The impact of coursework on attainment dependent on student characteristics

A study based on GCSE and A level outcomes between 2004 and 2017

ofqual

Commissioned by Ofqual, authored by: Anne Pinot de Moira

Contents

Background

Ofqual commissioned this work in 2019 in order to try to gain some deeper insight into the relationship between different forms of assessment within high stakes qualifications, to understand the extent to which different types of assessment – examination and non-examined assessment (coursework) - might be associated with different levels of performance for different groups or types of students. While there are often views or observations around, say, whether a particular gender 'does better' in examinations, or that students of lower socio-economic status may not have the same access to resources at home to support coursework, the research literature to support these views is not as substantial as it can be; and often does not adequately control for other factors.

This research aims to make the most of existing data around performance in different types of assessment. Over the last two decades, there have been changes in the proportion of examined and non-examined assessment (coursework) and this research utilises this data to see whether there is any evidence to show that differences in learner characteristics (gender, ethnicity, special educational needs, socio-economic class) are associated with different patterns of performance in relation to different types of assessment. The data is useful in that it comprises large, whole cohort data for the years analysed. The insights are likely to provide a useful evidence base and have a bearing on future decisions about use of different types of assessment in high stakes examinations.

Executive summary

When changes to the structure and content of an examination are made, it is important that the implications of such changes are clearly understood. This paper explores the impact of coursework on GCSE and A-level attainment over a period of considerable flux in the English secondary school examination system. Its intention is to provide evidence to support future discussions on the role of coursework.

The analysis compares specifications with and without coursework to investigate how attainment differs dependent on student characteristics. It includes data from six academic years between 2004 and 2017 and, within those years, it focuses on specifications in five different subject areas.

For the most part, the results support existing research evidence and so, because they are based on population data, they provide a firm foundation from which to understand the impact of coursework in qualifications. For the subjects and academic years under consideration, by far the greatest amount of variation in GCSE and A-level grade outcomes is explained by student prior and concurrent attainment. After controlling for academic and demographic differences, the key findings suggest:

- there is little evidence that coursework has any impact on outcomes for students of different socio-economic statuses (SES) or for students with special educational needs;
- male students perform better than female students in wholly examined GCSE specifications and also in GCSE specifications where there is a greater level of control in the coursework. Female students tend to have better outcomes than males where internally set, internally marked coursework is included;
- there is little indication of different outcomes for students of different ethnicities across different assessment types; except for students of Chinese ethnicity who despite performing well overall, perform relatively poorly when entered for specifications with coursework;
- in specifications where coursework was optional, the examined alternative appears to provide a safety net for less able students who failed to submit coursework.

Subject-level differences revealed by the analysis also point to the need for assessment solutions that are appropriate to each individual subject while also considering the validity and equity of any coursework units.

Introduction

Coursework, more recently referred to as non-examined assessment (NEA), is included in a qualification as a means of assessing elements that are difficult to assess through a written examination paper. In the past, coursework has been completed by students with very few limitations on the task set, the environment in which the task is taken and marking of the task. Following recommendations set out in QCA's review of coursework (Qualifications and Curriculum Authority, 2006), the rigour of coursework was increased for GCSE specifications between 2010 and 2017 and it was usually referred to as controlled assessment. Controlled assessment tasks require students to complete a task in controlled conditions at a time and place determined by their centre. Nevertheless, the central purpose of a coursework remains.

Over and above the level of control, courseworks differ in how they are specified and the way in which they are marked. By far the most common type of coursework is that where the task is set by the centre (internally set) and marked by the centre (internally marked). This type of coursework is normally sample-moderated by the awarding body to ensure that the correct standard has been applied by the centre. A coursework task can also be externally set by the awarding body and/or externally marked by the awarding body but this is less common. Some specifications have included optional coursework, flexible coursework formats or a mixture of formats.

The principle behind including coursework as part of a qualification is sound in that it is designed to increase the validity of the assessment. The role of coursework is to test skills, performances, achievements that might not be so readily assessed by a traditional timed examination and/or difficult logistically to be assessed by an external examiner. This might include a dramatic or sporting performance; writing, editing and crafting an extended written essay; undertaking some independent or group field work or research. However, as Johnson (2011) observes, the validity and concomitant reliability of teacher-assessed material is sparsely discussed in the literature despite the far-reaching practical and ethical implications of its use in high-stakes qualifications.

From an assessment or measurement perspective, coursework marks are often skewed towards the top end of the mark distribution. Where the mark distribution has a low standard deviation (ie is not spread out) then the coursework performance is under-weighted or under-represented in the final grade awarded (Elwood, 1999; Stringer, 2014).

From an ethical perspective, coursework (or high-stakes teacher assessment) can be seen to be problematic. Perceptions that coursework might favour different subgroups of the population are sometimes supported by research evidence of biases but, just as often, they are based on unsubstantiated claims. In his key speech about exam reforms, Michael Gove (then Secretary of State for Education) suggested that the expansion of coursework undermined the credibility of GCSE examinations (Hansard HC Deb, 2012a). In the following debate, coursework was said, by different speakers, to benefit middle class students whose parents could better support them and also to support students, generally from poorer backgrounds, who struggle with exams (Hansard HC Deb, 2012b).

Where research evidence on the impact of coursework is available, there is some indication that biases can be both implicit and explicit. In a study of teacher-assessed writing in Australia, Wyatt Smith & Castleton (2005) observed that contextual knowledge is instrumental in the judgements teachers make about work. Furthermore, that the context can extend beyond the classroom. Ready & Wright (2011) showed that teachers' perceptions of cognitive ability of young children are influenced by gender, social class and ethnicity. In a systematic review of research on the reliability and validity of teacher assessment, Harlen (2005) collected evidence from a range of studies indicating that teacher judgement might also be influenced by behaviour, special educational needs and prior academic achievement.

Some of the complications inherent in the valid assessment of coursework are, however, implicit in the design. Access to learning and resources may differ between different subgroups of the population. Not all students will have parental support (or influence), nor will they all have the same facilities to hand (Sammons, 1995; Scott, 1991). And the potential for bias against those at social disadvantage may even be exacerbated by the topic under consideration (Gipps, 1995). Task context and content are not necessarily universally accessible either at a practical or experiential level.

However, the effect of potential bias may have sometimes been overplayed. Elwood (1999), for example, concluded that coursework did little to increase the gender equity gap in GCSE performance. Indeed, there are many confounding factors complicating the evaluation of cause and effect with respect to examination performance; some are measurable and some are not. Even where quasi-measures of ability are available in the form of prior attainment statistics, these statistics are influenced by students' formative experiences. This paper aims to explore the relationship between student characteristics and GCSE or A-level grade outcome to understand any differential effect that coursework has on attainment.

Data

National Pupil Database

The data used in this study is taken from the National Pupil Database (NPD). It includes six separate academic years between 2004 and 2017 and five different subject areas. Where data was available, the years were selected to coincide with key points in reforms to GCSE and A-level qualifications (

Table 1). Not only did the period see changes to the curriculum but also to the structure of qualifications and the nature of the coursework element itself. For example, between 2007 and 2010, GCSE coursework was replaced with controlled assessment.

Table 1 Data included in the study of the impact of coursework on attainment

Year	Significance of Year
2004	First year in which prior attainment Key Stage 2 data was available
2007	Stable year bridging the gap; NEAs structured as coursework
2010	Last stable year before a significant change to GCSEs and A-levels; NEAs structured as controlled assessments
2011	First two-year modular GCSEs awarded in all subjects except English, mathematics and ICT; First major award of four-unit A-levels
2015	Stable year towards the end of the lifetime of many specifications and before a significant reform to GCSEs and A-levels. All modular GCSEs had to be taken in a linear fashion.
2017	First certification of new linear GCSEs in English language, English literature & mathematics; First certification of most mainstream A-levels

The subjects were selected because they all included coursework at some point over the period of interest. For GCSE, the subjects chosen were English language, English literature, geography, history and mathematics and, for A-level, English literature, geography and history. In these subjects, all examination results for all specifications taken by students in England were retained. There were, therefore, specifications offered by AQA, Edexcel, OCR and WJEC. Over 11.6 million GCSE and A-level results from 311 specifications were included in the study (

Table 2).

*Table 2 The distribution of entries across subjects included in the study*

Level	Subject	Entries	Specifications
GCSE	English Language	1,423,994	18
	English Literature	2,571,300	22
	Geography	1,131,448	39
	History	1,255,808	78
	Mathematics	3,550,372	75
A-level	English Literature	616,066	31
	Geography	426,568	26
	History	626,513	22
Total		**11,602,069**	**311**

Dependent variables

Grade outcome was used to measure the impact of coursework on attainment in different subgroups of the population. For A-level students, the A-level grade was converted to a numeric value and then assumed to be on a continuous scale. Grade A* was given a value of 6, grade A a value of 5 and so on down to grade U, which was given a value of 0.

Because of changes to the GCSE grading scheme in reformed qualifications in 2017, converting GCSE grade outcome to a single scale was slightly more complicated. Figure 1 shows the value given to each GCSE grade for the analysis presented in this paper. These values reflect the equivalences described by Ofqual[1].

Figure 1 Coding for GCSE grade outcomes

---

[1] https://ofqual.blog.gov.uk/2018/01/26/gcse-grade-boundaries-in-2018/
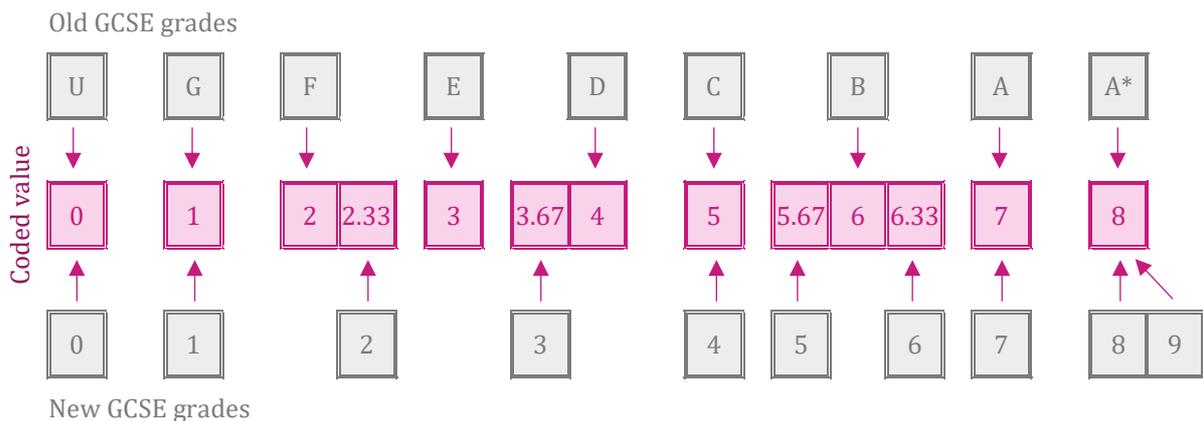
Missing data

The NPD includes a range of information about students from their examination performance to details of their socio-economic status. Not all of the data is complete and, furthermore, when the data is missing, it is not always missing at random. Table 3 lists the covariates available for use in analysis, their source and the percentage of missing values. Each covariate is cross-classified by centre type and by qualification level. The table clearly illustrates that there are differences in the availability of data dependent upon both centre type and qualification level.

Table 3 Covariates and the percentage of missing values

Source	Covariate	GCSE				A-level			
		State	Indep	Post16	Total	State	Indep	Post16	Total
NPD	Gender	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Academic year	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Special educational needs (SEN)	0.1	99.8	99.0	6.6	1.6	100.0	99.1	38.0
	Free school meals (FSM)	0.1	0.0	99.0	0.3	1.6	0.0	99.1	25.2
	Ethnicity	1.7	99.7	95.9	8.0	3.7	100.0	99.0	39.3
	English as an additional language (EAL)	0.3	99.7	95.8	6.7	2.0	99.9	99.0	38.2
	Concurrent attainment	0.0	0.0	0.0	0.0	42.8	26.3	46.0	41.4
	Prior attainment	4.5	76.0	76.8	9.1	18.0	15.9	21.3	18.6
Index of	IDACI 2015*	0.1	0.0	0.0	0.1	0.1	0.0	0.0	0.1
Multiple Deprivation	IDACI trend*	2.4	4.8	8.9	2.5	2.5	5.3	4.0	3.3
Documentation	Coursework	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total (000s)		**9,294**	**626**	**13**	**9,933**	**1,048**	**214**	**407**	**1,669**

* Income deprivation affecting children index

Covariates

Each of the covariates listed in Table 3 is described below and details of how they are formulated for use in the analysis are given in Appendix A.

Gender

For both the GCSE and A-level subjects, gender is recorded for all students.

Academic year

Students are entitled to take GCSE and A-level examinations at any age. These results all form part of the student record and are recorded against the student's academic year. However, only examination results related to the academic years of interest, listed in
Table 1, are included in the study.

Special educational needs

Special educational needs (SEN) data is only meaningfully available for students educated in the state sector. The covariate distinguishes two different types of provision. Those with the most complex needs are classified as SEN

(EHCP/Statement). This indicates that a student has been provided with interventions by the school with the support of external agencies and specialists (Department for Education and Skills, 2001). Those with less complex needs are classified as SEN (School). This indicates that a student has been provided with interventions that are additional to, or different from, those provided as part of a school's usual differentiated curriculum offer and strategies.

Free school meals

Students will be given a free school meal (FSM) if they or their guardians are eligible for income support or other such benefits. This covariate is used as a proxy for socio-economic status. Although FSM data is not collected for students attending independent schools, no student in an independent school is eligible and so, for state and independent schools, there is little missing data[2]. However, this information is almost completely missing from post-16 schools, meaning that any model including FSM necessarily excludes students in post-16 education.

Ethnicity and English as an additional language

Ethnicity and English as an additional language (EAL) data is only meaningfully available for students educated in the state sector. It is sparser for A-level students than GCSE students.

Concurrent attainment

For GCSE students, concurrent attainment is recorded on the NPD as the total GCSE (or equivalent) point score. There are no missing values but, because of the change in grading scale for the GCSE in 2017, values provided in the NPD do not have the same meaning across all years in the study. The measure of concurrent attainment has therefore been normalised for use as a covariate in the analysis.

Concurrent attainment for A-level students is not so well populated, with nearly 40% missing this measure. It is recorded on the NPD as the total A-level point score. As with all the other measures of attainment, there are differences in the way the statistic in the NPD has been configured over the years (

---

[2] It is probable that an extremely small proportion of the students attending independent schools would be eligible for free school meals were they educated in the state sector. However, no data is available because students only receive free school meals as a result of parent or guardian registration.

Figure 2). Therefore, it has been normalised within year to allow use in further analysis.

Figure 2 The distribution of total A-level (and equivalent) points score by academic year
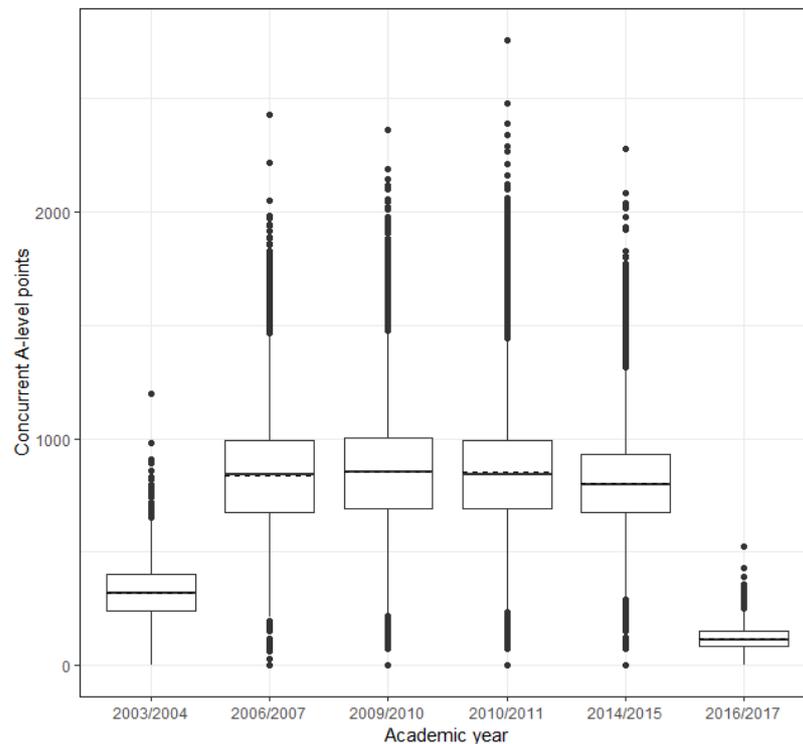

Prior attainment

Student prior attainment is measured differently for GCSE and A-level. For GCSE, the prior attainment data is a summary statistic created from the Key Stage 2, end of primary school, tests. In 2004, this statistic was recorded on the NPD as an average of test levels in reading, writing and mathematics. After that date, it was calculated using the actual test scores. Because of this discontinuity, the prior attainment score has been normalised within year. Key Stage 2 tests are only compulsory in state schools and so prior attainment data is missing for many GCSE students attending independent and post-16 centres.

For A-level, prior attainment is recorded on the NPD as the average GCSE (or equivalent) point score. The calculation has differed between years as illustrated in

Figure 3. Therefore, to enable more direct comparison across years, the scores have been normalised within year. The majority of the A-level students included in the study have a measure of prior attainment regardless of the centre type of entry.

Figure 3 The distribution of average GCSE point score for A-level students by academic year
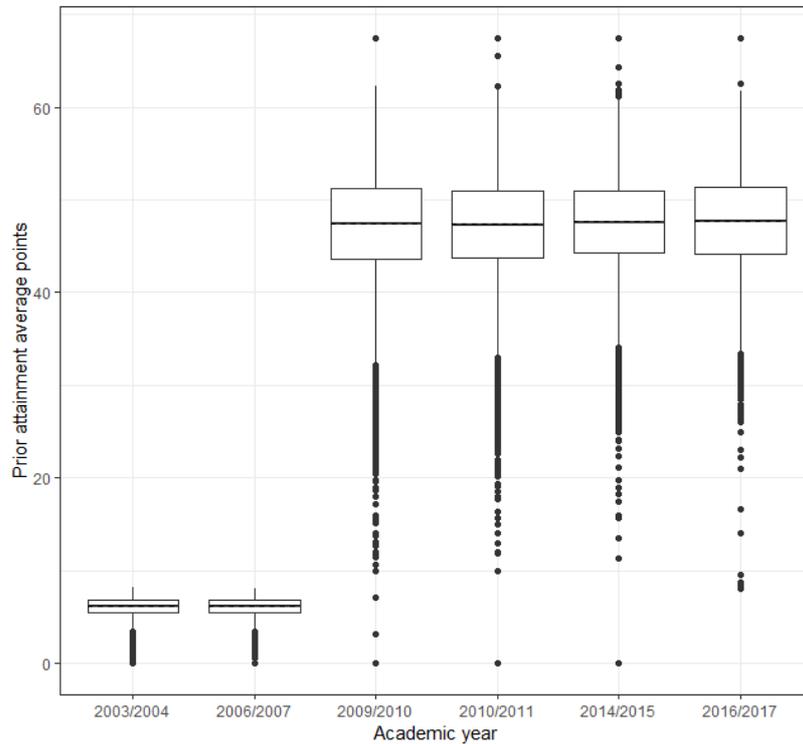
Income deprivation affecting children index (IDACI)

The IDACI 2015 covariate is an index of child deprivation. It is a measure of the percentage of children in a small group of postcodes who live in income-deprived families and is reported as a rank. A higher value means a greater level of deprivation. The covariate IDACI 2015 is taken from the 2015 dataset. Because student postcode is not available, the IDACI is matched to the centre postcode to give a proxy for the deprivation of the students in attendance. A weakness in this proxy lies in the fact that there will be some centres situated in deprived areas that attract privileged students and vice versa.

Over the period of interest for this study, several NPD statistical releases have included IDACI statistics. IDACI data is available from 2004, 2007, 2010 and 2015. This affords the opportunity to consider trends in deprivation and a covariate has been derived to measure these trends. Centre-level deprivation has been described as worsening if there has been a consistent increase in the percentile rank of more than 2% between each consecutive year. It has been described as improving if there has been a consistent decrease of more than 2%. Otherwise, centre-level deprivation has been described as stable. There are very few missing values for either IDACI 2015 or trend and those values that are missing do not appear to be systematic by centre type.

From 2015, the NPD has included a student-level IDACI statistic. The fact that this data is not available for the majority of the years covered in the study means that it cannot be used to provide meaningful information in the context of a longitudinal analysis as presented here. However, a brief analysis of its value as a measure of socio-economic status is provided in Appendix B.

Coursework

While student and qualification level information is included on the NPD, it does not carry details of units within qualifications, so these details were hand-coded for each specification from historic documentation. Four features of the coursework unit were recorded:

1. The percent that the coursework unit contributed to the final grade;
2. The type of coursework, coded as a three-level factor: No C/W, Internally set & Internally marked (I-I) or Other. Courseworks categorised in the Other category include: internally set/externally marked, externally set/internally marked, externally set/externally marked and specifications in which there is a choice of approaches;
3. The number of coursework units; and
4. Whether or not the coursework was optional.

Across the GCSE and A-level specifications included in the study, 35% of the entries were for specifications with no coursework element. The remaining entries included coursework; 55% with an internally set, internally marked (I-I) assessment and 10% with one of the less common formats. There were a few specifications where the coursework element was optional and, for these, the maximum possible coursework percentage was recorded. Coursework characteristics were unobtainable for a tiny percentage (0.06%) of small-entry, or outgoing, specifications. Where this was the case, the records were deleted.

Descriptive statistics

Of the GCSE specifications included in the study, 80% had at least one coursework unit and, when there was coursework, it accounted for an average of 26% of the final grade. The figures were slightly lower for the A-level where 61% of the specifications included coursework and, when there was coursework, it contributed an average of 26% to the final grade.

Table 4 shows the mean grade for the students entered for each subject included in the study. For the GCSE entries included in the study, the overall mean grade was close to the old grade C or the new grade 4 (see Figure 1 for details of the coding). It appeared slightly higher for those following an I-I coursework route although, within that summary statistic, there were differences between subjects. Despite the overall pattern, geography and mathematics students following a non-coursework path were the higher performers. For A-level, the mean grade was just over grade C and, on the whole, those following a coursework path performed better than their peers.

Table 4 The mean grade for GCSE and A-level entries included in the study (with and without a coursework element)

	GCSE			A-level		
	No C/W	**I-I**	**Other C/W**	**No C/W**	**I-I**	**Other C/W**
English Language	5.22	5.16	5.65			
English Literature	5.21	5.25	5.60	3.13	3.31	
Geography	5.16	5.16	4.90	3.24	3.14	3.33
History	4.99	5.19		3.06	3.46	3.39
Mathematics	4.88	4.51	4.53			
Overall mean	**4.98**	**5.18**	**4.58**	**3.16**	**3.34**	**3.37**

The means reported in

Table 4 take no account of any selection bias in specification entry. The raw data, however, suggests that GCSE specifications with coursework are taken in greater numbers by students with higher mean GCSE results and by GCSE students who are not in receipt of FSM (Table 5). For example, while 67% of the highest-attaining GCSE students entered specifications with coursework, the figure was 63% for the lowest-attaining students. These disparities may lend some weight to the literature that links access to resources and to parental support with concerns over the validity/fairness of coursework assessment.

Table 5 The percentage of entries for coursework and non-coursework specifications by (i) concurrent GCSE score for GCSE entries or (ii) prior GCSE score for A-level entries

	GCSE		A-level	
	No C/W	C/W	No C/W	C/W
Low GCSE	37.3	62.7	38.0	62.0
Mid-Low GCSE	36.7	63.3	39.2	60.8
Mid GCSE	34.1	65.9	39.6	60.4
Mid-High GCSE	31.7	68.3	39.1	60.9
High GCSE	33.3	66.7	37.1	62.9
No FSM	33.9	66.1	40.3	59.7
FSM	40.4	59.6	37.7	62.3

A closer look at this data, however, reveals that for GCSE this selection bias only exists in the specifications with optional coursework.

Table 6 shows that, where a GCSE coursework is compulsory, entry is uniform across the whole attainment range. In contrast, where it is optional, there is a considerable difference in the concurrent attainment of the students; there are fewer students in the lower attainment categories.

Table 6 The distribution of entries to coursework specifications dependent upon whether the coursework is optional (column percentages) by (i) concurrent GCSE score for GCSE entries or (ii) prior GCSE score for A-level entries

	GCSE		A-level	
	C/W	Opt C/W	C/W	Opt C/W
Low GCSE	19.4	14.3	20.2	20.2
Mid-Low GCSE	19.5	16.4	20.0	19.3
Mid GCSE	20.2	19.0	19.8	19.3
Mid-High GCSE	20.8	23.1	19.7	20.1
High GCSE	20.0	27.3	20.2	21.1

Coursework requires sustained effort over a period of time with a deadline before the start of examined assessment. If that deadline is missed, and the coursework is optional, it is possible to make late changes to a student's entry. In an analysis of examination absence, Chamberlain (2013) noted that students with unaccepted absence in one unit of an assessment were academically weaker than those with no absence. For GCSE, therefore, specifications with coursework optionality seem to provide a safety net for less able, and perhaps less motivated, students. For these students, failure to submit a credible coursework assignment is being addressed by a switch to the equivalent examined unit.

Even simple partitioning of the data illustrates the complex interrelationships between qualification design and attainment. Further graphs are provided in Appendix C. They suggest that students with special educational needs benefit more from coursework than those without (Appendix C.4); that the relationship between outcomes for coursework and non-coursework specifications does not change dependent upon school IDACI (Appendix C.7); and that the percentage a coursework contributes to a qualification is not related to grade outcome (Appendix C.11). The graphs also show that the relationship between the composite measures of student attainment and the grade awarded is non-linear (Appendix C.9 & C.10). The extent to which non-linear terms and more complex interactions are related to grade outcome is explored in greater detail in the next section using multilevel modelling.

Modelling

Design

The data held on the NPD is observational and so, for this study, there is no balance in terms of background variables between the coursework and non-coursework students. As seen in Table 5, there is evidence of selection bias in entry patterns, signalling some merit in statistical matching. However, the data has not been matched before analysis for three primary reasons:

1. The data on the NPD is population data and so findings from any analysis represent the real-world situation.
2. Not all legitimate covariates that distinguish between students are included on the NPD and some are not even measureable.
3. Computational limitations arising from the sheer volume of data included in the study preclude the use of matching.

Computational limitations also mean that linear, rather than logistic, multilevel models have been fitted to the GCSE and A-level grade outcomes. The models are as described generically in Box 1 and are fitted using the lme4 package for R (Bates, Mächler, Bolker, & Walker, 2015). β_n is the parameter estimate for the nth covariate (either a main effect or an interaction) and a list of each main effect covariate, and its formulation, is given in Appendix A.

$$
\begin{array}{lll}
GCSE/A\text{-}level\ grade_{ij} & = \beta_{0j} + \beta_1\ x_{ij} + \ldots + \beta_n\ x_{ij} + e_{ij} & \Rightarrow \quad \textit{Fixed effects} \\
\beta_{0j} & = \beta_0 + u_{0j} & \Rightarrow \quad \textit{Random intercepts} \\
\\
u_{0j} & \sim N(0,\ \sigma^2_{u0}) & \Rightarrow \quad \textit{School residual} \\
e_{0ij} & \sim N(0,\ \sigma^2_{e}) & \Rightarrow \quad \textit{Grade residual}
\end{array}
$$

Box 1

The use of a multilevel model exploits the inherent hierarchy in the data reflecting the fact that grade outcome (i) is nested within school of entry (j). Initial analysis suggests that at least 20% of the variation in GCSE and A-level results might be attributable to school influences. While this clustering in the data almost certainly arises from homogeneity of intake and similarity of teaching practices within schools, where specifications include internally marked coursework, it could also be an indication of marking biases.

For each model fitted, all educationally relevant covariates have been retained regardless of significance. In addition to the main effects, the models include interactions with coursework to allow evaluation of the impact of coursework on qualification outcome. Because the missing data is not missing at random, fitting models with all student demographic data means excluding most of those who attend independent or post-16 centres. To allow a greater understanding of the effect of centre type on outcome, two separate models have been derived (Table 7). The first includes all demographic information and it is dominated by the data from state

schools. The second excludes most demographic information but allows comparison between state and independent schools and between different types of coursework.

Table 7 Description of models fitted

	Model 1	Model 2
GCSE	• Includes demographic information => excludes most independent and post-16 centres • Coursework described simply as present or absent	• Excludes demographic information (except FSM) => includes state and independent • Coursework described as absent, I-I or other
A-level	• Includes demographic information => excludes most independent and post-16 centres • Coursework described simply as present or absent	• Excludes demographic information (except FSM) => includes state and independent • Coursework described as absent, I-I or other

Output from the modelling is presented in the next two sections and summarised in Table 8 and Table 9. Inevitably, because the models represent so many students, even small differences in outcomes between subgroups are statistically significant. Interpretation of the educational significance of the outcomes is supported by the provision of effect sizes for each covariate[3]. Fit for the GCSE models is very good, with around 70% of the variation in the data explained. For the A-level, this figure is lower (Snijders & Bosker, 1994).

Model 1 – Including demographic information

Table 8 shows the parameter estimates and effect sizes associated with the models of GCSE and A-level grade outcome. In both the GCSE and A-level models, the effect sizes for prior and concurrent attainment are completely dominant. However, both these measurements conceal a student's formative experiences; they represent the cumulative impact of all historic influences on performance. Thus, within the models, the effect of other covariates can be described as additional to that which is already entrenched. For example, in the GCSE subjects under consideration, the gap between students in receipt of free school meals and those who are not has slightly widened since the end of primary education (β = -0.085 (0.002)). Nevertheless, any significant interactions with coursework highlight conditions where the effect of coursework in a specification differs between subgroups of the population.

The models suggest that coursework has a different impact on GCSE attainment compared with A-level attainment. This is in contrast to what was seen in the raw data presented in

---

[3] Effect sizes are often reported as small (0.2), medium (0.5) or large (0.8) with reference to Cohen's original work in this area (Cohen, 1969). These broad classifications are not appropriate in the models of GCSE and A-level outcome because the large volume of data means "small" effect sizes could still have a widespread impact on students. Effect sizes are therefore included to clarify the *relative* importance of each covariate.

Table 4. When all other covariates are controlled, the inclusion of coursework has a negative effect on GCSE performance (β = -0.282 (0.007)) and a positive, but non-significant, effect on A-level performance (β = 0.021 (0.023)).

Figure 4 illustrates the interaction between academic year and coursework by using the parameter estimates to evaluate the mean grade, relative to a baseline candidate.[4,5] The dotted lines are included to emphasise the magnitude of difference in grade outcome between the coursework and non-coursework GCSEs. The findings seem to suggest that the staged assessment available in modular examinations might support GCSE students in distributing of coursework workload throughout the two years of study. However, as a note of caution, it should be noted that in 2015 and 2017, the non-coursework specifications included in the study were predominantly mathematics specifications and therefore the findings might simply be a function of the difference between mathematics and other subjects.

For the A-levels, where modular assessment was in place throughout the period of study, there is no step change between 2011 and 2015. Furthermore, the A-level model exhibits none of the subject confounding seen in the GCSE model because there are both coursework and non-coursework specifications available in all subject areas over the full time period of the study.

For GCSE in particular, once the presence of coursework is controlled, the number of coursework units and whether the coursework is optional seem to have a lesser effect on grade outcome. The coursework percentage, as a covariate in the model, accounts for almost no variation in the data. This is perhaps because, within subject, there is very little variation in coursework contribution to the qualification. For example, where coursework is included, it is weighted at 20% in all the GCSE mathematics specifications and at 25% for nearly all the GCSE geography and history specifications. Indeed, coursework contribution is often set by the regulator in the subject conditions and requirements.

Figure 4 Effect of the interaction between academic year and coursework on grade outcome
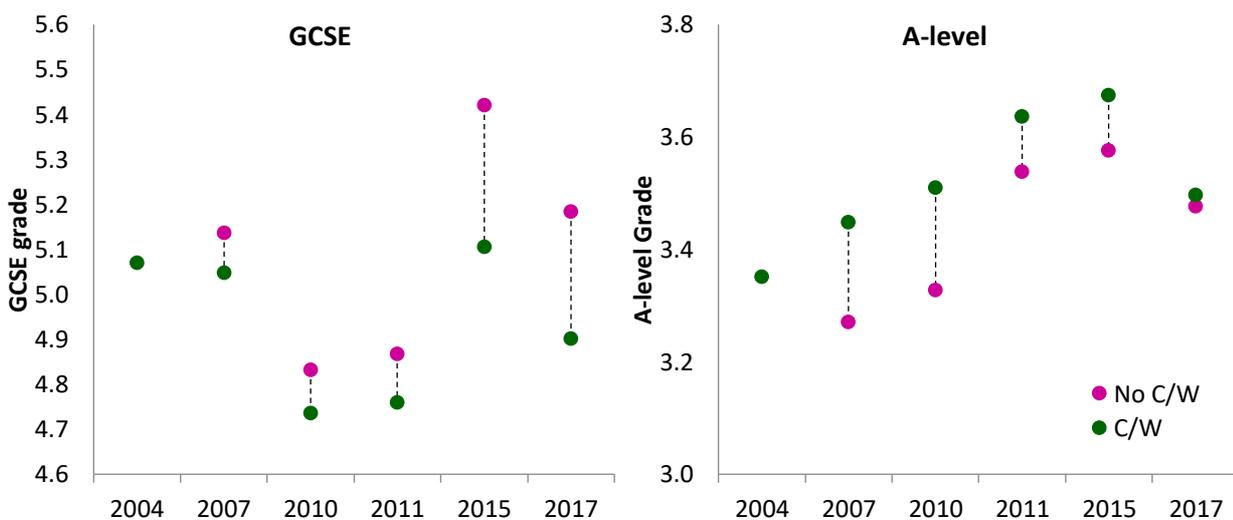

has a mean prior and concurrent attainment, whose first language is English, who has no SEN, is not in receipt of FSM, is of white ethnicity and is entered for GCSE English language or A-level English literature (as appropriate).

[5] Graphs do not include error bars as the standard errors associated with the parameter estimates discussed are so small as to be negligible.

After accounting for all other covariates, there is a gender difference in grade outcome for the GCSE subjects in the study. This difference alters depending upon whether the assessment includes a coursework unit. The interaction between coursework and gender is illustrated in Figure 5. For GCSE specifications with coursework, in general female students outperform their male peers whereas, for those without coursework, the reverse is true.

Figure 5 Effect of the interaction between gender and coursework on GCSE grade outcome
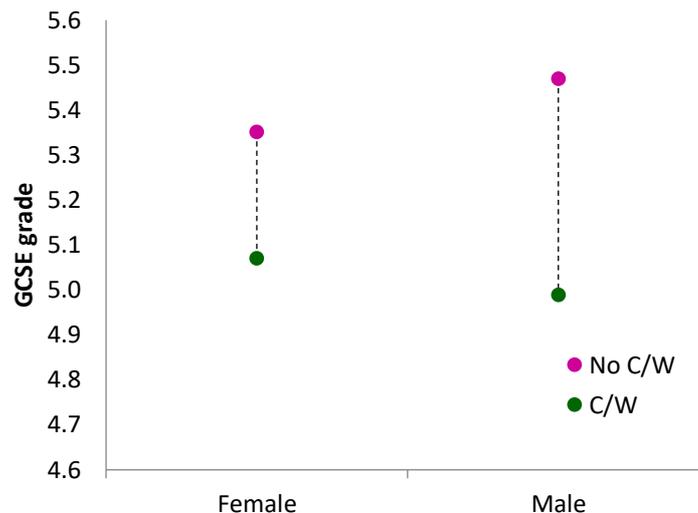
For GCSE, the indicators of socio-economic status (SES) – FSM and IDACI 2015 – both show that higher levels of deprivation are generally associated with lower grade outcomes. However, there is no evidence of an interaction with the inclusion of coursework in a specification. In other words, coursework offers neither advantage nor disadvantage to those of low SES. By the time students have progressed to A-level, there is no association between SES and grade outcome. This is probably because the entry hurdles applied to A-level study imply that, where social disadvantage may have played a part in prior attainment, students have either subsequently overcome this disadvantage or perhaps have not taken A-level qualifications.

While the evidence in the raw data suggests that students with special educational needs tend to benefit more from coursework than those without (Appendix C.4), the effect sizes once the models have been fitted show that the effect sizes associated with the interactions between SEN and coursework are close to zero. The effect sizes for ethnicity are also relatively small although there is some evidence that, despite performing well overall at GCSE, students of Chinese ethnicity perform relatively poorly compared with students of other ethnicities when entered for specifications with coursework.

While much has been written about the high performance of Chinese students, nothing has focused on coursework in particular (See, for example, Department for Education, 2015, 2018; Department for Education and Skills, 2006; Strand, 2015). High performance has been attributed to parental influence and the cultural value placed on education (see for example, Francis & Archer, 2005). The literature is

divided on how these influences might affect learning strategies. Woodrow & Sham (2001) suggest that students of Chinese ethnicity may be more inclined to work alone rather than in groups, not value peer discussions and learn best by memorising. These views, however, are dismissed as stereotypical, deficit constructions by Francis, Mau, & Archer (2017). Nevertheless, it is possible that the interactions between coursework and Chinese ethnicity observed in this study indicate that cultural differences proposed in some literature give more support to examined than non-examined assessment.

Finally, for GCSE there are some differences between subjects depending on whether or not the specification includes coursework. The interactions suggest that GCSE mathematics and English Literature students are less impacted by the inclusion of coursework in an assessment. For A-level, there is almost no difference in performance between subjects dependent on coursework.

Table 8 Parameter estimates and effect sizes for the multilevel linear regression analysis of GCSE and A-level grade outcomes (Model 1)

| Fixed Effects	GCSE			A-level		
	β	se	Effect Size	β	se	Effect Size
Intercept	5.352	0.004	0.000	3.330	0.022	0.000
Male	0.118	0.001	0.126	0.139	0.004	0.168
Worsening	-0.034	0.022	-0.036	0.067	0.033	0.081
Improving	0.044	0.016	0.047	-0.018	0.027	-0.021
English Literature	-0.037	0.002	-0.040			
Geography	-0.064	0.018	-0.069	0.000	0.015	0.000
History	-0.100	0.014	-0.107	-0.135	0.015	-0.163
Mathematics	-0.138	0.002	-0.147			
2007	-0.215	0.018	-0.229	-0.059	0.019	-0.071
2010	-0.520	0.003	-0.555	-0.003	0.014	-0.003
2011	-0.485	0.003	-0.518	0.208	0.015	0.251
2015	0.068	0.003	0.073	0.247	0.015	0.297
2017	-0.168	0.002	-0.179	0.146	0.015	0.176
SEN (School)	-0.146	0.002	-0.156	0.026	0.011	0.032
SEN (EHCP/Statement)	-0.031	0.004	-0.034	0.083	0.034	0.100
Free school meals	-0.085	0.002	-0.091	-0.016	0.013	-0.020
Any other ethnic group (AOEG)	0.140	0.005	0.150	-0.065	0.026	-0.079
Asian (ASIA)	0.142	0.003	0.151	0.044	0.012	0.054
Black (BLAC)	0.097	0.003	0.104	-0.036	0.015	-0.043
Chinese (CHIN)	0.216	0.010	0.230	-0.091	0.031	-0.110
Mixed (MIXD)	-0.007	0.003	-0.007	-0.026	0.013	-0.032
Other first language	0.119	0.002	0.127	-0.095	0.011	-0.115
IDACI Rank (2015)	-0.138	0.011	-0.083	-0.021	0.019	-0.014
Concurrent attainment (Norm)	1.113	0.001	1.189	0.639	0.003	0.771
Concurrent attainment (Norm) ^ 2	0.031	0.000	0.033	-0.085	0.001	-0.103
Prior attainment (Norm)	0.516	0.001	0.551	0.495	0.003	0.596
Prior attainment (Norm) ^ 2	0.014	0.000	0.015	0.033	0.002	0.040
Includes coursework	-0.282	0.007	-0.301	0.021	0.023	0.025
Coursework percentage	0.003	0.000	0.003	-0.006	0.000	-0.008
Number of coursework units	0.002	0.002	0.002	0.104	0.004	0.125
Coursework optional	-0.178	0.002	-0.190	-0.061	0.006	-0.073
Male * Includes coursework	-0.198	0.001	-0.211	-0.022	0.005	-0.027
Worsening * Includes coursework	0.021	0.005	0.023	-0.070	0.016	-0.085
Improving * Includes coursework	-0.040	0.004	-0.042	0.011	0.013	0.013
English Literature * Includes coursework	0.135	0.003	0.144			
Geography * Includes coursework	-0.136	0.019	-0.145	-0.007	0.016	-0.009
History * Includes coursework	-0.156	0.014	-0.166	-0.005	0.016	-0.006
Mathematics * Includes coursework	0.080	0.005	0.086			
2007 * Includes coursework	0.193	0.018	0.206	0.156	0.020	0.188
2010 * Includes coursework	0.186	0.003	0.199	0.162	0.016	0.195
2011 * Includes coursework	0.174	0.003	0.186	0.077	0.017	0.093
2015 * Includes coursework	-0.032	0.003	-0.034	0.076	0.017	0.092
2017 * Includes coursework				0.092	0.017	0.110
SEN (School) * Includes coursework	0.053	0.002	0.057	0.014	0.014	0.017
SEN (EHCP/Statement) * Includes coursework	0.104	0.006	0.111	0.076	0.040	0.092
FSM * Includes coursework	0.018	0.002	0.020	0.002	0.015	0.002
AOEG * Includes coursework	-0.160	0.007	-0.171	0.030	0.030	0.036
ASIA * Includes coursework	-0.104	0.003	-0.111	-0.042	0.014	-0.050
BLAC * Includes coursework	-0.067	0.004	-0.071	0.025	0.017	0.031
CHIN * Includes coursework	-0.267	0.012	-0.286	-0.048	0.038	-0.058
MIXD * Includes coursework	0.008	0.004	0.008	0.039	0.015	0.047
Other first language * Includes coursework	-0.139	0.003	-0.149	0.004	0.013	0.005
IDACI Rank (2015) * Includes coursework	0.024	0.003	0.014	0.006	0.010	0.004
Concurrent attainment (Norm) * Includes coursework	0.301	0.001	0.322	0.069	0.003	0.083
Concurrent attainment (Norm) ^ 2 * Includes coursework	0.083	0.000	0.088	-0.002	0.002	-0.002
Prior attainment (Norm) * Includes coursework	-0.265	0.001	-0.283	-0.065	0.004	-0.079
Prior attainment (Norm) ^ 2 * Includes coursework	-0.017	0.000	-0.018	0.016	0.002	0.019
Random Effects	Var	SD		Var	SD	
School residual	0.046	0.214		0.054	0.233	
Student residual	0.877	0.936		0.688	0.829	
Variance partition coefficient	5%			7%		
R_2C	0.715			0.609		

Notes: (i) The baseline subject for GCSE is English language and for A-level is English literature.
(ii) Figures in grey denote that the effect is not statistically significant.

Model 2 – Including centre type information

Table 9 shows the parameter estimates and effect sizes associated with the models of GCSE and A-level grade outcome (Model 2). In this second model, not only are independent schools fully represented, but also the coursework covariate is split to distinguish between I-I coursework and other coursework. In order to maximise the data available, the GCSE model includes concurrent attainment data while the A-level model includes prior attainment data. As with Model 1, for both GCSE and A-level, the effect sizes associated with prior and concurrent attainment are completely dominant.

Once the coursework covariate is split, it is apparent that the different types of coursework relate to grade outcome in different ways. For GCSE, the specifications with I-I coursework behave in a similar manner to those without any coursework at all. On the other hand, the other coursework formats[6] appear to elicit poorer performances. For A-level, coursework seems to have a beneficial effect on grade outcome no matter what the type. This pattern is particularly apparent when considering the interaction between coursework and centre type (Figure 6). It is no surprise to find that students from independent schools gain higher grades than those from state schools. The difference is nearly three quarters of a grade for GCSE but less than a quarter for A-level. Figure 6, suggests that GCSE students from independent schools perform better than their state educated peers when they enter for a specification with either no coursework or I-I coursework compared with when they enter for a specification with another type of coursework. This pattern is not evident for A-level.

*Figure 6 Effect of the interaction between centre type and coursework on grade outcome*
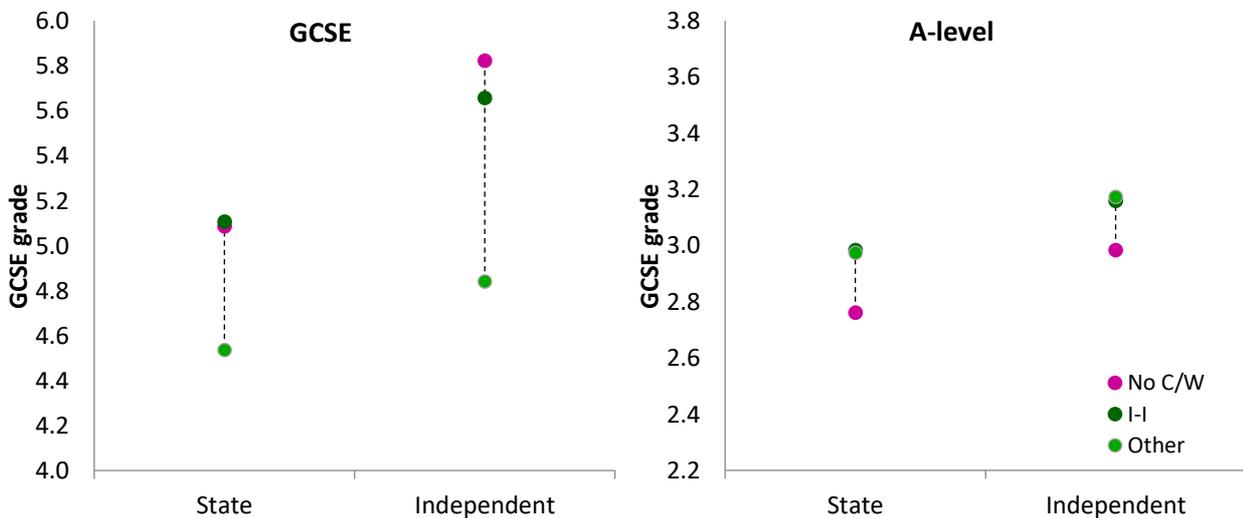
As far as gender is concerned, identifying the separate coursework types reveals an interesting difference between male and female GCSE students. Figure 7 shows that male GCSE outcomes are detrimentally affected whenever the specification includes

[6] Specifications classified in the other coursework category include: internally set/externally marked, externally set/internally marked, externally set/externally marked and specifications in which there is a choice of approaches

coursework, regardless of the coursework format. On the other hand, there is little difference in GCSE outcome for female students entered for specifications with no coursework and for specifications with I-I coursework. This finding is consistent with the existing, albeit dated, literature on type of coursework. Stobart, Elwood, & Quinlan (1992) noted that the greater the level of individual preparation outside the classroom, the better the performance of female students.

Figure 7 Effect of the interaction between gender and coursework on grade outcome
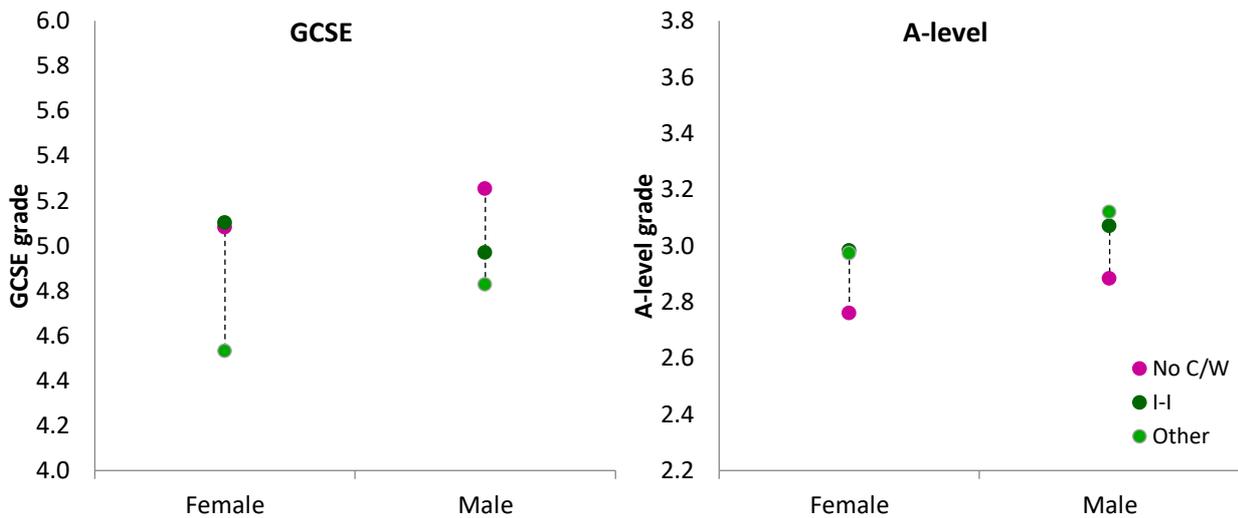
By the time students have progressed to A-level, the type of coursework seems to have little bearing on outcome: students tend to perform better in A-levels with a teacher assessed element regardless of its format. Furthermore, findings from the model, suggest they gain higher grades the greater the number of coursework units included in the A-level. Over the period of study, the number of coursework units in the subjects under investigation has decreased from up to three in history in 2007 to no more than one in any subject by 2017. The effect of multiple coursework units is, therefore, confounded with policy changes to assessment design and with the effects of staged assessment.

As with Model 1, there is very little evidence to suggest that SES, as measured by FSM and IDACI 2015, has any bearing on the relative performance in coursework and non-coursework assessments.

Table 9 Parameter estimate and effect sizes for the multilevel linear regression analysis of GCSE and A-level grade outcomes (Model 2)

Fixed Effects	GCSE			A-level		
	β	se	Effect Size	β	se	Effect Size
Intercept	5.084	0.019	0.000	2.762	0.012	0.000
Male	0.172	0.001	0.175	0.123	0.003	0.126
Independent	0.737	0.011	0.750	0.222	0.013	0.226
Worsening	-0.060	0.028	-0.061	0.048	0.029	0.049
Improving	0.031	0.021	0.031	0.006	0.026	0.006
English Literature	-0.030	0.002	-0.031			
Geography	0.023	0.018	0.023	0.074	0.007	0.075
History	-0.064	0.014	-0.065	-0.090	0.008	-0.092
Mathematics	-0.115	0.002	-0.117			
2007	0.019	0.002	0.019	0.152	0.013	0.155
2010	-0.373	0.018	-0.380	0.192	0.007	0.196
2011	-0.327	0.018	-0.333	0.161	0.007	0.164
2015	0.310	0.018	0.315	0.483	0.008	0.493
2017	0.085	0.018	0.087	0.269	0.008	0.274
Free school meals	-0.134	0.002	-0.136	-0.053	0.010	-0.054
IDACI Rank (2015)	-0.033	0.014	-0.019	-0.150	0.017	-0.082
Attainment (Norm)	1.495	0.001	1.523	0.912	0.002	0.930
Attainment (Norm) ^ 2	0.054	0.000	0.055	0.061	0.001	0.063
Internally set/ Internally marked	0.021	0.019	0.021	0.222	0.013	0.226
Other coursework	-0.550	0.024	-0.560	0.214	0.014	0.218
Coursework percentage	0.004	0.000	0.004	-0.013	0.000	-0.013
Number of coursework units	-0.015	0.002	-0.015	0.270	0.003	0.275
Coursework optional	-0.232	0.002	-0.236	-0.035	0.005	-0.036
Male * Internally set, internally marked	-0.304	0.001	-0.310	-0.034	0.004	-0.035
Male * Other coursework	0.124	0.002	0.126	0.022	0.008	0.022
Independent * Internally set, internally marked	-0.186	0.004	-0.189	-0.047	0.006	-0.048
Independent * Other coursework	-0.430	0.005	-0.438	-0.023	0.012	-0.024
Worsening * Internally set, internally marked	0.012	0.005	0.012	-0.061	0.014	-0.062
Improving * Internally set, internally marked	-0.038	0.004	-0.039	-0.006	0.011	-0.007
Worsening * Other coursework	0.022	0.008	0.022	-0.050	0.026	-0.051
Improving * Other coursework	-0.069	0.006	-0.070	-0.008	0.022	-0.008
English Literature * Internally set, internally marked	0.148	0.003	0.151			
Geography * Internally set, internally marked	-0.161	0.019	-0.164	-0.039	0.010	-0.040
History * Internally set, internally marked	-0.118	0.014	-0.120	0.030	0.008	0.031
Mathematics * Internally set, internally marked	-0.002	0.005	-0.002			
English Literature * Other coursework	0.119	0.010	0.121			
Geography * Other coursework	-0.070	0.025	-0.071	-0.049	0.010	-0.050
Mathematics * Other coursework	0.396	0.029	0.404			
2007 * Internally set, internally marked	-0.102	0.002	-0.104	-0.007	0.014	-0.007
2010 * Internally set, internally marked	0.010	0.018	0.010	0.110	0.009	0.112
2011 * Internally set, internally marked	-0.038	0.018	-0.039	0.178	0.009	0.182
2015 * Internally set, internally marked	-0.261	0.018	-0.266	0.026	0.010	0.027
2017 * Internally set, internally marked	-0.253	0.018	-0.257	0.196	0.010	0.200
2010 * Internally set, internally marked	0.487	0.567	0.496	-0.117	0.015	-0.119
2011 * Internally set, internally marked	0.450	0.033	0.459	0.166	0.098	0.169
2015 * Internally set, internally marked	0.030	0.022	0.031			
FSM * Internally set, internally marked	0.014	0.002	0.014	-0.003	0.013	-0.003
FSM * Other coursework	0.053	0.004	0.054	0.003	0.033	0.003
IDACI Rank (2015) * Internally set, internally marked	-0.014	0.003	-0.008	0.010	0.008	0.005
IDACI Rank (2015) * Other coursework	0.082	0.004	0.047	0.011	0.016	0.006
Attainment (Norm) * Internally set, internally marked	0.062	0.001	0.063	-0.045	0.002	-0.046
Attainment (Norm) * Other coursework	0.049	0.002	0.050	0.033	0.005	0.033
Attainment (Norm) ^ 2 * Internally set, internally marked	0.045	0.000	0.046	-0.007	0.001	-0.007
Attainment (Norm) ^ 2 * Other coursework	0.082	0.001	0.083			
Random Effects	**Var**	**SD**		**Var**	**SD**	
School residual	0.136	0.369		0.065	0.254	
Student residual	0.964	0.982		0.962	0.981	
Variance partition coefficient	12%			6%		
R₂C	0.699			0.497		

Notes: (i) The baseline subject for GCSE is English language and for A-level is English literature.
(ii) Figures in grey denote that the effect is not statistically significant.
(iii) For GCSE, the attainment covariate relates to concurrent attainment but for A-level it relates to prior attainment.

Discussion

The intention of this analysis is to provide evidence of how coursework interacts with grade outcome and how this interaction differs by student characteristics.

The models presented in this paper provide some evidence to support future discussions on the role of coursework. The patterns seen in the data do not imply cause and effect, they merely describe outcomes for different subgroups of the population. Furthermore, because of the inclusion of prior and current attainment in the model, they portray the patterns over and above formative experiences that are embedded in these proxy measures of ability. That said, in many cases the findings support existing research evidence and, because they are based on population data, they provide a firm foundation from which to understand the impact of coursework in qualifications.

Without fitting any models, it is clear that there is a difference in the entry patterns for coursework specifications. Where an assessment includes coursework, it seems to attract higher-attaining students. This is particularly noticeable in GCSE specifications where the coursework is optional. The evidence suggests that specifications with coursework optionality are providing an examined safety net for less able or, perhaps less motivated, students who fail to submit coursework.

By the time students enter for A-level, the patterns of coursework entry seem to change. There is much less evidence of entry bias dependent on prior attainment. Furthermore, the models show that those entering a specification with coursework are liable to do better than their peers. This contrasts with GCSE, where the inclusion of coursework tends to have a negative effect on grade outcome. In his work on validity, Stringer (2014) speculated that differences between GCSE and A-level might exist because of the introduction of controlled assessment. However, the current study spans the years before and after this policy change. It seems more likely that modularity, entry hurdles, maturity and motivation are all instrumental in the positive effect that coursework has on A-level grade outcomes.

Gender differences in attainment are never far from the news headlines. Indeed, in the annual results statistics, GCSE and A-level outcomes are reported separately for male and female students[7]. Throughout the period included in this study, female students have consistently performed better than male students except for at the very top A-level grades. The models, however, tell a different story from the raw data reported by the Joint Council for Qualifications (JCQ). While in absolute terms female students have higher attainment, the value added by secondary education is greater for male students; the gap is closing. The evidence suggests that, after accounting for prior and concurrent attainment, male students perform better than female students in wholly examined GCSE specifications and also in specifications where there is a greater level of control in the coursework. Female students tend to have better outcomes where an internally set, internally marked coursework is included. In her study of gender equity in teacher assessed tasks, Elwood (1999) concluded that 'although females may benefit from coursework in curriculum terms, it does not appear to be the sole factor that explains their better overall performance in relation to their male counterparts'. Based on the subjects included in this study,

---

[7] JCQ examination results

prior educational experience contributes to better female performance but this advantage is boosted by the inclusion of I-I coursework.

The extent to which student characteristics affect performance in GCSE and A-level coursework specifications seems relatively limited, despite previous research conducted elsewhere (Harlen, 2005; Ready & Wright, 2011; Wyatt Smith & Castleton, 2005). Gender aside, coursework has no impact on outcomes for students of different SES and close to no impact for those with SEN. Across all specifications, independent schools continue to out-perform state schools but, for the GCSE specifications included in the study, the extent of this advantage is less where non-I-I coursework forms part of the assessment.

One interesting outcome is that, while performing well overall at GCSE, students of Chinese ethnicity perform relatively poorly when entered for specifications with coursework. One possible explanation is that the cultural differences proposed in some literature give more support to examined than non-examined assessment.

Over time, changes to the content and structure of summative school-end assessments are inevitable and to be expected given the evolution of educational thinking. However, it is essential that these changes are not made in a vacuum; they should be supported by robust research evidence. There is no doubt that the inclusion of coursework in GCSE and A-level qualifications does favour some students over others. The findings from this study suggest those likely to be affected, either negatively or positively, by any future changes to coursework are: students that rely on an examined safety net when struggling with coursework; male and female students; and students of Chinese ethnicity. The subject-level differences revealed by the modelling also point to the need for assessment solutions that are appropriate to each individual subject while also considering the validity and equity of any coursework or non-examined assessment units.

Limitations

The findings from this study may be limited to the extent that they can be generalised beyond the subjects under consideration. Nevertheless, these subjects were chosen specifically because they have large entries and include specifications with coursework. The data used to derive the models account for over a quarter of the total GCSE entry and over a fifth of the total A-level entry. Because the data is population data over a 14-year period, it is likely to give a reasonably robust picture of the effect of coursework in core GCSE and A-level subjects. However, even though the data is longitudinal in nature, the findings do not describe time-dependent changes in grade outcome. This is because an academic year is almost completely confounded with key policy changes to coursework and with the specifications themselves. Moreover, the necessity to normalise the measures of prior and concurrent attainment means that year-on-year differences have been partially obscured.

The volume of data included in the analysis has presented significant computational challenges, precluding the use of matching and of logistic regression. Nevertheless, the arguments against the use of matching, in particular, are strong. Matching requires the data to be balanced on all covariates thought to affect differences in performance between the case and control variables; in this study coursework and non-coursework entries. The number of missing values for some of the covariates

undermines the procedure, as does the fact that not all covariates are measurable. The multilevel analysis presented suffers from some of the same limitations but the creation of two models makes explicit the weaknesses in the data. The models use population data and represent the real-world situation rather than attempting to infer what may have been the situation had the case and control variables been balanced.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models ssing lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Chamberlain, S. (2013). Sleeping in or selecting out? *Research in Education*, *79*(1), 53–66.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. London: Academic Press.

Department for Education. (2015). *Statistical first release GCSE and equivalent attainment by pupil* (Vol. SFR06/2015).

Department for Education. (2018). GCSE results ('Attainment 8') for children aged 14 to 16 (key stage 4).

Department for Education and Skills. (2001). *Speical educational needs code of practice*.

Department for Education and Skills. (2006). *Ethnicity and education: The evidence on minority ethnic pupils aged 5–16*.

Elwood, J. (1999). Equity issues in performance assessment: The contribution of teacher-assessed coursework to gender-related differences in examination performance. *Educational Research & Evaluation*, *5*(4), 321.

Francis, B., & Archer, L. (2005). British-Chinese pupils' and parents' constructions of the value of education. *British Educational Research Journal*, *31*(1), 89–108.

Francis, B., Mau, A., & Archer, L. (2017). The construction of British Chinese educational success: exploring the shifting discourses in educational debate, and their effects. *Journal of Ethnic and Migration Studies*, *43*(14), 2331–2345.

Gipps, C. (1995). What Do We Mean by Equity in Relation to Assessment? *Assessment in Education: Principles, Policy & Practice*, *2*(3), 271–281.

Hansard HC Deb. (2012a). 17 September Vol 550 Column 653.

Hansard HC Deb. (2012b). 17 September Vol 550 Column 660.

Harlen, W. (2005). Trusting teachers ' judgement: research evidence of the reliability and validity of teachers ' assessment used for summative purposes. *Research Papers in Education*, *20*(3), 245–270.

Ilie, S., Sutherland, A., & Vignoles, A. (2017). Revisiting free school meal eligibility as a proxy for pupil socio-economic deprivation. *British Educational Research Journal*, *43*(2).

Johnson, S. (2011). *A focus on teacher assessment reliability in GCSE and GCE*. *Ofqual Reliability Compendium*. Ofqual/11/4807.

Qualifications and Curriculum Authority. (2006). *A review of GCSE coursework*.

Ready, D. D., & Wright, D. L. (2011). Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities: The role of child background and classroom context. *American Educational Research Journal*, *48*(2), 335–360.

Sammons, P. (1995). Gender, ethnic and socio-economic differences in attainment

and progress: a longitudinal analysis of student achievement over 9 years. *British Educational Research Journal*, *21*(4), 465–485.

Scott, D. (1991). Issues and themes: Coursework and coursework assessment in the GCSE. *Research Papers in Education*, *6*(1), 3–19.

Snijders, T., & Bosker, R. (1994). Modeled variance in two-level models. *Sociological Methods & Research*, *22*(3), 342–363.

Stobart, G., Elwood, J., & Quinlan, M. (1992). Gender bias in examinations : How equal are the opportunities? *British Educational Research Journal*, *18*(3), 261–276.

Strand, S. (2015). *Ethnicity, deprivation and educational achievement at age 16 in England: trends over time: Annex to compendium of evidence on ethnic minority resilience to the effects of deprivation on attainment*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/att achment_data/file/439867/RR439B-Ethnic_minorities_and_attainment_the_effects_of_poverty_annex.pdf.pdf

Stringer, N. (2014). *The achieved weightings of assessment objectives as a source of validity evidence*. Ofqual/14/5375.

Woodrow, D., & Sham, S. (2001). Chinese pupils and their learning preferences. *Race Ethnicity and Education*, *4*(4), 377–394.

Wyatt Smith, C., & Castleton, G. (2005). Examining how teachers judge student writing: an Australian case study. *Journal of Curriculum Studies*, *37*(2), 131–154.

Appendix A – Covariate formulation

The baseline level of a factor is coloured GREEN AND PRESENTED IN SMALL CAPS

Covariate	GCSE	A-Level
Gender	FEMALE Male	FEMALE Male
Subject	ENGLISH LANGUAGE English Literature Geography History Mathematics	ENGLISH LITERATURE Geography History
Academic year	2004 2007 2010 2011 2015 2017	2004 2007 2010 2011 2015 2017
Special educational needs (SEN)	NO SEN SEN (School) SEN (EHCP/Statement)	NO SEN SEN (School) SEN (EHCP/Statement)
Free school meals (FSM)	NO FSM FSM	NO FSM FSM
Ethnicity	WHIT – WHITE AOEG – Any other ethnic group ASIA – Asian BLAC – Black CHIN – Chinese MIXD – Mixed race	WHIT – WHITE AOEG – Any other ethnic group ASIA – Asian BLAC – Black CHIN – Chinese MIXD – Mixed race
English as an additional language (EAL)	ENGLISH FIRST LANGUAGE English as an additional language	ENGLISH FIRST LANGUAGE English as an additional language
Concurrent attainment	Normalised total GCSE (or equivalent) point score	Normalised total A-level point score
Prior attainment	Normalised average KS2 score	Normalised total GCSE (or equivalent) point score
IDACI 2015	Centred IDACI rank (higher rank implies higher level of deprivation)	Centred IDACI rank (higher rank implies higher level of deprivation)
IDACI Trend	Worsening STABLE Improving	Worsening STABLE Improving
Coursework percentage	Uncentred percentage	Uncentred percentage
Coursework type	NO C/W I-I Other	NO C/W I-I Other
Number of coursework units	Uncentred number	Uncentred number
Coursework optional	NOT OPTIONAL OR INCLUDES NO COURSEWORK Optional	NOT OPTIONAL OR INCLUDES NO COURSEWORK Optional

Appendix B – Student-level IDACI

Model 1 has been refitted including student IDACI in order to evaluate this covariate as a measure of socio-economic status (Model 3). Student IDACI is only available for students with GCSE and A-level outcomes in the academic years ending 2015 and 2017. Findings, therefore, include no data from some years where key policy changes were made and are only based on about 35% of the GCSE data and 16% of the A-level data. Table B 1 shows the parameter estimates and effect sizes associated with Model 3.

Even though the model does not cover all years, the findings are broadly the same as those from Model 1. Measures of prior and concurrent attainment are still the dominant factors associated with grade. Gender differences persist, although they are somewhat more muted. And when all other covariates are controlled, SES has little bearing on grade outcome. It is worth remembering at this point that the effect of SES may well be wrapped up in the measures of prior and concurrent attainment. Nevertheless, what the new model does show is that, though educationally insignificant, student-level IDACI has a larger effect size for GCSE outcome than centre-level IDACI. However, the reverse is true at A-level.

For completeness, an additional model has been fitted (Table B 2). Model 4 includes only main effects and describes grade outcome when prior and concurrent attainment are not controlled. Naturally, this model provides a much poorer fit for the data. For GCSE outcomes, the effect sizes suggest that requirement for SEN has the greatest impact on grade outcome. This is followed by Chinese ethnicity, then eligibility for FSM and, only after that, student IDACI. The effect size for school IDACI is somewhat smaller. So, for GCSE, the stripped-back model of grade outcome throws up an interesting finding: eligibility for free school meals seems to be a better predictor of grade outcome than a student-specific measure of socio-economic status. This possibly warrants further research to understand why this may be the case and whether, for example, being eligible for free school meals has a greater impact upon concepts of self-efficacy in the context of education than immediate neighbourhood environment.

By the time students get to A-level, the picture is different and this might be because of the selection effect inherent in A-level entry. Socio-economic status – where it is measured by FSM, school IDACI or student IDACI – is still not the most important factor associated with grade outcome. In fact, for A-level, the inclusion of coursework on a specification appears to have the greatest impact. This is followed by black ethnicity (negative), then requirement for SEN (negative), gender and, only after that, student IDACI. In terms of relative impact, eligibility for free school meals has a smaller effect size than both student and school IDACI.

The use of FSM as a proxy for SES was discussed by Ilie, Sutherland, & Vignoles (2017). They concluded that this measure was sufficiently good that any disadvantages it had, compared with more sophisticated measures, were outweighed by the completeness of available data. Their conclusions, however, were based on Key Stage 4 data alone. While evidence presented here supports their Key Stage 4 findings, it casts some doubt on whether they are generalisable beyond this stage of education.

Table B 1 Parameter estimate and effect sizes for the multilevel linear regression analysis of GCSE and A-level grade outcomes (Model 3)

| Fixed Effects	GCSE			A-level		
	β	se	Effect Size	β	se	Effect Size
Intercept	5.612	0.004		3.789	0.032	
Male	0.027	0.001	0.030	0.091	0.007	0.116
Worsening	-0.048	0.022	-0.053	0.077	0.044	0.098
Improving	0.050	0.017	0.055	0.005	0.036	0.006
English Literature	-0.039	0.002	-0.043			
Geography	-0.401	0.003	-0.446	-0.206	0.031	-0.263
History	-0.496	0.003	-0.551	-0.148	0.032	-0.189
Mathematics	-0.136	0.002	-0.152			
2017	-0.238	0.002	-0.265	-0.072	0.007	-0.091
SEN (School)	-0.072	0.002	-0.080	0.040	0.019	0.052
SEN (EHCP/Statement)	0.018	0.005	0.020	0.082	0.053	0.105
Free school meals	*-0.044*	*0.002*	*-0.049*	*0.002*	*0.020*	*0.003*
Any other ethnic group (AOEG)	0.062	0.006	0.069	-0.023	0.037	-0.029
Asian (ASIA)	0.093	0.003	0.103	0.084	0.017	0.107
Black (BLAC)	0.072	0.004	0.080	0.012	0.022	0.015
Chinese (CHIN)	0.064	0.012	0.072	-0.040	0.048	-0.051
Mixed (MIXD)	-0.003	0.003	-0.003	0.028	0.019	0.036
Other first language	0.017	0.003	0.019	-0.077	0.016	-0.098
IDACI Rank (2015)	*0.012*	*0.011*	*0.008*	*-0.045*	*0.025*	*-0.031*
IDACI Rank (Pupil)	*-0.173*	*0.006*	*-0.058*	*-0.081*	*0.033*	*-0.028*
Concurrent attainment (Norm)	1.306	0.001	1.452	0.715	0.005	0.913
Concurrent attainment (Norm) ^ 2	-0.048	0.001	-0.054	-0.075	0.003	-0.095
Prior attainment (Norm)	0.246	0.001	0.273	0.454	0.005	0.579
Prior attainment (Norm) ^ 2	-0.011	0.000	-0.012	-0.005	0.003	-0.006
Includes coursework	-0.064	0.005	-0.071	-0.307	0.035	-0.392
Coursework percentage	0.006	0.000	0.007	0.002	0.001	0.002
Number of coursework units	-0.235	0.005	-0.262	0.050	0.011	0.064
Male * Includes coursework	-0.144	0.002	-0.160	0.025	0.009	0.032
Worsening * Includes coursework	0.019	0.007	0.021	-0.066	0.027	-0.084
Improving * Includes coursework	-0.042	0.006	-0.047	0.005	0.021	0.006
History * Includes coursework				0.078	0.033	0.099
2017 * Includes coursework	0.076	0.003	0.084	0.040	0.010	0.050
SEN (School) * Includes coursework	0.046	0.004	0.051	0.027	0.023	0.035
SEN (EHCP/Statement) * Includes coursework	0.135	0.009	0.151	0.077	0.063	0.098
*FSM * Includes coursework*	*0.008*	*0.003*	*0.009*	*-0.005*	*0.023*	*-0.006*
AOEG * Includes coursework	-0.108	0.010	-0.120	0.010	0.044	0.013
ASIA * Includes coursework	-0.069	0.005	-0.077	-0.050	0.020	-0.064
BLAC * Includes coursework	-0.044	0.005	-0.049	0.021	0.025	0.026
CHIN * Includes coursework	-0.192	0.018	-0.214	-0.048	0.063	-0.062
MIXD * Includes coursework	0.002	0.005	0.002	0.002	0.022	0.003
Other first language * Includes coursework	-0.113	0.004	-0.125	0.000	0.019	0.000
*IDACI Rank (2015) * Includes coursework*	*-0.001*	*0.004*	*-0.001*	*-0.028*	*0.017*	*-0.019*
*IDACI Rank (Pupil) * Includes coursework*	*0.093*	*0.008*	*0.031*	*0.096*	*0.038*	*0.034*
Concurrent attainment (Norm) * Includes coursework	0.278	0.002	0.308	0.038	0.006	0.048
Concurrent attainment (Norm) ^ 2 * Includes coursework	0.115	0.001	0.128	-0.003	0.003	-0.003
Prior attainment (Norm) * Includes coursework	-0.192	0.002	-0.213	-0.089	0.006	-0.113
Prior attainment (Norm) ^ 2 * Includes coursework	0.004	0.001	0.005	0.031	0.004	0.040
Random Effects	**Var**	**SD**		**Var**	**SD**	
School residual	0.034	0.184		0.063	0.251	
Student residual	0.809	0.900		0.614	0.784	
Variance partition coefficient	4%			9%		
R²C	**0.739**			**0.610**		

Notes: (i) The baseline subject for GCSE is English language and for A-level is English literature.
(ii) Figures in grey denote that the effect is not statistically significant.
(iii) SES covariates are in italics and bold.

Table B 2 Parameter estimate and effect sizes for the multilevel linear regression analysis of GCSE and A-level grade outcomes (Model 4)

Fixed Effects	β	se	Effect Size	β	se	Effect Size
Intercept	5.710	0.011		2.981	0.016	
Male	-0.334	0.002	-0.216	-0.178	0.005	-0.145
Worsening	0.211	0.063	0.137	0.115	0.071	0.094
Improving	-0.020	0.046	-0.013	0.035	0.058	0.028
English Literature	-0.021	0.003	-0.014			
Geography	-0.311	0.005	-0.201	0.156	0.009	0.127
History	-0.322	0.005	-0.208	0.064	0.008	0.052
Mathematics	-0.159	0.003	-0.103			
2017	-0.106	0.002	-0.069	0.051	0.006	0.041
SEN (School)	-1.390	0.003	-0.900	-0.199	0.013	-0.162
SEN (EHCP/Statement)	-1.728	0.007	-1.119	-0.128	0.034	-0.104
Free school meals	*-0.555*	*0.003*	*-0.359*	*-0.128*	*0.011*	*-0.104*
Any other ethnic group (AOEG)	0.164	0.008	0.106	-0.128	0.022	-0.104
Asian (ASIA)	0.307	0.004	0.199	-0.161	0.011	-0.131
Black (BLAC)	0.032	0.005	0.020	-0.245	0.012	-0.200
Chinese (CHIN)	0.696	0.014	0.450	0.036	0.035	0.029
Mixed (MIXD)	0.077	0.004	0.050	-0.023	0.011	-0.019
Other first language	-0.111	0.003	-0.072	-0.045	0.010	-0.037
IDACI Rank (2015)	*-0.626*	*0.030*	*-0.226*	*-0.275*	*0.039*	*-0.120*
IDACI Rank (Pupil)	*-1.594*	*0.007*	*-0.313*	*-0.636*	*0.022*	*-0.142*
Includes coursework	-0.010	0.008	-0.007	0.357	0.015	0.290
Coursework percentage	0.005	0.000	0.003	-0.018	0.001	-0.015
Number of coursework units	-0.086	0.009	-0.056	0.465	0.010	0.379
Random Effects	**Var**	**SD**		**Var**	**SD**	
School residual	0.364	0.603		0.276	0.526	
Student residual	2.387	1.545		1.512	1.229	
Variance partition coefficient	**13%**			**15%**		
R^2C	**0.263**			**0.193**		

Appendix C – Descriptive graphs

GCSE	1 Mean grade by gender	A-level
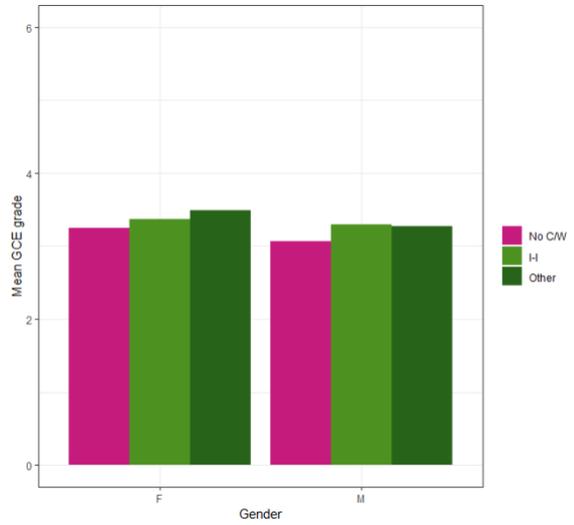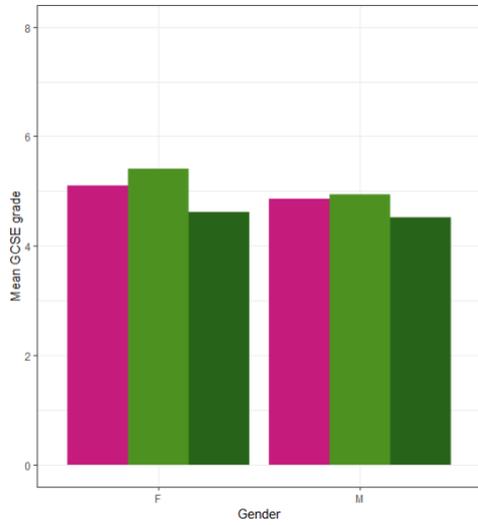
GCSE	2 Mean grade by free school meals	A-level
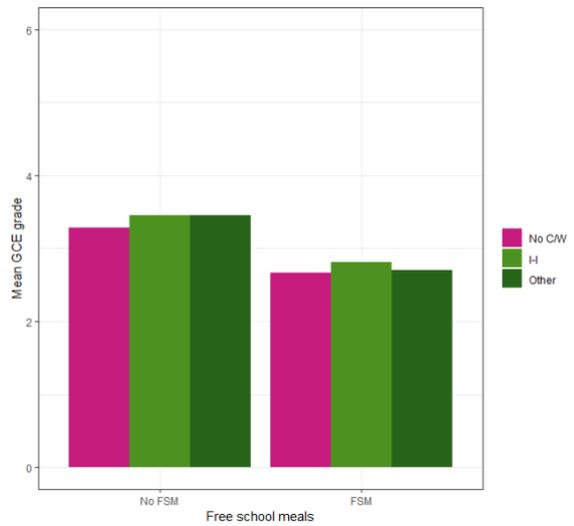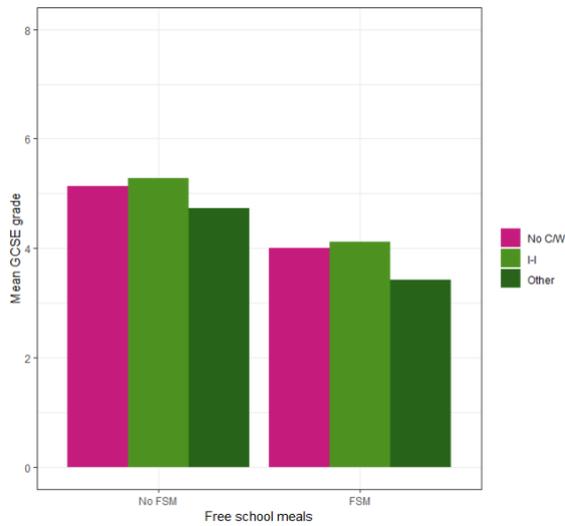
GCSE   **3 Mean grade by ethnicity**   **A-level**
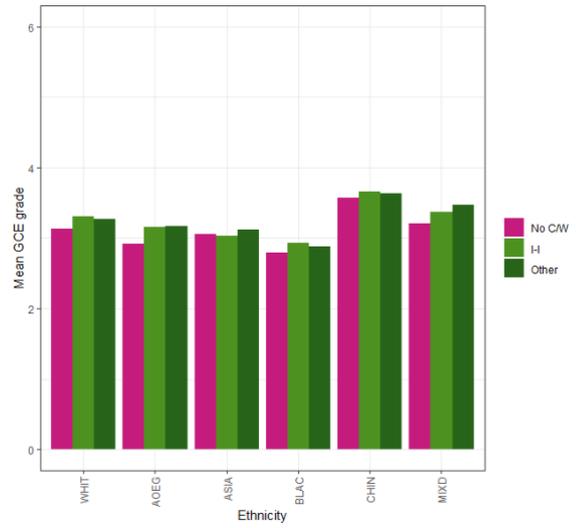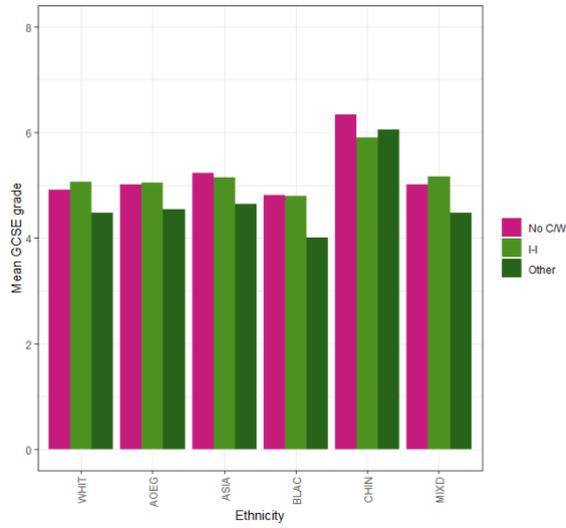
GCSE   **4 Mean grade by special educational needs requirements**   **A-level**
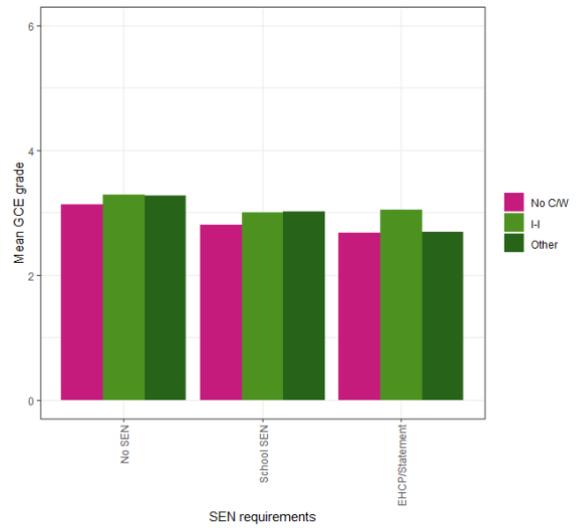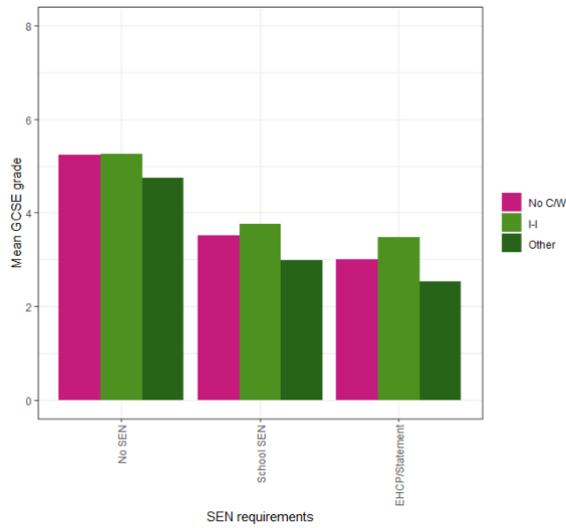
**GCSE**   **5 Mean grade by English as an additional language**   **A-level**
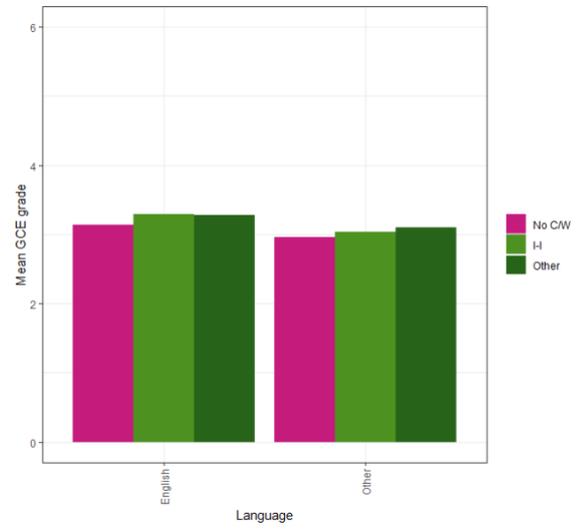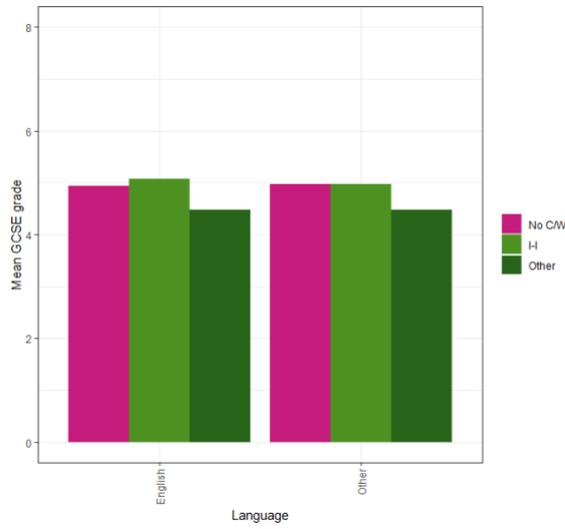
GCSE   **6 Mean grade by IDACI trend**   **A-level**
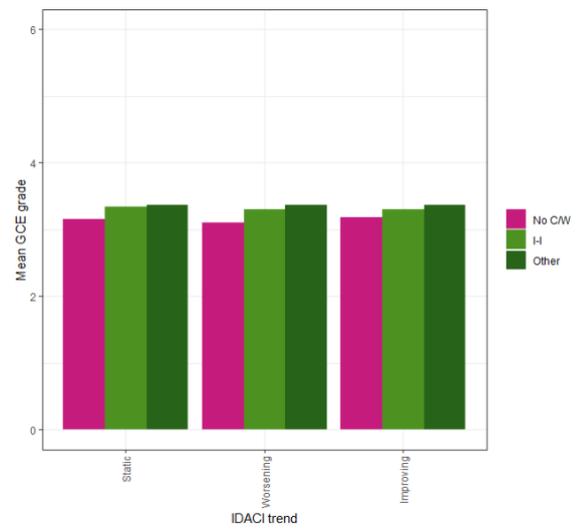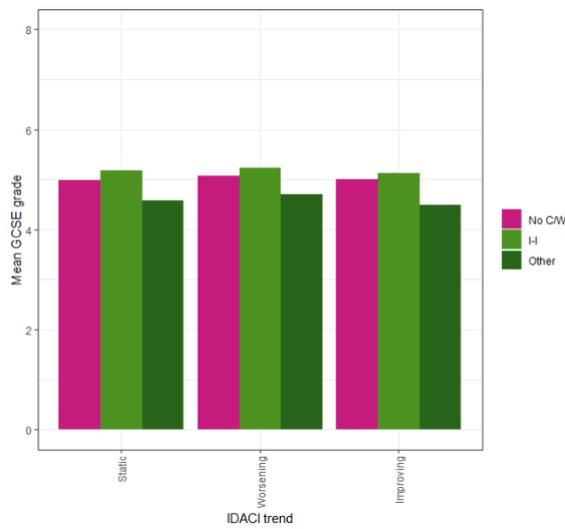
GCSE **7 Mean grade by school IDACI rank** **A-level**
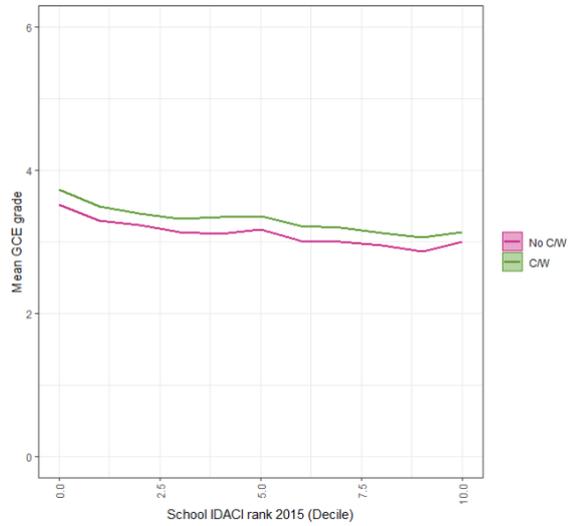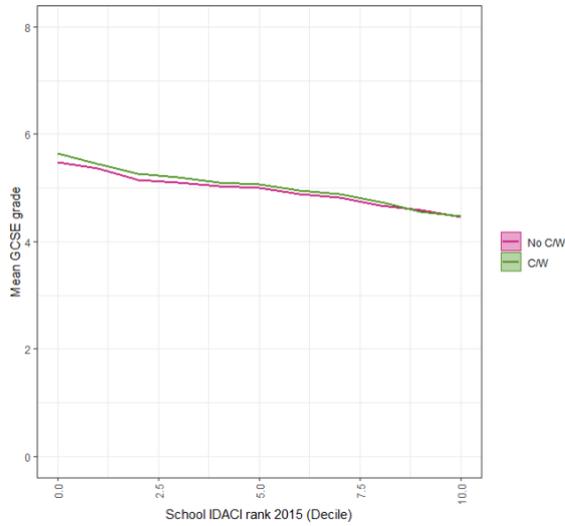
GCSE **8 Mean grade by student IDACI rank** **A-level**
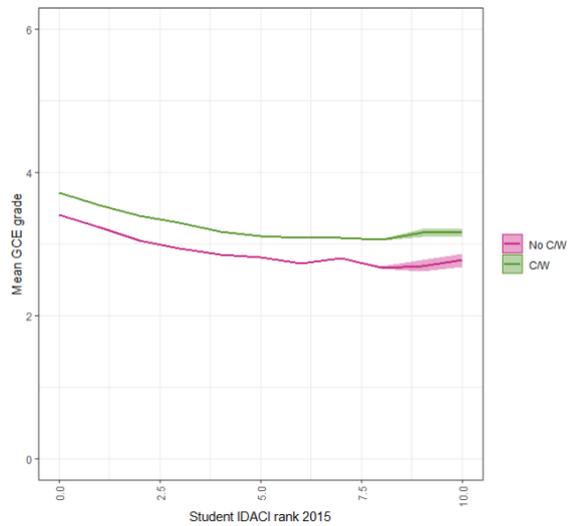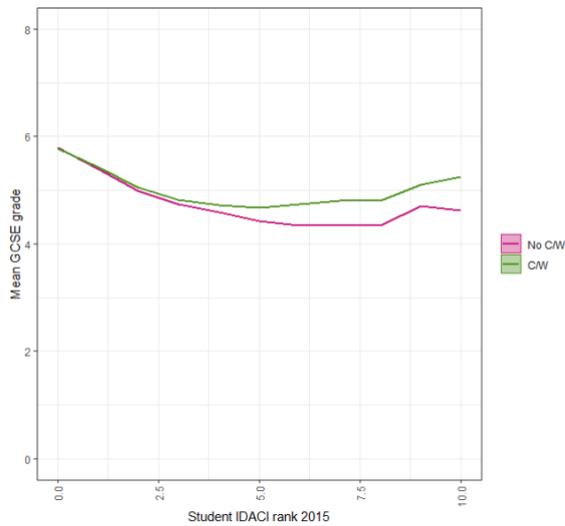
GCSE   **9 Mean grade by prior attainment**   **A-level**
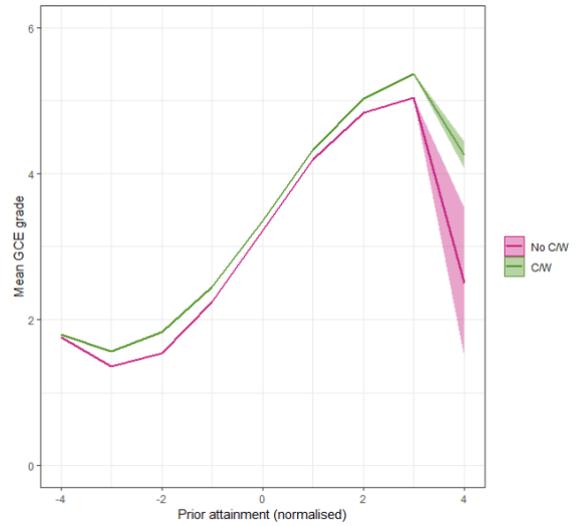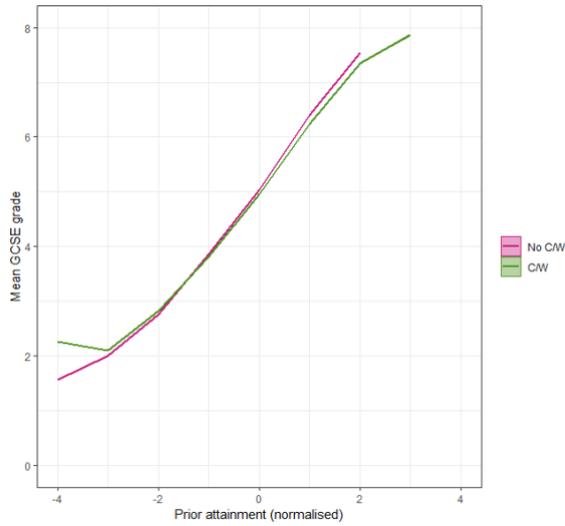
GCSE   **10 Mean grade by concurrent attainment**   **A-level**
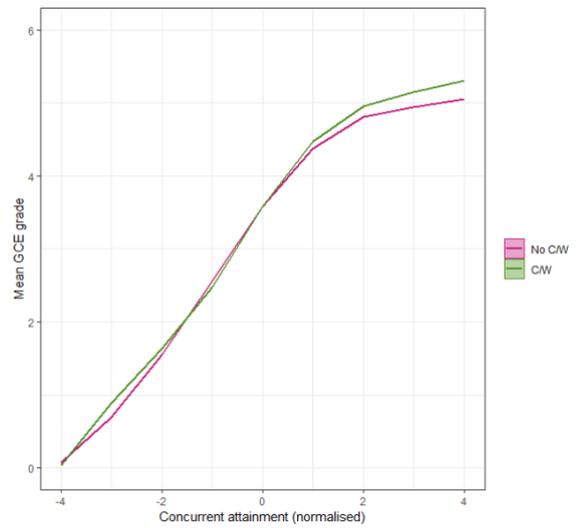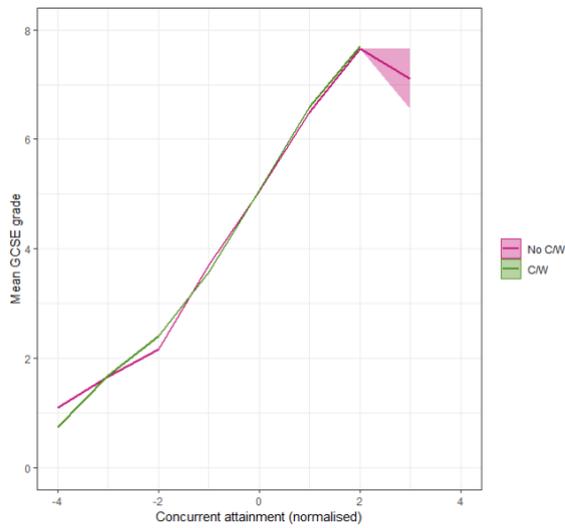
GCSE **11 Mean grade by percentage of coursework** **A-level**
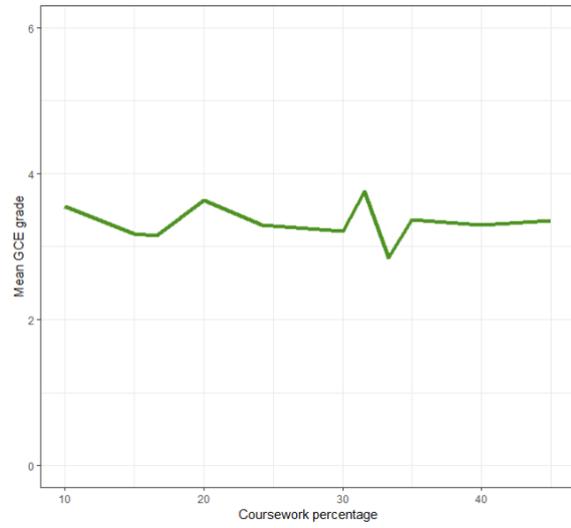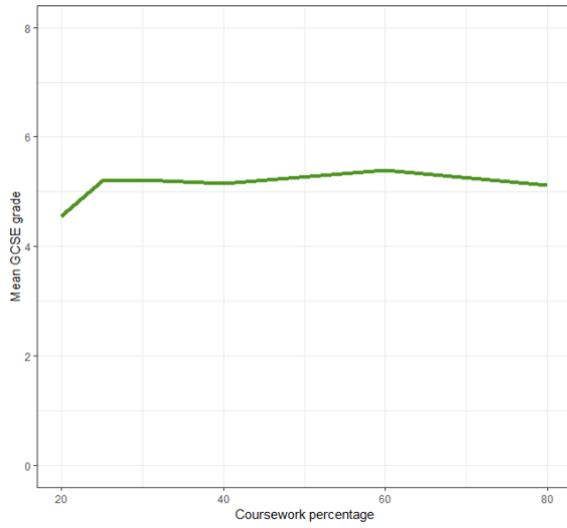
Published by:

ofqual

Earlsdon Park
53-55 Butts Road
Coventry
CV1 3BH

0300 303 3344
public.enquiries@ofqual.gov.uk
www.gov.uk/ofqual

June 2020   Ofqual/20/6629