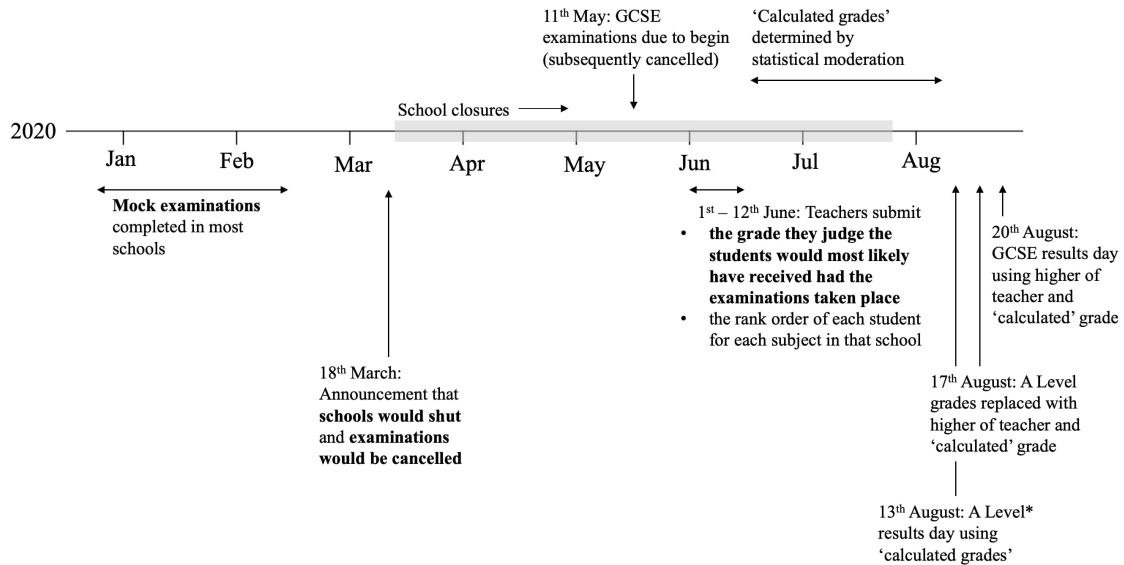
Students in England sit examinations for the General Certificate of Secondary Education (GCSEs) at the end of full-time compulsory schooling, the summer of the year in which they turn sixteen. The examinations are graded from one (low) to nine (high), with any scores below a one awarded a ‘U’ for ‘ungraded’. Grade boundaries are set by a national regulatory body once the distribution of raw marks is known so that the proportion of students achieving each grade is roughly comparable between years. School-level averages are then reported in league tables although, to disincentivise schools from prioritising students on the margin of achieving a pass grade (four), value-added measures are also given emphasis in national accountability frameworks. The grades that students receive in these examinations are highly determinate of the options available to them for post-16 education (Machin et al., 2020). A common benchmark is five GCSEs at a grade four or above, including English (either English Literature or English Language, whichever is higher) and maths.

There is little scope for grading bias in the English examination system. Grades for English Literature, English Language, and mathematics GCSE qualifications are entirely determined by students’ performance in a set of examinations sat at the end of a two-year course. The exam papers are externally and blindly marked by qualified teachers either from other schools or no longer in the profession, and identified by anonymous candidate numbers instead of names. Students’ handwriting is visible to the external marker and could give away a group identity – for example if handwriting differed by gender – yet grading biases associated with handwriting have not been supported by existing evidence (Baird, 1998). It is possible for students to sit the examinations early and this could give rise to bias if it is more likely to occur for certain groups. However early entry accounts for an average of only 2.4 percent of entries for English Literature, English Language, and mathematics (DfE, 2020a), and is no more likely to occur for certain ethnic groups than for others. Schools are also able to send exam scripts to be regraded once results have been received, though this occurrence is rare (.05 percent of students in 2019 (Ofqual, 2020a))

and uncorrelated with students' demographic characteristics (Machin et al., 2020).

2.2 Covid-19

Figure 1: Timeline showing details of the examination cancellations and teacher's grades in 2020



*A Levels are taken when students are age eighteen

Figure 1 is a timeline detailing the change in assessment methods in 2020. In England, the school year begins in September. Schools were instructed to close because of the Covid-19 pandemic on 18th March. The same announcement saw the cancellation of GCSE examinations scheduled for May and June that year. On 3rd April teachers received guidance indicating that they would be required to assign grades in place of the examinations. By mid-June schools were then asked to submit, for each student and for each subject in which they were entered, the grade that they judged the student would most likely have received had the examinations taken place. They were also asked to submit a rank order of each student in each subject. The grades and rankings were accompanied by evidence, mainly comprising marks and scripts from 'mock' examinations that the students will have sat prior to school closures, often in January or February. Although a statistical moderation process on these grades was initially implemented by a national regulatory body, it was

later discarded due to a loss of public confidence in the process. As a result, 95 percent of GCSE grades received by students in 2020 were the predicted grades assigned directly by teachers and schools, with the remaining five percent calculated through statistical modelling (Wei Lee et al., 2020).

Finally, certain precautions were taken to ensure that the grades that students received in 2020 were as comparable to those of previous cohorts as possible. Schools were instructed that their judgements should be based on existing evidence rather than any work completed by students after school closures. They were also asked not to discuss the grades with students or their parents or carers. We can therefore assume that the 2020 grades are not influenced by differences in school closure experiences between groups, or by manipulation to appease families.¹

¹In addition, judgements made by individual teachers were signed off by at least two members of staff – one of which was the lead teacher of the subject within the school – and head teachers required to submit a declaration that the grades had been generated according to the guidance.

3 Data

The National Pupil Database (NPD) contains administrative data on the universe school-age students in state schools in England. It includes demographic information about students, measures of their attainment at age seven, eleven, sixteen, and eighteen, and some information about schools and local authorities. As all state schools and examination boards in England are required to return these data by law, the NPD is both accurate and highly complete.

In this paper, our main analysis uses the cohorts of students in the NPD who turn sixteen in the academic years ending 2019 and 2020. However we also include the comparable cohorts for 2017 and 2018 to examine trends in the descriptive evidence, treating the data as a series of repeated cross-sections.² As the change to teacher predicted grades occurred for the end of year examinations in 2020, we remove any grades from examinations to which students were entered a year early. This inevitably includes a small proportion for whom this was their only recorded grade in a given subject (about 2.5 percent of students per cohort). There are also two GCSE English qualifications in England: English Language and English Literature. The vast majority of students in England sit both. If their grades differ, the higher of the two is used in headline performance measures for schools and, by students, to meet performance benchmarks (DfE, 2020b, p.53). We follow these national conventions and use the higher of the two grades in our main analyses. However we also present results for English Language and English Literature separately in robustness checks below.³

The initial sample comprises 2,252,437 students, or about 560,000 observations for each of the four cohorts. We drop students in independent schools for whom the data includes

²Cohorts prior to 2017 are incomparable because of significant examination reforms for GCSE maths and English qualifications which came into effect from 2017.

³In June 2020, 567,277 and 531,626 16-year-old students received English Language and English Literature grades respectively, which is similar in proportion to 2019 for which the figures are 546,607 and 514,191 (JCQ, 2020). The overall increase in entries is due to the 2020 cohort being approximately 3 percent larger than that of 2019 (Ofqual, 2020b). 5.4 percent of students were entered for English Literature a year early in 2019, which explains much of the discrepancy between the subjects (DfE, 2020a). As our sample is restricted to the end of year examinations in the year in which students turn sixteen, such ‘early entries’ are dropped. However these differences are not likely to cause any differences in results between English Literature and English Language. First, we are primarily interested in changes between 2019 and 2020, and patterns of entries appear consistent between these years (JCQ, 2020). Second, early entrants in English Literature are proportionate to the full sample with regards to ethnicity.

exam results but no background characteristics, as well as special schools, hospital schools, alternative provision providers, student referral units, and secure units, which together comprise 6.78 percent of the data. We also remove any students who are missing any of the explanatory variables or who have no recorded grade for GCSE maths or either GCSE English Literature or GCSE English Language (12.3 percent). This leaves a final sample size of 1,841,449 students in 3,396 schools across all four cohorts.

The student characteristics that we consider include gender, ethnicity, whether the student has been identified as having special educational needs (SEN), whether the student speaks English as an additional language (EAL), and their attainment at age eleven (KS2 subject-specific point scores).⁴ We use both eligibility for free school meals at any point in the last six years (FSM6) and a measure of income deprivation in the local area of residence (IDACI score) as proxies for students' household and neighbourhood deprivation. At the school-level, we consider a number of characteristics including the type of school, whether the school has selective admissions, and its region. School quality is proxied by a measure of school value-added from the end of primary school to examinations at age sixteen in eight prescribed subjects (Progress 8). Measures of school composition are constructed from the student data. We present a balance table for the sample in Appendix 1.

⁴For KS2 subject-specific point scores, we use the total marks in the KS2 maths tests for maths, and the marks in the KS2 English reading test for English. All of these tests are blindly and externally marked.

4 Descriptive evidence

We start by providing some basic descriptive evidence about GCSE grades and gaps in student performance over time. These clarify what significant changes occurred in 2020. Results for the 2017 and 2018 cohorts are included for the purpose of examining trends, though the decomposition analyses below will directly compare results for the 2020 cohort with the 2019 cohort only. The results using a combined 2018 and 2019 ‘control’ are very similar (see Appendix 2).⁵

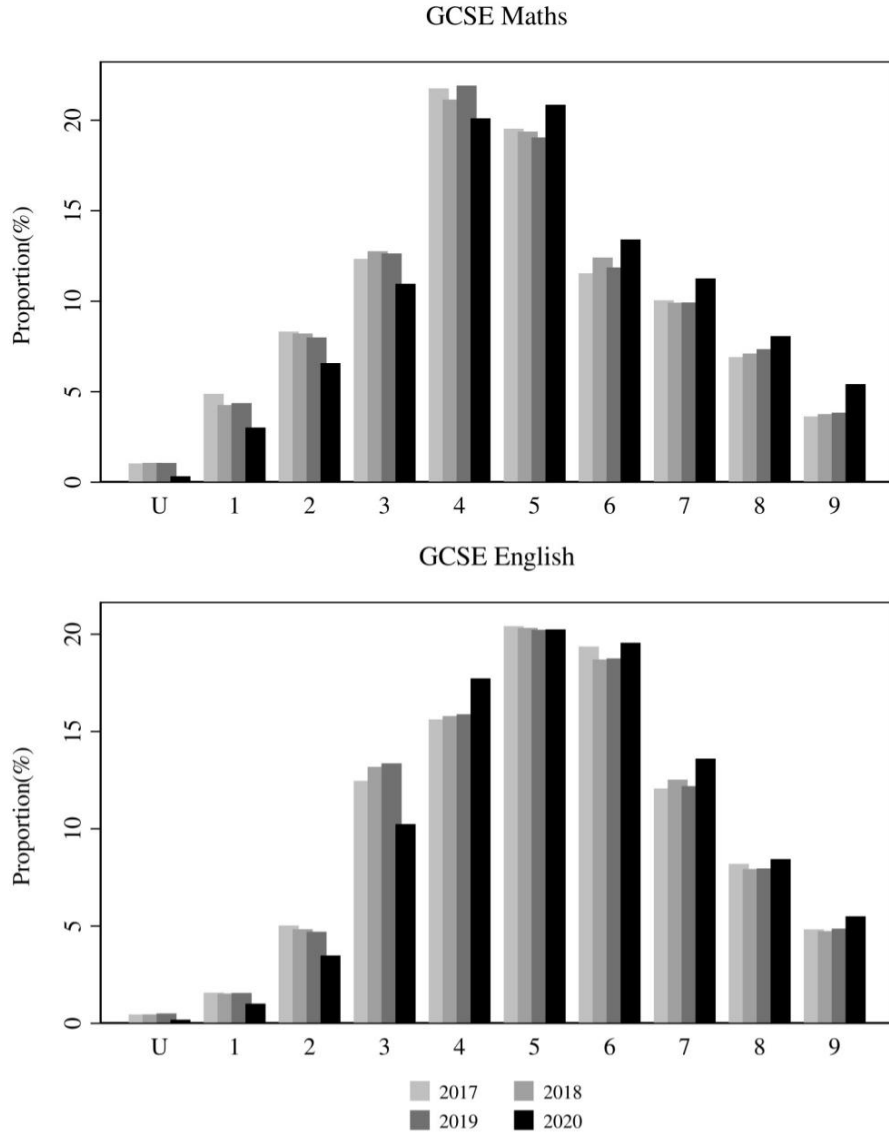
4.1 The distributions of grades

Figure 2 shows the grade distributions for maths and English between 2017 and 2020. The 2020 grades are considerably higher than those in previous years. This is unsurprising given the lack of grading criteria and the fact that teachers were likely to want to give their students the best possible opportunities for their future studies and work, for example by assuming that all students would have a ‘good’ exam day. The grade distributions in 2020 are also more left-skewed than those in previous years. This is because grades cannot exceed a grade nine. As a result, students have different growth potential when grades increase. For example, students at the top end of the grade distribution have less growth potential than those lower down. Such occurrences are known as ‘ceiling effects’ and may affect changes in grade gaps between ethnic groups which are positioned differently, on average, within the overall distribution of grades.

The first row of Table 1 shows the grade averages by year for all students. The grades for English are slightly higher on average than those for maths. As a result, ceiling effects are likely to at least partly explain the overall larger grade increases, between 2019 and 2020, for maths (0.31 grades) than for English (0.19 grades). The reduction in the standard deviations in 2020 is also illustrative of ceiling effects, yet it is interesting that this reduction is higher for English than maths; despite the grade increases being lower on average, the grades may yet be more greatly condensed, potentially driven by greater movement at the bottom end of the distribution.

⁵It is not possible to include 2017 as a control year due to the need to include lagged school value-added measures in the decompositions below. This is not possible for 2017 given changes in national assessments implemented that year.

Figure 2: Grade distributions by year



Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. Standard deviations in parentheses.

Table 1 also shows the mean grade achieved by students from different demographic groups. We include gaps for demographic groups which are not ethnicity for wider context. Female students consistently attain more highly than male students in English, but less highly than male students in maths. Indian and Black African students attain more highly than White British students in both subjects. Pakistani and Bangladeshi students attain more comparably to White British students, and Black Caribbean students considerably below. The grade gap between students who are and are not eligible to receive free school meals

Table 1: Mean grades by demographic group

	2017	GCSE Maths			2017	GCSE English		
		2018	2019	2020		2018	2019	2020
Total	4.71 (2.03)	4.75 (2.03)	4.75 (2.03)	5.06 (1.99)	5.23 (1.90)	5.21 (1.90)	5.21 (1.90)	5.40 (1.82)
Female	4.67 (2.00)	4.71 (1.98)	4.72 (2.00)	5.06 (1.95)	5.64 (1.83)	5.61 (1.83)	5.61 (1.84)	5.79 (1.78)
Male	4.75 (2.07)	4.79 (2.07)	4.78 (2.07)	5.07 (2.03)	4.82 (1.88)	4.82 (1.88)	4.81 (1.88)	5.02 (1.78)
White British	4.65 (2.00)	4.69 (2.00)	4.68 (2.00)	4.97 (1.96)	5.14 (1.90)	5.13 (1.90)	5.12 (1.90)	5.33 (1.83)
Pakistani & Bangladeshi	4.61 (2.04)	4.65 (2.04)	4.73 (2.06)	5.11 (2.01)	5.33 (1.81)	5.29 (1.82)	5.34 (1.84)	5.48 (1.75)
Indian	5.79 (2.06)	5.80 (2.04)	5.90 (2.07)	6.29 (1.96)	6.12 (1.77)	6.08 (1.76)	6.10 (1.79)	6.27 (1.68)
Black African	4.80 (1.99)	4.81 (1.98)	4.78 (2.00)	5.18 (1.91)	5.56 (1.76)	5.52 (1.77)	5.49 (1.78)	5.59 (1.68)
Black Caribbean	3.94 (1.88)	3.82 (1.86)	3.80 (1.83)	4.28 (1.78)	4.90 (1.79)	4.75 (1.79)	4.70 (1.76)	4.98 (1.65)
Multiethnic	4.67 (2.10)	4.72 (2.10)	4.73 (2.11)	5.04 (2.04)	5.40 (1.96)	5.36 (1.93)	5.36 (1.94)	5.50 (1.86)
No free school meals	5.01 (1.98)	5.05 (1.97)	5.05 (1.98)	5.35 (1.93)	5.49 (1.84)	5.47 (1.84)	5.46 (1.84)	5.66 (1.76)
Free school meals	3.82 (1.93)	3.86 (1.92)	3.85 (1.93)	4.17 (1.91)	4.47 (1.87)	4.44 (1.86)	4.44 (1.88)	4.63 (1.80)
Observations	4524551	443818	455614	489766	4524551	443818	455614	489766

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. Standard deviations in parentheses.

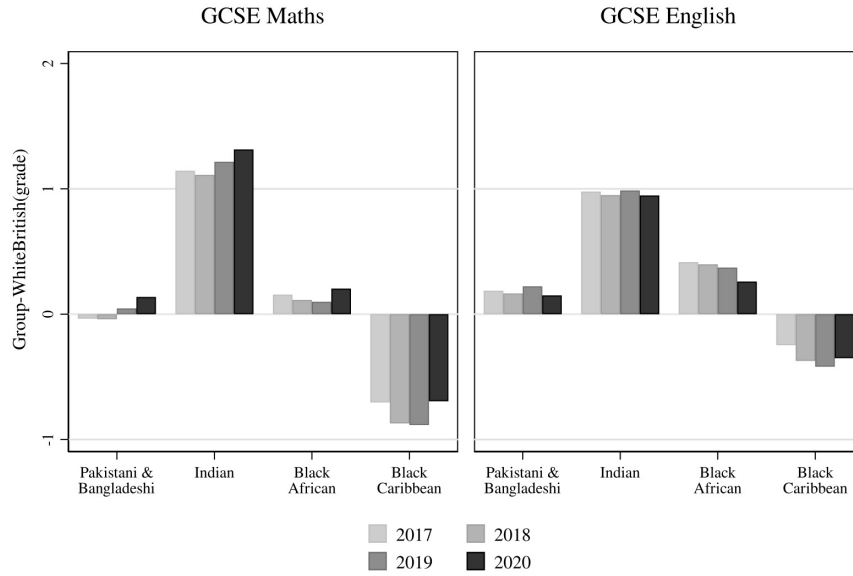
(a binary measure of household income deprivation) is both large and persistent.

4.2 Grade gaps between groups

Figure 3 shows the grade gaps for students from ethnic minority backgrounds in the years 2017, 2018, 2019, and 2020. Each bar shows the difference in the average grade that students received in relation to the reference group, White British students. In 2020, when the grades are from teacher predictions rather than from blindly marked examinations, there appear to be positive relative grade changes for ethnic minority students in maths and negative relative grade changes for ethnic minority students in English. Black Caribbean students, for whom there appear to be positive relative grade changes in both subjects, are an exception.

The grade gap changes do not appear to follow a consistent trend over time. This is important as it could be that grade gaps in 2020 simply differ from those in previous years because students in ethnic minority groups are attaining increasingly highly relative to White British students. This could occur, for example, if gradual changes in teaching practices are resulting in more beneficial teaching for these groups, or if there are changes to attitudes towards school and examinations which occur for some groups and not others.

Figure 3: Grade gaps by demographic group



Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both.

However a basic appraisal of the raw gaps displayed in Figure 2 does not immediately support such explanations.

Tests for the significance of the grade gaps, and their changes, are shown in Table 2. The standard errors in the ‘2019’ and ‘2020’ columns test whether the gaps are different from zero, and those in the ‘Change’ columns the change in the gaps themselves. Across both subjects, changes in the grades of ethnic minority relative to White British students are consistently larger than changes according to gender or between students who are and are not eligible for free school meals. For maths, the relative grade changes for all ethnic minority groups are both positive and significant at the one percent level. For English, changes for Pakistani and Bangladeshi and Black African students are both negative and significant at the one percent level. The point estimate for Indian students is negative, and for Black Caribbean students positive, though neither reach significance at either the one or five percent level.

Table 2: Grade gaps by demographic group

	GCSE Maths			GCSE English		
	2019	2020	Change	2019	2020	Change
Male	0.058*** (0.018)	0.004 (0.017)	-0.054*** (0.010)	-0.798*** (0.016)	-0.767*** (0.015)	0.031*** (0.009)
Pakistani & Bangladeshi	0.045 (0.036)	0.137*** (0.034)	0.092*** (0.023)	0.222*** (0.034)	0.149*** (0.032)	-0.073*** (0.020)
Indian	1.216*** (0.050)	1.314*** (0.048)	0.098*** (0.028)	0.987*** (0.043)	0.947*** (0.039)	-0.040 (0.027)
Black African	0.099*** (0.032)	0.204*** (0.028)	0.105*** (0.023)	0.372*** (0.031)	0.260*** (0.027)	-0.112*** (0.023)
Black Caribbean	-0.884*** (0.034)	-0.694*** (0.031)	0.190*** (0.035)	-0.419*** (0.036)	-0.350*** (0.030)	0.069* (0.035)
Free school meals	-1.202*** (0.014)	-1.177*** (0.013)	0.025** (0.010)	-1.028*** (0.014)	-1.025*** (0.013)	0.003 (0.010)
Observations	455614	489766		455614	489766	

Notes: Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. Free school meals indicates if a student is known to have been eligible for free school meals in the past six years. The reference groups are female, White British, and no free school meals. Standard errors in parentheses, clustered at the school-level. * Significant at 10%; ** significant at 5%; *** significant at 1%.

4.3 Changes in the compositions of ethnic groups

Table 3 compares the observable characteristics of each the ethnic groups of focus in 2019 and 2020, and tests for the significance of the differences between those years. This is important because grade gaps in 2020 may differ from those in previous years because the cohorts, and the groups within them, are unbalanced on observable (and unobservable) characteristics. It could be, for example, that ethnic minority students in 2020 are simply higher attaining in maths and lower attaining in English than their 2019 equivalents, or that they attend different types of schools.

There is a slight increase in the proportion of White British students identified as having a special educational need in 2020 compared to 2019. All ethnic minority groups are also slightly less likely to speak English as an additional language in 2020 than in 2019. Regarding deprivation, all ethnic minority groups experience a significant reduction in their eligibility for free school meals in 2020 while White British students do not. Finally, whilst prior attainment in maths is remarkably balanced between the years for each group, Indian, Black African and Black Caribbean students have higher English prior attainment in 2020 than in 2019.

Changes in school characteristics are generally more balanced across the groups. There is

a slight decline in the proportion of Indian, Black African, and (to a lesser extent) White British students living in London in 2020. There are also large changes in the proportion of students eligible for free school meals in ethnic minority students' schools between these years, concurrent with the changes in individual status noted above. (This is likely to be because most ethnic minority groups have a higher eligibility for free school meals than White British students, and a higher likelihood of being educated in schools with high numbers of ethnic minority peers (Dustmann et al., 2010).) Finally, the lagged value-added measures of White British and Black African students' schools are slightly lower in 2020 than in 2019, and slightly higher for Pakistani and Bangladeshi students.

Overall, these differences do not intuitively explain the positive relative grade changes for ethnic minority students in maths and negative relative grade changes for ethnic minority students in English described above. Rather, the greater prevalence of special educational needs among White British students in 2020, coupled with the reduction in free school meals eligibility among ethnic minority groups, would tentatively suggest that ethnic minority students should attain more highly relative to their White British peers in 2020. We might also expect this difference to be greater in the case of English, given that the English prior attainment of ethnic minority students in the 2020 cohort is also generally higher.

It is possible, however, that it is not only observed characteristics between cohorts, but also unobserved characteristics, which differ in 2020 compared to 2019. Relationships between characteristics and grades may also change. One characteristic, for example, may have much greater returns for attainment in one subject in 2020 compared to 2019. If this characteristic happens to be more greatly represented by some ethnic groups than others then this group will see greater relative grade changes overall due to factors other than ethnicity. Accounting for these other possible contributions to the grade gaps is our next step.

Table 3: Student and school characteristics by ethnic group and cohort

	White British			Pakistani & Bangladeshi			Indian			Black African			Black Caribbean		
	2019	2020	Change	2019	2020	Change	2019	2020	Change	2019	2020	Change	2019	2020	Change
Pupil characteristics															
Male	0.501 (0.500)	0.503 (0.500)	0.002 (0.001)	0.500 (0.500)	0.496 (0.500)	-0.004 (0.004)	0.514 (0.500)	0.509 (0.500)	-0.005 (0.006)	0.479 (0.500)	0.484 (0.500)	0.005 (0.006)	0.492 (0.500)	0.485 (0.500)	-0.006 (0.009)
Free school meals	0.213 (0.409)	0.213 (0.409)	0.000 (0.001)	0.376 (0.484)	0.364 (0.481)	-0.012*** (0.004)	0.143 (0.350)	0.130 (0.336)	-0.013*** (0.004)	0.495 (0.500)	0.458 (0.498)	-0.037*** (0.006)	0.481 (0.500)	0.451 (0.498)	-0.030*** (0.009)
Neighbourhood deprivation	0.172 (0.134)	0.173 (0.134)	0.001 (0.000)	0.274 (0.114)	0.275 (0.116)	0.001 (0.001)	0.193 (0.112)	0.192 (0.113)	-0.001 (0.001)	0.317 (0.126)	0.316 (0.128)	-0.001 (0.001)	0.301 (0.123)	0.306 (0.125)	0.004* (0.002)
Special educational needs	0.120 (0.325)	0.128 (0.335)	0.008*** (0.001)	0.094 (0.292)	0.096 (0.295)	0.002 (0.002)	0.059 (0.236)	0.058 (0.234)	-0.001 (0.003)	0.094 (0.291)	0.094 (0.292)	0.001 (0.003)	0.166 (0.372)	0.173 (0.378)	0.007 (0.007)
First language not English	0.004 (0.064)	0.004 (0.063)	0.000 (0.000)	0.687 (0.464)	0.672 (0.469)	-0.015*** (0.004)	0.588 (0.492)	0.569 (0.495)	-0.019*** (0.006)	0.523 (0.499)	0.497 (0.500)	-0.026*** (0.006)	0.025 (0.156)	0.024 (0.154)	-0.001 (0.003)
Maths prior attainment (age 11)	0.004 (0.990)	0.005 (0.994)	0.001 (0.002)	-0.121 (1.044)	-0.108 (1.025)	0.012 (0.009)	0.307 (0.957)	0.306 (0.935)	0.000 (0.012)	-0.091 (1.006)	-0.106 (0.986)	-0.014 (0.011)	-0.344 (1.009)	-0.341 (1.004)	0.003 (0.018)
English prior attainment (age 11)	0.042 (0.992)	0.041 (0.987)	-0.001 (0.002)	-0.261 (0.991)	-0.257 (1.000)	0.003 (0.008)	0.097 (0.971)	0.135 (0.968)	0.038*** (0.012)	-0.127 (0.967)	-0.097 (0.976)	0.029*** (0.011)	-0.263 (0.940)	-0.219 (0.944)	0.044*** (0.017)
School characteristics															
Selective admissions	0.041 (0.197)	0.040 (0.195)	-0.001* (0.000)	0.039 (0.195)	0.040 (0.195)	0.000 (0.002)	0.156 (0.363)	0.159 (0.366)	0.003 (0.005)	0.048 (0.214)	0.052 (0.221)	0.003 (0.002)	0.013 (0.113)	0.012 (0.111)	0.000 (0.002)
Region:															
London	0.057 (0.233)	0.055 (0.228)	-0.003*** (0.001)	0.262 (0.440)	0.266 (0.442)	0.004 (0.004)	0.294 (0.456)	0.280 (0.449)	-0.015** (0.006)	0.587 (0.492)	0.576 (0.494)	-0.011* (0.006)	0.625 (0.484)	0.635 (0.482)	0.010 (0.009)
East of England	0.128 (0.334)	0.124 (0.330)	-0.004*** (0.001)	0.070 (0.254)	0.072 (0.258)	0.002 (0.002)	0.058 (0.235)	0.067 (0.250)	0.009*** (0.003)	0.078 (0.268)	0.076 (0.265)	-0.002 (0.003)	0.059 (0.235)	0.053 (0.223)	-0.006 (0.004)
North East	0.169 (0.375)	0.170 (0.375)	0.001 (0.001)	0.177 (0.382)	0.180 (0.384)	0.003 (0.003)	0.062 (0.241)	0.062 (0.240)	-0.001 (0.003)	0.044 (0.205)	0.044 (0.206)	0.000 (0.002)	0.031 (0.174)	0.027 (0.162)	-0.004 (0.003)
North West	0.158 (0.365)	0.156 (0.363)	-0.003*** (0.001)	0.161 (0.368)	0.152 (0.359)	-0.010*** (0.003)	0.101 (0.301)	0.102 (0.303)	0.002 (0.004)	0.061 (0.239)	0.056 (0.231)	-0.004* (0.003)	0.028 (0.166)	0.029 (0.168)	0.001 (0.003)
South East	0.181 (0.385)	0.178 (0.383)	-0.003*** (0.001)	0.081 (0.273)	0.077 (0.267)	-0.004* (0.002)	0.132 (0.339)	0.136 (0.342)	0.003 (0.004)	0.084 (0.278)	0.092 (0.289)	0.007** (0.003)	0.043 (0.203)	0.044 (0.203)	0.001 (0.004)
South West	0.102 (0.302)	0.113 (0.316)	0.011*** (0.001)	0.010 (0.099)	0.011 (0.102)	0.001 (0.001)	0.026 (0.158)	0.029 (0.169)	0.004* (0.002)	0.021 (0.144)	0.023 (0.150)	0.002 (0.002)	0.022 (0.147)	0.018 (0.134)	-0.004 (0.002)
East Midlands	0.098 (0.297)	0.096 (0.295)	-0.001* (0.001)	0.038 (0.192)	0.043 (0.202)	0.004*** (0.002)	0.133 (0.340)	0.136 (0.343)	0.003 (0.004)	0.044 (0.205)	0.046 (0.210)	0.002 (0.002)	0.032 (0.177)	0.037 (0.188)	0.004 (0.003)
West Midlands	0.106 (0.308)	0.108 (0.311)	0.002** (0.001)	0.201 (0.401)	0.200 (0.400)	-0.001 (0.003)	0.193 (0.395)	0.188 (0.391)	-0.005 (0.005)	0.081 (0.272)	0.086 (0.280)	0.005* (0.003)	0.159 (0.366)	0.158 (0.365)	-0.001 (0.006)
Urban	0.843 (0.364)	0.839 (0.368)	-0.004*** (0.001)	0.986 (0.117)	0.987 (0.115)	0.000 (0.001)	0.978 (0.146)	0.976 (0.153)	-0.002 (0.002)	0.986 (0.117)	0.982 (0.134)	-0.004*** (0.001)	0.989 (0.103)	0.988 (0.109)	-0.001 (0.002)
Size of cohort	191 (63.8)	196 (64.2)	4.64*** (0.156)	199 (58.1)	202 (59.0)	3.10*** (0.486)	197 (65.1)	201 (64.0)	4.57*** (0.806)	184 (55.0)	189 (55.7)	4.56*** (0.618)	180 (56.5)	182 (55.8)	2.68*** (0.996)
School governance type:															
Local Authority Maintained	0.251 (0.433)	0.230 (0.421)	-0.021*** (0.001)	0.336 (0.472)	0.311 (0.463)	-0.026*** (0.004)	0.243 (0.429)	0.222 (0.415)	-0.021*** (0.005)	0.339 (0.473)	0.314 (0.464)	-0.026*** (0.005)	0.324 (0.468)	0.307 (0.461)	-0.017** (0.008)
Single Academy Trust	0.193 (0.395)	0.189 (0.391)	-0.004*** (0.001)	0.155 (0.362)	0.155 (0.361)	-0.001 (0.003)	0.254 (0.435)	0.256 (0.436)	0.002 (0.005)	0.209 (0.407)	0.203 (0.402)	-0.006 (0.005)	0.197 (0.398)	0.180 (0.385)	-0.017** (0.007)
Multi Academy Trust	0.522 (0.500)	0.556 (0.497)	0.034*** (0.001)	0.493 (0.500)	0.526 (0.499)	0.033*** (0.004)	0.486 (0.500)	0.508 (0.500)	0.022*** (0.006)	0.429 (0.495)	0.468 (0.499)	0.038*** (0.006)	0.448 (0.497)	0.488 (0.500)	0.040*** (0.009)
Other	0.034 (0.182)	0.026 (0.158)	-0.009*** (0.000)	0.016 (0.124)	0.008 (0.091)	-0.007*** (0.001)	0.017 (0.131)	0.015 (0.121)	-0.003* (0.002)	0.023 (0.150)	0.016 (0.127)	-0.007*** (0.002)	0.030 (0.171)	0.025 (0.155)	-0.005* (0.003)
Proportion free school meals	0.220 (0.127)	0.219 (0.125)	-0.002*** (0.000)	0.369 (0.156)	0.357 (0.153)	-0.011*** (0.001)	0.240 (0.142)	0.203 (0.137)	0.014*** (0.002)	0.354 (0.169)	0.336 (0.162)	-0.019*** (0.002)	0.374 (0.158)	0.363 (0.154)	-0.011*** (0.003)
Average neighbourhood deprivation	0.179 (0.083)	0.179 (0.083)	0.000** (0.000)	0.264 (0.082)	0.266 (0.082)	0.002** (0.001)	0.207 (0.076)	0.205 (0.075)	-0.003*** (0.001)	0.269 (0.082)	0.270 (0.083)	0.001 (0.001)	0.275 (0.077)	0.279 (0.077)	0.003** (0.001)
School value-added (lagged)	-0.024 (0.404)	-0.030 (0.409)	-0.006*** (0.001)	0.164 (0.449)	0.190 (0.455)	0.026*** (0.004)	0.336 (0.486)	0.343 (0.474)	0.007 (0.006)	0.192 (0.451)	0.178 (0.462)	-0.014*** (0.005)	0.079 (0.447)	0.084 (0.465)	0.005 (0.008)
Observations	327424	348296		28414	29623		12515	13180		15116	17041		6291	6419	

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Free school meals indicates if a student is known to have been eligible for free school meals in the past six years. Neighbourhood deprivation is a rank based on the proportion of children under age 16 in a local area living in a low-income household. School value-added refers to how many average grades higher or lower the students in that school achieve across eight qualifying GCSE subjects compared to students across the country who score comparatively at age 11. Standard deviations in parentheses in the ‘2019’ and ‘2020’ columns and standard errors in ‘Change’, clustered at the school level. * Significant at 10%; ** significant at 5%; *** significant at 1%.

5 Empirical framework

We use decomposition analysis to decompose the gap in outcomes between ethnic minority groups and their White British peers into that explained by differences in the groups’ observed characteristics and an unexplained part. The most widely used decomposition analysis in the literature is the Oaxaca-Blinder decomposition (Oaxaca, 1973; Blinder, 1973). In Oaxaca-Blinder, group-identifying coefficients are often assumed to equate to the proportion of the gap being ‘explained’ by a covariate (Gelbach, 2016). However, when the estimated contribution of each covariate to the gap is calculated by progressively adding additional covariates to a base model, the order in which covariates are added can change the reported results (path dependency).

An alternative type of decomposition is that of Gelbach (2016). Rather than using a sequential procedure to estimate the contribution of different covariates, and therein building in the issue of path dependency, Gelbach’s decomposition uses the omitted variable bias formula to yield consistent estimates which are order-invariant. As a result, the extent to which variation in a covariate increases or reduces the size of a gap in outcomes depends on simply: (1) whether the covariate has a positive or negative effect on the outcome in a model in which the groups are combined, and; (2) in which group the covariate has a higher mean. The main assumption is that the returns to all covariates are the same across the different ethnic groups.

In this paper we use Gelbach decompositions to investigate changes in grade gaps in 2020 compared to 2019.⁶ We derive our empirical specification from a simple education production function which assumes that the GCSE grades received by individual students in the year in which they turn sixteen are directly related to student characteristics – including student prior-attainment – and school characteristics. We estimate our empirical specification as follows.

In equation (1), Y_{it} is the grade received by student i at time t . The vector X_{it}^P includes student characteristics, specifically gender, ethnicity, indicators for special educational needs, indicators of English being an additional language, and measures of family income

⁶Our approach is similar to Blanden et al. (2015).

and neighbourhood deprivation. The vector X_{it}^S includes school characteristics such as the school governance type, quality of educational provision, and the deprivation level of the school's student body. The vector Y_{it-1} refers to subject-specific attainment measures when students are eleven years old. Time t is either 2019, when grades were assigned through blindly marked examinations, or 2020, when they were predicted by teachers.

$$Y_{it} = X_{it}^P \beta_t^P + X_{it}^S \beta_t^S + Y_{it-1} \beta_{t-1} + \epsilon_{it} \quad (1)$$

We use equation (1) to decompose the difference between the average grades received by students in different ethnic minority groups ($group_{it}$) and the reference group, namely their White British peers. Equation (2) estimates the raw grade gap at time t as α_t^{raw} , and equation (3) the conditional gap controlling for student and school characteristics as α_t^{full} .

$$Y_{it} = \alpha_t^{raw} \cdot group_{it} + \epsilon_{it} \quad (2)$$

$$Y_{it} = \alpha_t^{full} \cdot group_{it} + X_{it}^P \beta_t^P + X_{it}^S \beta_t^S + Y_{it-1} \beta_{t-1} + \epsilon_{it} \quad (3)$$

We then decompose the difference in the raw and conditional gaps ($\hat{\alpha}_t^{full} - \hat{\alpha}_t^{raw}$), giving the contributions of different factors to the grade gaps between students from White British and ethnic minority backgrounds. Equation (4) shows that the contribution of the k^{th} group to the grade gaps can be written as the omitted variable bias resulting from excluding this covariate or group of covariates from the full regression in equation (3).

$$\hat{\alpha}_t^{full} - \hat{\alpha}_t^{raw} = \sum_{k=X^P, X^S, Y^{t-1}} (X_k' X_k)^{-1} X_k' group_i \beta_t^k \quad (4)$$

We estimate this decomposition for grade gaps in 2019 and 2020 separately. We then estimate an additional decomposition in which the 2019 and 2020 data is pooled and a 2020 interaction is added to each group indicator and covariate. The corresponding

coefficients for each year – either as the reference group alone, for 2019, or combined with the interaction coefficients, for 2020 – are identical to those estimated using the individual decompositions outlined above. However the interaction coefficients themselves allow additional tests for the significance of the changes in the contributions made by different covariates between these years. We report these interaction coefficients – in the column labelled ‘Change’ – in the results below.

The effects of using teacher predicted grades rather than those from blindly marked examinations will be present in two parts of the results. First, there may be different returns to either school or student characteristics in 2020 in comparison to 2019 due to the use of teacher prediction rather than blindly marked examinations. If this is the case, it will be reflected in the size of the change of the contribution of that covariate or group of covariates between the years. However differences in the mean levels of covariates across ethnic groups between the years will also affect the size of changes in these contributions. It is possible to distinguish between the two possible channels by considering how the known covariate levels, as seen in Table 3, differ for the relevant ethnic groups between 2019 and 2020. The individual regression coefficients for each decomposition are reported in the fine-grained results, available on request.

The second part of the results containing the effects of using teacher predicted grades rather than those from blindly marked examinations will be the unexplained components. This is because the unexplained components contain contributions to the gap in grades between ethnic groups which are not explained by the observed student and school characteristics. If the error term is truly random each year, it follows that any changes in these unexplained components can be attributed to effects of the method of assessment. In many decomposition analyses, these unexplained components are also interpreted as measures of bias or discrimination (e.g. Duncan and Sandy, 2013). The extent to which these interpretations are valid here will be discussed further below.

6 Results

6.1 Changes in the explained and unexplained components of the gaps

The first panel of each of Tables 4 and 5 shows the raw gap in grades, the contribution to that gap accounted for our observed covariates, and the unexplained component. The raw gaps are identical to those in Table 2. For both maths and English, the negative explained components indicate that most ethnic minority groups would be expected to attain lower grades than White British students with whom they were matched on all covariates. (As these components are largely driven by gaps in prior attainment, the exception is students of an Indian background, who would be expected to attain considerably higher.) The unexplained components are consistently positive for most groups. This indicates that most ethnic minority groups attain higher grades than would be expected given their observed characteristics. The exception is Black Caribbean students, who attain lower grades.

Of particular note for the present paper is the fact that changes in the unexplained components differ by subject; they are positive and significant for all groups for maths, but negative and significant for all groups apart from Black Caribbean students for English. This not only means that differences in the levels of covariates are insufficient to explain the changes to the raw gaps in 2020 compared to 2019, but also that these changes remain insufficiently accounted for by additionally considering any differences in the returns to these covariates. In other words, the positive relative grade changes for ethnic minority students in maths and negative relative grade changes for ethnic minority students in English in 2020 cannot be fully explained by differences in either the levels or returns to the observed covariates between 2020 and 2019.

6.2 Changes in the contributions of the explained components

The second panels of Tables 4 and 5 show the estimated contributions made by different covariates or groups of covariates to the explained components each year, as well as the changes between those contributions. Positive estimates indicate a covariate (or group of covariates) which contributes positively to the grades received by the ethnic minority

group relative to White British students, and negative estimates a covariate (or group of covariates) which contributes negatively.

Changes in the contributions of the explained components primarily reflect differences in the levels of covariates across the groups and cohorts. The large contributions of prior attainment go some way to explaining why the aggregate explained component changes are greater in the case of English, as English prior attainment for Indian, Black African, and Black Caribbean students is considerably higher in 2020 than in 2019. As expected, average levels of family and neighbourhood deprivation have a negative effect on the average attainment ethnic minority students compared to their White British peers, but to a lesser extent in 2020 than in 2019. Similarly, the increase in White British students with special educational needs in 2020 – as well as the decrease in ethnic minority students speaking English as an additional language – has a positive effect relative to 2019 for most ethnic minority groups. The contributions of school characteristics and lagged measures of school performance are sizeable but generally smaller than those of deprivation or prior attainment. However the changes in the contributions of school characteristics are relatively large, and more so in the case of English than maths.

To further explore the changes in the contributions of school characteristics, we present a more detailed decomposition for English in Table 6 (see Appendix 3 for the same decomposition for maths). This decomposition is identical to Table 4 but reports the contribution of each school characteristic separately rather than in aggregate. It shows that two covariates account for considerable proportions of the change associated with school characteristics: region, and the proportion of students eligible for free school meals. In the latter case, the positive effects are accounted for by reductions in the levels of eligible peers among ethnic minority groups in 2020 which, as mentioned above, occur to a far lesser degree for White British students. The proportions of ethnic minority students living in different regions, however, do not change considerably between the cohorts. This means that the changes associated with region originate not in different levels but in different returns to the region covariate in 2020 compared to 2019, and that there may be some regional effect which at least partly explains the lower relative grades of ethnic minority groups in English in 2020 compared to 2019.

Table 4: Gelbach decomposition of grade gaps by ethnic groups for GCSE maths

	Pakistani & Bangladeshi			Indian			Black African			Black Caribbean		
	2019	2020	Change	2019	2020	Change	2019	2020	Change	2019	2020	Change
Raw gap	0.046 (0.036)	0.138*** (0.034)	0.092*** (0.023)	1.217*** (0.050)	1.315*** (0.048)	0.098*** (0.028)	0.099*** (0.031)	0.205*** (0.028)	0.105*** (0.023)	-0.884*** (0.034)	-0.693*** (0.031)	0.191*** (0.035)
Explained gap	-0.153*** (0.032)	-0.099*** (0.029)	0.054** (0.021)	0.812*** (0.049)	0.805*** (0.046)	-0.007 (0.025)	-0.186*** (0.030)	-0.188*** (0.027)	-0.002 (0.021)	-0.720*** (0.031)	-0.673*** (0.028)	0.047 (0.029)
Unexplained gap	0.199*** (0.017)	0.238*** (0.017)	0.039** (0.019)	0.405*** (0.019)	0.510*** (0.018)	0.105*** (0.020)	0.286*** (0.017)	0.393*** (0.015)	0.107*** (0.018)	-0.164*** (0.022)	-0.020 (0.020)	0.144*** (0.026)
<i>Amount explained by:</i>												
Male	0.000 (0.001)	0.002 (0.003)	0.002 (0.002)	-0.001 (0.001)	-0.001 (0.003)	0.000 (0.002)	0.002* (0.001)	0.005* (0.003)	0.003 (0.002)	0.001 (0.001)	0.005* (0.003)	0.004* (0.002)
Family & neighbourhood deprivation	-0.191*** (0.006)	-0.182*** (0.006)	0.009** (0.004)	-0.017*** (0.006)	-0.002 (0.006)	0.015*** (0.003)	-0.262*** (0.006)	-0.246*** (0.006)	0.016*** (0.005)	-0.246*** (0.006)	-0.237*** (0.006)	0.009* (0.006)
Special educational needs	0.008*** (0.001)	0.013*** (0.001)	0.004*** (0.001)	0.020*** (0.001)	0.027*** (0.001)	0.008*** (0.001)	0.008*** (0.001)	0.013*** (0.001)	0.005*** (0.001)	-0.015*** (0.002)	-0.017*** (0.003)	-0.003 (0.003)
First language not English	0.102*** (0.007)	0.108*** (0.007)	0.006 (0.008)	0.087*** (0.006)	0.091*** (0.006)	0.004 (0.007)	0.077*** (0.006)	0.080*** (0.005)	0.002 (0.006)	0.003*** (0.000)	0.003*** (0.000)	0.000 (0.001)
Prior attainment (age 11)	-0.168*** (0.022)	-0.156*** (0.020)	0.012 (0.015)	0.496*** (0.036)	0.468*** (0.035)	-0.028 (0.020)	-0.143*** (0.020)	-0.172*** (0.019)	-0.029* (0.016)	-0.521*** (0.022)	-0.505*** (0.021)	0.016 (0.025)
School characteristics	-0.015* (0.008)	-0.008 (0.007)	0.007 (0.008)	0.016** (0.007)	0.008 (0.007)	-0.007 (0.007)	0.004 (0.010)	0.014* (0.009)	0.011 (0.010)	-0.003 (0.010)	0.014 (0.010)	0.016 (0.011)
School value-added (lagged)	0.111*** (0.012)	0.124*** (0.012)	0.014 (0.008)	0.213*** (0.016)	0.214*** (0.015)	0.001 (0.009)	0.127*** (0.011)	0.118*** (0.011)	-0.009 (0.008)	0.061*** (0.013)	0.065*** (0.012)	0.004 (0.009)

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. The gap is the gap in average grades between the ethnic minority group and White British students. Family & neighbourhood deprivation includes free school meals indicator and a rank based on the proportion of children under age 16 in a local area living in a low-income household score (cubic). Prior attainment (age 11) includes subject-specific age 11 attainment score, standardised by year (quadratic). School characteristics include indicators for selectivity, sixth form, religious affiliation, urban, region (9 categories), and school governance type (4 categories), as well as cohort size, proportion of students eligible for free school meals (cubic) and average neighbourhood deprivation score of the school (cubic). School value-added refers to how many average grades higher or lower the students in that school achieve across eight qualifying GCSE subjects compared to students across the country who score comparatively at age 11 (cubic). Standard errors in parentheses, clustered at the school level. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Gelbach decomposition of grade gaps by ethnic groups for GCSE English

	Pakistani & Bangladeshi			Indian			Black African			Black Caribbean		
	2019	2020	Change	2019	2020	Change	2019	2020	Change	2019	2020	Change
Raw gap	0.222*** (0.034)	0.151*** (0.032)	-0.072*** (0.020)	0.987*** (0.043)	0.948*** (0.039)	-0.039 (0.027)	0.372*** (0.031)	0.261*** (0.027)	-0.111*** (0.023)	-0.419*** (0.036)	-0.349*** (0.030)	0.070** (0.035)
Explained gap	-0.194*** (0.030)	-0.129*** (0.027)	0.066*** (0.018)	0.480*** (0.042)	0.530*** (0.039)	0.050** (0.022)	-0.073** (0.029)	-0.019 (0.025)	0.054*** (0.019)	-0.403*** (0.030)	-0.322*** (0.026)	0.080*** (0.024)
Unexplained gap	0.416*** (0.020)	0.279*** (0.018)	-0.137*** (0.020)	0.506*** (0.022)	0.418*** (0.018)	-0.088*** (0.024)	0.445*** (0.019)	0.280*** (0.016)	-0.164*** (0.021)	-0.017 (0.025)	-0.027 (0.022)	-0.010 (0.030)
<i>Amount explained by:</i>												
Male	0.001 (0.007)	0.004 (0.006)	0.003 (0.003)	-0.008 (0.007)	-0.003 (0.006)	0.005 (0.004)	0.014* (0.007)	0.010* (0.005)	-0.004 (0.004)	0.006 (0.007)	0.009* (0.005)	0.003 (0.005)
Family & neighbourhood deprivation	-0.210*** (0.007)	-0.187*** (0.006)	0.023*** (0.004)	-0.016** (0.007)	0.001 (0.006)	0.018*** (0.004)	-0.289*** (0.006)	-0.255*** (0.006)	0.035*** (0.005)	-0.273*** (0.007)	-0.245*** (0.006)	0.027*** (0.006)
Special educational needs	0.015*** (0.002)	0.021*** (0.002)	0.005** (0.002)	0.036*** (0.002)	0.045*** (0.002)	0.009*** (0.002)	0.015*** (0.002)	0.022*** (0.002)	0.006** (0.002)	-0.027*** (0.004)	-0.028*** (0.004)	-0.001 (0.005)
First language not English	0.118*** (0.008)	0.119*** (0.007)	0.001 (0.010)	0.100*** (0.007)	0.100*** (0.006)	0.000 (0.008)	0.089*** (0.006)	0.088*** (0.006)	-0.002 (0.007)	0.003*** (0.001)	0.003*** (0.001)	0.000 (0.001)
Prior attainment (age 11)	-0.300*** (0.013)	-0.307*** (0.014)	-0.007 (0.010)	0.053** (0.023)	0.099*** (0.025)	0.046*** (0.014)	-0.174*** (0.013)	-0.150*** (0.014)	0.023** (0.011)	-0.312*** (0.014)	-0.285*** (0.015)	0.027* (0.016)
School characteristics	0.073*** (0.010)	0.101*** (0.009)	0.028*** (0.009)	0.112*** (0.010)	0.083*** (0.009)	-0.029*** (0.008)	0.146*** (0.012)	0.152*** (0.011)	0.006 (0.012)	0.139*** (0.013)	0.161*** (0.012)	0.022* (0.013)
School value-added (lagged)	0.108*** (0.013)	0.120*** (0.012)	0.012 (0.009)	0.203*** (0.016)	0.204*** (0.015)	0.001 (0.010)	0.125*** (0.011)	0.114*** (0.011)	-0.011 (0.008)	0.061*** (0.013)	0.063*** (0.012)	0.002 (0.009)

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. The gap is the gap in average grades between the ethnic minority group and White British students. Family & neighbourhood deprivation includes free school meals indicator and a rank based on the proportion of children under age 16 in a local area living in a low-income household score (cubic). Prior attainment (age 11) includes subject-specific age 11 attainment score, standardised by year (quadratic). School characteristics include indicators for selectivity, sixth form, religious affiliation, urban, region (9 categories), and school governance type (4 categories), as well as cohort size, proportion of students eligible for free school meals (cubic) and average neighbourhood deprivation score of the school (cubic). School value-added refers to how many average grades higher or lower the students in that school achieve across eight qualifying GCSE subjects compared to students across the country who score comparatively at age 11 (cubic). Standard errors in parentheses, clustered at the school level. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Detailed Gelbach decomposition of grade gaps by ethnic groups for GCSE English

	Pakistani & Bangladeshi			Indian			Black African			Black Caribbean		
	2019	2020	Change	2019	2020	Change	2019	2020	Change	2019	2020	Change
Raw gap	0.222*** (0.034)	0.151*** (0.032)	-0.072*** (0.020)	0.987*** (0.043)	0.948*** (0.039)	-0.039 (0.027)	0.372*** (0.031)	0.261*** (0.027)	-0.111*** (0.023)	-0.419*** (0.036)	-0.349*** (0.030)	0.070** (0.035)
Explained gap	-0.194*** (0.030)	-0.129*** (0.027)	0.066*** (0.018)	0.480*** (0.042)	0.530*** (0.039)	0.050** (0.022)	-0.073** (0.029)	-0.019 (0.025)	0.054*** (0.019)	-0.403*** (0.030)	-0.322*** (0.026)	0.080*** (0.024)
Unexplained gap	0.416*** (0.020)	0.279*** (0.018)	-0.137*** (0.020)	0.506*** (0.022)	0.418*** (0.018)	-0.088*** (0.024)	0.445*** (0.019)	0.280*** (0.016)	-0.164*** (0.021)	-0.017 (0.025)	-0.027 (0.022)	-0.010 (0.030)
<i>Amount explained by:</i>												
Male	0.001 (0.007)	0.004 (0.006)	0.003 (0.003)	-0.008 (0.007)	-0.003 (0.006)	0.005 (0.004)	0.014* (0.007)	0.010* (0.005)	-0.004 (0.004)	0.006 (0.007)	0.0092* (0.005)	0.003 (0.005)
Family & neighbourhood deprivation	-0.210*** (0.007)	-0.187*** (0.006)	0.023*** (0.004)	-0.016** (0.007)	0.001 (0.006)	0.018*** (0.004)	-0.289*** (0.006)	-0.255*** (0.006)	0.035*** (0.005)	-0.273*** (0.007)	-0.245*** (0.006)	0.027*** (0.006)
Special educational needs	0.015*** (0.002)	0.021*** (0.002)	0.005** (0.002)	0.036*** (0.002)	0.045*** (0.002)	0.009*** (0.002)	0.015*** (0.002)	0.022*** (0.002)	0.006** (0.002)	-0.027*** (0.004)	-0.028*** (0.004)	-0.001 (0.005)
First language not English	0.118*** (0.008)	0.119*** (0.007)	0.001 (0.010)	0.100*** (0.007)	0.100*** (0.006)	0.000 (0.008)	0.089*** (0.006)	0.088*** (0.006)	-0.002 (0.007)	0.003*** (0.001)	0.003*** (0.001)	0.000 (0.001)
Prior attainment (age 11)	-0.300*** (0.013)	-0.307*** (0.014)	-0.007 (0.010)	0.053** (0.023)	0.099*** (0.025)	0.046*** (0.014)	-0.174*** (0.013)	-0.150*** (0.014)	0.023** (0.011)	-0.312*** (0.014)	-0.285*** (0.015)	0.027* (0.016)
<i>School characteristics:</i>												
Selective admissions	0.000 (0.002)	0.000 (0.001)	0.000 (0.001)	0.034*** (0.006)	0.022*** (0.005)	-0.012** (0.005)	0.002 (0.002)	0.002 (0.001)	0.000 (0.001)	-0.008*** (0.001)	-0.005*** (0.001)	0.003*** (0.001)
Region	0.049*** (0.007)	0.035*** (0.006)	-0.014** (0.005)	0.049*** (0.007)	0.028*** (0.006)	-0.020*** (0.006)	0.106*** (0.011)	0.075*** (0.009)	-0.031*** (0.010)	0.113*** (0.012)	0.083*** (0.011)	-0.030** (0.012)
Urban	0.003 (0.002)	0.001 (0.002)	-0.002 (0.003)	0.003 (0.002)	0.001 (0.002)	-0.002 (0.003)	0.003 (0.002)	0.001 (0.002)	-0.002 (0.003)	0.003 (0.002)	0.001 (0.002)	-0.002 (0.003)
Size of cohort	-0.001 (0.001)	-0.003* (0.002)	-0.002 (0.001)	-0.001 (0.001)	-0.003 (0.002)	-0.002 (0.002)	0.001 (0.001)	0.003*** (0.001)	0.002** (0.001)	0.002 (0.001)	0.007*** (0.002)	0.005*** (0.002)
School governance type	-0.001 (0.002)	-0.001 (0.001)	0.000 (0.002)	0.000 (0.001)	0.001 (0.001)	0.000 (0.001)	-0.003* (0.002)	-0.002 (0.001)	0.001 (0.001)	-0.003*** (0.001)	-0.002* (0.001)	0.000 (0.001)
Proportion free school meals	-0.061*** (0.015)	-0.002 (0.013)	0.059*** (0.017)	-0.007 (0.005)	0.004 (0.003)	0.011** (0.004)	-0.050*** (0.014)	0.000 (0.011)	0.050*** (0.014)	-0.062*** (0.015)	-0.002 (0.013)	0.060*** (0.017)
Average neighbourhood deprivation	0.084*** (0.015)	0.071*** (0.012)	-0.013 (0.016)	0.034*** (0.006)	0.030*** (0.005)	-0.004 (0.006)	0.088*** (0.015)	0.073*** (0.013)	-0.015 (0.017)	0.094*** (0.017)	0.080*** (0.014)	-0.014 (0.018)
School value-added (lagged)	0.108*** (0.013)	0.120*** (0.012)	0.012 (0.009)	0.203*** (0.016)	0.204*** (0.015)	0.001 (0.010)	0.125*** (0.011)	0.114*** (0.011)	-0.011 (0.008)	0.061*** (0.013)	0.063*** (0.012)	0.002 (0.009)

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. The gap is the gap in average grades between the ethnic minority group and White British students. Family & neighbourhood deprivation includes free school meals indicator and a rank based on the proportion of children under age 16 in a local area living in a low-income household score (cubic). Prior attainment (age 11) includes subject-specific age 11 attainment score, standardised by year (quadratic). Region has 9 categories. School governance type has 4 categories. Proportion free school meals is a cubic. Average neighbourhood deprivation is a cubic. School value-added refers to how many average grades higher or lower the students in that school achieve across eight qualifying GCSE subjects compared to students across the country who score comparatively at age 11 (cubic). Standard errors in parentheses, clustered at the school level. * Significant at 10%; ** significant at 5%; *** significant at 1%.

6.3 Investigating the changes associated with region

To investigate the changes associated with region, we calculate the average grades in maths and English by year and region. The results are displayed in Table 7. These show that, across the regions, grade increases in 2020 are smaller for English than for maths. Of particular note is the case of London, which, despite having higher average grades in both subjects in 2019, has the second largest grade increase in maths but the second lowest in English in between 2019 and 2020.

Table 7: Mean grades by demographic group

	GCSE Maths			GCSE English		
	2019	2020	Change	2019	2020	Change
All	4.752 (2.034)	5.063 (1.992)	0.312*** (0.006)	5.210 (1.902)	5.405 (1.824)	0.195*** (0.006)
London	5.051 (2.121)	5.401 (2.035)	0.350*** (0.016)	5.620 (1.895)	5.775 (1.795)	0.155*** (0.016)
East of England	4.755 (2.007)	5.063 (1.976)	0.308*** (0.015)	5.172 (1.884)	5.380 (1.807)	0.208*** (0.016)
North East	4.619 (2.016)	4.923 (1.981)	0.305*** (0.016)	5.064 (1.892)	5.273 (1.829)	0.208*** (0.015)
North West	4.609 (1.996)	4.930 (1.976)	0.321*** (0.014)	5.107 (1.907)	5.311 (1.829)	0.204*** (0.015)
South East	4.905 (2.040)	5.161 (2.005)	0.256*** (0.014)	5.299 (1.925)	5.442 (1.853)	0.142*** (0.015)
South West	4.712 (1.991)	5.067 (1.943)	0.355*** (0.027)	5.122 (1.896)	5.391 (1.812)	0.270*** (0.027)
East Midlands	4.675 (2.001)	4.966 (1.961)	0.291*** (0.019)	5.087 (1.848)	5.298 (1.778)	0.211*** (0.019)
West Midlands	4.587 (2.012)	4.916 (1.977)	0.329*** (0.017)	5.078 (1.868)	5.289 (1.801)	0.211*** (0.018)
Observations	455614	489766		455614	489766	

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. Standard deviations in parentheses in the ‘2019’ and ‘2020’ columns, standard errors in ‘Change’. * Significant at 10%; ** significant at 5%; *** significant at 1%.

One possible explanation of this pattern of results is ceiling effects. As the London grade average is considerably higher than that in other regions for English, it could be that students educated in London have less growth potential in 2020 and therefore make smaller gains. Such effects would have a negative contribution to the attainment of ethnic minority students in aggregate because they are significantly more likely to live in London than their White British peers – this is the case for about 30 percent of Pakistani and Bangladeshi and Indian students, and about 60 percent of Black African and Black Caribbean students.

A second possible explanation is that ethnic minority groups receive systematically lower grades in English in 2020 and that this lowers London’s average grade increase in aggregate.

To explore this possibility, in Appendices 4a and 4b we repeat the analyses from Table 7 but further broken down by ethnic group. The results indeed show that, across virtually all groups and regions, grade increases are higher for ethnic minority students than for White British students in maths, and lower in English. Moreover, this occurs despite Pakistani and Bangladeshi, Indian, and Black African students often achieving higher grades than White British students in both subjects. Nonetheless the grade increases for ethnic minority students in London are still at the lower end of those across all regions in English, and their average grades high. This suggests that a combination of ethnic minority groups receiving systematically lower grades and ceiling effects may be accounting for some of patterns that we see in English in 2020 compared to 2019.

6.4 Summary of results

The decomposition results indicate that there are significant changes in the unexplained components of the grade gaps between ethnic minority students and their White British peers in 2020 compared to 2019. These changes are positive for all ethnic minority groups for maths but negative for all ethnic minority groups apart from Black Caribbean students for English. This means that a large contribution of the pattern of positive relative grade changes in maths and negative relative grade changes in English for ethnic minority groups in 2020 cannot be explained by differences in either the levels or returns to the observed covariates across the years. Instead, these unexplained changes are likely to be at least partly effects of the use of teacher predictions of student examination performance in place of blindly marked examinations.

Region is the only covariate to contribute to the grade gap changes due to differences in returns rather than differences in the levels of the covariates between groups across cohorts, with negative returns for ethnic minority students in English. These regional effects appear to be due to a combination of ethnic minority groups receiving systematically lower grades in English across regions, and ceiling effects occurring in some high-achieving regions such as London. However, as ethnic minority students would have been expected to attain more highly without them, these regional effects do not diminish the consistently negative and significant unexplained components of the English gap changes.

7 Robustness

First, rather than restricting the sample to complete cases, we consider a sample allowing for missing covariates. When we run the descriptive analyses on this sample we see the same pattern of raw gaps but with generally larger changes than those presented above. It is well established that disadvantaged students are more likely to be missing information in administrative data (e.g. Burgess and Greaves, 2013). Given that prior attainment is one of the considered covariates, excluded observations will also contain students who entered the English schooling system after the age of eleven. Although we cannot establish the link concretely, it seems plausible that students with missing data are also those for whom any differential teacher predictions of examination performance may be more potent.

Second, we re-run the decomposition analyses separately for English Language and English Literature, rather than the higher of the two. Appendices 5a and 5b show the detailed decomposition results. These are very similar to the combined results presented in Table 6 above, with the signs and significance of the explained and unexplained components across both outcomes as expected albeit with some variation in magnitude. However, the results do differ in two aspects. First, the changes in the raw grade gaps are smaller for English Language than English Literature, and in some cases no longer statistically significant. Second, the contribution of region is smaller than that in Table 6, and more so for English Language. Appendix 6 shows the mean grades for both outcomes and by demographic group. The grades for White British students across the two outcomes are comparable. The grades for all ethnic minority groups, however, are higher for English Literature than for English Language. This is likely to explain why the grade gap changes using the combined outcome are more similar to those for English Literature than for English Language, as this is the outcome which more likely to be the higher of the two grades for ethnic minority students.⁷

⁷Establishing why there are differences in ethnic minority students' grades for English Literature and English Language is beyond the limitations of the data available to us. However a few options are plausible. First, the skills assessed in English Literature may be different to those assessed in English Language in a way that matters for teachers' predictions of students' examination performance across ethnic groups. Second, teachers may have been more tentative to award ethnic minority students considerably higher grades in English Literature than in English Language in 2020, for example because they felt that it would appear an unusual result when the grades were checked by the national regulatory body. Finally, the English Language examinations may be more culturally biased than English Literature examinations.

Third, as most of our ethnic minority groups of focus generally attain more highly than their White British peers, it could be that ceiling effects are driving our results. However we present two pieces of evidence to suggest that ceiling effects alone are not driving the overall findings. First, the tables in Appendix 4 show that the patterns of positive relative grade changes in maths and negative relative grade changes in English are not generally sensitive to whether the average grade attainment of ethnic groups in 2019 was higher, similar to, or lower than that for White British students across both subjects and regions. Second, following Wyness and Murphy (2020), we partition the sample into three groups according to their prior attainment. Appendix 7 shows graphs of the raw grade gaps for these groups, and illustrates that the patterns presented above remain across these them. Last, we briefly consider why the pattern of negative relative grade changes for ethnic minority groups in English is not apparent for Black Caribbean students; this group instead experience a slightly positive relative grade change in 2020 compared to 2019, significant at the ten percent level. Appendix 8 shows the distribution of grades by subject and year and additionally by ethnic group. The distributions for the grades received by Black Caribbean students in 2017, 2018, and 2019 are slightly more variable than those of the other ethnic minority groups. As there are between two and five times fewer Black Caribbean students than any other ethnic minority group of focus, sample size and imprecision could be a contributing factor to the anomalous English result in 2020. Black Caribbean students are also the only ethnic group to be right-skewed in their grades in both subjects. Most ethnic minority groups appear to experience smaller grade increases than White British students at the top end of the grade distribution for English, and larger ones lower down. As Black Caribbean students' grades are more likely to be in these lower parts of the grade distribution, they may therefore experience more positive grade changes than the other groups in aggregate. Both sample size and factors related to where Black Caribbean students sit in the grade distributions, then, may contribute to their positive relative grade changes across both subjects.

8 Alternative explanations

We find that teachers in England seem to have higher predictions of ethnic minority students' examination performance in maths and lower predictions of ethnic minority students' examination performance in English compared to their White British peers. In the literature, subject- and group-specific stereotyping is found to be the most likely explanation for differences in ethnic minority students' grades in teacher assessments compared to blindly marked examinations (Campbell, 2015; Burgess and Greaves, 2013). It is beyond the scope of the data available to us to establish whether stereotype formation is the correct mechanism for any of the findings presented in the context considered here. However, we briefly consider three alternative explanations.

First, the grades that teachers predict students based on information from low-stakes classroom assessments may be systematically different to those that students receive from one-off high-stakes examinations. In 2020, although asked to predict how their students would have performed had the examinations gone ahead, teachers also had to supply evidence for their predictions based on marks from coursework, class work, class tests, and mock examinations (Holmes et al., 2021). Studies using data gathered from undergraduate students in the United States indicate that non-white students may perform worse on high-stakes compared to low-stakes assessments (e.g. Kader, 2016). If this occurred in the case of GCSE students in England, we would see ethnic minority students receive higher teacher predicted grades in 2020 compared to White British students and compared to those that they had received in 2019. Although this may go some way to explaining the results that we see here for maths, it seems unlikely that such effects would work in contrary directions across subjects. We therefore do not think that this is a compelling explanation for the findings.

Second, teachers' use of prior attainment data to predict grades in 2020 may have informed the trends presented here. The school attainment of most ethnic minority groups in England improves at a greater rate relative to White British students throughout compulsory schooling, with speaking English as an additional language the single most important factor in this differential (Dustmann et al., 2010). All ethnic minority groups apart from Black Caribbean students also attain lower in age 11 English assessments than in maths. If

ethnic minority students generally make greater progress in English-related subjects than in maths between the ages of eleven and sixteen, then schools or specific teachers who treat age eleven attainment measures as accurate information for age sixteen attainment across subjects and across student groups may therefore be more likely to predict ethnic minority students lower grades in English than in maths. However, if language is the main driver of these relative improvements, we would expect the covariate indicating whether students have a first language other than English to account for a significant proportion in the changes to grade gaps between 2019 and 2020. Contrary to this, Table 4 and Table 5 indicate that language actually plays a very minor role in accounting for the grade gap changes.

Finally, there is a large and growing literature on the effect of in-group preferences – also known as student-teacher matching effects – on teachers’ judgements. Studies from the US, for example, have consistently found that racial, ethnic, and gender differences between students and teachers inform teachers’ perceptions about students’ potential (Papageorge et al., 2018; Gershenson et al., 2016) and behaviour (Lindsay and Hart, 2017; Dee, 2007). Although we did have access to school-level data including a measure of the proportion of teachers at a school who identified as belonging to an ethnic minority group, we did not include this measure in the final analyses for two reasons. First, school-level teacher workforce proportions are a very imprecise proxy for the precise student-teacher matching indicators that are used in the papers discussed above. The data did not, for example, allow us to discern whether students were more likely to be taught by an ethnic minority teacher in maths or in English, which is information which we would need to argue that any such effects are related to our present findings. Second, when we did run the analyses using these aggregate measures, we found no substantial results.

As none of these alternatives present a compelling mechanism for the results presented above, we conclude that differential teacher predictions for the examination performance of different ethnic groups across subjects – perhaps driven by subject- and group-specific stereotyping – is the most likely explanation for at least some part of the grade gap changes observed.

9 Conclusion

In this paper we exploit an exogenous change in national assessment methods in England to examine whether teachers have different predictions for the examination performance of students with different characteristics. In particular, we examine whether there may be evidence of differential teacher predictions for students from different ethnic groups. The literature comparing teacher assessments with blindly marked examination grades for students in England paints a remarkably consistent story, with students from ethnic minority groups found to receive lower teacher assessed grades than examination grades in English and either similar or higher teacher assessed grades than examination grades in maths. Our results are not directly comparable with this literature as they examine teachers' predictions of students' examination performance rather than a separate, intentionally teacher assessed measure of students' skills. Nonetheless they are congruent with this general pattern, with raw grade gaps in 2020 indicating positive relative grade changes for all of our ethnic minority groups of focus in maths and, for all apart from Black Caribbean students, negative relative grade changes in English.

We explore different explanations which may account for these raw gap changes. They do not appear to be the result of trends in gap changes over time, or differences in the levels of observed covariates between the cohorts. In fact, changes in the characteristics of groups across the cohorts would suggest that – *ceteris paribus* – ethnic minority students would have received higher relative grades in 2020, and more so in the case of English than maths. Contrary to this, the ethnic minority groups generally experience negative relative grade changes in English. We do find some indication that ceiling effects may play a part in explaining the slightly smaller grade increases for ethnic minority groups than White British students in English, related to ethnic minority students' high achievement in 'normal' years. Nonetheless large parts of the gap changes remain unexplained. We consider alternative mechanisms but conclude that patterns in differential teacher predictions for the examination performance of different ethnic groups across subjects is a convincing explanation of at least some part of the results.

Due to the observational nature of this study, we do not claim causal identification of differential teacher predictions in the results above. In addition, this study inevitably has

significant limitations; we cannot, for example, control for non-cognitive characteristics related to attitudes to learning that may be correlated with grades for both blindly marked examinations and teacher predictions, nor are we able to fully account for the possible impacts of ceiling effects on the grade distributions in 2020. Nonetheless from a policy perspective these results are at the very least a useful exploration of some of the likely effects of using teacher predictions in the context of national examinations in England. Overall the results suggest that, due to a combination of grade distribution restrictions and differential teacher predictions of examination performance across ethnic groups, the use of teacher predictions may see ethnic minority students receive higher grades relative to White British students in maths but lower grades relative to White British students in English than would otherwise be the case.

References

- Alesina, A., Carlana, M., La Ferrara, E. and Pinotti, P. (2019). ‘Revealing stereotypes: Evidence from immigrants in schools’. CEPR Press Discussion Paper 13555.
- Baird, J. (1998). ‘What’s in a name? Experiments with blind marking in A-level examinations’. *Educational Research* 40(2): 191-202.
- Blanden, J., Greaves, E., Gregg, P., Macmillan, L. and Sibiet, L. (2015). Understanding the improved performance of disadvantaged pupils in London. LSE CASE Working Paper 21.
- Blinder, A. (1973). ‘Wage discrimination: Reduced form and structural estimates’. *Journal of Human Resources* 8: 436-455.
- Blunkett, D., Rowan, K., Nargund, P. and Sandby-Thomas, R. (2022). Learning and skills for economic recovery, social cohesion and a more equal Britain.
- Borcan, O., Lindahl, M. and Mitrut, A. (2017). ‘Fighting corruption in education: What works and who benefits?’. *American Economic Journal: Economic Policy* 9(1): 180-209.
- Burgess, S. and Greaves, E. (2013). ‘Test scores, subjective assessment, and stereotyping of ethnic minorities’. *Journal of Labor Economics* 31(3): 535-76.
- Campbell, T. (2015). ‘Stereotyped at Seven? Biases in Teacher Judgement of Students’ Ability and Attainment’. *Journal of Social Policy* 44(3): 517-47.
- Dee, T. (2007). ‘Teachers and the Gender Gaps in Student Achievement’. *The Journal of Human Resources* 42(3): 528-54.
- Department for Education [DfE] (2020a). Ad-hoc notice: Early entry into GCSE examinations in England.
- Department for Education [DfE] (2020b). Secondary accountability measures: Guide for maintained secondary schools, academies and free schools.
- Dhillon, D. (2005). ‘Teachers’ estimates of candidates’ grades: Curriculum 2000 advanced level qualifications’. *British Educational Research Journal* 31(1): 69- 88.
- Diamond, R., Persson, P. (2016). The long-term consequences of teacher discretion in grading of highstakes tests. NBER Working Paper 22207.
- Dorsey, J. and J. Colliver (1995). ‘Effect of Anonymous Test Grading on Passing Rates as Related to Gender and Race’. *Academic Medicine* 70(4): 321-23.
- Duncan, K. and Sandy, J. (2013). ‘Using the Blinder-Oaxaca decomposition method to measure racial bias in achievement tests’. *The Review of Black Political Economy* 40(2): 185-206.
- Dustmann, C., Machin, S. and Schonberg, U. (2010). ‘Ethnicity and educational achievement in compulsory schooling’. *The Economic Journal* 120: 272-297.
- Gelbach, J. (2016). ‘When do covariates matter? And which ones, and how much?’. *Journal of Labor Economics* 34(2): 509-543.
- Gershenson, S., Holt, S. and Papageorge, N. W. (2016). ‘Who Believes in Me? The Effect of Student-Teacher Demographic Match on Teacher Expectations’. *Economics of*

Education Review 52: 209-24.

Gibbons, S. and Chevalier, A. (2008). 'Assessment and age 16+ education participation'. Research Papers in Education 23(2): 113–23.

Holmes, S., Churchward, D., Howard, E., Keys, E., Leahy, F. and Tonin, D. (2021). Centre Assessment Grades: Teaching Staff Interviews, Summer 2020. Ofqual Report.

Hunt, E., Tuckett, S., Robinson, D., Hutchinson, J. and Coleman, C. (2022) Covid-19 and Disadvantage Gaps in England 2020. Education Policy Institute.

Joint Council for Qualifications [JCQ] (2020). GCSE (Full Course) outcomes for main grade set for each jurisdiction (Age 16): Results Summer 2020.

Kader, A. (2016). 'Debilitating and facilitating test anxiety and student motivation and achievement in principles of microeconomics'. International Review of Economics Education 23: 40-46.

Lavy, V. and Megalokonomou, R. (2019). Persistency in Teachers' Grading Biases and Effect on Longer Term Outcomes: University Admission examinations and Choice of Field of Study. NBER Working Paper 26021.

Lavy, V. and Sand, E. (2018). 'On the Origins of Gender Human Capital Gaps: Short and Long Term Consequences of Teachers' Biases'. Journal of Public Economics 167(C): 263–79.

Lindahl, E. (2007). Comparing teachers' assessments and national test results—Evidence from Sweden. IFAU working paper: 24.

Lindsay, C. and Hart, C. (2017). 'Exposure to same-race teachers and student disciplinary outcomes for Black students in North Carolina'. Educational Evaluation and Policy Analysis. 39(3): 485-510.

Machin, S., McNally, S., and Ruiz-Valenzuela, J. (2020). 'Entry through the narrow door: The costs of just failing high stakes examinations'. Journal of Public Economics 190.

Oaxaca, R. (1973). 'Male-female wage differentials in urban labor markets'. International Economic Review 14: 693-709.

Office of Qualifications and Examinations Regulation [Ofqual] (2020a). Appeals for GCSE, AS, A Level and Project: 2018 to 2019 academic year.

Office of Qualifications and Examinations Regulation [Ofqual] (2020b). Provisional Entries for GCSE, AS and A level: Summer 2020 exam series.

Papageorge, N., Gershenson, S. and Kang, K. (2018). Teacher expectations matter. NBER Working Paper 25255.

Reeves, D. J., Boyle, W. F. and Christie, T. (2001). 'The Relationship between Teacher Assessments and Student Attainments in Standard Test Tasks at Key Stage 2, 1996 – 98'. British Educational Research Journal 27(2): 141-60.

Sammons, P. (2007). School Effectiveness and Equity: Making Connections. CfBT Education Trust.

Wei Lee, M., Stringer, N. and Zanini, N. (2020). Student-level equities analyses for GCSE and A Level. Ofqual Research and Analysis.

Wyness, G. and Murphy, E. (2020). 'Minority Report: the impact of predicted grades on university admissions of disadvantaged groups'. *Education Economics* 28(4): 333-50.

Appendices

Appendix 1

Covariate balance table by year

	2017	2018	2019	2020
Pupil characteristics				
Male	0.498 (0.500)	0.500 (0.500)	0.501 (0.500)	0.502 (0.500)
Ethnicity:				
White British	0.749 (0.434)	0.732 (0.443)	0.719 (0.450)	0.711 (0.453)
Pakistani & Bangladeshi	0.056 (0.230)	0.061 (0.240)	0.062 (0.242)	0.060 (0.238)
Indian	0.024 (0.155)	0.027 (0.162)	0.027 (0.163)	0.027 (0.162)
Black African	0.029 (0.169)	0.031 (0.174)	0.033 (0.179)	0.035 (0.183)
Black Caribbean	0.013 (0.114)	0.013 (0.115)	0.014 (0.117)	0.013 (0.114)
Multiethnic	0.044 (0.205)	0.046 (0.210)	0.049 (0.217)	0.052 (0.222)
Any Other Ethnic Group	0.075 (0.263)	0.078 (0.269)	0.084 (0.277)	0.089 (0.285)
Free school meals	0.253 (0.435)	0.249 (0.433)	0.248 (0.432)	0.244 (0.430)
Neighbourhood deprivation	0.192 (0.138)	0.192 (0.137)	0.195 (0.139)	0.196 (0.139)
Special educational needs	0.109 (0.311)	0.110 (0.312)	0.114 (0.318)	0.120 (0.325)
First language not English	0.131 (0.337)	0.137 (0.344)	0.141 (0.348)	0.141 (0.348)
Maths prior attainment (age 11)	0.000 (1.000)	0.000 (1.000)	0.000 (1.000)	0.000 (1.000)
English prior attainment (age 11)	0.000 (1.000)	0.000 (1.000)	0.000 (1.000)	0.000 (1.000)
School characteristics				
Selective admissions	0.044 (0.206)	0.047 (0.212)	0.046 (0.209)	0.046 (0.209)
Region:				
London	0.138 (0.345)	0.141 (0.348)	0.144 (0.351)	0.143 (0.351)
East of England	0.117 (0.322)	0.117 (0.321)	0.117 (0.322)	0.115 (0.319)
North East	0.150 (0.357)	0.149 (0.356)	0.147 (0.355)	0.148 (0.355)
North West	0.141 (0.348)	0.139 (0.346)	0.142 (0.349)	0.139 (0.346)
South East	0.160 (0.367)	0.161 (0.368)	0.164 (0.370)	0.162 (0.368)
South West	0.098 (0.297)	0.095 (0.293)	0.084 (0.277)	0.092 (0.289)
East Midlands	0.084 (0.278)	0.085 (0.279)	0.088 (0.283)	0.087 (0.282)
West Midlands	0.112 (0.315)	0.113 (0.317)	0.113 (0.317)	0.114 (0.318)
Urban	0.876 (0.330)	0.877 (0.328)	0.876 (0.329)	0.874 (0.332)
Size of cohort	190 (62.7)	188 (63.8)	191 (62.6)	196 (63.0)
School governance type:				
Local Authority Maintained	0.335 (0.472)	0.298 (0.457)	0.264 (0.441)	0.244 (0.429)
Single Academy Trust	0.185 (0.388)	0.192 (0.393)	0.196 (0.397)	0.192 (0.394)
Multi Academy Trust	0.417 (0.493)	0.456 (0.498)	0.509 (0.500)	0.541 (0.498)
Other	0.062 (0.242)	0.055 (0.229)	0.031 (0.174)	0.023 (0.151)
Proportion free school meals	0.252 (0.151)	0.247 (0.148)	0.244 (0.144)	0.240 (0.140)
Average neighbourhood deprivation	0.193 (0.088)	0.193 (0.088)	0.195 (0.089)	0.196 (0.089)
School value-added (lagged)	0.020 (0.337)	0.029 (0.409)	0.028 (0.430)	0.026 (0.435)
Observations	452451	443818	455614	489766

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Free school meals indicates if a student is known to have been eligible for free school meals in the past six years. Neighbourhood deprivation is a rank based on the proportion of children under age 16 in a local area living in a low-income household. School value-added refers to how many average grades higher or lower the students in that school achieve across eight qualifying GCSE subjects compared to students across the country who score comparatively at age 11. Standard deviations in parentheses.

Appendix 2

Grade gaps by demographic group with combined 2018/19 control

	GCSE Maths			GCSE English		
	2018/19	2020	Change	2018/19	2020	Change
Male	0.067*** (0.018)	0.004 (0.017)	-0.063*** (0.009)	-0.795*** (0.015)	-0.767*** (0.015)	0.028*** (0.008)
Pakistani & Bangladeshi	0.003 (0.034)	0.137*** (0.034)	0.134*** (0.021)	0.194*** (0.032)	0.149*** (0.032)	-0.045** (0.019)
Indian	1.165*** (0.049)	1.314*** (0.048)	0.149*** (0.025)	0.969*** (0.041)	0.947*** (0.039)	-0.022 (0.024)
Black African	0.105*** (0.029)	0.204*** (0.028)	0.098*** (0.021)	0.384*** (0.029)	0.260*** (0.027)	-0.124*** (0.021)
Black Caribbean	-0.878*** (0.030)	-0.694*** (0.031)	0.184*** (0.030)	-0.397*** (0.030)	-0.350*** (0.030)	0.046 (0.031)
Free school meals	-1.196*** (0.013)	-1.177*** (0.013)	0.019** (0.009)	-1.028*** (0.013)	-1.025*** (0.013)	0.003 (0.009)
Observations	899432	489766		899432	489766	

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. Free school meals indicates if a student is known to have been eligible for free school meals in the past six years. The reference groups are female, White British, and no free school meals. Standard errors in parentheses, clustered at the school-level. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix 3

Detailed Gelbach decomposition of grade gaps by ethnic groups for GCSE maths

	Pakistani & Bangladeshi			Indian			Black African			Black Caribbean		
	2019	2020	Change	2019	2020	Change	2019	2020	Change	2019	2020	Change
Raw gap	0.046 (0.036)	0.138*** (0.034)	0.092*** (0.023)	0.1217*** (0.050)	0.1315*** (0.048)	0.098*** (0.028)	0.099*** (0.031)	0.205*** (0.028)	0.105*** (0.023)	-0.884*** (0.034)	-0.693*** (0.031)	0.191*** (0.035)
Explained gap	-0.153*** (0.032)	-0.099*** (0.029)	0.054** (0.021)	0.812*** (0.049)	0.805*** (0.046)	-0.007 (0.025)	-0.186*** (0.030)	-0.188*** (0.027)	-0.002 (0.021)	-0.720*** (0.031)	-0.673*** (0.028)	0.047 (0.029)
Unexplained gap	0.199*** (0.017)	0.238*** (0.017)	0.039** (0.019)	0.405*** (0.019)	0.510*** (0.018)	0.105*** (0.020)	0.286*** (0.017)	0.393*** (0.015)	0.107*** (0.018)	-0.164*** (0.022)	-0.020 (0.020)	0.144*** (0.026)
<i>Amount explained by:</i>												
Male	0.000 (0.001)	0.002 (0.003)	0.002 (0.002)	-0.001 (0.001)	-0.001 (0.003)	0.000 (0.002)	0.002*** (0.001)	0.005* (0.003)	0.003 (0.002)	0.001 (0.001)	0.005* (0.003)	0.004* (0.002)
Family & neighbourhood deprivation	-0.191*** (0.006)	-0.182*** (0.006)	0.009** (0.004)	-0.017*** (0.006)	-0.002 (0.006)	0.015*** (0.003)	-0.262*** (0.006)	-0.246*** (0.006)	0.016*** (0.005)	-0.246*** (0.006)	-0.237*** (0.006)	0.009* (0.006)
Special educational needs	0.008*** (0.001)	0.013*** (0.001)	0.004*** (0.001)	0.020*** (0.001)	0.027*** (0.001)	0.008*** (0.001)	0.008*** (0.001)	0.013*** (0.001)	0.005*** (0.001)	-0.015*** (0.002)	-0.017*** (0.003)	-0.003 (0.003)
First language not English	0.102*** (0.007)	0.108*** (0.007)	0.006 (0.008)	0.087*** (0.006)	0.091*** (0.006)	0.004 (0.007)	0.077*** (0.006)	0.080*** (0.005)	0.002 (0.006)	0.003*** (0.000)	0.003*** (0.000)	0.000 (0.001)
Prior attainment (age 11)	-0.168*** (0.022)	-0.156*** (0.020)	0.012 (0.015)	0.496*** (0.036)	0.468*** (0.035)	-0.028 (0.020)	-0.143*** (0.020)	-0.172*** (0.019)	-0.029* (0.016)	-0.521*** (0.022)	-0.505*** (0.021)	0.016 (0.025)
<i>School characteristics:</i>												
Selective admissions	0.000 (0.001)	0.000 (0.001)	0.000 (0.000)	0.018*** (0.005)	0.013*** (0.004)	-0.005 (0.004)	0.001 (0.001)	0.001 (0.001)	0.000 (0.001)	-0.004*** (0.001)	-0.003*** (0.001)	0.001 (0.001)
Region	-0.001 (0.005)	0.004 (0.005)	0.005 (0.005)	0.006 (0.005)	0.003 (0.005)	-0.004 (0.005)	0.015 (0.009)	0.022*** (0.008)	0.007 (0.009)	0.016 (0.010)	0.023** (0.010)	0.007 (0.010)
Urban	0.001 (0.002)	0.003 (0.002)	0.002 (0.002)	0.001 (0.002)	0.003 (0.002)	0.002 (0.002)	0.001 (0.002)	0.003 (0.002)	0.002 (0.002)	0.001 (0.002)	0.003 (0.002)	0.002 (0.002)
Size of cohort	-0.001 (0.001)	-0.003* (0.002)	-0.002 (0.001)	-0.001 (0.001)	-0.003 (0.002)	-0.002 (0.002)	0.001 (0.001)	0.004*** (0.001)	0.003** (0.001)	0.002* (0.001)	0.007*** (0.002)	0.005*** (0.002)
School governance type	0.000 (0.001)	0.000 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.000 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.000 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.000 (0.001)
Proportion free school meals	-0.008 (0.015)	0.014 (0.011)	0.022 (0.015)	0.000 (0.003)	0.003* (0.002)	0.003 (0.003)	-0.008 (0.013)	0.011 (0.009)	0.020 (0.013)	-0.009 (0.015)	0.013 (0.011)	0.022 (0.016)
Average neighbourhood deprivation	-0.006 (0.014)	-0.025** (0.011)	-0.019 (0.014)	-0.008 (0.005)	-0.009** (0.004)	-0.001 (0.005)	-0.005 (0.014)	-0.026** (0.012)	-0.021 (0.015)	-0.006 (0.015)	-0.029** (0.013)	-0.022 (0.016)
School value-added (lagged)	0.111*** (0.012)	0.124*** (0.012)	0.014 (0.008)	0.213*** (0.016)	0.214*** (0.015)	0.001 (0.009)	0.127*** (0.011)	0.118*** (0.011)	-0.009 (0.008)	0.061*** (0.013)	0.065*** (0.012)	0.004 (0.009)

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. The gap is the gap in average grades between the ethnic minority group and White British students. Family & neighbourhood deprivation includes free school meals indicator and a rank based on the proportion of children under age 16 in a local area living in a low-income household score (cubic). Prior attainment (age 11) includes subject-specific age 11 attainment score, standardised by year (quadratic). Region has 9 categories. School governance type has 4 categories. Proportion free school meals is a cubic. Average neighbourhood deprivation is a cubic. School value-added refers to how many average grades higher or lower the students in that school achieve across eight qualifying GCSE subjects compared to students across the country who score comparatively at age 11 (cubic). Standard errors in parentheses, clustered at the school level. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix 4

4a. Mean grades by region, group and year for GCSE maths

	White British			Pakistani & Bangladeshi			Indian			Black African			Black Caribbean		
	2019	2020	Change	2019	2020	Change	2019	2020	Change	2019	2020	Change	2019	2020	Change
All	4.684 (1.995)	4.974 (1.959)	0.290*** (0.007)	4.729 (2.059)	5.111 (2.011)	0.382*** (0.022)	5.900 (2.073)	6.287 (1.964)	0.388*** (0.027)	4.782 (1.997)	5.177 (1.908)	0.395*** (0.023)	3.799 (1.828)	4.279 (1.784)	0.480*** (0.034)
London	5.010 (2.100)	5.287 (2.031)	0.276*** (0.029)	5.247 (2.048)	5.628 (1.967)	0.381*** (0.037)	6.076 (2.109)	6.432 (1.982)	0.356*** (0.047)	4.811 (1.998)	5.203 (1.901)	0.392*** (0.029)	3.840 (1.844)	4.350 (1.782)	0.510*** (0.043)
East of England	4.702 (1.967)	4.977 (1.936)	0.276*** (0.016)	4.653 (2.084)	5.007 (2.056)	0.354*** (0.074)	6.392 (2.013)	6.881 (1.845)	0.489*** (0.110)	4.921 (2.033)	5.395 (1.908)	0.473*** (0.088)	3.846 (1.724)	4.519 (1.773)	0.673*** (0.141)
North East	4.620 (2.001)	4.914 (1.967)	0.294*** (0.016)	4.366 (1.988)	4.796 (1.967)	0.430*** (0.056)	5.723 (1.949)	5.972 (1.874)	0.248*** (0.105)	4.544 (1.940)	4.881 (1.783)	0.337*** (0.102)	3.684 (1.801)	3.914 (1.818)	0.230 (0.190)
North West	4.565 (1.979)	4.871 (1.958)	0.306*** (0.017)	4.624 (1.973)	4.970 (1.948)	0.346*** (0.046)	5.731 (1.990)	6.308 (1.847)	0.578*** (0.077)	4.537 (1.945)	5.091 (1.930)	0.553*** (0.091)	3.408 (1.865)	4.253 (1.778)	0.845*** (0.241)
South East	4.820 (2.004)	5.063 (1.973)	0.243*** (0.015)	4.919 (2.036)	5.309 (2.052)	0.390*** (0.064)	6.434 (2.015)	6.754 (1.856)	0.320*** (0.070)	5.005 (2.057)	5.349 (1.965)	0.344*** (0.079)	4.055 (1.900)	4.336 (1.754)	0.281* (0.148)
South West	4.696 (1.978)	5.044 (1.929)	0.348*** (0.027)	4.718 (2.010)	5.334 (2.011)	0.616*** (0.150)	5.701 (2.219)	6.202 (2.004)	0.501*** (0.173)	4.516 (1.965)	4.888 (1.962)	0.372*** (0.127)	3.345 (1.731)	4.203 (1.752)	0.858*** (0.224)
East Midlands	4.638 (1.975)	4.923 (1.937)	0.285*** (0.020)	4.769 (2.055)	5.078 (1.972)	0.309*** (0.083)	5.514 (2.054)	5.853 (1.987)	0.339*** (0.063)	4.639 (1.988)	4.909 (1.881)	0.270*** (0.096)	3.714 (1.694)	4.043 (1.694)	0.328** (0.142)
West Midlands	4.561 (1.973)	4.876 (1.941)	0.315*** (0.019)	4.399 (2.063)	4.767 (1.982)	0.368*** (0.049)	5.552 (2.013)	5.944 (1.972)	0.391*** (0.066)	4.666 (1.933)	5.062 (1.896)	0.396*** (0.078)	3.727 (1.797)	4.031 (1.790)	0.304*** (0.080)

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. Standard deviations in parentheses in the '2019' and '2020' columns, standard errors in 'Change'. * Significant at 10%; ** significant at 5%; *** significant at 1%.

4b. Mean grades by region, group and year for GCSE English

	White British			Pakistani & Bangladeshi			Indian			Black African			Black Caribbean		
	2019	2020	Change	2019	2020	Change	2019	2020	Change	2019	2020	Change	2019	2020	Change
All	5.116 (1.900)	5.326 (1.827)	0.210*** (0.007)	5.338 (1.837)	5.475 (1.754)	0.137*** (0.019)	6.103 (1.793)	6.273 (1.679)	0.170*** (0.027)	5.488 (1.777)	5.585 (1.683)	0.098*** (0.022)	4.697 (1.759)	4.976 (1.646)	0.279*** (0.035)
London	5.566 (1.970)	5.748 (1.881)	0.182*** (0.028)	5.861 (1.785)	5.986 (1.698)	0.124*** (0.032)	6.191 (1.822)	6.368 (1.675)	0.177*** (0.051)	5.550 (1.767)	5.624 (1.677)	0.074*** (0.028)	4.771 (1.777)	5.059 (1.639)	0.287*** (0.046)
East of England	5.109 (1.869)	5.317 (1.795)	0.207*** (0.017)	5.184 (1.864)	5.380 (1.755)	0.196*** (0.060)	6.493 (1.749)	6.577 (1.676)	0.084 (0.097)	5.603 (1.802)	5.720 (1.722)	0.117 (0.081)	4.595 (1.634)	5.095 (1.610)	0.500*** (0.137)
North East	5.049 (1.896)	5.267 (1.833)	0.218*** (0.017)	4.984 (1.786)	5.158 (1.709)	0.173*** (0.039)	5.852 (1.707)	6.047 (1.689)	0.195** (0.088)	5.278 (1.754)	5.285 (1.605)	0.007 (0.093)	4.531 (1.702)	4.483 (1.579)	-0.048 (0.161)
North West	5.061 (1.918)	5.265 (1.836)	0.204*** (0.017)	5.154 (1.789)	5.306 (1.697)	0.152*** (0.049)	6.136 (1.741)	6.370 (1.638)	0.234*** (0.076)	5.174 (1.734)	5.451 (1.680)	0.276*** (0.093)	4.352 (1.697)	4.892 (1.617)	0.541 (0.181)
South East	5.213 (1.915)	5.371 (1.852)	0.158*** (0.015)	5.410 (1.840)	5.519 (1.779)	0.109* (0.063)	6.492 (1.749)	6.541 (1.638)	0.049 (0.070)	5.660 (1.776)	5.706 (1.669)	0.046 (0.078)	4.787 (1.870)	4.947 (1.722)	0.160 (0.145)
South West	5.104 (1.893)	5.373 (1.811)	0.269*** (0.028)	5.342 (1.855)	5.439 (1.820)	0.098 (0.147)	5.960 (1.775)	6.408 (1.633)	0.449*** (0.141)	4.940 (1.679)	5.102 (1.705)	0.162 (0.130)	4.273 (1.752)	4.915 (1.588)	0.642*** (0.207)
East Midlands	5.036 (1.846)	5.254 (1.777)	0.217*** (0.021)	5.311 (1.749)	5.452 (1.663)	0.141* (0.076)	5.752 (1.761)	5.955 (1.647)	0.204*** (0.077)	5.322 (1.793)	5.471 (1.683)	0.149 (0.095)	4.660 (1.705)	4.800 (1.593)	0.140 (0.155)
West Midlands	4.988 (1.860)	5.243 (1.800)	0.255*** (0.021)	5.145 (1.844)	5.234 (1.766)	0.089* (0.046)	5.906 (1.787)	6.058 (1.694)	0.152*** (0.057)	5.325 (1.809)	5.517 (1.691)	0.192** (0.076)	4.577 (1.716)	4.757 (1.666)	0.180** (0.085)

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. Standard deviations in parentheses in the '2019' and '2020' columns, standard errors in 'Change'. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix 5

5a: Detailed Gelbach decomposition of grade gaps by ethnic groups for GCSE English Language

	Pakistani & Bangladeshi			Indian			Black African			Black Caribbean		
	2019	2020	Change	2019	2020	Change	2019	2020	Change	2019	2020	Change
Raw gap	0.078** (-0.031)	0.070** (-0.030)	-0.008 (-0.020)	0.837*** (-0.042)	0.857*** (-0.038)	0.020 (-0.028)	0.181*** (-0.029)	0.156*** (-0.026)	-0.026 (-0.023)	-0.534*** (-0.033)	-0.420*** (-0.029)	0.113*** (-0.034)
Explained gap	-0.238*** (-0.027)	-0.159*** (-0.026)	0.078*** (-0.018)	0.411*** (-0.040)	0.491*** (-0.038)	0.079*** (-0.021)	-0.139*** (-0.027)	-0.066*** (-0.024)	0.073*** (-0.019)	-0.444*** (-0.028)	-0.363*** (-0.025)	0.081*** (-0.023)
Unexplained gap	0.316*** (-0.017)	0.230*** (-0.017)	-0.086*** (-0.019)	0.426*** (-0.020)	0.366*** (-0.018)	-0.060** (-0.024)	0.320*** (-0.018)	0.222*** (-0.016)	-0.098*** (-0.020)	-0.090*** (-0.023)	-0.058*** (-0.021)	0.032 (-0.029)
<i>Amount explained by:</i>												
Male	-0.001 (-0.007)	0.003 (-0.005)	0.003 (-0.003)	-0.007 (-0.007)	-0.003 (-0.006)	0.005 (-0.004)	0.013* (-0.007)	0.009* (-0.005)	-0.004 (-0.004)	0.004 (-0.006)	0.009 (-0.005)	0.004 (-0.005)
Family & neighbourhood deprivation	-0.195*** (-0.006)	-0.180*** (-0.006)	0.015*** (-0.004)	-0.014** (-0.006)	0.004 (-0.006)	0.018*** (-0.004)	-0.271*** (-0.006)	-0.248*** (-0.006)	0.023*** (-0.005)	-0.255*** (-0.006)	-0.239*** (-0.006)	0.015** (-0.006)
Special educational needs	0.014*** (-0.002)	0.020*** (-0.002)	0.006*** (-0.002)	0.033*** (-0.002)	0.043*** (-0.002)	0.011*** (-0.002)	0.014*** (-0.002)	0.021*** (-0.002)	0.007*** (-0.002)	-0.024*** (-0.004)	-0.028*** (-0.004)	-0.003 (-0.005)
First language not English	0.093*** (-0.007)	0.112*** (-0.007)	0.019** (-0.009)	0.079*** (-0.006)	0.095*** (-0.006)	0.015** (-0.008)	0.071*** (-0.006)	0.083*** (-0.005)	0.012* (-0.007)	0.003*** (0.000)	0.003*** (0.000)	0.001 (-0.001)
Prior attainment (age 11)	-0.295*** (-0.013)	-0.306*** (-0.014)	-0.011 (-0.010)	0.055** (-0.023)	0.103*** (-0.025)	0.048*** (-0.014)	-0.174*** (-0.013)	-0.149*** (-0.014)	0.024** (-0.011)	-0.311*** (-0.014)	-0.286*** (-0.015)	0.025 (-0.016)
<i>School characteristics:</i>												
Selective admissions	0.000 (-0.002)	0.000 (-0.001)	0.000 (-0.001)	0.032*** (-0.006)	0.021*** (-0.005)	-0.011** (-0.005)	0.002 (-0.002)	0.002 (-0.001)	0.000 (-0.001)	-0.008*** (-0.001)	-0.005*** (-0.001)	0.003** (-0.001)
Region	0.032*** (-0.005)	0.024*** (-0.005)	-0.008 (-0.005)	0.026*** (-0.006)	0.014*** (-0.005)	-0.012** (-0.005)	0.062*** (-0.010)	0.044*** (-0.009)	-0.018* (-0.010)	0.065*** (-0.011)	0.049*** (-0.011)	-0.016 (-0.011)
Urban	0.003 (-0.002)	0.001 (-0.002)	-0.003 (-0.003)	0.003 (-0.002)	0.001 (-0.002)	-0.003 (-0.002)	0.003 (-0.002)	0.001 (-0.002)	-0.003 (-0.003)	0.003 (-0.002)	0.001 (-0.002)	-0.003 (-0.003)
Size of cohort	-0.001 (-0.001)	-0.004* (-0.002)	-0.003* (-0.002)	-0.001 (-0.001)	-0.003 (-0.003)	-0.003 (-0.002)	0.001 (-0.001)	0.005*** (-0.002)	0.004*** (-0.001)	0.001 (-0.001)	0.008*** (-0.002)	0.007*** (-0.002)
School governance type	-0.001 (-0.001)	0.000 (-0.001)	0.001 (-0.001)	0.001 (-0.001)	0.001 (-0.001)	0.000 (-0.001)	-0.002 (-0.001)	-0.001 (-0.001)	0.001 (-0.001)	-0.002 (-0.001)	-0.001 (-0.001)	0.001 (-0.001)
Proportion free school meals	-0.064*** (-0.014)	0.000 (-0.013)	0.064*** (-0.016)	-0.006 (-0.005)	0.005* (-0.003)	0.011** (-0.004)	-0.053*** (-0.012)	0.001 (-0.011)	0.055*** (-0.014)	-0.065*** (-0.014)	-0.001 (-0.013)	0.065*** (-0.016)
Average neighbourhood deprivation	0.081*** (-0.013)	0.063*** (-0.012)	-0.019 (-0.015)	0.032*** (-0.006)	0.025*** (-0.005)	-0.007 (-0.006)	0.085*** (-0.014)	0.064*** (-0.013)	-0.021 (-0.016)	0.092*** (-0.015)	0.070*** (-0.014)	-0.021 (-0.018)
School value-added (lagged)	0.095*** (-0.011)	0.109*** (-0.011)	0.014* (-0.008)	0.180*** (-0.013)	0.185*** (-0.013)	0.006 (-0.009)	0.110*** (-0.010)	0.103*** (-0.010)	-0.008 (-0.007)	0.053*** (-0.011)	0.056*** (-0.011)	0.003 (-0.008)

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The gap is the gap in average grades between the ethnic minority group and White British students. Family & neighbourhood deprivation includes free school meals indicator and a rank based on the proportion of children under age 16 in a local area living in a low-income household score (cubic). Prior attainment (age 11) includes subject-specific age 11 attainment score, standardised by year (quadratic). Region has 9 categories. School governance type has 4 categories. Proportion free school meals is a cubic. Average neighbourhood deprivation is a cubic. School value-added refers to how many average grades higher or lower the students in that school achieve across eight qualifying GCSE subjects compared to students across the country who score comparatively at age 11 (cubic). Standard errors in parentheses, clustered at the school level. * Significant at 10%; ** significant at 5%; *** significant at 1%.

5b: Detailed Gelbach decomposition of grade gaps by ethnic groups for GCSE English Literature

	Pakistani & Bangladeshi			Indian			Black African			Black Caribbean		
	2019	2020	Change	2019	2020	Change	2019	2020	Change	2019	2020	Change
Raw gap	0.322*** (-0.036)	0.204*** (-0.035)	-0.119*** (-0.022)	1.065*** (-0.044)	1.001*** (-0.042)	-0.064** (-0.029)	0.453*** (-0.033)	0.315*** (-0.029)	-0.138*** (-0.024)	-0.360*** (-0.038)	-0.324*** (-0.033)	0.036 (-0.037)
Explained gap	-0.156*** (-0.032)	-0.106*** (-0.030)	0.050** (-0.021)	0.518*** (-0.043)	0.560*** (-0.041)	0.042* (-0.023)	-0.048 (-0.031)	0.001 (-0.027)	0.049** (-0.020)	-0.386*** (-0.033)	-0.316*** (-0.029)	0.070*** (-0.025)
Unexplained gap	0.479*** (-0.021)	0.310*** (-0.020)	-0.169*** (-0.021)	0.547*** (-0.023)	0.441*** (-0.020)	-0.106*** (-0.026)	0.500*** (-0.021)	0.314*** (-0.018)	-0.187*** (-0.022)	0.026 (-0.027)	-0.007 (-0.024)	-0.033 (-0.031)
<i>Amount explained by:</i>												
Male	-0.002 (-0.008)	0.002 (-0.006)	0.004 (-0.003)	-0.012 (-0.008)	-0.005 (-0.006)	0.007 (-0.005)	0.016** (-0.008)	0.010* (-0.006)	-0.005 (-0.004)	0.007 (-0.008)	0.010* (-0.006)	0.003 (-0.006)
Family & neighbourhood deprivation	-0.218*** (-0.007)	-0.195*** (-0.007)	0.023*** (-0.005)	-0.020*** (-0.007)	-0.002 (-0.006)	0.018*** (-0.004)	-0.301*** (-0.007)	-0.264*** (-0.006)	0.037*** (-0.006)	-0.285*** (-0.007)	-0.256*** (-0.006)	0.029*** (-0.007)
Special educational needs	0.015*** (-0.003)	0.019*** (-0.002)	0.004 (-0.002)	0.036*** (-0.002)	0.041*** (-0.002)	0.005** (-0.002)	0.013*** (-0.002)	0.019*** (-0.002)	0.006** (-0.003)	-0.029*** (-0.005)	-0.030*** (-0.004)	-0.001 (-0.005)
First language not English	0.127*** (-0.009)	0.128*** (-0.008)	0.001 (-0.010)	0.109*** (-0.008)	0.108*** (-0.007)	-0.001 (-0.009)	0.097*** (-0.007)	0.094*** (-0.006)	-0.003 (-0.008)	0.003*** (-0.001)	0.003*** (-0.001)	0.000 (-0.001)
Prior attainment (age 11)	-0.291*** (-0.014)	-0.309*** (-0.015)	-0.018* (-0.011)	0.050** (-0.023)	0.092*** (-0.026)	0.042*** (-0.015)	-0.176*** (-0.013)	-0.158*** (-0.014)	0.018 (-0.012)	-0.309*** (-0.014)	-0.294*** (-0.015)	0.015 (-0.016)
School characteristics:												
Selective admissions	0.000 (-0.002)	0.000 (-0.001)	0.000 (-0.001)	0.034*** (-0.006)	0.022*** (-0.005)	-0.012** (-0.005)	0.002 (-0.002)	0.002 (-0.001)	0.000 (-0.001)	-0.008*** (-0.002)	-0.005*** (-0.001)	0.003** (-0.001)
Region	0.058*** (-0.007)	0.047*** (-0.007)	-0.011* (-0.006)	0.054*** (-0.008)	0.036*** (-0.007)	-0.018*** (-0.006)	0.111*** (-0.012)	0.090*** (-0.011)	-0.021* (-0.011)	0.122*** (-0.013)	0.101*** (-0.012)	-0.021* (-0.013)
Urban	0.003 (-0.003)	0.001 (-0.002)	-0.002 (-0.003)	0.003 (-0.003)	0.001 (-0.002)	-0.002 (-0.003)	0.003 (-0.003)	0.001 (-0.002)	-0.002 (-0.003)	0.003 (-0.003)	0.001 (-0.002)	-0.002 (-0.003)
Size of cohort	-0.001 (-0.001)	-0.003* (-0.002)	-0.001 (-0.001)	-0.001 (-0.001)	-0.003 (-0.002)	-0.002 (-0.002)	0.002 (-0.001)	0.004*** (-0.001)	0.002* (-0.001)	0.003 (-0.002)	0.007*** (-0.002)	0.004** (-0.002)
School governance type	-0.002 (-0.002)	0.000 (-0.002)	0.001 (-0.002)	0.001 (-0.002)	0.001 (-0.001)	0.000 (-0.001)	-0.004** (-0.002)	-0.001 (-0.001)	0.003* (-0.002)	-0.003* (-0.002)	-0.002 (-0.001)	0.002 (-0.001)
Proportion free school meals	-0.072*** (-0.018)	-0.023 (-0.014)	0.049*** (-0.018)	-0.009 (-0.006)	0.003 (-0.004)	0.012*** (-0.004)	-0.061*** (-0.016)	-0.018 (-0.012)	0.043*** (-0.016)	-0.074*** (-0.018)	-0.024* (-0.015)	0.050*** (-0.019)
Average neighbourhood deprivation	0.106*** (-0.017)	0.090*** (-0.014)	-0.016 (-0.017)	0.042*** (-0.008)	0.034*** (-0.006)	-0.008 (-0.007)	0.112*** (-0.018)	0.093*** (-0.014)	-0.018 (-0.018)	0.121*** (-0.019)	0.103*** (-0.016)	-0.018 (-0.020)
School value-added (lagged)	0.122*** (-0.015)	0.138*** (-0.015)	0.015 (-0.010)	0.232*** (-0.018)	0.232*** (-0.017)	0.001 (-0.011)	0.140*** (-0.013)	0.129*** (-0.012)	-0.011 (-0.009)	0.065*** (-0.015)	0.069*** (-0.015)	0.004 (-0.011)

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The gap is the gap in average grades between the ethnic minority group and White British students. Family & neighbourhood deprivation includes free school meals indicator and a rank based on the proportion of children under age 16 in a local area living in a low-income household score (cubic). Prior attainment (age 11) includes subject-specific age 11 attainment score, standardised by year (quadratic). Region has 9 categories. School governance type has 4 categories. Proportion free school meals is a cubic. Average neighbourhood deprivation is a cubic. School value-added refers to how many average grades higher or lower the students in that school achieve across eight qualifying GCSE subjects compared to students across the country who score comparatively at age 11 (cubic). Standard errors in parentheses, clustered at the school level. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix 6

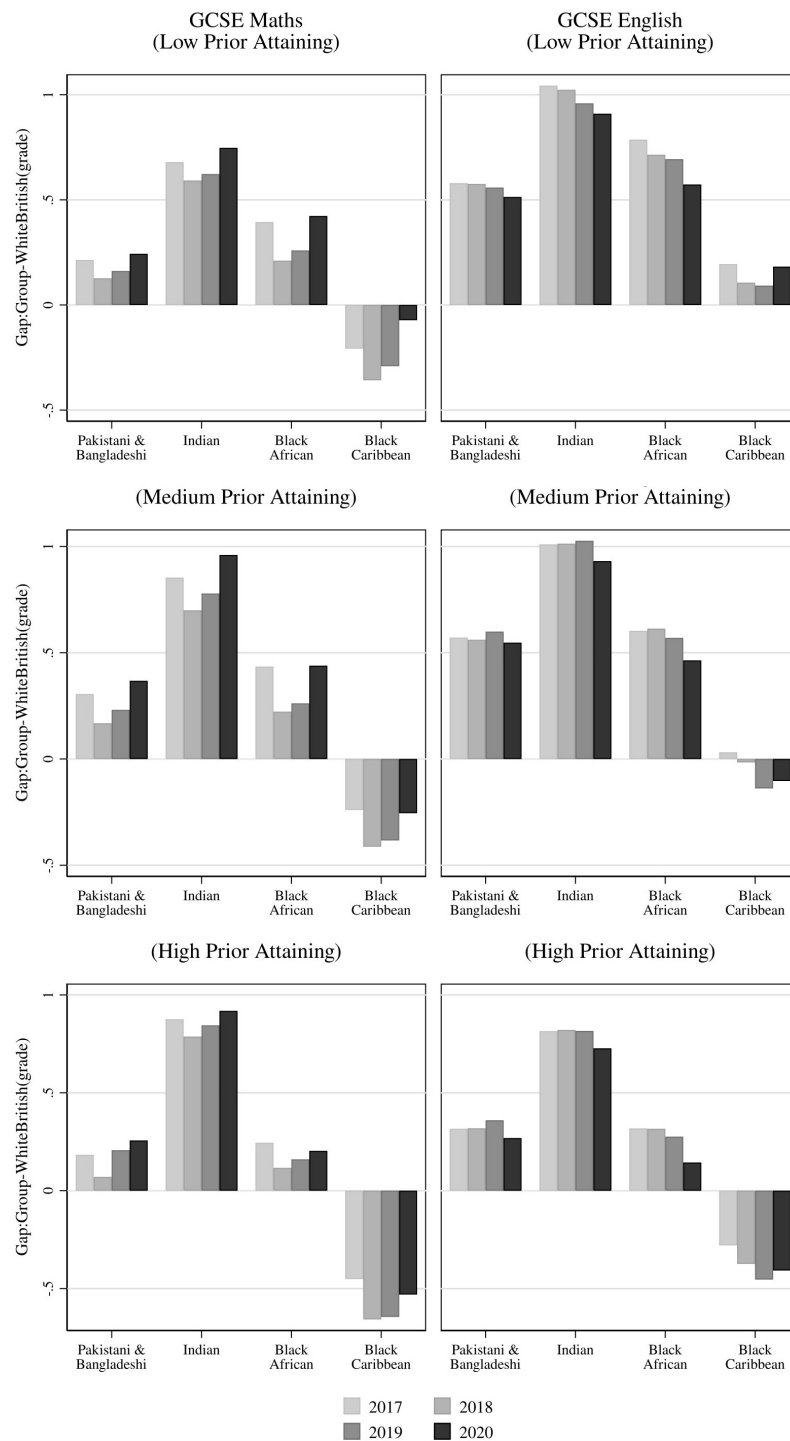
Mean grades by demographic group for GCSE English Literature and GCSE English Language

	GCSE English Language				GCSE English Literature			
	2017	2018	2019	2020	2017	2018	2019	2020
Total	4.81 (1.84)	4.80 (1.83)	4.80 (1.84)	5.19 (1.80)	4.82 (1.98)	4.85 (1.98)	4.89 (1.97)	5.18 (1.88)
Female	5.22 (1.79)	5.17 (1.78)	5.18 (1.79)	5.56 (1.77)	5.23 (1.92)	5.28 (1.90)	5.31 (1.90)	5.58 (1.84)
Male	4.40 (1.79)	4.44 (1.80)	4.42 (1.80)	4.82 (1.75)	4.42 (1.96)	4.43 (1.96)	4.47 (1.95)	4.78 (1.84)
White British	4.76 (1.84)	4.75 (1.84)	4.74 (1.84)	5.13 (1.80)	4.72 (1.97)	4.74 (1.97)	4.77 (1.96)	5.09 (1.88)
Pakistani & Bangladeshi	4.81 (1.73)	4.78 (1.72)	4.82 (1.77)	5.20 (1.72)	4.98 (1.91)	5.01 (1.90)	5.10 (1.90)	5.29 (1.81)
Indian	5.55 (1.74)	5.51 (1.74)	5.58 (1.76)	5.99 (1.67)	5.81 (1.86)	5.80 (1.84)	5.84 (1.85)	6.09 (1.72)
Black African	4.94 (1.69)	4.93 (1.70)	4.92 (1.73)	5.29 (1.65)	5.25 (1.84)	5.25 (1.85)	5.23 (1.83)	5.40 (1.72)
Black Caribbean	4.38 (1.68)	4.24 (1.66)	4.21 (1.66)	4.71 (1.61)	4.55 (1.89)	4.44 (1.90)	4.42 (1.83)	4.76 (1.70)
Multiethnic	4.94 (1.89)	4.91 (1.86)	4.91 (1.88)	5.27 (1.83)	5.00 (2.05)	5.02 (2.02)	5.05 (2.01)	5.29 (1.92)
No free school meals	5.06 (1.79)	5.05 (1.79)	5.04 (1.79)	5.44 (1.74)	5.08 (1.92)	5.10 (1.91)	5.14 (1.90)	5.43 (1.82)
Free school meals	4.07 (1.76)	4.05 (1.75)	4.05 (1.77)	4.43 (1.75)	4.06 (1.97)	4.09 (1.96)	4.12 (1.96)	4.39 (1.86)
Observations	451688	437198	450144	484919	449587	423708	430350	461105

Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. Standard deviations in parentheses.

Appendix 7

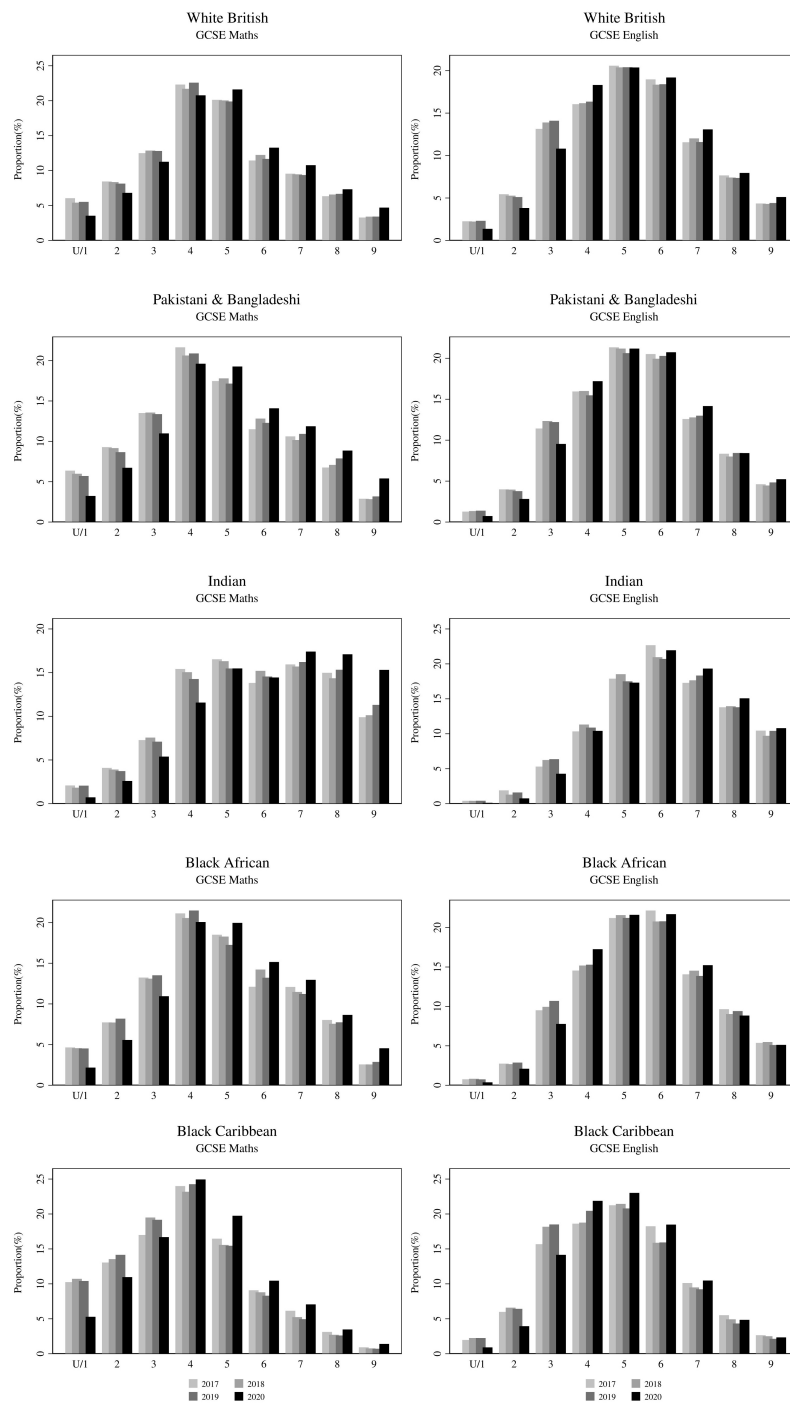
Gap graphs by prior attainment group



Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. Prior attainment groups use exact thirds of the prior attainment scores, standardised by year.

Appendix 8

Grade distributions by year and ethnic group



Notes: Notes: Sample comprises students turning age sixteen in mainstream schools in England with no missing covariates. Exam grades are the May and June (end of year) examinations only. The English grade is the highest of English Literature and English Language if students received grades for both. Grades 'U' and '1' combined due to low cell counts in some groups.