

REPORT

An investigation of the variability in grade outcomes in Extended Project Qualification

Results from analysis of data from National Pupil Database

ofqual

Authors

Qingping He and Beth Black from Ofqual's Strategy Risk and Research directorate

Contents

<u>Executive summary</u>	4
<u>Introduction</u>	4
<u>Methodology</u>	5
<u>Data</u>	5
<u>Analysis</u>	6
<u>Relationships between EPQ outcomes and prior/concurrent attainment</u>	6
<u>Multinomial logistic regression</u>	7
<u>Propensity score matching (PSM)</u>	7
<u>Results and discussion</u>	7
<u>Changes in EPQ entries and grade outcomes over time</u>	7
<u>Relationship between EPQ outcome and prior/concurrent attainment</u>	9
<u>Correlations with other performance indicators and change in ability over time</u> ..	9
<u>Correlations with A level subjects</u>	11
<u>Relationships with prior/concurrent attainment – grade characteristic curves</u> ...	12
<u>Multinomial logistic regression against prior/concurrent attainment</u>	15
<u>Change in other characteristics over time</u>	28
<u>Propensity score matching analysis</u>	28
<u>Stability of EPQ grade boundaries over time</u>	30
<u>Conclusion</u>	32
<u>References</u>	33

Executive summary

This report presents findings from an analysis of the variability over time in the relationships between grade outcomes of the Extended Project Qualification (EPQ) and other variables that could potentially influence students' performance on EPQ. This involved analysis of data extracted from the National Pupil Database (NPD) from 2009 to 2016. The variables used in this study included students' prior and concurrent attainment (attainment at KS2, GCSE and A level), type of school attended, gender, number of years that the school has entered students for EPQ consecutively (NCY), social economic status (as represented by whether the student is eligible for free school meals - FSM), special educational needs (SEN), ethnicity and first language used at home. The analysis involved empirical examination of the data and statistical modelling using multinomial logistic regression and propensity score matching (PSM).

It has been found that there have been changes in the relationships between EPQ and the other variables over time, particularly since 2014. Students from 2014 onwards generally had better EPQ grade outcomes than similar students (in terms of values on the set of variables used for comparison) from 2013 and before. Empirical analysis of the relationship between EPQ grade outcome and performance at GCSE, A level and KS2 indicated that there was a noticeable increase in mean EPQ grade of about 0.20 grade from 2013 to 2014 and an increase of about 0.46 grade from 2010 to 2016 for students with similar prior or concurrent attainment. Multinomial logistic regression analysis suggested that grade C in 2016 is about half a grade easier than the same grade in 2010 and grade A a third of a grade easier. PSM analysis indicated that from 2013 to 2014, there could be an increase of about 0.21 grade in mean EPQ grade for similar students from 2013 to 2014 and about 0.48 grade from 2010 to 2016. Findings from this investigation indicate that there might have been modest grade inflation in EPQ grade outcomes, particularly since 2014.

Introduction

The Extended Project Qualification was first introduced in 2008. The entries have gone up steadily, from slightly over 1,700 in 2008 to over 38,000 in 2016 (see Gill, 2016; also Table 1 in the next section). This has now made it a significant qualification. It could grow further as schools and colleges may move to offer this qualification alongside A levels, especially as AS entries are declining as a result of decoupling from A level (Ofqual, 2017). There has also been some suggestion that the qualification is increasingly valued by learners as well as higher education (HE) institutions, because of the learning and research skills it helps to develop. This may drive greater student demand and greater value placed on results achieved (Stephenson, 2017).

During the reform of GCSEs and A levels, Ofqual has paid particular attention to the proportions of non-exam assessment (NEA) in these. We have required that the amount of NEA is the minimum necessary to assess the content requirements. This is due largely to the issues with authenticity and the reliability of marking. This has resulted in a significant reduction in (and in some cases removal of) NEA from reformed GCSEs and A levels. Where it remains, we have generally put in place significant new controls. This was driven by concerns about how NEA was operating in many unreformed GCSEs and A levels. The Extended Project is therefore unusual

in that it is assessed entirely through non-exam assessment. Given the nature of the qualification and the wide range of potential projects that students may produce, the nature and level of regulatory controls that can be imposed are limited.

There has been a steady, but noticeable, increase in middle to higher EPQ grades achieved since its launch in 2008 (see Gill, 2016; Table 2 below) which we want to explore to understand further. In particular, we want to explore through the available data any possible explanations for the increase which may include changes in the cohort taking these qualifications, teachers' increased familiarity with the assessment requirements and issues with awarding bodies' approaches to maintaining standards. Without further analysis we cannot be sure what is causing these changes. In particular, we want to know whether there is evidence of grade inflation in EPQ. To achieve this, work was planned to answer the following questions:

- How is students' achievement in EPQ related to their performance in GCSE/A level qualifications and other performance indicators, and whether this has changed over time?
- Has the ability (e.g. mean GCSE score) and other characteristics of the cohort taking EPQ changed over time?
- What is the nature of correlations between EPQ and other A level subjects over time?
- To what extent have grade boundaries been moved over time (in order to control outcomes)?

To answer these questions, this report presents findings from work looking at cohort data from the National Pupil Database (NPD) and reviewing raw mark and grade boundary data from awarding bodies.

Methodology

Data

In this study, candidate level data, which include some basic background information and EPQ grade, and their prior attainment at KS2, GCSE and GCE A level from 2008 to 2016, were collected from the NPD for analysis. Table 1 below shows the EPQ entries from 2008 to 2016 (these are the number of students with a valid EPQ grade). The entries have increased over 20 fold to over 38,000 since its introduction in 2018, making it a significant Level 3 qualification. However, the data from 2008 was not included in the analysis, as this was the first year EPQ was introduced, A* was not awarded in that year and the entry was small.

Table 1 EQP entries from 2008 to 2016.

Year	EPQ entry
2008	1,706
2009	5,511
2010	17,604
2011	26,539
2012	31,549
2013	34,155
2014	36,620
2015	38,096
2016	38,548

In addition to the data collected from the NPD, grade boundary marks were obtained from the awarding bodies' websites.

Analysis

As the focus of this study was to provide tentative explanations about the improvement in EPQ outcomes over time, particularly whether such improvement is attributable to possible grade inflation, analysis of the data was focused on variation over time of the relationships between EPQ grade outcomes and other variables that could potentially affect students' EPQ performance through empirical examination of data and statistical modelling. These variables include a broad range of factors such as students' ability (as represented by their prior and concurrent attainment), types of schools attended by the students, and other characteristics. Any increase in EPQ grade outcomes over time, after controlling for these variables, could potentially be an indication of possible grade inflation.

The prior and concurrent attainment used in this study includes the performance of the EPQ students at GCSE, A level and KS2. To facilitate the analysis, the GCSE, A level and EPQ letter grades were converted into numerical values representing ordered category scores: U→0, G→1, F→2, E→3, D→4, C→5, B→6, A→7, A*→8 for GCSEs, and U→0, E→1, D→2, C→3, B→4, A→5, A*→6 for A levels and EPQ. KS2 numerical level was retained. The mean GCSE score for a candidate was calculated as the average of the numerical grades of the GCSE subjects taken by the candidate, and the mean A level score was calculated as the average of the grades of all A level subjects taken. The mean KS2 score was calculated as the mean of the numerical KS2 maths level and English level.

Relationships between EPQ outcomes and prior/concurrent attainment

Students' ability (as represented by their prior or concurrent attainment) is one of the primary factors that can affect their performance in examinations. The relationships between EPQ grade outcomes and attainment at GCSE, KS2 and A level were investigated empirically first. Assuming that the standards of the prior or concurrent attainment have been maintained, changes in the relationships over time could suggest changes in EPQ grade standards.

Multinomial logistic regression

Analysis using multinomial logistic regression involving the use of prior and concurrent attainment was conducted, in order to investigate how the relationships between EPQ grade outcome and prior/concurrent attainment vary over time in more detail. For multinomial logistic regression, the relationship between each EPQ outcome grade (performance category) for a specific year and the prior or concurrent attainment is modelled separately. The lowest EPQ outcome (U) was taken as the reference category (category 0) in the present analysis. In a specific year, the logarithm of the ratio of the probability of an EPQ student being classified into a specific category to the probability of being classified into the reference category is modelled as a linear function of the prior or concurrent attainment. Grade difficulty could be defined using the model parameters. Variation of grade difficulty or values of model parameters over time would suggest differences in EPQ grade standards for students with similar prior or concurrent attainment distribution.

Propensity score matching (PSM)

Many characteristics of the EPQ cohort in one year that could potentially affect the performance of the students might be different from those of cohorts in other years. Such characteristics might include students' prior attainment (for example KS2 and GCSE results) and/or the concurrent performance measure, such as A level results, gender, socio-economic status, types of school attended and many other factors. If 2 EPQ cohorts are different with respect to these characteristics, it would be difficult to attribute any difference in EPQ outcomes to difference in grading standards between the 2 years. To make the comparison of EPQ outcomes between 2 years more meaningful, students from different years should ideally be similar in terms of those characteristics. Propensity score matching (PSM) could be used to achieve this. Propensity score (PS, the probability of being included in one of the 2 years – control group or treatment group) for each student was estimated based on their scores on the characteristics used for comparison using a statistical model. The distributions of the background characteristics were made similar for the students from the 2 years by assigning appropriate weights calculated using propensity scores. Difference in EPQ grade distribution between the 2 groups that were matched on a broad range of characteristics that could influence EPQ outcomes, would suggest a difference in EPQ grade standards between the years. The PSM R package *twang* was used for PS analysis (see Ridgeway et al., 2017) in this study.

Results and discussion

Changes in EPQ entries and grade outcomes over time

Table 2 shows EPQ grade outcomes from 2009 to 2016 (also see Figure 1). There has been a steady increase in the percentage of candidates achieving higher grades (grade B and above) from 2009. For example, the percentage of candidates achieving A* increased from about 10.5% in 2009 to 17.8% in 2016. Cumulative percentages of students receiving C or above increased from about 68.7% in 2009 to 81.3% in 2016. Changes in grade outcomes over time could be caused by a range of

factors such as changes in performance, changes in the national cohort in terms of ability, familiarity of the assessment requirement, improvement in teaching, grade inflation, and others.

Table 2 EPQ grade outcomes (cumulative %) from 2009 to 2016.

	A*	A	B	C	D	E	U
2009	11.66	31.93	53.66	70.22	82.12	90.93	100.00
2010	12.97	30.94	49.85	67.92	81.48	91.96	100.00
2011	13.85	31.89	51.52	69.35	83.08	93.23	100.00
2012	14.90	34.74	54.66	72.73	85.39	93.54	100.00
2013	15.20	36.02	55.85	73.64	85.87	93.81	100.00
2014	17.34	40.69	62.77	80.08	90.97	96.92	100.00
2015	17.73	41.08	63.17	81.40	91.73	96.86	100.00
2016	18.71	43.52	65.67	83.25	93.06	97.90	100.00

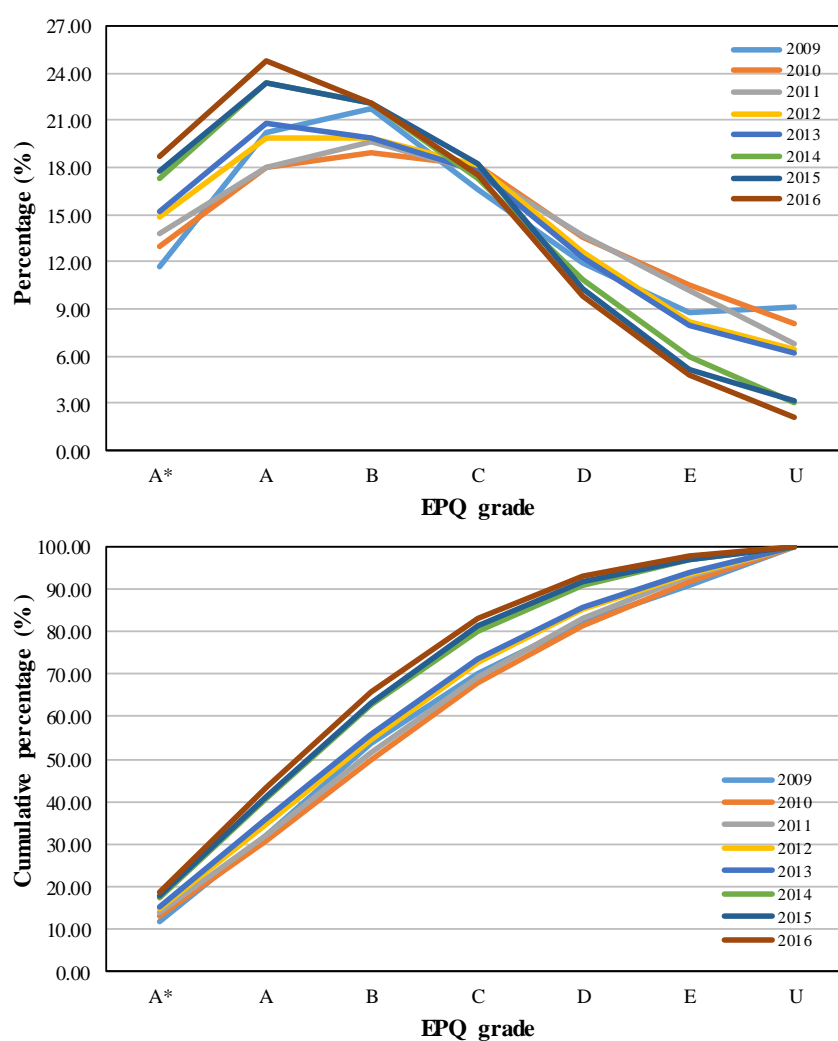


Figure 1 Distributions of EPQ grade outcomes from 2009 to 2016.

The graph at the top of Figure 2 shows the distribution of cumulative percentages of centres that had entered students for EPQ for up to 3 years consecutively from 2009. The graph below shows the cumulative percentages of students from these centres. From 2013 onwards, over 85% of centres and nearly 90% of students came from centres that had entered their students for EPQ for 3 or more years continuously.

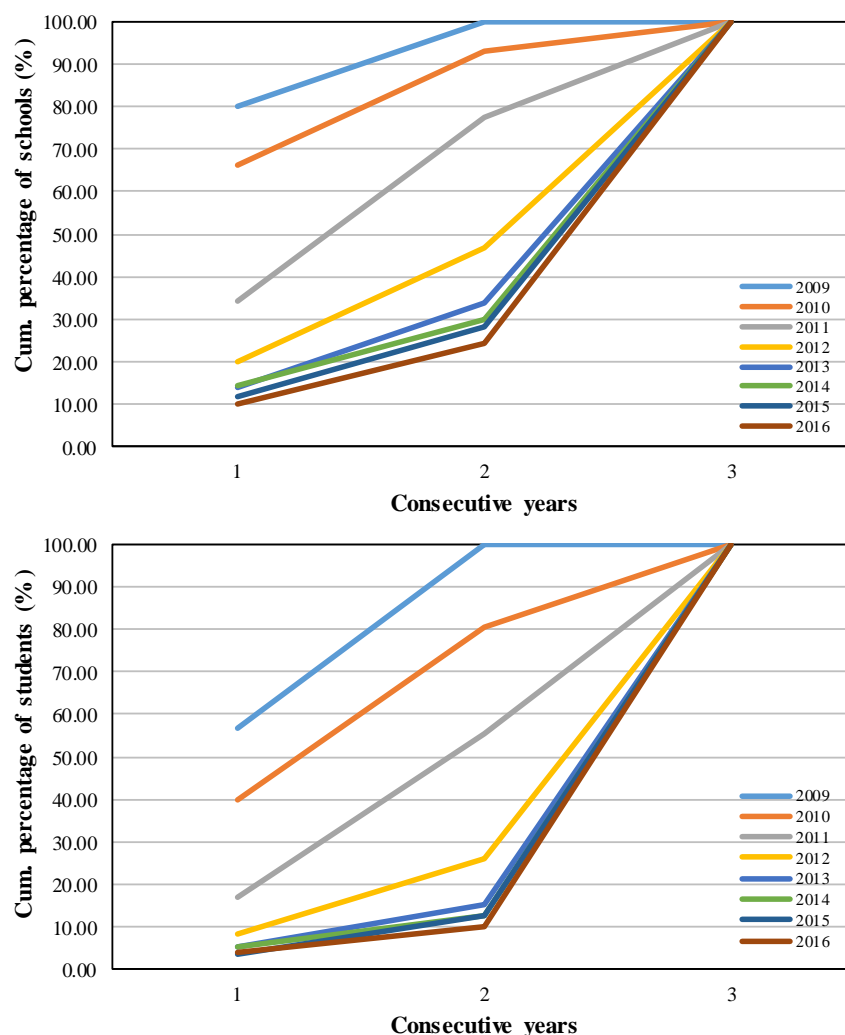


Figure 2 Distributions of cumulative percentage of centres entering students for up to 3 consecutive years (top) and cumulative percentage of students (bottom).

Relationship between EPQ outcome and prior/concurrent attainment

Correlations with other performance indicators and change in ability over time

Correlations were computed to examine relationships between EPQ grades and students' achievement at GCSE, A level and KS2 over the period of study. These

are shown in Table 3. As expected, correlations of EPQ with mean KS2 level are the lowest, ranging from 0.30 in 2009 to 0.37 in 2013. Correlations with mean GCSE grade vary from 0.50 in 2009 to 0.60 in 2010 and are generally slightly lower than those with mean A level grade. This may be expected as A levels were generally taken at the same time as EPQ, but GCSEs were taken 2 years before. It appears that (ignoring 2009) the correlations decreased very slightly over time. In general, about 10% of the variance in EPQ grades could be accounted for by variation in mean KS2 level. A third of the variance could be explained by variation in mean GCSE score or mean A level grade.

Table 3 Correlation of EPQ grade outcome with performance at GCSE, A level and KS2.

Year	Correlation coefficient of EPQ grade with		
	Mean GCSE grade	Mean A level grade	Mean KS2 level
2009	0.50 (3,451)	0.50 (4,894)	0.30 (4,659)
2010	0.57 (12,135)	0.60 (16,066)	0.37 (15,002)
2011	0.57 (18,834)	0.59 (24,698)	0.36 (22,920)
2012	0.55 (23,248)	0.57 (29,558)	0.35 (27,223)
2013	0.55 (26,367)	0.59 (32,225)	0.37 (29,571)
2014	0.53 (29,351)	0.56 (34,751)	0.33 (31,720)
2015	0.52 (30,410)	0.55 (31,625)	0.33 (32,634)
2016	0.52 (30,441)	0.56 (35,273)	0.32 (31,848)

Table 4 shows how the mean GCSE score, mean A level grade and the average KS2 level of the EPQ entries vary from 2009 to 2016. In most of the years, the match rate with GCSE and KS2 was over 90% and higher than that with A level, which varied from below 70% in 2009 to 82% in 2015. It appears that average GCSE score of the cohorts went up noticeably from about 6.2 in 2009 to 6.4 in 2016, while the mean A level grade and mean KS level went up very slightly. Noticeably, 2013 and 2014 saw the largest increase in mean GCSE score. This increase in the ability of the EPQ cohort over time may have partly contributed to the improvement in EPQ grade outcomes over time.

Table 4 Variation of attainment of EPQ students at GCSE, A level and KS2 over time.

Year	Performance Indicator	Mean	Std. Deviation	N	Match rate (%)
2009	mean GCSE score	6.20	1.06	5413	98.22
	mean A level grade	3.67	1.15	3780	68.59
	Mean KS2 level	4.53	0.52	5141	93.29
2010	mean GCSE score	6.16	1.10	17329	98.44
	mean A level grade	3.81	1.29	12887	73.20
	Mean KS2 level	4.49	0.53	16204	92.05
2011	mean GCSE score	6.22	1.09	26117	98.41
	mean A level grade	3.86	1.26	19795	74.59
	Mean KS2 level	4.51	0.52	24252	91.38
2012	mean GCSE score	6.26	1.08	31072	98.49
	mean A level grade	3.84	1.25	24268	76.92
	Mean KS2 level	4.52	0.51	28623	90.73
2013	mean GCSE score	6.37	1.04	33658	98.54
	mean A level grade	3.87	1.22	27400	80.22
	Mean KS2 level	4.59	0.49	30873	90.39
2014	mean GCSE score	6.45	0.99	36058	98.47
	mean A level grade	3.88	1.21	30268	82.65
	Mean KS2 level	4.62	0.47	32919	89.89
2015	mean GCSE score	6.46	0.98	32722	85.89
	mean A level grade	3.88	1.20	31356	82.31
	Mean KS2 level	4.59	0.46	33858	88.88
2016	mean GCSE score	6.47	0.95	36083	93.61
	mean A level grade	3.85	1.19	31062	80.58
	Mean KS2 level	4.61	0.46	32610	84.60

Correlations with A level subjects

Table 5 below shows the correlations of EPQ with 10 of the most popular A level subjects taken by EPQ candidates in 2010, 2013, 2014 and 2016. These are generally slightly lower than the correlations between EPQ and average A level grade. Highest correlations of EPQ are with A level history and English literature, while the lowest is with mathematics. For most of the subjects, correlations decreased very slightly over time. The correlations presented in Table 5 are generally slightly higher than those reported by Gill (2016) for 2015. This is probably because the EPQ candidates included in our analysis had achieved an EPQ grade or were unclassified.

Table 5 Correlations of EPQ with a selection of A level subjects over time (numbers of candidates indicated in brackets)

	2009	2010	2011	2012	2013	2014	2015	2016
Maths	0.28 (966)	0.43 (3322)	0.41 (5587)	0.39 (7107)	0.41 (8383)	0.39 (9555)	0.37 (9639)	0.39 (9563)
Biology	0.37 (767)	0.48 (2863)	0.48 (4668)	0.46 (5957)	0.46 (6951)	0.44 (8101)	0.46 (8029)	0.45 7798
History	0.51 (694)	0.56 (2617)	0.55 (4187)	0.53 (5180)	0.52 (6155)	0.50 (6986)	0.48 (7325)	0.49 (7470)
Chemistry	0.32 (678)	0.46 (2332)	0.44 (3880)	0.44 (4956)	0.44 (5864)	0.41 (6956)	0.41 (6814)	0.43 (6411)
Eng. Lit.	0.53 (700)	0.56 (2565)	0.56 (3975)	0.52 (4650)	0.52 (5414)	0.52 (5981)	0.51 (6528)	0.49 (6561)
Psychology	0.38 (632)	0.49 (2519)	0.48 (3553)	0.45 (4549)	0.45 (5180)	0.43 (5670)	0.43 (6037)	0.43 (6223)
Geography	0.45 (339)	0.50 (1357)	0.48 (2044)	0.49 (2594)	0.48 (3059)	0.46 (3715)	0.45 (4340)	0.46 (4204)
Physics	0.37 (360)	0.49 (1355)	0.46 (2243)	0.44 (2904)	0.45 (3365)	0.44 (3874)	0.43 (3834)	0.43 (3715)
Economics	0.48 (171)	0.52 (771)	0.48 (1353)	0.47 (1718)	0.42 (2280)	0.48 (2577)	0.48 (2745)	0.44 (3004)
Sociology	0.39 (279)	0.48 (1095)	0.46 (1669)	0.43 (2004)	0.44 (2155)	0.40 (2493)	0.38 (2678)	0.43 (2937)
Eng. Lang.	0.41 (410)	0.47 (1045)	0.51 (1675)	0.49 (2079)	0.51 (2269)	0.47 (2409)	0.48 (2417)	0.48 (2225)
English	0.43 (203)	0.53 (760)	0.47 (1170)	0.48 (1287)	0.53 (1446)	0.41 (1374)	0.44 (1338)	0.47 (1194)

Relationships with prior/concurrent attainment – grade characteristic curves

Figure 3 shows the relationship between average EPQ grades and average performance at KS2, GCSE and A level of the candidates from 2009 to 2016. To produce these curves, the mean GCSE score (or mean A level grade or mean KS2 level) for each candidate is calculated, and the mean GCSE score range was divided into 20 equal intervals. A candidate was assigned to one of the 20 intervals based on his/her mean GCSE scores. For candidates in each mean GCSE score interval, their average numerical grades in EPQ were calculated to derive the curves. These curves may be called EPQ grade characteristic curves (GCCs). The difference in the relationship between average EPQ grade and mean GCSE score (or mean A level grade or mean KS2 level) between years could be a reflection of inconsistency in grading standards, as different level of mean GCSE score was required to achieve the same EPQ grade (assuming consistency in GCSE standards over time is maintained). Grades on the top curves of the graph may be considered easier than

those on the curves below, since for similar mean GCSE performance, candidates on the top curves achieved higher EPQ grades than those on curves below. For mean GCSE, the curves from 2009 to 2013 almost overlap with each other (from EPQ grade D to grade A*), which may suggest that the grades from these years were of similar performance standards. The 2014 curve is above those for 2009 to 2013, which may suggest that the 2014 grades were easier than the grades from 2009 to 2013. The 2016 curve almost overlaps with that for 2015 for lower or middle ability students, but is above the 2015 curve for high ability students. In contrast, the 2015 curve is above the 2014 curve for the lower to middle ability students, but almost overlaps with the 2014 curve for high ability students. This suggests that there is differentiated change in EPQ grade difficulty over time. It is possible to estimate a measure of difference in difficulty or standard between 2 years using these graphs. Take 2013 and 2014 as an example, the differences in average EPQ grade between 2014 and 2013 at specific mean GCSE score intervals can be weighted using the proportions of students from 2013 that fall into the intervals to make the GCSE profile of the students similar for both years. This weighted difference is estimated to be 0.21 ± 0.01 (in EPQ grade unit), which reflects difference in EPQ grade outcomes between 2013 and 2014 if the distribution of the mean GCSE score for 2014 is similar to that for 2013 and may be used as a measure of difference in difficulty between the 2 years. For mean A level grade and average KS2 level, the GCCs show similar patterns over time as those with mean GCSE score. The weighted difference of the average EPQ grade between 2014 and 2013 was estimated to be 0.19 ± 0.01 , based on mean A level grade, and 0.25 ± 0.02 based on mean KS2 level. These graphs suggest that from 2014 onwards, the EPQ grades were easier than the same grades from 2009 to 2013. It also appears that the grades from 2014 to 2016 became progressively easier. The difference in difficulty of EPQ grades from 2010 to 2016 was estimated to be about 0.46 grade, based on mean GCSE score; 0.47 grade, based on mean A level grade; and 0.55 grade based on mean KS2 level.

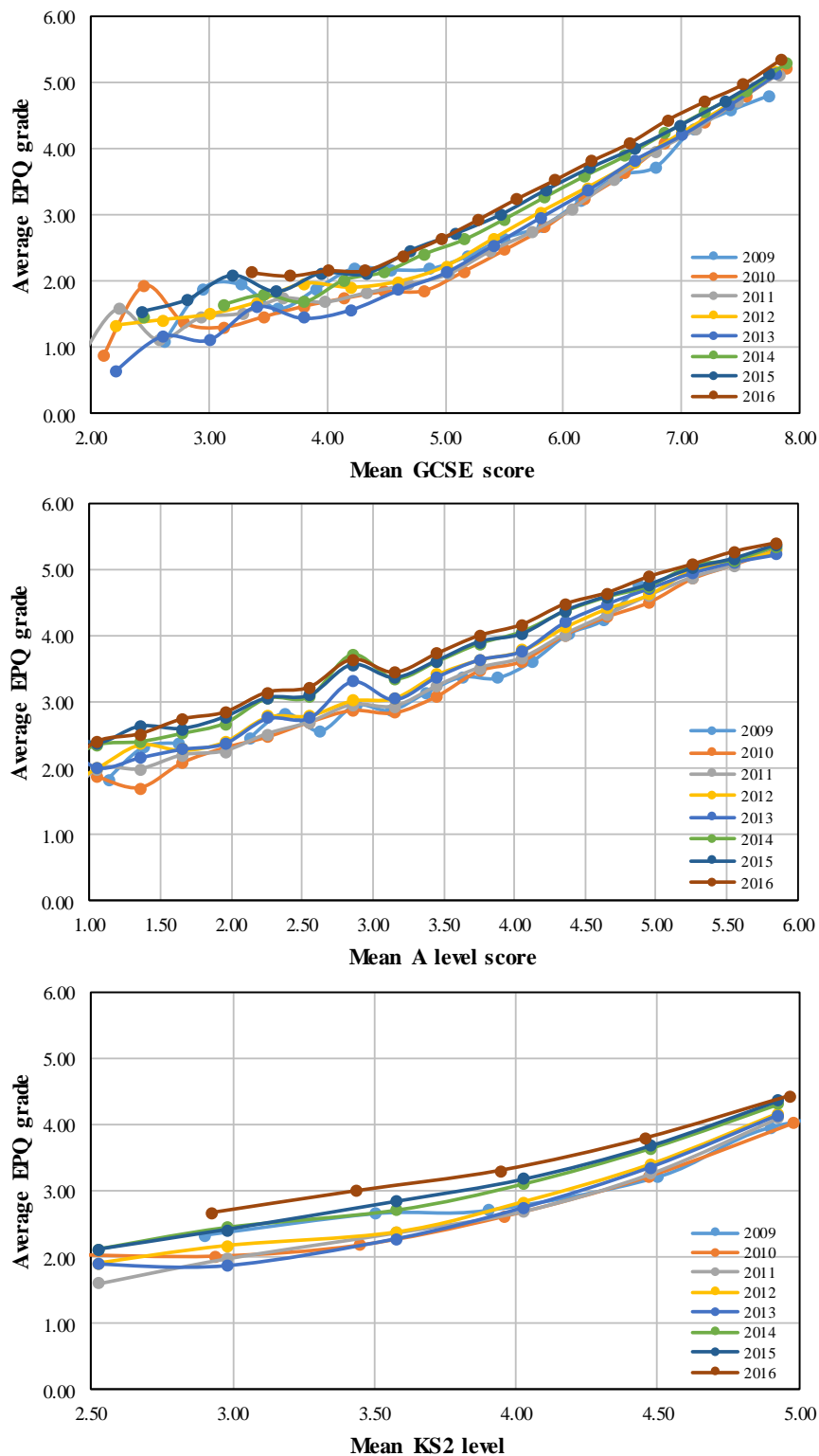


Figure 3 Relationship between average EPQ grade and mean GCSE score (top), mean A level grade (middle) and mean KS2 level (bottom).

Multinomial logistic regression against prior/concurrent attainment

EPQ grade difficulty based on multinomial logistic regression

With multinomial logistic regression, the relationship between an EPQ grade (performance category) relative to the reference grade (category 0, which is set as the lowest EPQ outcome U) and the attainment at GCSE, A level and KS2 is modelled separately. For a specific EPQ grade or performance category k , the logarithm of the ratio of the probability of a student being classified into this category, P_k , to the probability of being classified into the reference category P_0 , given his/her prior or concurrent attainment x , is expressed as a linear function of x :

$$\ln(P_k/P_0) = \alpha_k + \beta_k x \quad (1)$$

where β_k is the regression coefficient and α_k is the intercept. For 2 adjacent categories k and $k - 1$, the logarithm of the odds ratio can be expressed as:

$$\ln(P_k/P_{k-1}) = (\alpha_k - \alpha_{k-1}) + (\beta_k - \beta_{k-1})x \quad (2)$$

The difficulty of category k , δ_k , can be defined as the value of x where the probability of being classified into category k is the same as the probability of being classified into category $k - 1$, that is $P_k = P_{k-1}$:

$$\delta_k = -(\alpha_k - \alpha_{k-1})/(\beta_k - \beta_{k-1}) \quad (3)$$

Variation in α_k and β_k or δ_k over time would suggest differences in grade distributions for students with similar prior or concurrent attainment distribution and could be interpreted as variation in grade standard.

Regression with mean GCSE score

Once the model parameters are estimated, the probability of a student with a fixed measure of prior or concurrent attainment being classified into each of the performance categories can be calculated. As an example, Figure 4 depicts how the probabilities of receiving individual EPQ grades varies with mean GCSE score for students from 2013 (category probability curves – CPCs), with the distributions of the actual observed proportions of students in individual EPQ grades superimposed. The model appears to fit the data reasonably well, particularly for the middle and higher grades. The probability of receiving the lowest EPQ outcome (U) decreases monotonically with the increase in mean GCSE score, while the probability of receiving the highest EPQ grade (A*) increases monotonically with the increase in mean GCSE score. For example, the probability of being unclassified (U) decreases from 0.55 at a mean GCSE score of 2.0 (\approx average GCSE grade F) to 0.06 at the mean GCSE score of 6.0 (\approx average GCSE grade B). Similarly, the probability of receiving EPQ A* increases from 0.01 at a mean GCSE score of 5.0 (\approx average GCSE grade C) to 0.32 at a mean GCE score of 7.5 (\approx average above GCSE grade A). For the intermediate grades, the CPCs have a unimodal distribution; the probability of being awarded a grade increases with the increase in mean GCSE score before the peak, but decreases after the peak.

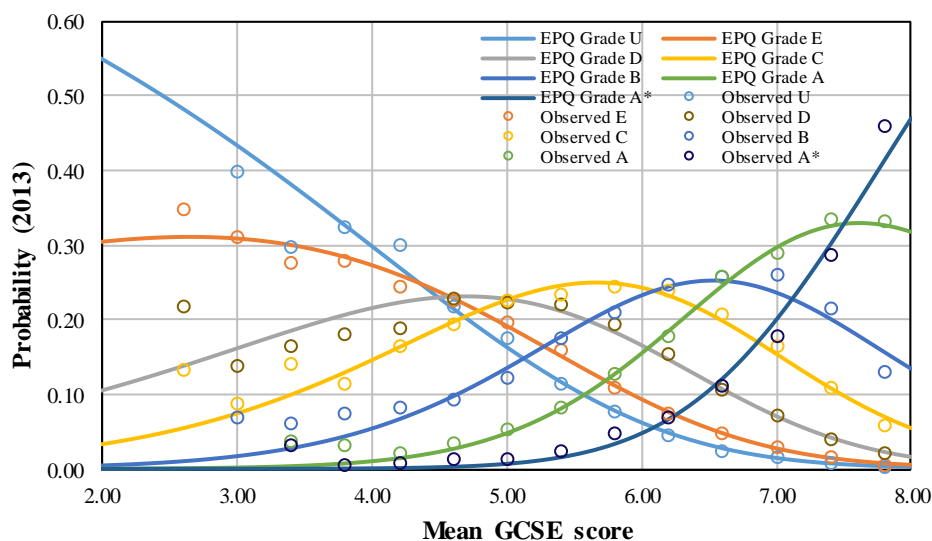


Figure 4 Model predicted probability distributions of being classified into EPQ grades and observed proportions against mean GCSE score in 2013.

Assuming that GCSE standards are being maintained over time, the CPCs for different years should be similar and close to each other if the standards of EPQ grades and the other factors that affect EPQ outcomes have been maintained the same. Differences between in the CPCs between years may suggest difference in grade standards. If the EPQ category probability curves from 2 different years are compared, the year with the curves on the left will generally have better EPQ outcomes than the other year if they have similar mean GCSE score distribution. That is, the grades from the year in the left can be said to be “easier” than the grades from the year in the right when the prior GCSE attainment measure is used as a basis for comparison. In contrast, the grades for the curves on the right could be seen to be more “difficult”. Figure 5 compares the CPCs for 2013 and those for 2016. As can be seen, the CPCs from 2016 were considerably to the left of those from 2013. This may suggest that the EPQ grades from 2016 are easier than those from 2013. More detailed examination of how the CPCs vary with mean GCSE score between the 2 years will demonstrate this further. For the lowest and highest categories (U and A*), the CPCs from the 2 years do not cross. As can be seen, the area under the CPC of EPQ grade U for 2016 is substantially smaller than that for 2013. In contrast, the area under the CPC of grade A* for 2016 is considerably larger than that for 2013. For the intermediate grades, the CPCs of the same EPQ grade from the 2 years cross each other and the order of probability changes direction at the mean GCSE score of the intersection point. For example, for the CPCs of EPQ grade C, the probability increased from about 0.15 in 2013 to 0.23 in 2016 at a mean GCSE score of 4.0 (GCSE grade D). The probability of achieving a grade C was about the same (0.25) at a mean GCSE score of 6.0 (GCSE grade B) for both years, and decreased from about 0.16 in 2013 to 0.13 in 2016, at a mean GCSE score of 7.0 (GCSE grade A). For the CPCs of EPQ grade B, the curves cross at a mean GCSE score of about 6.8, and the probability for 2016 is larger than that for 2013 on the left of the intersection point, but smaller on the right of the intersection point. As long as the intersection point of the CPCs of 2 adjacent EPQ grades of 2016 is to the

left of the corresponding intersection point of 2013, then 2016 will have a better EPQ outcome than 2013 if the GCSE profile of the students for both years was the same. It is noted that the difference in the CPCs between the 2 years is larger on the left side of the intersection points than on the right side. This may partly reflect the fact that the distributions of the EPQ grades in both years were skewed towards the highest grades.

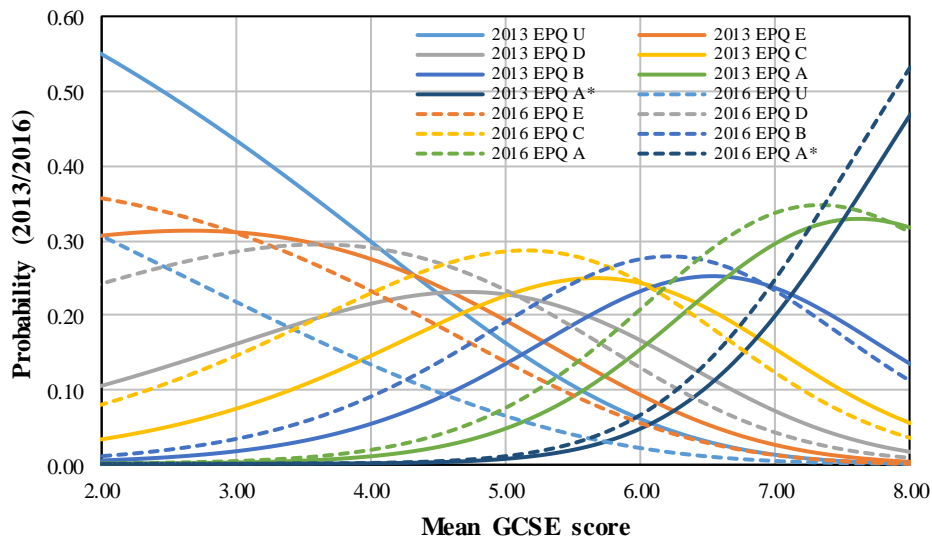


Figure 5 Modelled probability distributions of being classified into EPQ categories against mean GCSE score for students from 2013 and 2016 based on logistic regression.

To examine how the CPCs at individual grades vary with mean GCSE score across the years of study in detail, Figure 6 compares the CPCs of the EPQ students at grades C, B, A and A* over time. The CPCs show greater variability at grade C between the years in the region of lower to middle mean GCSE score. As is clear, at grade C, the CPCs from the different years are different in terms of the location of the peak and the shape. However, most of the CPCs (except for 2009) cross at an average GCSE grade of about 6.1 (\approx GCSE grade B). The left side of the curves spreads more widely than the right side. The CPCs from 2010, 2011 and 2013 are similar and closely bundled together, which may suggest that the standard at this grade in these years is similar when compared based on mean GCSE score. The CPC for 2009, which is the second year of the introduction of the EPQ, is flatter than those from 2010 to 2013 and is to the left of these curves, suggesting that this grade is easier than that from the other years. The CPC for 2014 is considerably to the left of the CPCs from 2010 to 2013. The CPCs from 2015 and 2016 are also very similar and close and are further to the left of the CPC from 2014. At grade B, most of the CPCs cross at a mean GCSE score of about 6.8. The CPCs from 2014 to 2016 are close and are to the left of the CPCs from 2009 to 2013, which are close and similar. At EPQ grades A and A*, the CPCs from 2010 to 2013 are similar and close. The CPCs for 2014 and 2015 are also very similar and close together and are to the left of the CPCs from 2010 to 2013. The CPCs for 2016 are further to the left of the other CPCs, suggesting that grade A and A* from 2016 are considerably easier than those from the other years, particularly those from 2013 and before.

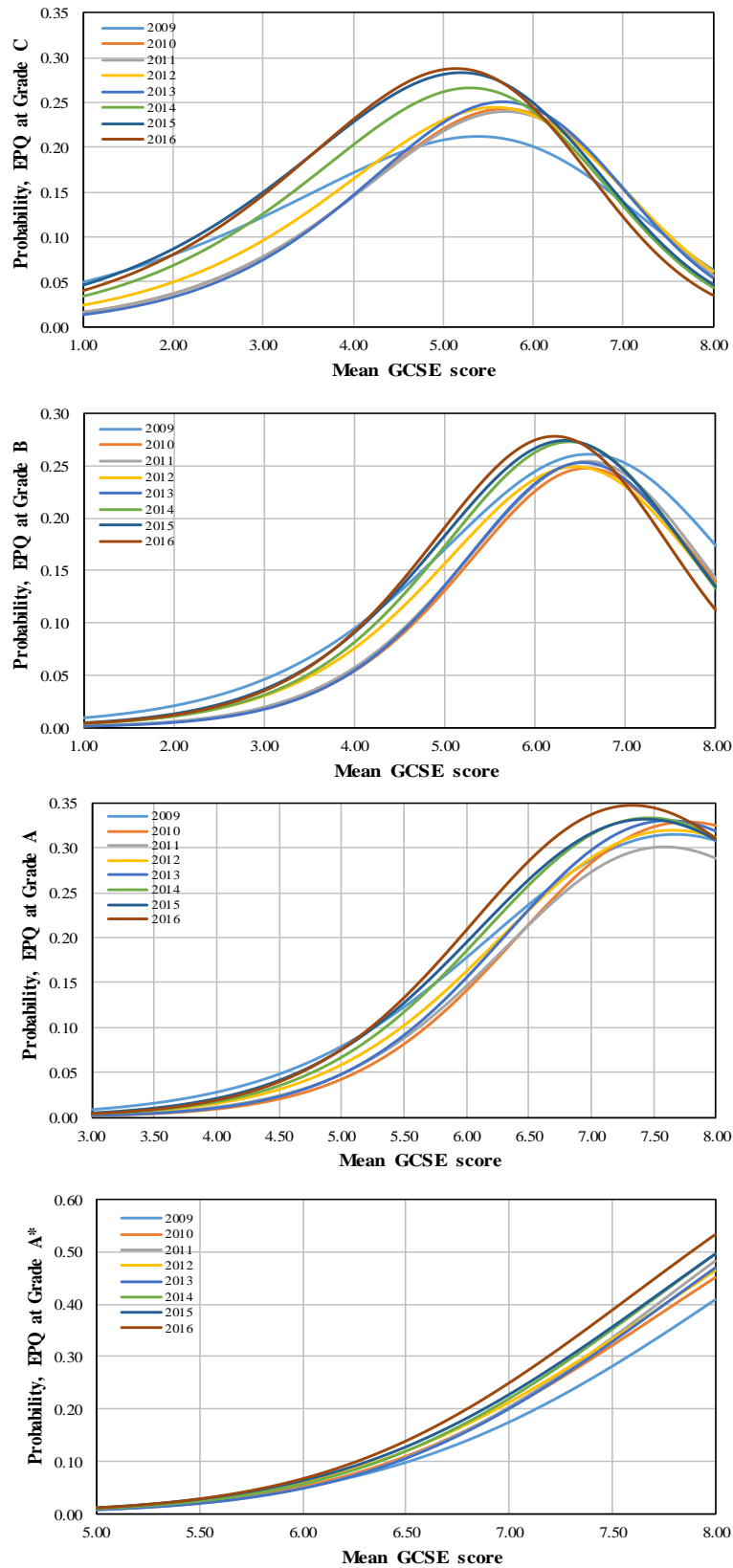


Figure 6 Model predicted probability distributions at individual EPQ grades against mean GCSE score for students from 2009 to 2016.

The difficulties of individual EPQ grades in a particular year can be estimated using equation (3). The relative difficulties of the EPQ grades from a particular year (relative to a reference year) can be defined as the differences of the grade difficulties between that year and the reference year. Using 2010 as the reference year, Table 6 shows the relative difficulties (in unit of GCSE grade) of EPQ grades C to A* from different years. Negative values indicate that the grades were easier than the reference year, while positive values harder. These values of relative grade difficulties are consistent with the curves shown in Figure 6. There was a step decrease in grade difficulty between 2014 and 2013, particularly at grades C and B. Overall, grade difficulties in 2016 are substantially lower than those in 2010, with grade C being nearly half a GCSE grade and grade B a third of a GCSE grade easier.

Table 6 Relative EPQ grade difficulties from 2009 to 2016 (relative to 2010, in unit of GCSE grade)

	2009	2010	2011	2012	2013	2014	2015	2016
C	-0.32	0.00	0.19	-0.06	0.00	-0.30	-0.47	-0.48
B	-0.59	0.00	-0.08	-0.12	-0.01	-0.29	-0.25	-0.33
A	-0.01	0.00	0.07	-0.09	-0.07	-0.13	-0.16	-0.29
A*	0.13	0.00	-0.14	-0.08	0.00	-0.07	-0.10	-0.15

Based on the relationship between EPQ grade and mean GCSE score shown in Figure 3, the relative EPQ grade difficulties listed in Table 6 can be expressed in unit of EPQ grade, which are shown in Table 7. The C and A grades in 2016 were estimated to be about 0.45 grade and 0.27 grade easier than the same grades in 2010, respectively.

Table 7 Relative EPQ grade difficulties from 2009 to 2016 (relative to 2010, in unit of EPQ grade) (based on mean GCSE score)

	2009	2010	2011	2012	2013	2014	2015	2016
C	-0.32	0.00	0.20	-0.06	0.00	-0.29	-0.42	-0.45
B	-0.59	0.00	-0.09	-0.12	-0.01	-0.28	-0.22	-0.31
A	-0.01	0.00	0.08	-0.09	-0.08	-0.12	-0.14	-0.27
A*	0.13	0.00	-0.16	-0.08	0.00	-0.07	-0.09	-0.14

The results presented above indicated that the EPQ grades from 2014 to 2016 are easier than those from 2013 and before.

A simple logistic regression was also carried out to look at the comparability of the aggregated EPQ grade outcomes over the years of study. Figure 7 shows how the probabilities of achieving EPQ grade C or above, B or above, and A and A* vary with mean GCSE score based on simple logistic regression. It is clear from the figure that for students who had similar mean GCSE scores, the probability of being awarded a grade C or above at EPQ, for those from 2014 to 2016, was considerably higher than those from 2013 and before, particularly for students with middle to higher mean GCSE scores. For students receiving a grade B or above in EPQ, those from 2014 to 2016 also had higher probabilities than those from 2009 to 2013 with similar mean

GCSE scores, particularly for students with higher mean GCSE scores. For students receiving A and A* in EPQ, the probability curves for students from 2014 and 2015 are similar and closely bundled and are considerably to the left of the curves for students from 2009 to 2013. The probability curve for students from 2016 is the leftmost curve and is considerably further to the left of the curves from 2009 to 2013 than the curves from 2015 and 2016, particularly for those with high mean GCSE scores. Again, these curves suggest that the EPQ grades from 2014 to 2016 are easier than those from 2013 and before.

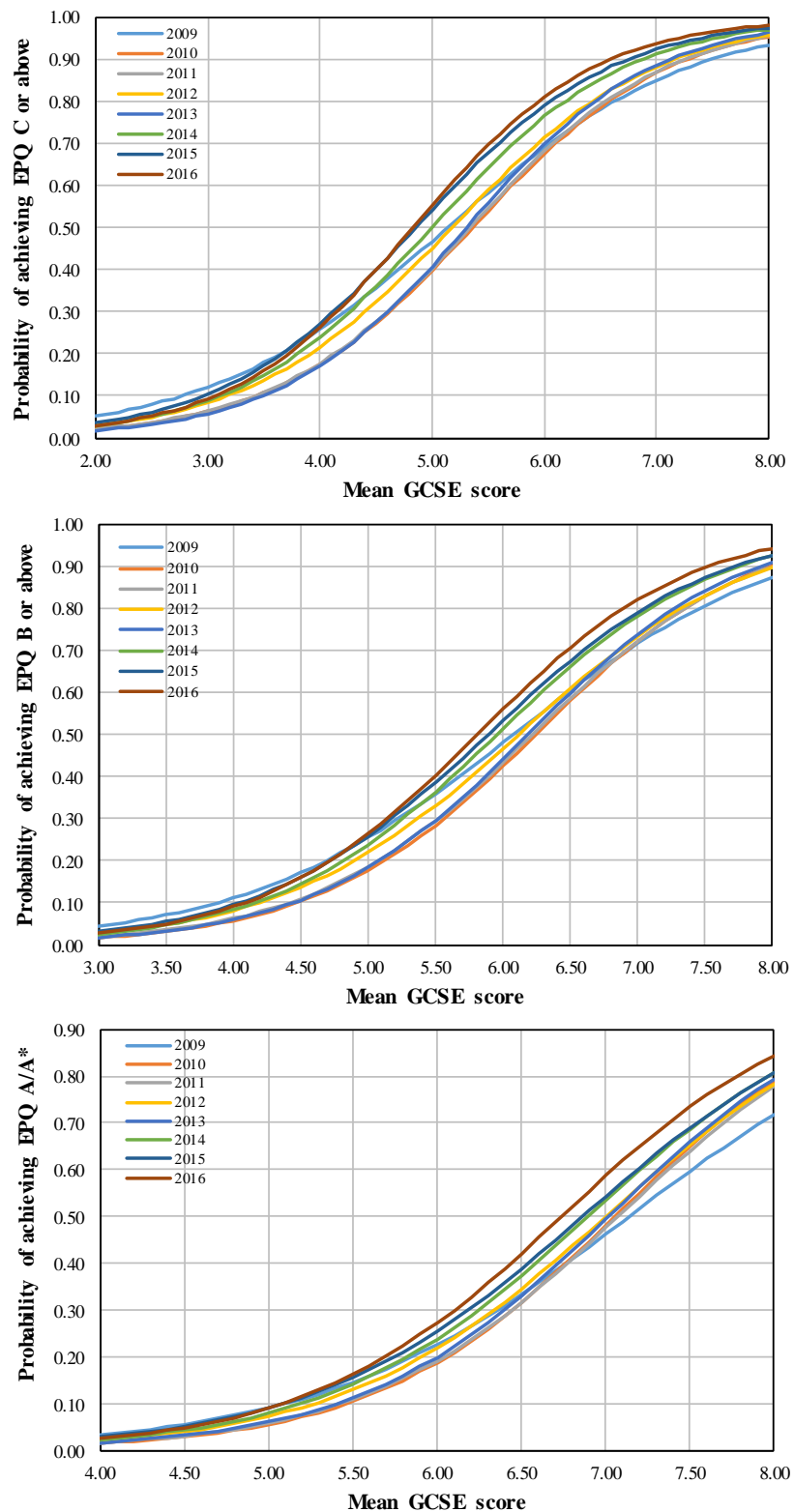


Figure 7 Simple logistic regression predicted probability distributions of receiving individual grades against mean GCSE score for students from 2009 to 2016.

Regression with mean A level grade

Figure 8 depicts the EPQ category probability distributions against mean A level grade for students from 2013, with the distributions of the observed proportions of students receiving individual EPQ grades superimposed. These curves are similar to those for mean GCSE score in terms of shape. The probability of receiving the lowest EPQ grades (grades U and E) decreases monotonically with the increase in mean A level score, while that of receiving the highest grade (A*) increase monotonically with the increase in mean A level score. For a student with an average A level grade of 3.0 (A level grade C), the probability of being awarded an EPQ grade A* is about 0.04, while that for a student with an average A level of 5.0 (A level grade A) is about 0.32. Again, for the intermediate grades, the CPCs have a unimodal distribution, with the probability of being awarded a grade increases with the increase in mean A level score before the peak but decreases after.

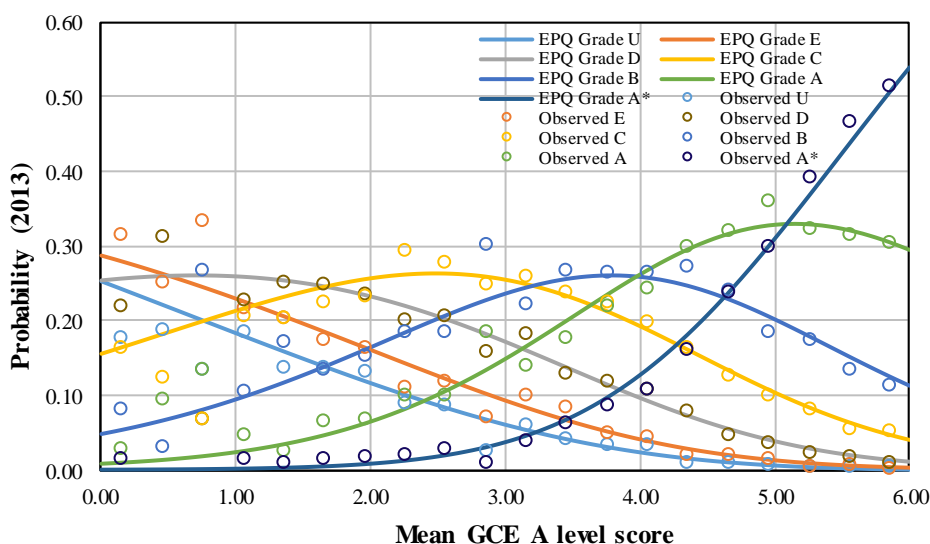


Figure 8 Model predicted probability distributions of being classified into EPQ grades and observed proportions against mean A level score in 2013 based on logistic regression.

Similar to the use of mean GCSE grade discussed above, if it is assumed that A level standards are being maintained over time, the CPCs from different years should be similar and close to each other if the standards of EPQ grades and the factors that influence EPQ outcomes have been maintained the same. Any difference in the mean A level based CPCs between years may suggest a difference in EPQ grade standards. Similar to mean GCSE score, if the mean A level grade based EPQ category probability curves from 2 different years are compared, the year with the curves on the left will have better EPQ outcomes than the other year if they have similar mean A level grade distribution. The EPQ grades from the year with CPCs on the left could be said to be “easier” than the grades from the year on the right. Figure 9 shows how the CPCs of the EPQ students for 2013 and those for 2016 vary with mean A level grade. The CPCs from 2016 were considerably to the left of those from 2013, suggesting that the EPQ grades from 2016 are easier than those from 2013, as judged by mean A level score.

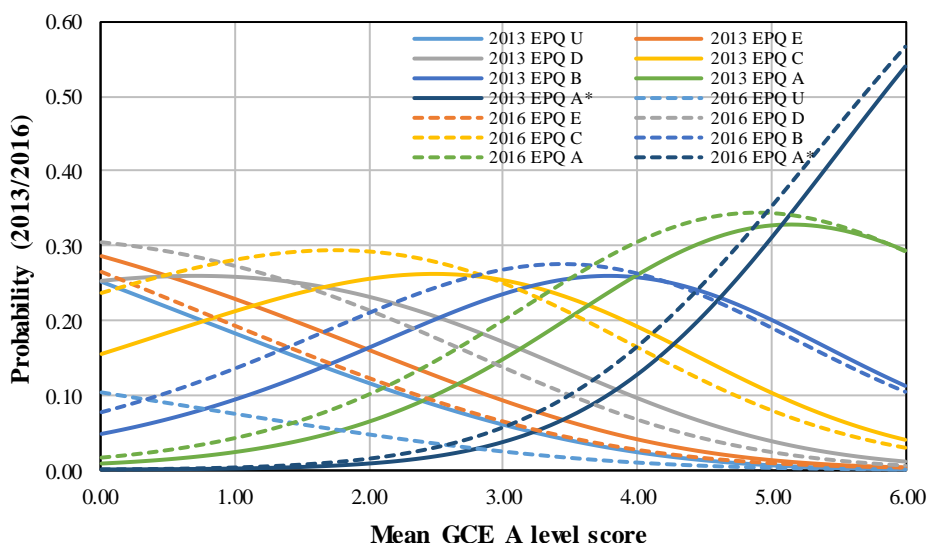


Figure 9 Modelled probability distributions of being classified into EPQ categories against mean A level grade for students from 2013 and 2016.

Figure 10 shows how the CPCs at EPQ grades A*, A, B and C vary with mean A level score from 2009 to 2016. These are similar to the curves depicted in Figure 6. Similar to the use of mean GCSE score, relative EPQ grade difficulties based on mean A level grade can also be defined and estimated. Again, using 2010 as the reference year, Table 8 shows the relative difficulties (in unit of EPQ grade) of grades C to A* from different years. These values are slightly higher than those derived using mean GCSE score and are consistent with the curves shown in Figure 10. Again, grade difficulties in 2016 are substantially lower than those in 2010, with grades C and A being about 0.69 and 0.37 grade easier.

Table 8 Relative EPQ grade difficulties from 2009 to 2016 (in unit of EPQ grade) (based on mean A level grade)

	2009	2010	2011	2012	2013	2014	2015	2016
C	-1.90	0.00	0.22	-0.25	-0.15	-0.47	-0.62	-0.69
B	-0.53	0.00	-0.10	-0.17	-0.15	-0.46	-0.37	-0.40
A	-0.13	0.00	0.09	-0.13	-0.17	-0.24	-0.24	-0.37
A*	-0.10	0.00	-0.11	-0.10	-0.09	-0.12	-0.15	-0.17

These findings suggest that that the EPQ grades from 2014 to 2016 are easier than those from 2013 and before when mean A level grade is used as a basis for comparison.

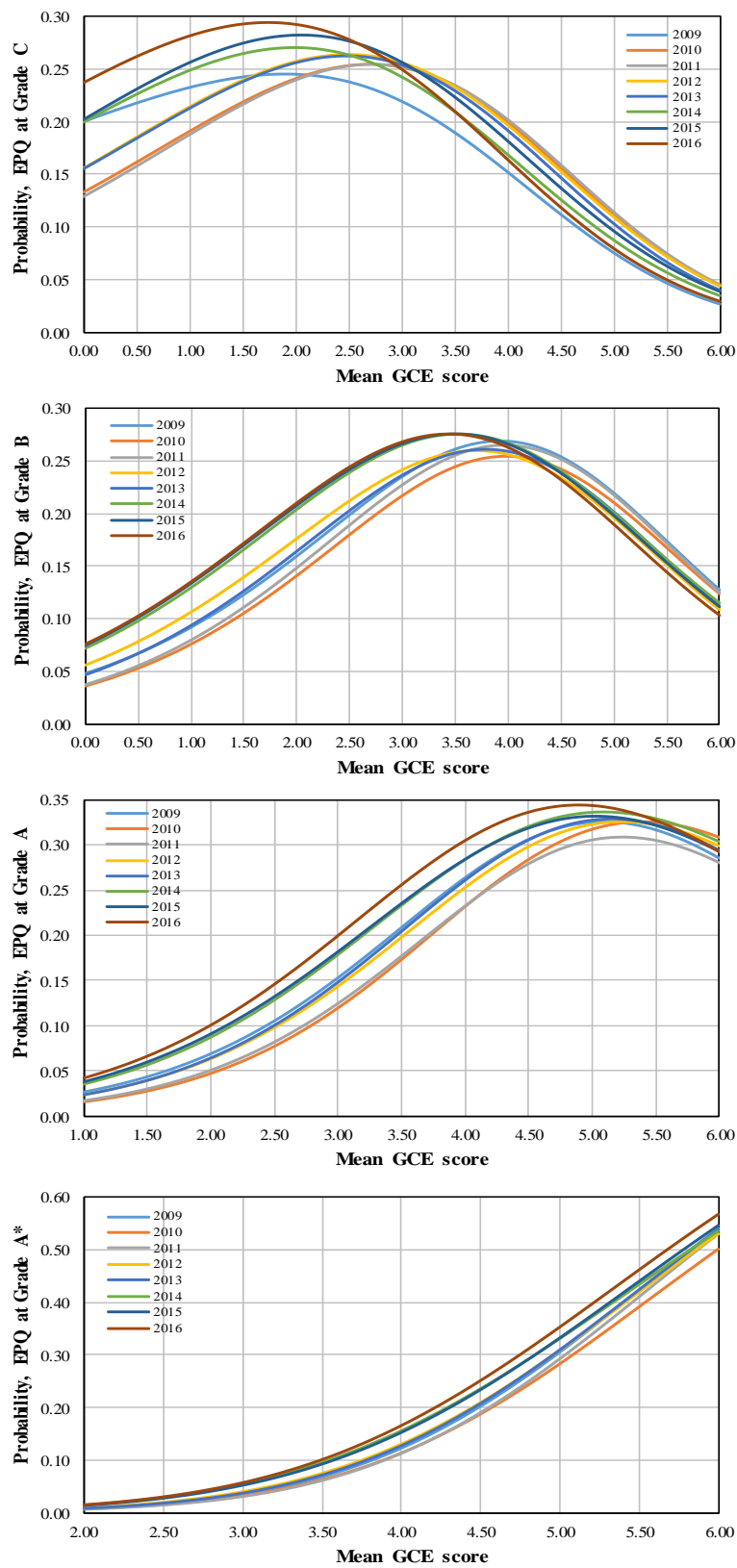


Figure 10 Model predicted probability distributions at individual EPQ grades against mean A level score for students from 2009 to 2016.

Regression with mean KS2 level

Figure 11 compares the model predicted EPQ category probability distributions and the observed proportions of EPQ students receiving corresponding EPQ grades in 2013 against mean KS2 level. Again, the model appears to fit the data reasonably well, particularly for the higher grades.

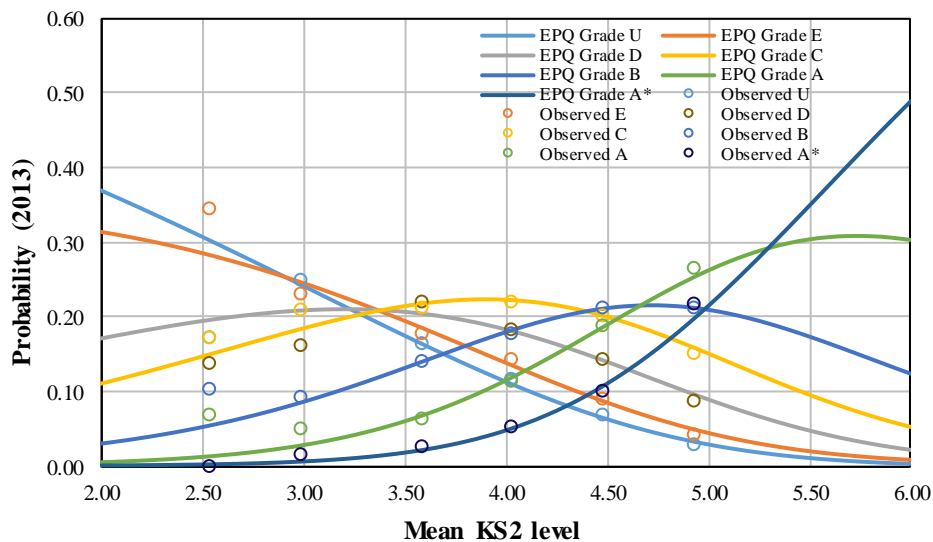


Figure 11 Model predicted probability distributions of being classified into EPQ grades and observed proportions against mean KS2 level in 2013 based on logistic regression.

Figure 12 compares the CPCs of the EPQ students from 2013 and those of students from 2016. Similar to the mean GCSE score and mean A level grade, the CPCs from 2016 were generally to the left of those from 2013. If it is assumed that KS2 standards are being maintained over time, the CPCs from different years should be similar and close to each other if the standards of EPQ grades and the factors that influence EPQ outcomes have been maintained the same. When the mean KS2 based EPQ category probability curves from 2 different years are compared, the year with the curves on the left will have better EPQ outcomes than the other year, if they have similar distribution of mean KS2 level, and could therefore be viewed as easier. Figure 12 suggests that the EPQ grades from 2016 are easier than those from 2013 as judged by mean KS2 level.

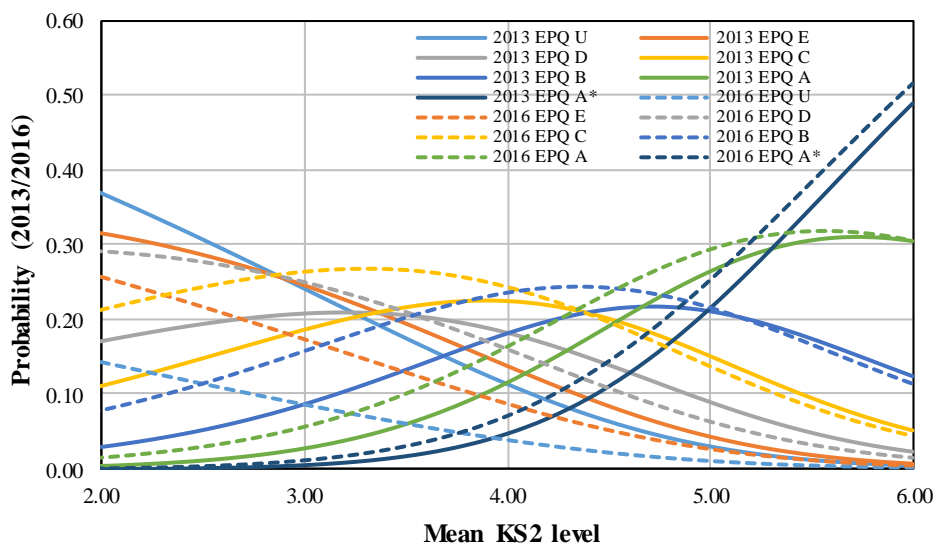


Figure 12 Model probability distributions of being classified into EPQ categories against mean KS2 level for students from 2013 and 2016.

Figure 13 compares the mean KS2 level based CPCs of the EPQ students at grades C, B, A and A* from 2009 to 2016. These are similar to the curves shown in Figure 6 for mean GCSE score and Figure 10 for mean A level grades. The CPC from 2009 is rightmost, while that from 2016 is leftmost. Table 9 shows the relative difficulties of grades C to A* (in unit of EPQ grade) from different years base multinomial regression with mean KS2 level. These values are generally slightly lower than those derived, using mean GCSE score and mean A level grade, and are consistent with the curves shown in Figure 13. Again, grade difficulties in 2016 are substantially lower than those in 2010, particularly for grades C and B.

Table 9 Relative EPQ grade difficulties from 2009 to 2016 (in unit of EPQ grade) (based on mean KS2 level)

	2009	2010	2011	2012	2013	2014	2015	2016
C	0.20	0.00	0.16	0.14	-0.12	-0.05	-0.12	-0.55
B	-0.60	0.00	-0.05	-0.10	-0.03	-0.24	-0.22	-0.28
A	0.03	0.00	0.07	-0.09	-0.09	-0.11	-0.15	-0.19
A*	0.29	0.00	-0.20	-0.06	0.01	-0.03	-0.09	-0.06

Figure 13 and Table 9 suggest that the EPQ grades from 2014 to 2016 are generally easier than those from 2013 and before based on mean KS2 level.

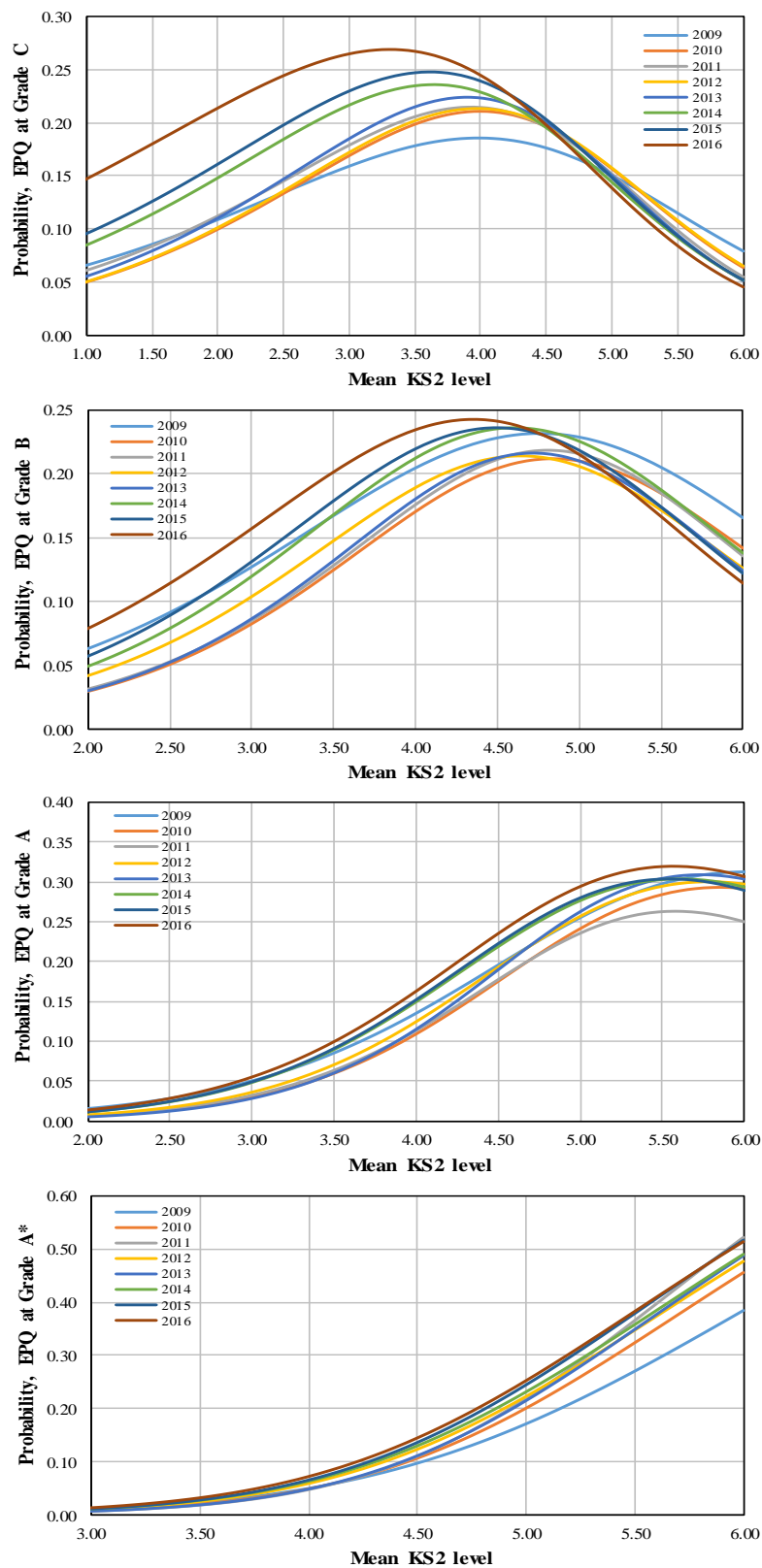


Figure 13 Modelled probability distributions of being classified into EPQ categories against mean KS2 level for students from 2013 and 2016.

Change in other characteristics over time

Table 10 shows the percentages of EPQ students from selective and independent schools (School Type as Yes or 1, others as 0), eligible for free school meals (FSM as Yes or 1, others as 0), requiring special education need (SEN as Yes or 1, others as 0), gender (Female as 1 and Male as 0), first language not English (Language as 1, English as 0) and from non-white families (Ethnicity non-White as 1, White as 0) from 2009 to 2016. It is to be noted, except for school type and gender, all the other variables had missing values for over 40% of the cohort. The data show some small changes over time for some variables eg about 60% of the EPQ students are female. Other variables show more change over time eg the percentage of students from independent and selective schools has gone up from less than 10% in 2009 to over 23% in 2015.

Table 10 Percentage of students with respect to a selection of background variables.

	Percentage (%)					
	Ind. and sel. school	FSM(Y)	SEN(Y)	Male / Female	Ethnicity (non-White)	Language (non-Eng.)
2009	9.65	1.65	1.92	0.41	6.22	3.36
2010	16.81	1.89	2.65	0.40	9.47	5.10
2011	18.94	2.25	3.38	0.42	10.72	6.21
2012	19.27	2.54	3.09	0.41	10.63	6.23
2013	20.78	2.63	3.12	0.41	11.16	6.58
2014	22.51	2.71	3.26	0.40	12.92	7.30
2015	23.87	2.98	2.90	0.40	13.62	7.80
2016		2.92	2.50	0.40	14.51	8.39

Propensity score matching analysis

As has been shown, there has been variability in many of the characteristics of the EPQ entries over time. Table 11 below shows the correlations of some of the characteristics listed in Table 10 and NCY (taking values 1, 2 or 3) with EPQ grade from 2009 to 2016. Most of these correlations are small but significant. School type and gender correlated better with EPQ grade than the other factors. The correlation between EPQ grade and the number of years that schools have entered students for EPQ consecutively is negative in 2009 but positive after 2010. This may partly reflect the gradual increase over time in the number of students from such schools. Variability in these factors over time needs to be accounted for when comparing EPQ outcomes between years. PSM analysis is one of the approaches that can be used for this purpose.

Table 11 Correlation of EPQ grade outcome with background variables.

Year	FSM	Ethnicity	Language	SEN	School Type	Gender	NCY
2009	-0.07	0.03	0.00	-0.02	0.22	0.10	-0.08
2010	-0.05	-0.04	-0.02	-0.07	0.24	0.12	0.00
2011	-0.07	-0.04	-0.05	-0.08	0.25	0.09	0.02
2012	-0.07	-0.04	-0.05	-0.06	0.22	0.10	0.05
2013	-0.08	-0.05	-0.07	-0.07	0.23	0.11	0.05
2014	-0.09	-0.05	-0.06	-0.04	0.23	0.09	0.06
2015	-0.08	-0.03	-0.04	-0.05	0.23	0.10	0.06
2016	-0.08	-0.03	-0.03	-0.03	0.16	0.11	0.10

Table 12 shows the effect (difference in mean EPQ grade estimated using a weighted regression model containing a treatment indicator) between 2 years (the first is the control group and the second is the treatment group) based on propensity score matching with respect to 3 sets of covariates. The first set of covariates (covariate set1) include mean GCSE score, mean GCE score and mean KS2 level. In addition to those in the first set, the second set also include gender, special education needs status, eligible for free school meals, and school type. The third set of covariates further include the number of consecutive years (NCY) that the school has entered students for EPQ, ethnicity and first language. The matching was done in 2 ways, including and excluding students with missing values on the relevant variables. In most cases, the difference is larger when students with missing values were excluded from the matching than when they were included, particularly for the third set of covariates. This is probably due to more variables in the third set, which would exclude more cases when students with missing values were excluded. Table 12 suggests that students in 2010 would achieve higher EPQ grades than similar students (in terms of mean GCSE score, mean A level grade and mean KS2 level and other characteristics) in 2009. There may be fluctuations in standards in the first few years when a qualification is introduced (e.g. the saw tooth effect due to increased familiarity with the qualification and its assessment requirements). Further year-on-year comparison suggested that before 2013 the difference in mean EPQ grades was small, but larger afterwards. In particular, there is an increase of about 0.21 grade in mean EPQ grade from 2013 to 2014 and an increase of about 0.13 grade from 2015 to 2016. Comparing the students from 2013 with those from 2016, the EPQ outcome in 2016 was estimated to be about 0.35 grade higher than that in 2013. This difference is close to the sum of the PSM-estimated year-on-year changes in these years. Since EPQ outcome correlates with average A level grade, mean GCSE score and mean KS2 level much better than with the other factors (see Tables 3 and 11), including additional variables in covariate set1 made no differences or only very small differences in the estimated changes in grade outcome between 2 years. The results shown in Table 12 are also generally consistent with those from the analyses discussed in previous sections, suggesting that there may have been grade inflation in EPQ from 2014.

Table 12 Mean EPQ grade difference between years after controlling for difference in different factors that could potentially influence EPQ performance.

	Mean EPQ grade difference					
	Covariate set 1		Covariate set 2		Covariate set 3	
	Inc Miss	Exc Miss	Inc Miss	Exc Miss	Inc Miss	Exc Miss
2009-2010	0.02	0.04				
2010-2011	0.02	0.04				
2011-2012	0.10	0.09	0.10	0.02	0.12	0.08
2012-2013	-0.03	0.01	-0.04	0.09	-0.05	0.09
2013-2014	0.21	0.19	0.20	0.15	0.19	0.15
2014-2015	0.04	0.03	0.03	0.05	0.05	0.05
2015-2016	0.13	0.10	0.14	0.11	0.14	0.11
2013-2016	0.35	0.31	0.34	0.33	0.34	0.33

Covariate Set 1: Mean GCSE, mean GCE, mean KS2

Covariate Set 2: Mean GCSE, mean GCE, mean KS2, Gender, SEN, FSM, School type

Covariate Set 3: Mean GCSE, mean GCE, mean KS2, Gender, SEN, FSM, School type, Number of consecutive years, Ethnicity, Language

Inc Miss: Students with missing values on the variables were included in the analysis.

Exc Miss: Students with missing values on the variables were excluded in the analysis.

Values coloured in green have $p > 0.05$. All others have $p < 0.05$

Stability of EPQ grade boundaries over time

Figure 14 shows how EPQ grade boundaries have changed over time (AQA and WJEC at the overall qualification level, while OCR and Pearson at unit level). For all awarding organisations, grade boundaries have been moved at some point, most likely to control outcomes. At the overall qualification level, EPQ grade boundaries from AQA at grades A and below increased in the first few years of the qualification, with larger changes associated with lower grades. In contrast, for the 3 optional EPQ routes offered by WJEC, changes in grade boundaries occurred in more recent years. The reason behind this is not clear. Further, Options 1 and 2 had grade boundaries lowered, while Option 3 increased. For OCR and Pearson, their EPQ *qualification level* grade boundaries have remained unchanged over time (24, 20, 16, 12, 8, 4 for A*, A, B, C, D and E respectively), even though some unit boundaries have altered. However, the grade boundaries of some of the units (Pearson) and for all OCR units have gone up over recent years, probably indicating attempts to control grade outcomes. Comparing the quality of students' work over time at the boundary marks could illustrate the extent to which these changed grade boundaries reflect maintenance or changes in qualification performance standards.

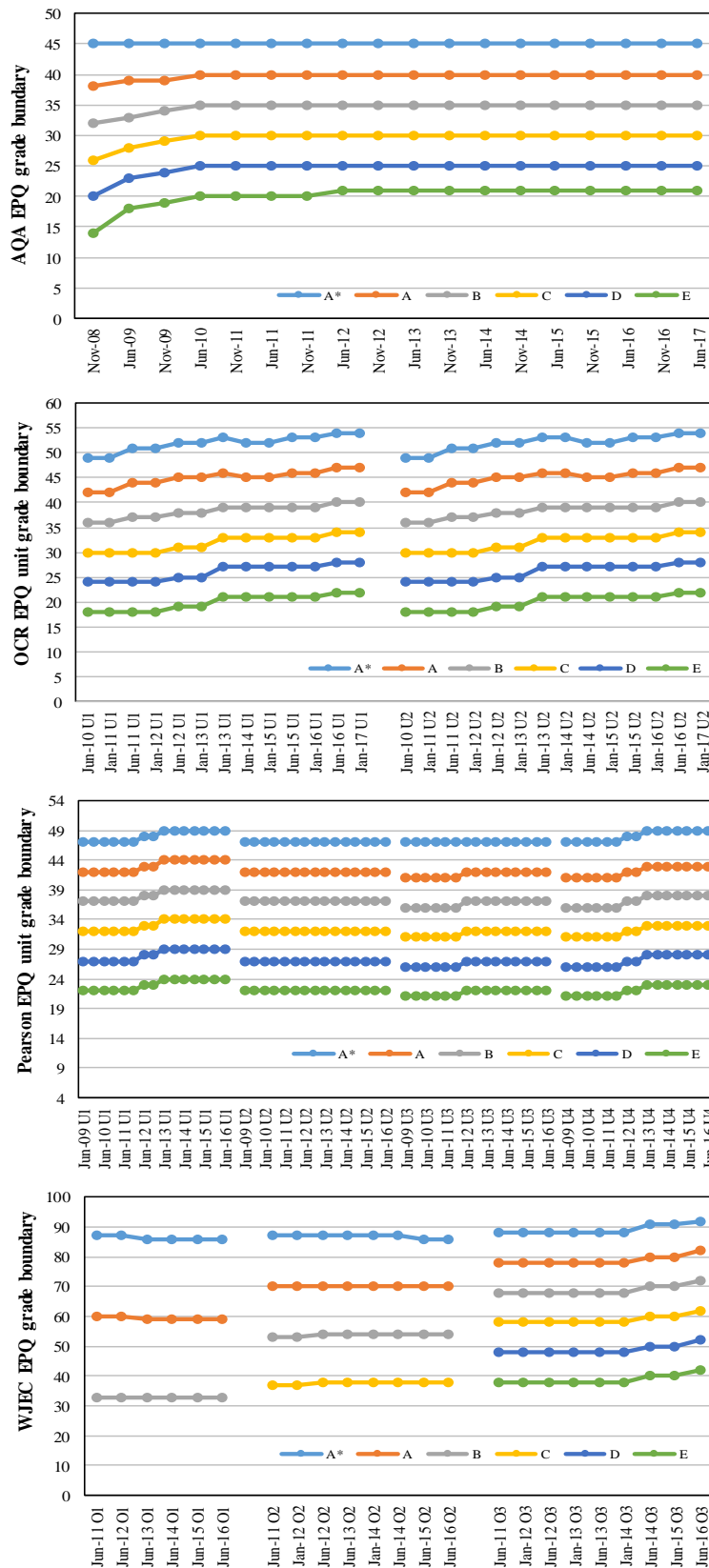


Figure 14 Changes in EPQ grade boundaries over time at the overall qualification level (AQA and WJEC) and unit level (OCR and Pearson).

Conclusion

The grade outcomes of EPQ have increased considerably from its introduction in 2008, with more and more students achieving higher grades, particularly from 2014 onwards. The work reported here was intended to provide some sense of the scale of this increase for EPQ outcomes and some tentative possible explanations.

Empirical analysis of the relationships between EPQ grade outcome and performance at GCSE, A level and KS2 indicated that there was a noticeable increase in mean EPQ grade of about 0.20 grade from 2013 to 2014 and an increase of about 0.46 grade from 2010 to 2016 for students with similar prior or concurrent attainment. Multinomial logistic regression analysis suggested that grade C in 2016 is about half a grade easier than the same grade in 2010 and grade A a third of a grade easier. PSM analysis indicated that from 2013 to 2014, there could be an increase of about 0.21 grade in mean EPQ grade for similar students from 2013 to 2014 and about 0.48 grade from 2010 to 2016. Findings from this investigation indicate that there might have been modest grade inflation in EPQ grade outcomes, particularly since 2014.

There is also some evidence that boards display different patterns of managing outcomes through setting of grade boundaries.

References

- Gill, T. (2016). Uptake and results in the Extended Project Qualification 2008-2015. Cambridge Assessment. Available on line at:
<http://www.cambridgeassessment.org.uk/Images/306859-uptake-and-results-in-the-extended-project-qualification-2008-2015.pdf>
- Ofqual (2015). Setting GCSE, AS and A Level Grade Standards in Summer 2014 and 2015. Ofqual: Coventry. Available on line at:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/451321/2015-08-05-summer-series-gcse-as-and-a-level-grade-standards.pdf
- Ofqual (2017). Provisional summer 2017 exam entries: GCSEs. AS and A levels. Ofqual: Coventry. Available on line at:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/619387/Summer_2017_exam_entries_GCSEs_AS_and_A_levels_in_England.pdf
- Ridgeway, G., McCaffrey, D., Morral, A., Burgette, L. and Griffin, B. (2017). Toolkit for Weighting and Analysis of Nonequivalent Groups: A tutorial for the twang package. RAND. Available online at: <https://cran.r-project.org/web/packages/twang/vignettes/twang.pdf>
- Stephenson, C. (2017). Exploring students' experiences of the Extended Project Qualification. Paper presented at the 14th AEA-Europe Conference in Prague, November 9-11, 2017.



© Crown Copyright 2018

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated.

To view this license, visit

www.nationalarchives.gov.uk/doc/open-government-licence/

or write to

Information Policy Team, The National Archives, Kew, London TW9 4DU

Published by:



Earlsdon Park
53-55 Butts Road
Coventry
CV1 3BH

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