RESEARCH AND ANALYSIS

An exploration of grade inflation in ‘older style’ level 3 BTEC Nationals

2006 to 2016

ofqual

Authors

This report was written by Benjamin M. P. Cuff, Nadir Zanini, and Beth Black, from Ofqual's Strategy Risk and Research directorate.

Contents

Tables

Figures

Executive summary

Applied General and Level 3 Tech Levels qualifications represent around 6% of all vocational and other qualifications, with over 300k certifications each year (Ofqual, 2018). These qualifications represent a key pathway at key stage 5 (16-19 year olds), offering a more applied pathway than A levels, and can count towards school performance tables. Many of these qualifications involve a high proportion of internal assessment (sometimes referred to as 'teacher-based assessment' or 'non-examined assessment' [NEA]). In recent years, the profile of grade outcomes has been changing, with increases seen in top grades. It is possible that, given the methods of control deployed around standard setting in these qualifications that the increases in outcomes represent unwarranted grade inflation ('illegitimate increases in outcomes over time, which do not reflect a genuine increase in the ability of cohorts of students').

The purpose of this report is to investigate whether there is evidence of unwarranted grade inflation in greater depth, using a combination of datasets provided by the Department for Education (the 'National Pupil Database') and the Higher Education Statistics Agency (HESA). This research necessarily focuses upon those qualifications with sufficient certifications to support the analyses. While there are a number of awarding organisations (AOs) providing Applied Generals and Level 3 Tech Levels qualifications, the majority of certifications are in Pearson's Applied General qualifications, in level 3 BTECs. In the academic year 2016-17, Pearson was the largest provider with 81% of all Applied General certificates. The next highest provider, University of the Arts London, had 9% (Ofqual, 2018).

Pearson's level 3 (L3) BTECs (also known as BTEC Nationals) have steadily risen in popularity over time, now being offered by the majority of schools and colleges. There have been concerns raised by research (Gill, 2016a; Thomson, 2017, 2018), and stakeholders, suggesting that it may be 'easier' to achieve equivalent outcomes in 'older style' L3 BTECs compared to A levels, and that equivalent outcomes in these course types do not offer equivalent levels of preparation for university. If there is grade inflation, the implication is that this 'gap' will have widened over time. The purpose of this report is to investigate these issues in greater depth. While the same issues may well apply to other similar vocational qualifications, we have focussed on 'older style' L3 BTECs here as they have sufficient data, longevity, the largest market share, and thus have attracted the greatest attention from stakeholders. While a newer suite of BTECs has recently been introduced, the focus of this report is on the 'older style' L3 BTECs (which are still being delivered), analysing the profile, outcomes and subsequent performance of 4 cohorts of students between 2005/2006 and 2015/2016.

Our first set of analyses focussed on level 3 outcomes. While grade distributions for A levels remained stable over time, grade distributions for 'older style' L3 BTECs have become progressively more skewed over time, towards the upper end of the grade range. In particular, the proportion of students achieving 'top grades' (distinction or distinction star) has increased substantially over time. For 'older style' L3 BTEC subsidiary diplomas (equivalent in size to an A level) the proportion of students achieving top grades has increased from 21% in 2005/2006 to 61% in 2015/2016. Similar patterns were found for other types of 'older style' L3 BTECs. These increases occurred under the context of stable prior attainment, suggesting no

particular change in the overall ability of the cohort. Further analyses showed that after controlling for differences in various characteristics of 'older style' L3 BTEC and A level students (eg demographics/prior attainment), 'older style' L3 BTEC students were still significantly more likely to achieve top grades than A level students, with this gap increasing over time. Students taking a mixture of 'older style' L3 BTEC and A level courses were also much more likely to achieve top grades in their BTEC qualifications, compared to their A level qualifications. Findings suggest the existence of unwarranted grade inflation, as no factors could be identified to support the legitimacy of these large increases in attainment (eg changes in the overall ability of cohorts, or their characteristics). This first set of analyses in themselves provide sound evidence of unwarranted grade inflation. Subsequent analyses were to further verify what this first set of analyses have indicated.

Our next set of analyses focussed on the subsequent degree outcomes of these students. If the increases in level 3 outcomes were legitimate, and standards had been maintained, one would expect 'older style' L3 BTEC students to perform similarly over time at university (and in comparison with A level students over time). However, findings show that successive cohorts of 'older style' L3 BTEC students with the same outcomes were increasingly less likely to achieve an upper 2nd or 1st class degree than A level students with equivalent level 3 outcomes. The same findings were apparent across different degree subject areas, and for graduates who completed degrees in the same subject area as their level 3 studies. Again, these findings seem consistent with the existence of grade inflation, as they suggest that 'older style' L3 BTEC outcomes have offered progressively lower levels of preparation for university over time, compared to equivalent A level outcomes. In other words, this indicates that 'older style' L3 BTEC outcomes have become devalued over time. Note that demographics, and changes in demographics (eg socio-economic background), have been controlled for in the analyses.

Given the applied nature of 'older style' L3 BTECs, it was also felt important to consider the performance of 'older style' L3 BTEC students in entering employment. For this, we used the Destinations of Leavers from Higher Education (DLHE) survey collected by HESA, which includes information on the employment status of university graduates, 6 months after graduation. It was found that, after controlling for differences in attainment at university, as well as university type and degree subject studied, 'older style' L3 BTEC students exhibited a progressively lower likelihood over time of being full-time employed, in a 'highly skilled' occupation (managerial or professional occupations), and having a 'good' salary (£20k p/a), relative to A level students with equivalent level 3 outcomes. While not necessarily being strong metrics for employment 'performance', these findings again provide support that 'older style' L3 BTEC outcomes have become devalued over time, calling into question the legitimacy of the increases in level 3 outcomes observed (suggesting unwarranted grade inflation).

It is possible that this apparent inflation of grades over time is due to the methods used in the standard maintaining of 'older style' L3 BTECs, meaning that these issues may also apply to other qualification types operating similar methods. Given that A level outcomes were also likely subject to a small amount of grade inflation prior to the introduction of tighter statistical controls in 2011, the fact that some grade inflation has apparently occurred for 'older style' L3 BTECs is perhaps unsurprising (BTEC awards are not subject to the same statistical controls). However, the difference in the size of these changes is notable: the percentage of A level students

achieving a grade A rose by 10% in 14 years (1996-2010; Ofqual, 2015), while the percentage of 'older style' L3 BTEC subsidiary diploma students achieving a distinction or distinction star rose by 40% in just 10 years (2006-2016; this study). Careful consideration is needed moving forward as to how such issues might be mitigated in the future, including for these 'older style' L3 BTECs still available to students and the newer suite of BTECs recently introduced. Indeed, while this research has focused upon 'older style' L3 BTECs, it is possible that these findings may generalise to other similar qualifications showing grading increases (ie those operating similar models of standards maintenance in similar contexts).

1 Introduction

Applied General and Level 3 Tech Levels qualifications represent around 6% of all vocational and other qualifications, with over 300k certifications each year (Ofqual, 2018). These qualifications represent a key pathway at key stage 5 (16-19 year olds), offering a more applied pathway than A levels, and can count towards school performance tables. Many Applied General and Tech Levels qualifications involve a high proportion of teacher-based assessment (sometimes referred to as 'internal assessment' or 'non-examined assessment' [NEA]). In recent years, the profile of grade outcomes has been changing, with increases seen in top grades. It is possible that, given the methods of control deployed around standard setting in these qualifications that the increases in outcomes represent unwarranted grade inflation ('illegitimate increases in outcomes over time, which do not reflect a genuine increase in the ability of cohorts of students').

Internal assessment can, in the right context and with the right mechanisms of control, allow standards to be maintained so that, over time, the same grade represents the same level of achievement. However, it is possible that the controls in place are not sufficient to manage pressures of grade inflation. Table 1 briefly describes the controls in place for some Applied Generals and Level 3 Tech Levels qualifications offered by 5 of the largest providers. While this report does not offer a complete analysis of the strengths and weaknesses of particular features of internal assessment, it highlights how some features, in combination with one another, may mean that there is insufficient control. In particular, where teachers directly grade work, any subsequent moderation or verification activity might only offer 'retrospective' advice to individual centres around the suitability of a grade awarded. In contrast, systems which require teachers to mark rather than grade work, can provide greater opportunity to intervene before certification, eg for moderation processes to 'scale' the marks prior to certification should a centre be judged to have been lenient or severe. Furthermore, for internal assessments which are marked, AOs can set grade boundaries in a way which appropriately maintains standards.

Other aspects of the context and quality assurance processes, such as resitting/resubmission policies, sampling regime, centre accreditation models (and Direct Claims Status), arrangements and requirements around standardisation of assessors/verifiers will also impact upon the degree of control an Awarding Organisation will have over standards maintenance.

Appendix A shows grading outcomes over time for some Applied General qualifications and Level 3 Tech Levels qualifications which are or have been on performance tables, where there is sufficient entry. A summary of the same information is also shown in Table 2. In most cases, outcomes show increases in top grades over time. However, readers should note that no assertions can be made at this stage as to whether these increases are legitimate or otherwise (ie due to grade inflation) – but provides background and context to the subsequent analyses.

Table 1. A summary of the main quality assurance methods for a number of Applied General and level 3 Tech Level qualifications.

Broad type of qualification	**Internal assessment percentage**	**Grading or marking of internal assessments**	**Pre-or post- certification / moderation?**	**Frequency of verification / moderation**	**Resubmission policy**
C&G Level 3 Extended Diploma - Land	100%	Most assessments directly graded.	Pre-certification and post-certification verification for centres with Direct Claims status. Sample of work varies from 0% to 100% depending on risk.	Annually, with additional activities where required.	Students may re-submit once only.
C&G Level 3 Extended Diploma - Travel	Approx. 65%	Around half of assessments are marked and half are graded.	As above	As above	Re-submission allowed. Grade capped at Pass
RSL Level 3 for Music Practitioners	60%	Directly graded	Pre-certification sampling of usually around 10% of candidates per centre.	Annually	One re-submission for internal assessment (uncapped).
UAL Level 3 Applied General Extended Diploma in Art & Design	50%	Directly graded	Pre-certification sampling of around 10% of candidates per centre.	Annually	One re-submission for internal assessment (uncapped).
UAL Level 3 Extended Diploma in Art & Design	100%	Directly graded	Pre-certification sampling of around 10% of candidates per centre.	Annually	One re-submission. Grade capped at Pass.
OCR Nationals Level 3	Varies by qualification – approx. 50% to 75%.	Marked	By unit. Sample size dependent on unit entry. A minimum of six candidates reviewed within the sub-sample (entry dependent)	Every series (2-3 per year)	Legacy Nationals – resits unlimited. New Nationals – learners may resubmit each assessment once.

Broad type of qualification	**Internal assessment percentage**	**Grading or marking of internal assessments**	**Pre-or post- certification / moderation?**	**Frequency of verification / moderation**	**Resubmission policy**
OCR Cambridge Technicals Level 3	Varies by qualification – approx. 30% to 58%.	Directly graded	By claim. Sample size dependent on units entered, assessors involved, grades submitted	On demand. As claims are submitted	Resubmission and re-moderated
Pearson BTEC Level 3 National (2016)	Varies by qualification – approx. 33% to 58%	Directly graded	Pre-certification sampling. Sample size depends upon size of cohort and number of units.	Twice, prior to results certification.	One resubmission on the same task.
Pearson BTEC Level 3 National (2010) ('older style')	100%	Directly graded	For verification, pre-certification sampling of a minimum of four learners dependent on number of candidates.	For verification, twice prior to results certification. Random sampling of centres with accreditation.	One resubmission on original task.

Table 2. The proportion of top grades awarded over time for a number of Applied General and level 3 Tech Level qualifications (as a % of the total number of grades awarded). NB data sourced from National Pupil Database.

	The proportion of top grades awarded in each academic year (as a % of all grades awarded)			
Qualification type (top grade)	2005/2006	2009/2010	2012/2013	2015/2016
'Older style' level 3 BTEC Subsidiary diplomas (D/D*)[a]	20.8	32.5	41.6	61.0
'Older style' level 3 BTEC Diplomas (DD+)[a]	12.9	29.8	39.3	57.8
'Older style' level 3 BTEC Extended diplomas (DDD+)[a]	15.4	30.8	37.2	48.9
City & Guilds Extended Diploma at level 3 (D*)	-	-	3.3	10.5
OCR Cambridge Technical certificates (D*)	-	-	28.2	48.1
OCR Cambridge Technical introductory diplomas (D*)	-	-	35.2	53.7
OCR Cambridge Technical subsidiary diplomas (D*D*)	-	-	20.3	30.0
OCR National certificates (D)	18.8	30.4	44.4	-
OCR National diplomas (D)	-	30.4	43.5	-
OCR National extended diplomas (D)	-	32.4	45.8	-
Rock School Limited level 3 VRQs (all types; D*)	-	-	7.0	11.2
University of the Arts London level 3 VRQs (all types; D)	-	-	21.7	22.3

Notes. Rows reflect each of the graphs presented in Figure 2 (see Section 4.1.1) and Appendix A. Only 4 academic years are shown, as we only collected data relating to those years – blank cells indicate where less than 100 (or 0) grades were awarded in that year. Figures only reflect data held within the National Pupil Database. More detail on the data underlying these figures can be found in Section 2 and Appendix A. Entry sizes for these qualification types can be found in Appendix A.

[a] Because Distinction star (D*) grades were only introduced for BTECs in 2010, D and D* grades have been combined for BTECs, so as to allow for an observation of historical trends.

The purpose of this report is to investigate whether there is evidence of unwarranted grade inflation in greater depth, using a combination of datasets provided by the Department for Education (the National Pupil Database) and the Higher Education Statistics Agency (HESA). Inevitably, this research focuses upon those qualifications with sufficient certifications to support the analyses. While there are a number of awarding organisations (AOs) providing Applied Generals and Tech Levels qualifications, the vast majority of certifications are in Pearson's Applied General qualifications, in their level 3 BTECs. In the academic year 2016-17, Pearson was the largest provider with 81% of all Applied General certificates; the next highest provider, University of the Arts, London, had 9% (Ofqual, 2018).

Level 3 BTECs are "career-based qualifications designed to give students the skills they need to move on to higher education or go straight into employment" (Pearson, 2017a). These BTECs offer a more applied pathway for students, as an alternative to the more traditional (and more academically focussed) A levels. While new BTECs have been introduced for first teaching from 2016 (see Pearson, 2017b) and have different sets of controls in place (see Table 1), the focus of this report is on the 'older style' level 3 (L3) qualifications delivered under the NQF and QCF[1]. These 'older style' L3 BTECs are still being offered at the time of writing. BTECs have traditionally been assessed and marked internally by centres[2] (with some external verification being carried out by Pearson – see Pearson, 2015).

It is important to note that while 'older style' L3 BTECs sit alongside other similar vocational qualifications, the focus here is on 'older style' L3 BTECs, which continue to hold a large market share, and have therefore received the greatest attention from stakeholders. While some other qualifications in this context show increases in top outcomes (Table 2), and have some similar controls in place (Table 1), we are not able to conduct the same analyses for qualifications – the volume of certifications is too small to support similar analyses as those in this report. However, any issues raised by the analyses conducted on older style L3 BTEC Nationals might raise questions about similar qualifications with similar design features.

BTECs are offered at levels 1-7, from the equivalent of a functional skills qualification to a postgraduate degree. We focus here on level 3 BTECs (equivalent to A levels), as these have the greatest importance for school leavers wishing to use their qualifications to gain entry into higher education or employment. Level 3 BTECs (also known as 'BTEC Nationals') can be further divided according to their size (ie the number of guided learning hours – see Table 3). For comparison, each A level qualification has 360 guided learning hours, meaning that a BTEC subsidiary diploma is equivalent in size to 1 A level, and an extended diploma is equivalent in size to 3 A levels. Level 3 BTECs are broadly graded pass, merit, distinction, and distinction star (D* - first awarded from 2010). The grades a candidate receives upon successful completion of the course depends on the size of the qualification (see Table 3). Certificates and subsidiary diplomas are awarded a single-letter grade, 90 credit diplomas and diplomas are awarded double-letter grades, and extended diplomas are awarded triple-letter grades.

[1] National Qualifications Framework (NQF) and Qualifications and Credit Framework (QCF).
[2] Schools and colleges.

Table 3. Size and grading of 'older style' L3 BTEC Nationals

Level 3 qualification type	Number of guided learning hours	Equivalence of size to 1 A level	Grading
Certificate	180	0.5x	P – D*
Subsidiary diploma	360	1x	P – D*
90 credit diploma	540	1.5x	PP – D*D*
Diploma	720	2x	PP – D*D*
Extended Diploma	1080	3x	PPP – D*D*D*

Note. 'P' = pass, 'D*' = distinction star. D* was first introduced for courses first taught in 2010. Prior to that, distinction (D) was the highest grade possible.

It is worth noting that candidates do not certificate for their 'older style' L3 BTEC course until they are ready. It is for this reason that almost no 'older style' L3 BTEC candidates achieve a 'fail' outcome, but rather they would simply not certificate if they would not be able to secure at least a pass (another option would be to certificate in a smaller sized 'older style' L3 BTEC for which they could secure a pass). This is different to A levels, where most candidates certificate in the summer of their second year of study, and may indeed achieve a 'fail' result.

While A levels are offered by almost all schools and colleges in England and remain the most popular level 3 qualification taken by students, BTECs have risen in popularity over time. Level 3 BTEC uptake rose by 74% between 2008 and 2015 (Gill, 2016c), and the percentage of schools and colleges offering level 3 BTECs also rose from 37% in 2007/2008 (Gill, 2013) to 65% in 2015 (Gill, 2016b). An increasingly popular route is now also to take a combination of A level and level 3 BTEC courses, which was the case for 7% of level 3 students in 2015 (Gill, 2016c)[3]. The number of students entering university with at least 1 BTEC qualification has also substantially risen over time (a 300% increase between 2010/2011 and 2012/2013 according to Allan, 2017).

Alongside this rise in popularity, level 3 BTECs have also gained increased importance as indicators of student ability for employers and university admissions officers alike.[4] However, stakeholders have expressed perceptions that it may be easier to achieve top grades in 'older style' L3 BTECs compared to A levels, and that 'older style' L3 BTECs may give an inflated estimate of student ability (eg Nelson, 2016b). These concerns call into question the validity of 'older style' L3 BTEC outcomes, especially in terms of their use for selection purposes into higher education and/or employment.

As part of our usual regulatory activities, and in response to these concerns, the purpose of the current report is to investigate the validity of 'older style' L3 BTEC

[3] To avoid over-complicating matters, the remainder of the current report shall mainly focus upon BTEC and A levels as separate routes. However, readers, and researchers wishing to further explore these themes, should be aware of the fact that many students do take a combined route at level 3.

[4] For example see https://www.telegraph.co.uk/education/2018/06/04/universities-now-admitting-twice-many-btec-students-dida-decade/

outcomes (relative to A level outcomes) and how this has changed over time, therefore providing empirical evidence on the potential existence of grade inflation. For the purposes of this report, we define grade inflation as 'illegitimate increases in outcomes over time, which do not reflect a genuine increase in the ability of cohorts of students'. We shall begin by reviewing some of the concerns that have been raised, before presenting the results of our own analyses into the existence (or not) of grade inflation in 'older style' L3 BTECs, focussing on 4 cohorts of school leavers between 2006 and 2016. Again, the focus throughout is on the 'older style' suite of L3 BTECs, not those more recently introduced.

1.1 Review of the existing evidence suggesting grade inflation

1.1.1 Evidence relating to level 3 (key stage 5) outcomes

It is worth noting to start that very little published research has been carried out into potential grade inflation in 'older style' L3 BTECs. In particular, only 1 report seems to have addressed the over-time element of our definition of grade inflation. Specifically, the Higher Education Funding Council for England (HEFCE, 2015) showed that the proportion of 'older style' L3 BTEC students achieving 'top grades' (defined as 3 distinctions or better) or 'high grades' (at least 2 distinctions and a merit or better) steadily rose between 2006 and 2013, despite the fact that the prior attainment of these cohorts actually decreased over this time (measured in this case by the proportion of candidates achieving 5 or more A*-C grades at GCSE). This would suggest grade inflation as it is unclear why outcomes would legitimately increase in this manner under the context of decreasing prior attainment, which is often thought of as a measure of the general ability of the cohort. In comparison, both outcomes and prior attainment for A levels remained stable over time, which is more aligned with what one would expect in the absence of grade inflation.

These findings might be explained by differences in the marking/awarding processes that exist for 'older style' L3 BTECs and A levels. For example, while a compensatory approach is taken for A levels, 'older style' L3 BTECs are graded according to firm criterion referencing (firm in the sense that candidates must be deemed to have achieved all pass criteria to achieve a pass, and all merit criteria to achieve a merit, etc.). As this approach does not allow for any adjustment of grade boundaries (there are no 'marks'), these criteria set the standard, and so become the method for standards maintenance. Arguably, because of accountability measures, teachers involved in grading have a vested interest in increasing outcomes over time, which this method cannot control for. Ultimately, this method is vulnerable to pressures of grade inflation. While 'older style' L3 BTEC awards are internally assessed and only subject to external verification by visiting subject experts (see Pearson, 2018), A levels are mainly externally assessed, and awards are subject to statistical controls (the 'comparable outcomes' approach – see Ofqual, 2017). The latter has been shown to be successful in curbing grade inflation (Ofqual, 2015), while the adequacy of the former has not, and has even been called into question (eg by the Wolf Report – Wolf, 2011). As noted before, 'older style' L3 BTECs are not alone in operating these processes, which are shared by other similar vocational and technical qualifications.

Stakeholder perceptions also seem to align with the idea that 'older style' L3 BTEC outcomes may have become inflated compared to A levels. For example, some media sources have argued that it may be easier to achieve equivalent outcomes in 'older style' L3 BTECs than A levels (Nelson, 2016a, 2016b). In the 2016 public perceptions survey, Ofqual (2016b) reported that only a third of young people (32%) and teachers (31%) surveyed believed that 'older style' L3 BTECs were of equivalent challenge to A levels (although these figures had risen somewhat from the same survey in 2015 – from 19% and 28% respectively). Allan (2017, sec. 4) also reported stakeholder concerns about the numbers of top grades (3 distinctions or above) being awarded for 'older style' L3 BTECs, again contributing to the perception 'older style' L3 BTECs may be 'easier' than A levels.

It is important to bear in mind that while these descriptive statistics and more anecdotal sources of evidence may point to the existence of grade inflation, they do not take into account various potentially confounding variables, which could offer an explanation for why outcomes could legitimately rise in this manner. For example, it is possible that there may have been a change in the characteristics of cohorts not reflected in their prior attainment at GCSE (prior attainment was taken into account by HEFCE, 2015), which could result in outcomes legitimately rising over time. The current report aims to address this gap, to build upon the existing evidence by exploring this issue in greater depth.

1.1.2 Evidence relating to level 6 (undergraduate degree) outcomes

Another concern that has been raised is that 'older style' L3 BTEC students may be relatively less well prepared for university compared to A level students with apparently equivalent levels of attainment at level 3, as measured by UCAS tariff scores (eg see Gill, 2015, 2016a, 2017). While this may not necessarily be problematic, so long as stakeholders recognise any such differences, any changes in equivalencies over time (eg due to grade inflation) would be a cause for concern. Prior analyses in this area focus on single years, and so cannot attest to any grade inflation over time, however findings do suggest that that 'older style' L3 BTEC outcomes may offer inflated estimates of ability compared to A level outcomes.

To first give some context, grades achieved at level 3 are each allocated a number of points according to The University and Colleges Admissions Service's (UCAS) tariff system. These points are aligned in such a way so as to allow university admissions officers to make comparisons of attainment between different qualifications (UCAS, 2017)[5]. It should be noted that while only about a third of higher education courses actually have tariff entry requirements (UCAS, 2017 – the rest use grades), the tariff nevertheless offers a useful metric of attainment at level 3. The idea is that students with the same number of tariff points should be equally prepared for university, regardless of the qualifications that were taken at level 3. Tariff equivalencies between A levels and 'older style' L3 BTECs over the period of

[5] The process for determining the alignment of tariff scores has been described by UCAS (2003). It is worth noting that while tariff points can be achieved via BTEC routes, not all universities/degrees accept BTECs in their entry requirements (Russell Group, 2017, pp. 18–19).

interest[6] are shown in Table 4. As this shows, 'older style' L3 BTEC distinctions are deemed equivalent with A level A grades, merits with C grades, and passes with E grades. 'older style' L3 BTEC diplomas and extended diplomas are aligned in the same manner as in Table 4, but are equivalent to 2 and 3 A level qualifications respectively. For example, D*D* (BTEC diploma) receives the same number of points as A*A* for A levels (280). D*D*D* (BTEC extended diploma) receives the same number of points as A*A*A* for A levels (420). Certificates are equivalent to half an A level, and 90 credit diplomas are equivalent to 1.5 A levels.

It is worth noting that Pearson also promotes 'older style' L3 BTECs as having parity with A levels in the UCAS tariff, in the manner outlined in Table 4 (Pearson, 2018b) (other awarding organisations do the same for their qualifications).[7]

Table 4. A level / 'older style' L3 BTEC UCAS tariff alignment

'older style' L3 BTEC Subsidiary Diploma Grade	A level Grade	Tariff points
D*	A*	140
D	A	120
	B	100
M	C	80
	D	60
P	E	40

Gill (2015)[8] assessed the equivalency of the tariff points earned by 'older style' L3 BTEC and A level students in terms of the probability of those students going on to achieve a first-class or upper second class undergraduate degree (in 2013). He concluded that the tariff did appear to be misaligned between the 2 course types, as the inclusion of course type (ie 'older style' L3 BTECs vs A levels) seemed to improve the predictive validity of a model of the relationship between tariff points and outcomes at university. In a second study, Gill (2016a) demonstrated that university students holding 'older style' L3 BTEC qualifications were less likely to gain a good degree result (a first or upper second class degree) compared to A level students with the same number of UCAS tariff points in 2012 or 2013 (also see Gill, 2017). These studies seemed to suggest that 'older style' L3 BTEC outcomes were not equivalent to A level outcomes in the manner described in Table 4, because it was usually assumed (at least under the old tariff system – see Footnote 6) that students holding the same number of tariff points should perform equally well at university. Other research has suggested that those following a vocational route are also more likely to drop out of their studies mid-degree compared to A level students, again possibly suggesting that they are potentially less well prepared for university (Bailey

[6] A new UCAS tariff has been introduced for university entries from 2017 onwards (UCAS, 2016), but this is beyond the period of interest for the current report.
[7] While this reference refers to alignment within the new tariff introduced in 2017, it does note that alignment was the same in the old tariff system.
[8] Gill (2015) actually replicated an earlier study by Green and Vignoles (2012), who tested the equivalence of tariff scores between A levels and the International Baccalaureate (IB). Findings suggested that the tariff also appeared to be misaligned between A levels and the IB.

& Bekhradnia, 2008). However, this study does not take into account other possible explanations, such as social factors.

While these reports suggest that 'older style' L3 BTEC grades may offer an inflated estimate of ability in relation to equivalent grades in A levels, each report only provides a cross-sectional account in a single year. In the current study, we include over-time analyses, to explore whether a change in equivalencies has occurred over time (ie grade inflation), or whether these findings reflect a more stable issue with misalignment between these course types (ie inter-qualification comparability).

1.1.3 Evidence relating to employment outcomes

Finally, it is important to remember that one purpose of BTEC courses is to prepare students for future employment (Pearson, 2017a). Because of their applied focus, one needs to consider that 'older style' L3 BTECs might not fully prepare students for university (see the previous section), but they might prepare students well for employment. Indeed, statistics reported by London Economics (2013) do suggest that those following an 'older style' L3 BTEC route were in employment to a similar level after graduating from university than those holding A level qualifications, and occupied a greater number of managerial positions (this data covered 1996 to 2011 – all years were combined into one dataset). However, employer perceptions of these qualifications appear to be less positive. In the 2016 Ofqual perceptions survey (Ofqual, 2016b, fig. 30), 48% of employers disagreed with the statement that they valued 'older style' L3 BTECs as highly as A levels (up from 42% in 2015). Just 28% agreed with this statement (down from 31% in 2015); 24% stated that they did not know whether or not they agreed/disagreed. While these are not necessarily strong metrics for performance in the workplace, this is certainly worth investigating.

1.2 Research questions

In accordance with the preceding discussions, the 3 main research questions that drove the current research were as follows:

1. Do changes in outcomes over time at level 3 suggest that grade inflation has occurred for 'older style' L3 BTECs?
2. Do changes in level 6 outcomes over time suggest the same?
3. Do changes in employment outcomes over time suggest the same?

In answering these questions, we shall consider the relative validity of 'older style' L3 BTEC outcomes over time compared to A level outcomes over the same period of time. By targeting these issues from the multiple angles above, we aim to draw together and expand upon the existing evidence to investigate this issue of potential grade inflation in more depth. In particular we combine the use of a statistical method that allows us to control for factors potentially affecting attainment, with an over-time research design. This will allow us to inform our regulatory position on 'older style' L3 BTEC qualifications, and to inform whether efforts are needed to address grade inflation in these qualifications, and possibly any other qualifications which operate similar models of standard maintaining and quality assurance (eg those offered by other awarding organisations).

2 Data

Data for all analyses came in the form of a linked dataset of the National Pupil Database (NPD), managed by the Department for Education, and data provided by the Higher Education Statistics Agency (HESA). Where missing demographic data existed in the NPD (census information for some types of post-16 centres are not captured by the NPD), this was populated from the Individualised Learner Record (ILR), managed by the Skills Funding Agency (however, we did not have matching ILR data for the latest cohort in the dataset – 2015/2016). Figure 1 shows the final structure of the linked datasets, which is described in more detail below.

NPD data (supplemented by the ILR where available) included data on level 3 students' age, gender, ethnicity, free school meal eligibility, IDACI[9] score, centre number, centre type, qualifications taken, outcomes at level 3, and prior attainment (average point score at key stage 4). We requested this data to cover 4 academic years: 2005/2006, 2009/2010, 2012/2013, and 2015/2016. The dataset was filtered to only include A level students that both took their exams and completed their level 3 studies in the same academic year, to focus on typical candidates (some had taken some of their exams several years before completing their studies). We also filtered the dataset to include only those that certificated in the summer (very few 'older style' L3 BTEC candidates complete their studies in the winter). As many of our analyses included prior attainment, students with missing data for prior attainment were excluded from the dataset. We focussed only on 'older style' L3 BTEC subsidiary diplomas, diplomas, and extended diplomas, as certificates and 90 credit diplomas only first appeared in the dataset in 2012/2013, limiting the over-time exploration that can be made for these course types.

HESA data included variables on degree subject area, university attended, university mission group[10], and the degree classification achieved: a first, an upper second (2:1), lower second (2:2) or a third. Data was linked to the aforementioned NPD data via an anonymised student identifier, and any non-linking records were excluded. Tariff scores on entry were calculated by the authors according to Table 4. Records were included where students started their undergraduate degree in the same or following year after leaving school (to include those taking gap years), completed an undergraduate degree, were on full time courses lasting 3 years only, were between the ages of 20 and 25 when they completed their degree, and were home students (international students were excluded, as few records would be held for their level 3 outcomes). Where multiple outcomes were recorded for a single graduate (eg due to retakes or appeals), the most recent outcome was selected. So as to not confuse the analyses on subject area, only those that completed a degree with the main subject weighted as 75% or more of the overall degree were included.

HESA (the agency) also periodically collects information on 'graduate destinations', as part of its Destinations of Leavers from Higher Education (DLHE) survey. This survey targets samples of university graduates twice yearly, covering information

[9] Income Deprivation Affecting Children Index (IDACI). Students are assigned an IDACI score based upon where they live. This is a measure of the proportion of children living in that area that are in income-deprived families (DCLG, 2015, para. 3.2.17).

[10] Many universities belong to one of five mission groups: the 1994 group (disbanded in 2013, but retained for the purposes of this report), GuildHE, MillonPlus group, Russell group, and the University Alliance. Not all universities belong to one of these groups – these have been categorised as 'other'.

such as what individuals were doing post-graduation (eg whether they were in full- or part-time employment, voluntary work, or further study), and if employed, their salary and occupation. The January survey covers those that graduated between January and July in the previous year; the April survey covers those that graduated between August and December in the previous year.

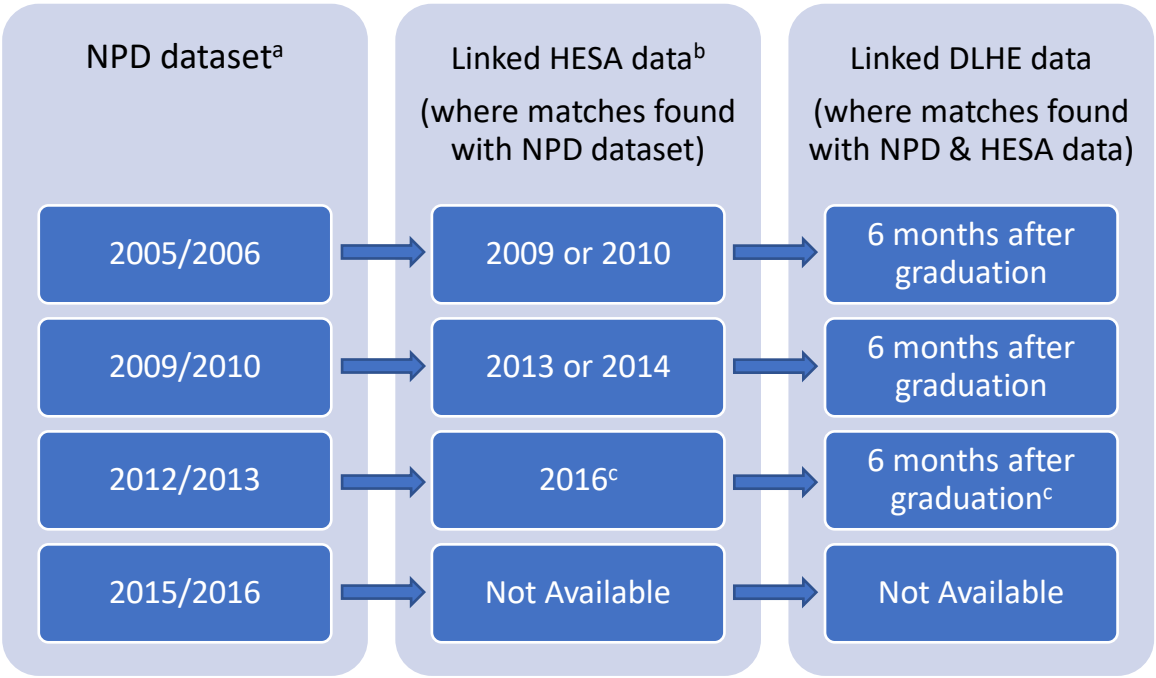


Figure 1. Data linking process

[a] Academic year in which the student completed their level 3 studies

[b] Year in which the student completed their degree

[c] Data covered 2009 to 2016. We therefore had no data for those who started their degree in 2014 (and graduated in 2017).

3 Propensity score weighting methodology

As our analyses spanned multiple years, it was important to take into account the fact that cohorts will differ over time in a number of ways such as their demographic makeup. It is important to control for such factors, as they could support the legitimacy of any changes in outcomes, which might otherwise be attributed to grade inflation. In other words, any changes in outcomes over time for 'older style' L3 BTECs may simply reflect a change in the characteristics of the cohorts taking those qualifications, rather than being a reflection of the qualifications themselves (ie grade inflation).

Ideally, in cases such as these where randomised allocation into groups is not possible (ie to conduct a 'randomised controlled trial', or RCT), one might wish to match pairs of candidates from the 2 different course types that each hold exactly the same characteristics. That way, one could model the relationship between course type (ie 'older style' L3 BTECs vs. A levels) and outcomes, under the absence of any differences in these other variables. In practice, however, it would be very difficult to identify enough pairs of exactly matching candidates to be able to perform any meaningful analyses.

One method of circumventing this issue is via the calculation of 'propensity scores' (for reviews, see Austin, 2011; Brookhart, Wyss, Layton, & Stürmer, 2013). These are defined as the conditional probability of being in a 'treatment' group (in this case, taking a 'older style' L3 BTEC vs. taking an A level), based on the background characteristics observed for each case. Propensity scores can be calculated using background characteristics to predict group membership for the cohort. It was shown by Rosenbaum and Rubin (1983) that by balancing those scores between the 2 groups, one in effect also balances the impacts of background characteristics on the outcome variable, allowing one to better establish the causal relationship between group membership and the outcome of interest. Propensity score methods rely on the assumption that all potential characteristics affecting group membership are included in the propensity score estimation.

A 'propensity score *weighting*' methodology was adopted.[11] Following this method, each case is assigned a weighting based upon its propensity score, so that when these weightings are combined across the cohort, there is no longer an association between background characteristics and group membership. One can then conduct additional analyses to assess the relationship between group membership and the outcome of interest. Interested readers are directed to Brookhart et al. (2013) for a fuller description of how exactly these weights are calculated.

[11] 'Propensity score *matching*' is perhaps a more commonly used method. However, this method usually results in a loss of data when groups are unbalanced in size (unmatched cases are usually dropped from the analysis). Given the current context, where there were many more candidates taking A levels than BTECs, propensity score weighting was deemed more appropriate, as this method does not lead to any loss of data.

3.1 PSW method (for level 3, level 6, and employment analyses)

For the current research, weights were calculated separately for each different set of comparisons that were made between 'older style' L3 BTEC and A level students (ie separately for each of the level 3, level 6, and employment analyses, and for each cohort/year). This was necessary to effectually balance between each group of 'older style' L3 BTEC and A level candidates. Weights were calculated using the *'twang'* package for *R* (see Ridgeway, McCaffrey, Morral, Burgette, & Griffin, 2017; Ridgeway, McCaffrey, Morral, Griffin, & Burgette, 2017). This package makes use of gradient boosted regression models, which was deemed to better predict the course type taken on the basis of background characteristics than would be achievable via more traditional regression methods. Interested readers are directed elsewhere for more comprehensive descriptions of these machine-learning methods (eg Elith, Leathwick, & Hastie, 2008; Lee, Lessler, & Stuart, 2009).

In our propensity score models, we included course type (BTEC vs. A level) as the dependent (predicted) variable, and gender, ethnicity, free school meal eligibility, IDACI score, average key stage 4 (KS4) point score, and centre type as independent (predictor) variables. These predictor variables were included, as they have been shown to correlate with students' choice of courses (eg see Thomson & Keshwani, 2017; Vidal Rodeiro, Sutch, & Zanini, 2013). Such variables may also have an impact on some of our outcomes variables, if not controlled for. For example, family background (eg IDACI and free school meal eligibility) has been shown to correlate with employment outcomes, even after controlling for educational achievement (Macmillan, Tyler, & Vignoles, 2015). Candidates with any missing data for any of the above variables were excluded from the propensity score analyses. For each model produced, diagnostics were performed which confirmed that an appropriate balancing of covariates had been successful (diagnostic results can be found in Appendix B).

One issue that we initially faced was that very large weights were produced for some individual cases, meaning that they had a very large influence on the analyses that followed. This issue has been described previously by Robins, Hernán, & Brumback (2000), who recommended calculating 'stabilised' weights (SW), which are still able to balance covariates, but reduce the weighting of very influential cases. We used the formulas shown by Xu et al. (2010 – based upon Robins et al., 2000), where *p* is the proportion of the overall sample in the 'older style' L3 BTEC group (unweighted probability), and π_i is the propensity score that was calculated for each case:

$$\text{'older style' L3 BTEC group: } SW_i = \frac{p}{\pi_i}$$

$$\text{A level group: } SW_i = \frac{1-p}{1-\pi_i}$$

4 Results

4.1 Level 3 (school) outcomes

We shall begin with a simple exploration of grade distributions over time for 'older style' L3 BTECs and A levels, before presenting more in depth analyses that control for various confounding variables. While the focus of this report is on 'older style' L3 BTECs, A level outcomes are included as they offer a useful benchmark against which 'older style' L3 BTECs can be compared. To offer a sense of the size of the dataset, Table 5 shows the number of grades that were awarded each year per qualification type, for those that were included in the dataset.

Table 5. Number of grades awarded per qualification type (after exclusion criteria)

	Total number of grades awarded			
Qualification type	2005/2006	2009/2010	2012/2013	2015/2016
'older style' L3 BTEC sub. dip.	9,600	50,765	88,125	63,410
'older style' L3 BTEC diploma	9,465	27,200	36,110	17,985
'older style' L3 BTEC ext. dip.	34,425	63,970	76,750	43,325
A level	758,535	850,560	838,525	644,665

Note. Values have been rounded to the nearest 5. "Sub. dip." = subsidiary diploma; "ext. dip." = extended diploma.

4.1.1 Grade distributions

Figure 2 shows the grade distributions for each course, in each academic year.[12] Because A level A* grades and BTEC D* grades were only awarded from 2009/2010, 'top grades' (defined as A or above in A levels, or D/DD/DDD or above in BTECs) have been combined in all analyses to allow for comparability over time.

As one can see, distributions for A levels have remained relatively stable over time. 'older style' L3 BTEC outcomes, however, have noticeably changed over time, with all course types progressively shifting from a positively skewed distribution in 2005/2006 (ie skewed towards the lower end of the range) to a negatively skewed distribution in 2015/2016 (ie skewed towards the upper end of the range). The number of top grades for 'older style' L3 BTECs (ie D/D*; DD+; DDD+) in particular have increased over time, reaching 61% for subsidiary diplomas in 2015/2016. While the proportions of top grades in A levels (A/A*) remained below 30% in all years, the proportion of top grades in 'older style' L3 BTECs have tripled between 2005/2006 and 2015/2016 for subsidiary diplomas and extended diplomas, and quadrupled for

[12] Grade distributions for other qualifications which have similar structures and similar mechanisms for maintaining standards show similar patterns – these can be found in Appendix A.

diplomas. Figure 3 shows that these changes occurred under the context of stable prior attainment for all course types.

It is worth noting here that A levels are not *necessarily* inherently 'stable' qualifications – rather, one way their 'stability' has been ensured is via the awarding methodology used to maintain standards (the comparable outcomes approach). 'older style' L3 BTECs are not subject to the same control.

Figure 4 allows for more direct comparisons of top grades. This shows that the proportion of students achieving top grades increased for all course types between 2005/2006 and 2009/2010. However, following the introduction of tighter statistical controls for A levels in 2010 (known as the 'comparable outcomes approach' – see Ofqual, 2017), the proportion of students achieving top grades in A levels remained stable over time. Proportions for 'older style' L3 BTECs, which do not share the same statistical controls, have continued to steadily increase over time.

Taken together, Figures 2-4 seem consistent with the proposition that 'older style' L3 BTEC outcomes have been subject to grade inflation. In particular, it seems surprising that the attainment of top grades should rise so notably, given that the overall ability of the cohort (indicated via prior attainment) does not seem to have changed in this time. The patterns observed for 'older style' L3 BTECs are in stark contrast to those observed for A levels, which have remained much more stable over the period of interest.

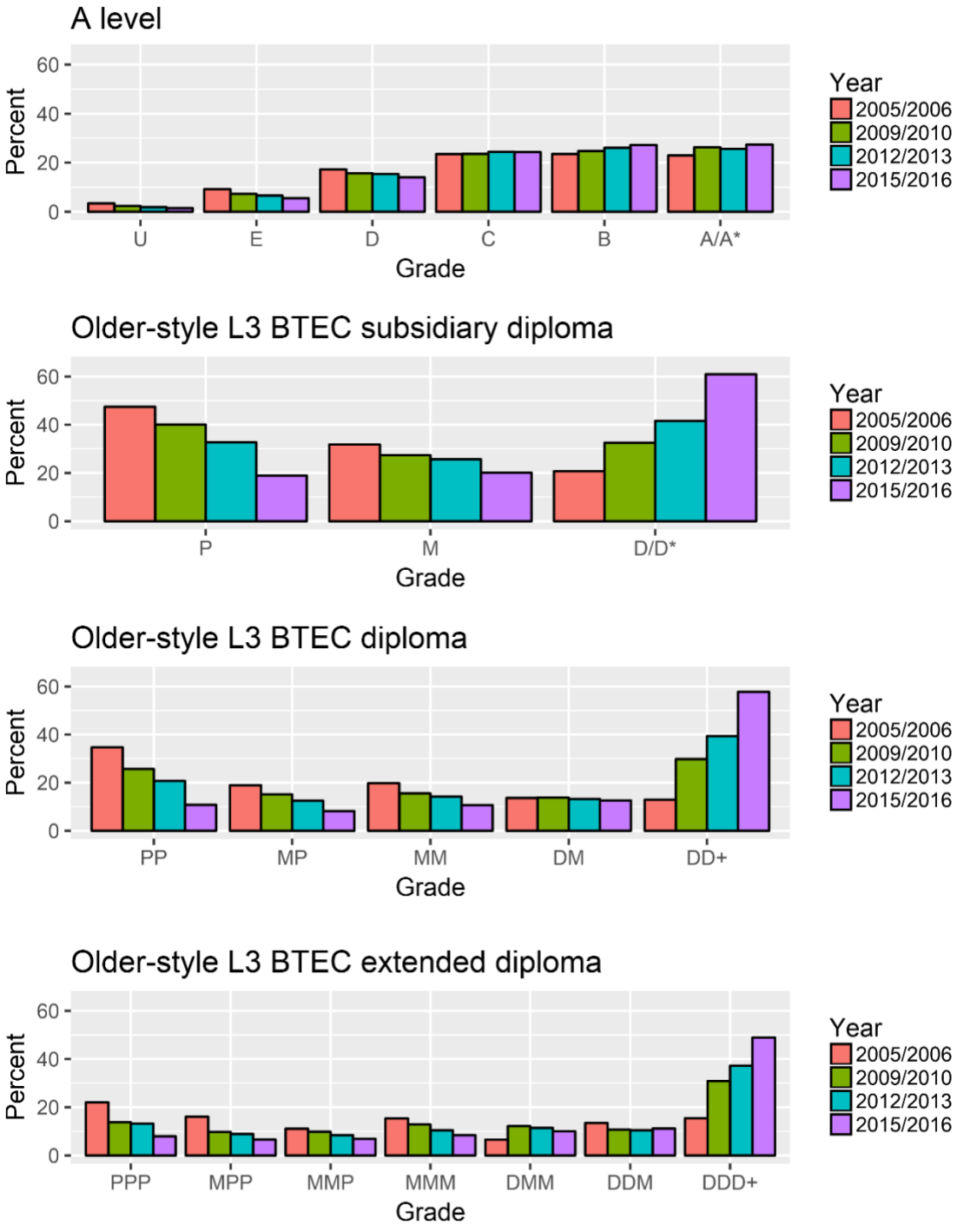


Figure 2. Grade distributions over time

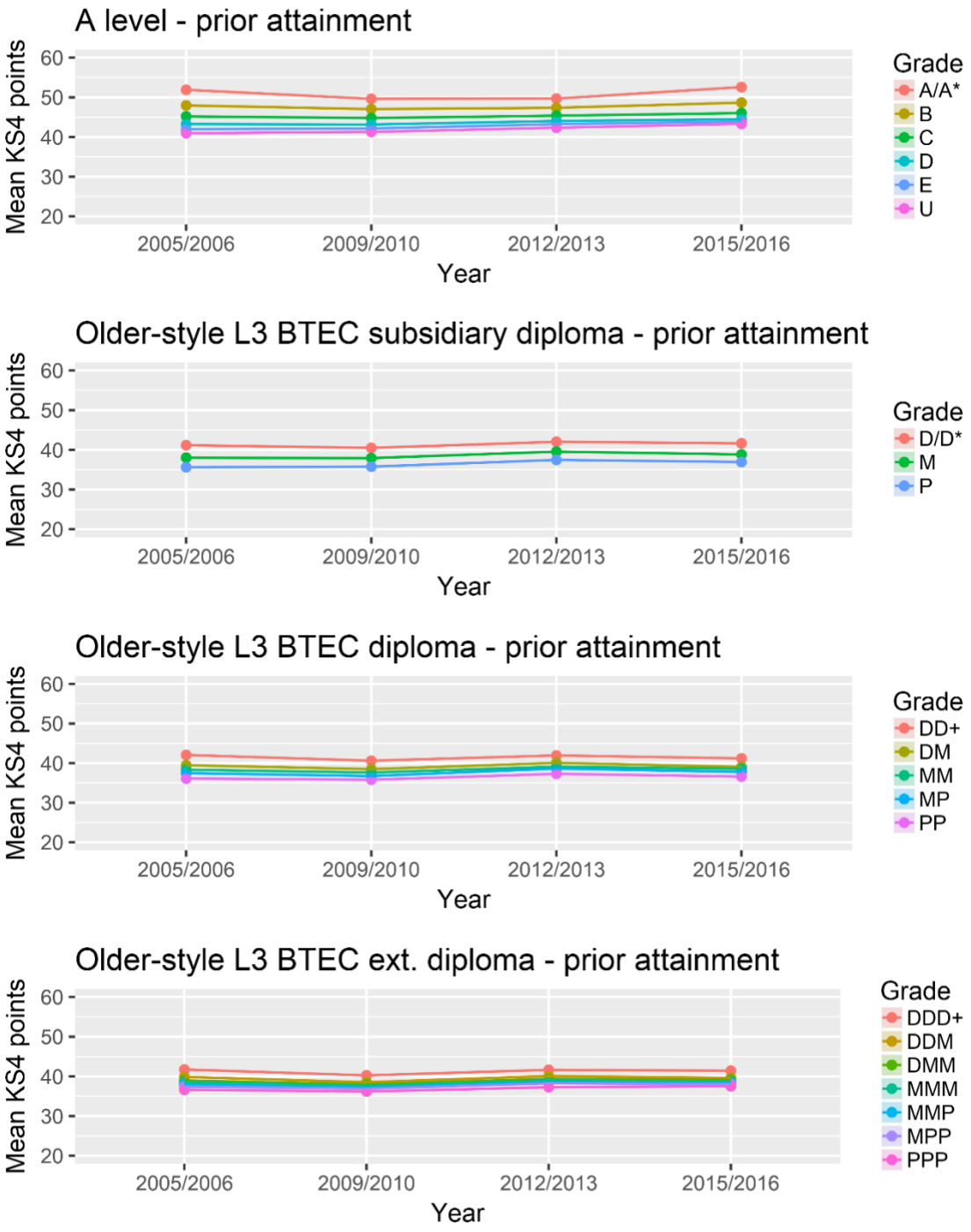


Figure 3. Prior attainment over time, by outcomes at level 3

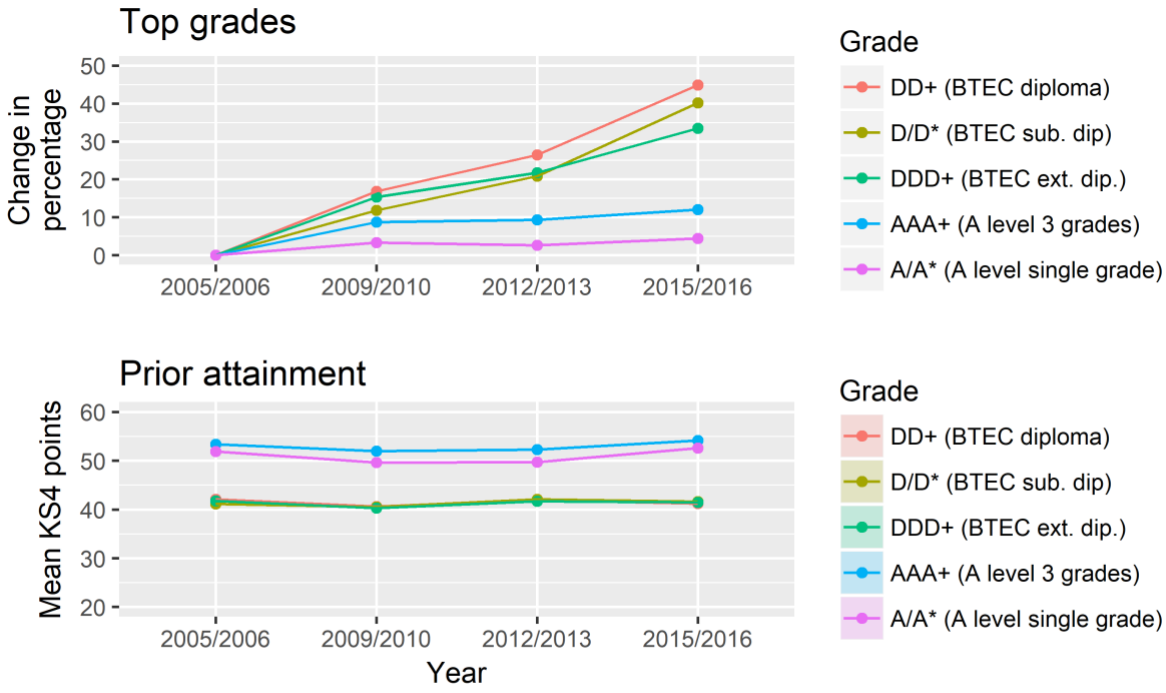


Figure 4. Attainment of top grades over time

Notes. To aid readability, the y-axis of the top graph has been transformed to show the change in percentage from 2005/2006 (thus the 1st point of each line shows zero change; the 2nd point shows the increase in percentage points from 2005/2006; etc.). In addition, to provide a comparison against 'older style' L3 BTEC extended diplomas, candidates have been identified that achieved 3 A level grades. The combined attainment for these candidates have also been plotted (ie those achieving 3 top grades: AAA+), labelled 'A level 3 grades'. "Sub. dip." = subsidiary diploma; "ext. dip." = extended diploma.

4.1.2 Predicted probabilities of achieving top grades in 'older style' L3 BTECs vs. A levels

It is important to remember that a change in prior attainment is not the only factor that can have an impact on outcomes at level 3. Other factors, such as a change in the demographic characteristics of cohorts, could support the legitimacy of the increases in 'older style' L3 BTEC outcomes observed. It was important, therefore, to control for such factors in our analyses. This was done via the propensity score weighting method described earlier. We began with a comparison of 'older style' L3 BTEC subsidiary diplomas and A levels. As each subsidiary diploma is equivalent in size to 1 A level, each grade (or entry) was considered 1 'case' here (rather than each candidate's aggregated grades across multiple qualifications). After calculating propensity score weights (see Section 3), the relationship between the achievement

of top grades (A/A* or D/D*) and course type was modelled separately for each year. This was done via logistic regressions, with each case weighted by their stabilised weight. Updated sample sizes are shown in Table 6 (ie following the exclusions of those with missing data). More students were excluded in 2015/2016 due to missing demographic data than in other years, as we did not have matching ILR data for that cohort. Nevertheless, sample sizes were still more than sufficient for the analyses conducted.

Table 6. Sample sizes for level 3 regressions

Qualification type	2005/2006	2009/2010	2012/2013	2015/2016
'older style' L3 BTEC sub. dip.	9,100	49,285	86,400	33,590
A level (individual grades)	638,715	725,515	699,000	391,630
'older style' L3 BTEC ext. dip.	32,085	62,050	74,790	7,610
A level (3 grades combined)	93,100	114,850	125,560	89,940

Note. Values have been rounded to the nearest 5. "Sub. dip." = subsidiary diploma; "ext. dip." = extended diploma.

To take into account the fact that outcomes are clustered within schools, standard error estimation was clustered by centre number. Failing to do so might have resulted in under-estimated standard errors (eg see Cameron & Miller, 2015). In the current case, one would have ideally also wished to account for the fact that grades are clustered within candidates. However, given that this would have translated in a very computationally intensive operation which was deemed to be impractical, readers should bear in mind that standard errors may be somewhat underestimated for this analysis. As all other analyses were conducted at a candidate level, and so nesting was not relevant, this limitation is restricted only to this current analysis.

Odds ratios (a measure of effect size) are shown in Table 7. In this case, these would be interpreted as the ratio of the odds of achieving top grades in an 'older style' L3 BTEC compared to the odds of achieving top grades in an A level. For example, in 2005/2006, the odds of achieving top grades in an 'older style' L3 BTEC subsidiary diploma were 3.18 times greater than the odds of achieving top grades in an A level (note that this is not the same as saying you are 3.18 times more likely to achieve top grades in a 'older style' L3 BTEC compared to an A level). All contrasts were statistically significant.

Table 7. Logistic regression: course type and the achievement of a top grades

Academic year	Odds ratio	95% confidence interval
'older style' L3 BTEC subsidiary diploma		
2005/2006	3.18***	2.57 – 3.92
2009/2010	5.70***	5.13 – 6.34
2012/2013	7.44***	6.81 – 8.13
2015/2016	24.24***	22.08 – 26.60
'older style' L3 BTEC extended diploma		
2005/2006	8.00***	6.57 – 9.75
2009/2010	16.71***	14.86 – 18.78
2012/2013	22.46***	19.65 – 25.67
2015/2016	46.70***	29.73 – 73.35

Note. Reference category = A levels (each grade or 3 grades combined). ***$p < .001$

A more easily understandable metric to odds ratio is the difference in predicted probabilities (ie on a scale where 0 implies no probability of an event occurring, and 1 implies that an event will certainly occur). Using the regression model that had been produced, in which the influence of background characteristics has been removed, one can predict the probability that a student will achieve top grades, based upon the type of course that a student could take. These predicted probabilities are shown in Figure 5. As one can see, while probabilities remained stable over time for A levels, the probability for 'older style' L3 BTEC students is consistently higher in each year, and this difference has increased over time. Again, the reader is reminded that background characteristics have been controlled for in this figure, via the propensity score weighting method, and so these differences in probabilities cannot be accounted for by differences in prior attainment, for example. This would therefore call into question the legitimacy of the increases observed.

Table 7 and Figure 5 also show the results relating to 'older style' L3 BTEC extended diplomas. As extended diplomas are equivalent in size to 3 A levels, it did not make sense in this case to compare them against individual A level grades. Comparisons were instead made against candidates who had achieved 3 A level grades (only), meaning that the bottom panel of Figure 5 compares the probability of getting 3 top grades in either course type, at a candidate level. Sample sizes for 'older style' L3 BTEC extended diplomas are again shown in Table 6. These results again suggest the presence of grade inflation for 'older style' L3 BTECs, as the observed increases in attainment over time (relative to A levels) cannot be attributed to a change in the characteristics of cohorts.

Analyses were not performed for 'older style' L3 BTEC diplomas (equivalent in size to 2 A levels), as there is no obvious group against which to compare outcomes for these students (most A level candidates take 3 courses).

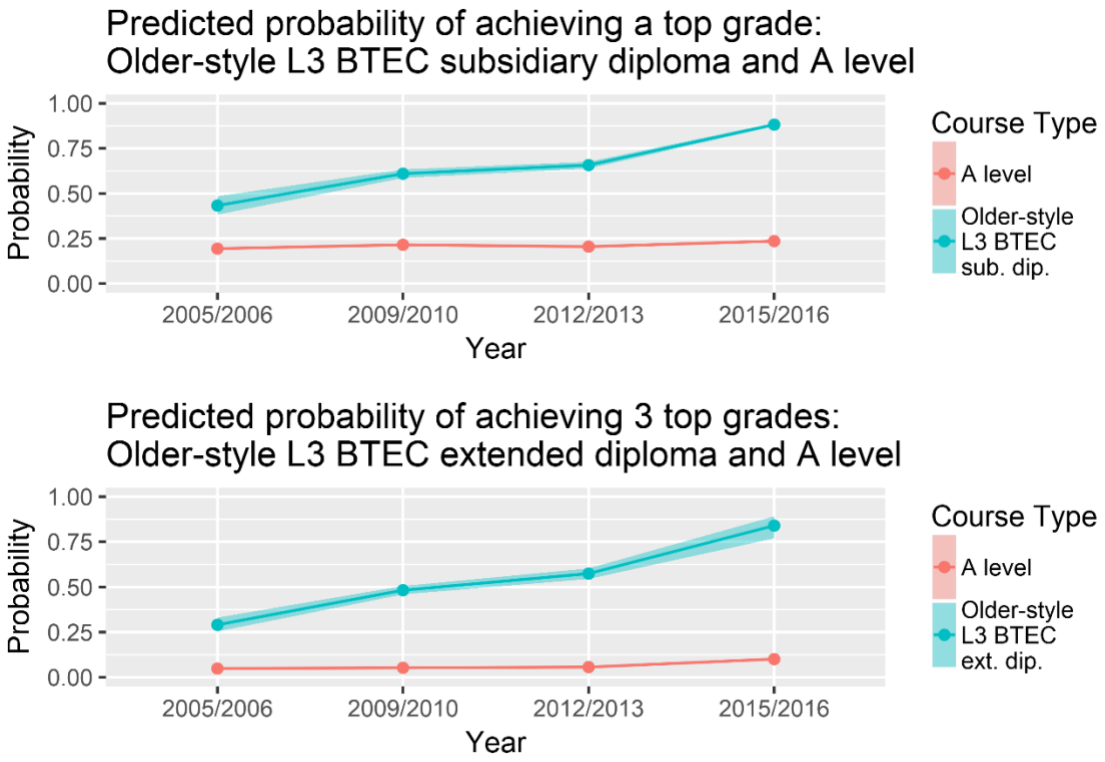


Figure 5. Predicted probability of achieving top grades over time

Note. "Sub. dip" = subsidiary diploma; "ext. dip." = extended diploma. Shaded ribbons show the 95% confidence interval of the predicted probability. Probabilities are for candidates with 'average' prior attainment. As 'older style' L3 BTEC students tend to have lower than average levels of prior attainment (see Figure 4), the probabilities shown here are somewhat higher than the proportions actually achieving top grades (see Figure 2). In other words, when the differences in prior attainment between the 2 groups are taken into account, the difference in outcomes between 'older style' L3 BTEC and A level students becomes all the more apparent.

4.1.3 Candidates taking a mixture of 'older style' L3 BTECs and A levels

Despite our efforts, it is still possible that other, unaccounted for variables might explain these differences. Some variables, such as a difference in personalities, are not captured by the data, and so would not necessarily have been balanced through the propensity score weighting. Another way to tackle this issue is to compare the performance of students who took a combination of 'older style' L3 BTECs and A levels at level 3. As noted in the introduction, taking a combined route is an increasingly popular option. By comparing grades achieved for different courses within the same candidates, far fewer concerns of confounding variables exist[13]. We therefore identified candidates who had achieved at least 1 'older style' L3 BTEC, and at least 1 A level grade. Sample sizes for these analyses are shown in Table 8.

[13] Some possibilities do still exist, however, such as differences in the classroom environment of different course types. However, one would not expect such differences to have a large impact.

Table 8. Sample sizes for candidates taking a mixture of 'older style' L3 BTECs and A levels

	Number of students			
Qualification type (in addition to A levels)	2005/2006	2009/2010	2012/2013	2015/2016
'older style' L3 BTEC subsidiary diploma	1,255	9,770	20,905	22,455
'older style' L3 BTEC diploma	910	5,365	7,945	5,590
'older style' L3 BTEC extended diploma	1,285	1,740	2,155	1,320

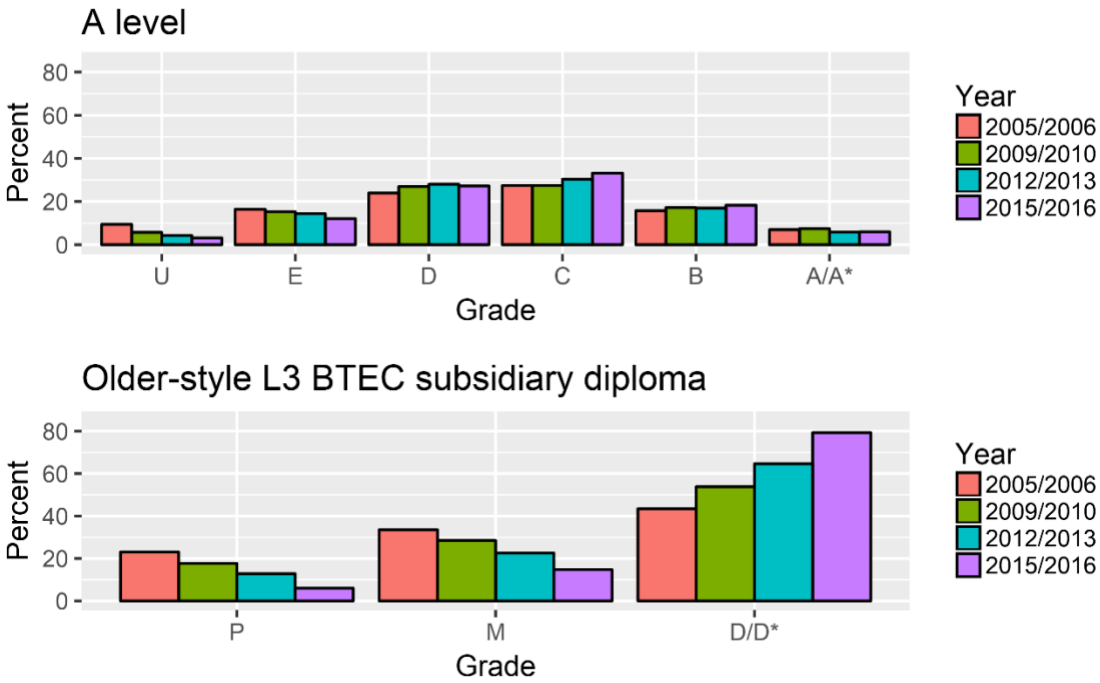
Note. Values have been rounded to the nearest 5.

The same graphs as before were plotted: Figure 6 shows the grade distributions and Figure 7 shows the mean prior attainment. As before, findings show that while outcomes for these students' A levels remained stable over time, outcomes for their 'older style' L3 BTECs became progressively more negatively skewed over time (towards the upper end). To use subsidiary diplomas as an example, in 2015/2016, 6% of this group's A levels were graded A/A*, but 79% of 'older style' L3 BTECs were graded D/D*, despite these being the same candidates taking these qualifications. Again, changes over time occurred under a context of stable prior attainment, which would appear consistent with the existence of grade inflation.

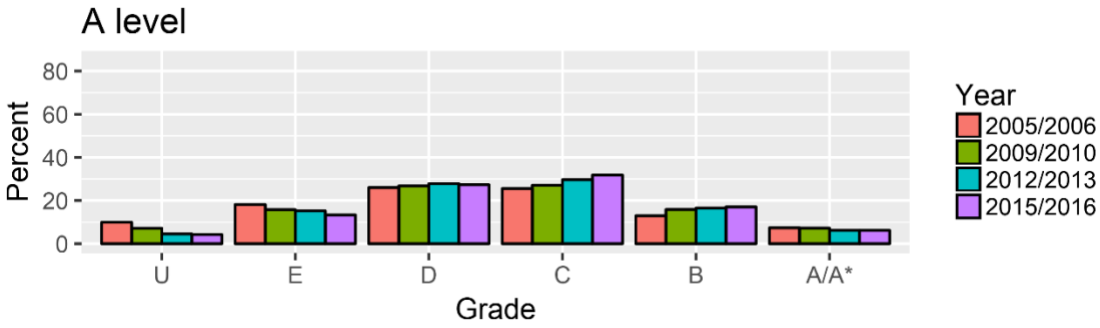
Figure 8 presents a comparison of the achievement of top grades over time. Although we did not need to calculate propensity score weights here (no balancing of background characteristics was necessary), some weighting of outcomes was still needed. This is because some candidates had, for example, achieved 1 'older style' L3 BTEC grade and 2 A level grades, and others vice-versa, which created an unequally weighted influence of candidates on the results for each course type. As such, we calculated weights for each candidate, such that the combined weights of all their outcomes summed to 1, and the sum of their weights for their 'older style' L3 BTECs were equal to the sum of their weights for their A levels (50% for each course type). For example, a candidate with 2 'older style' L3 BTEC grades and 1 A level grade had the former weighted 25% each, and the latter weighted 50%. Thus, they had an equally weighted influence on the findings relating to each course, and had an equal overall weighting to other candidates with different grade profiles.

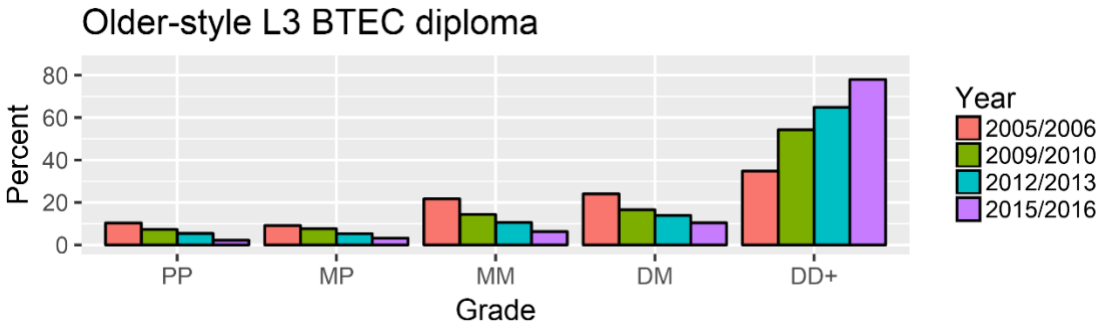
The observations that can be drawn from these figures are the same as before: attainment of top grades in 'older style' L3 BTECs has steadily risen over time, even though the prior attainment of the cohort has not changed. No such increases in the same candidates' A level outcomes were observed. Again, this seems consistent with the proposition that 'older style' L3 BTEC grades may have been subject to grade inflation over this period of time.

Candidates taking a mix of A levels and 'older style' L3 BTEC subsidiary diplomas



Candidates taking a mix of A levels and 'older style' L3 BTEC diplomas





Candidates taking a mix of A levels and 'older style' L3 BTEC extended diplomas

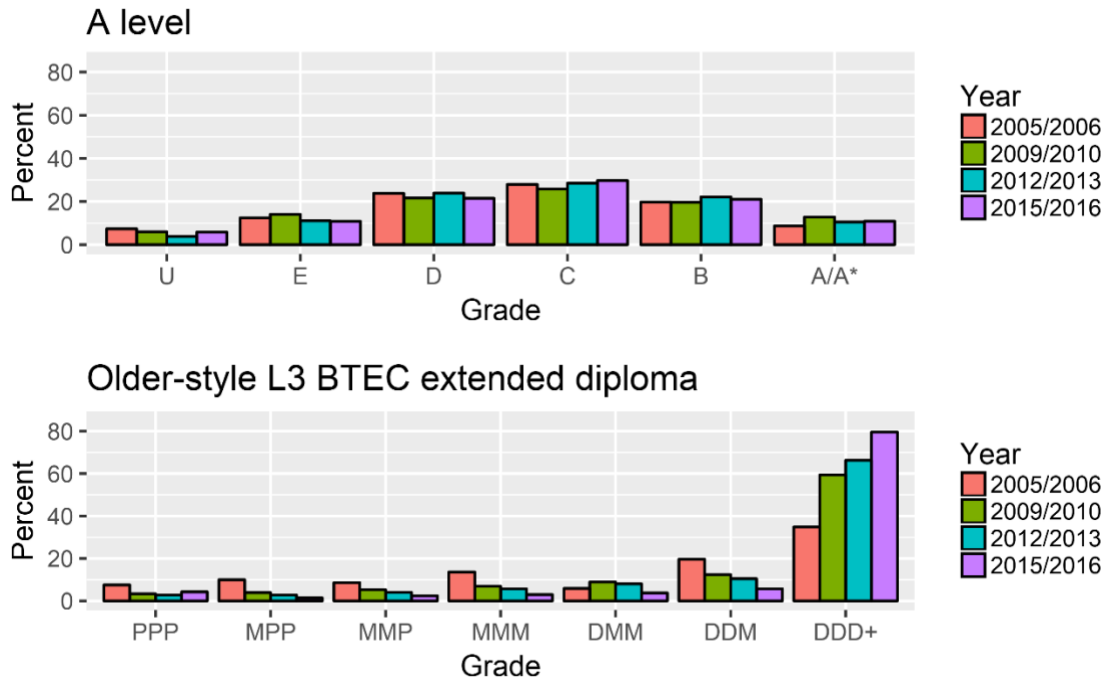
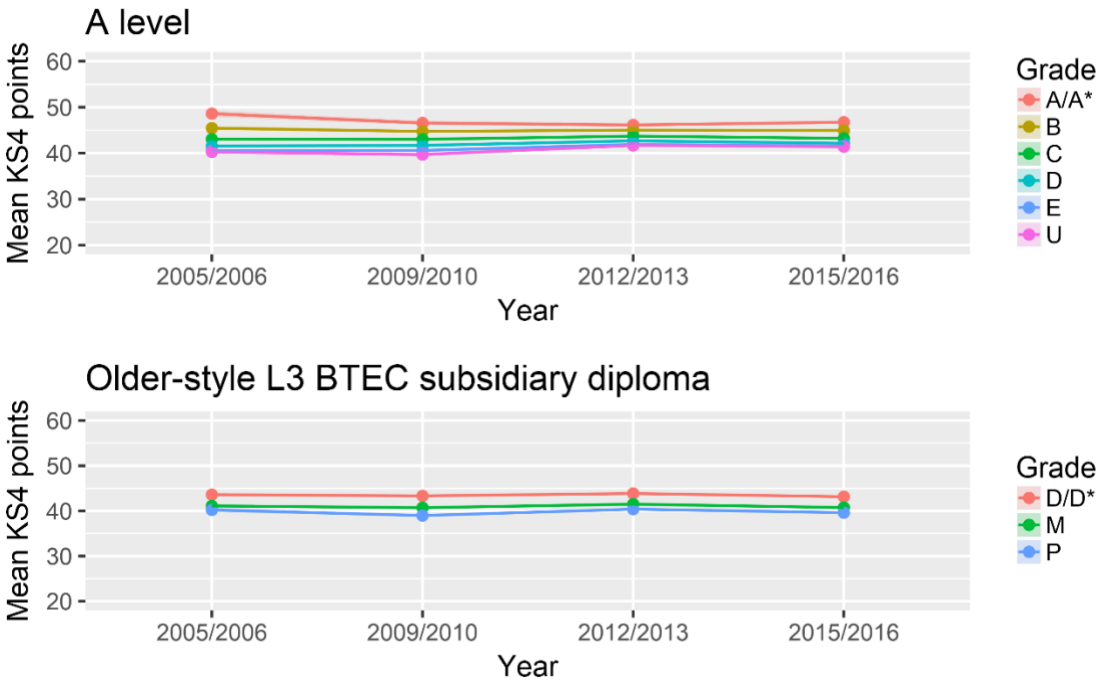
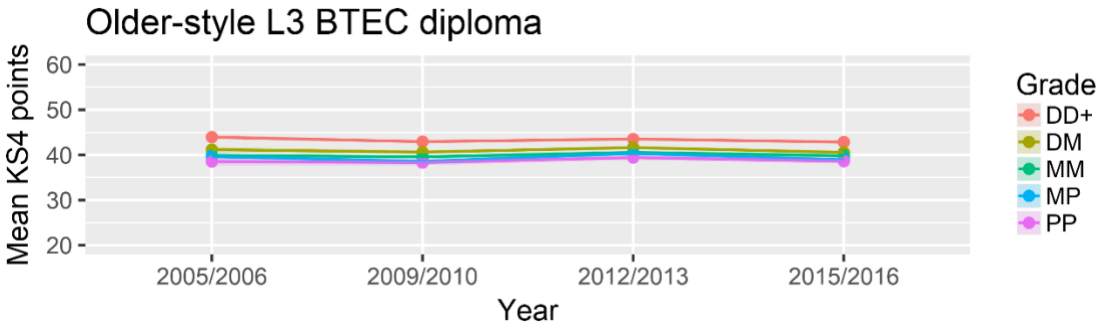


Figure 6. Grade distributions over time for candidates taking a mix of courses

Candidates taking a mix of A levels and 'older style' L3 BTEC subsidiary diplomas



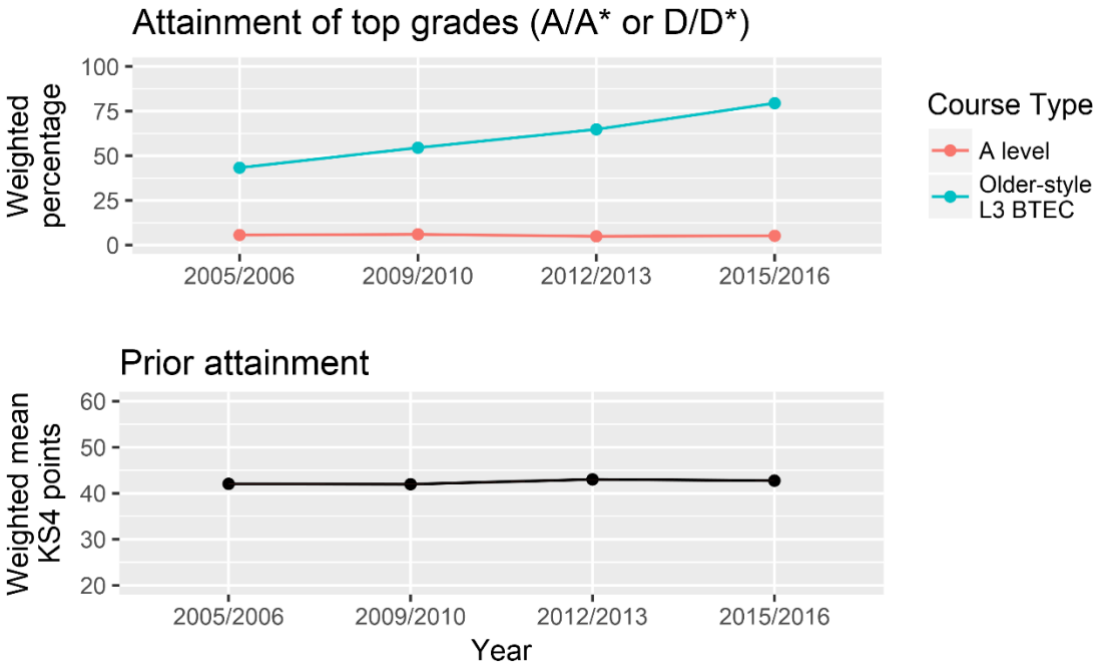
Candidates taking a mix of A levels and 'older style' L3 BTEC diplomas



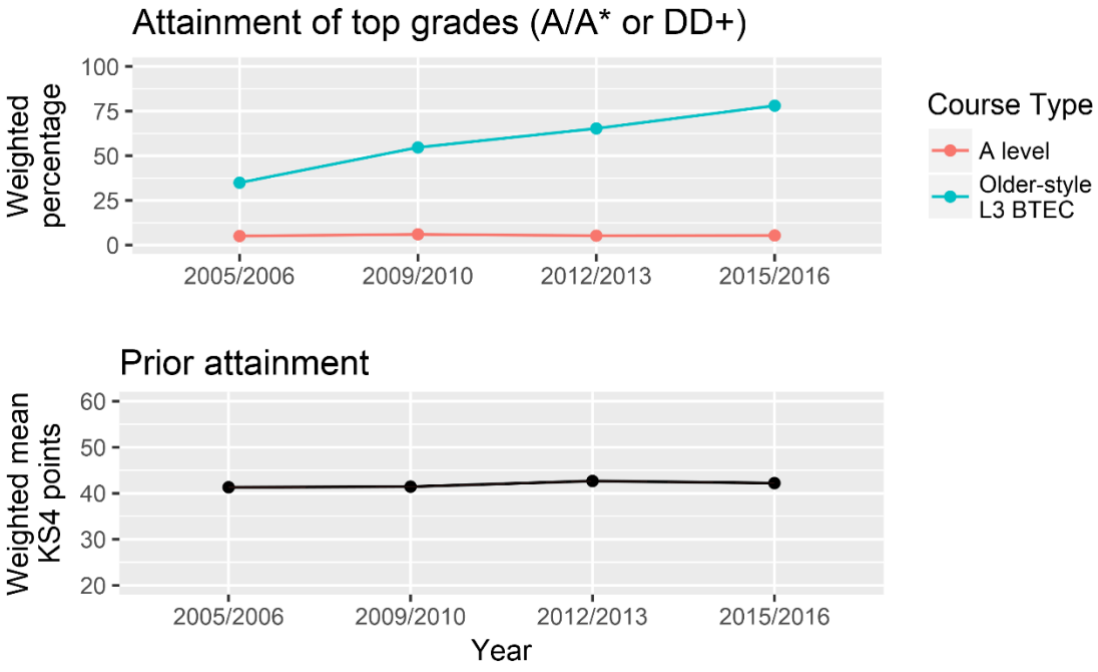
Candidates taking a mix of A levels and 'older style' L3 BTEC extended diplomas

Figure 7. Prior attainment over time for candidates taking a mix of courses, by grade achieved

Candidates taking a mix of A levels and 'older style' L3 BTEC subsidiary diplomas



Candidates taking a mix of A levels and 'older style' L3 BTEC diplomas



Candidates taking a mix of A levels and 'older style' L3 BTEC extended diplomas

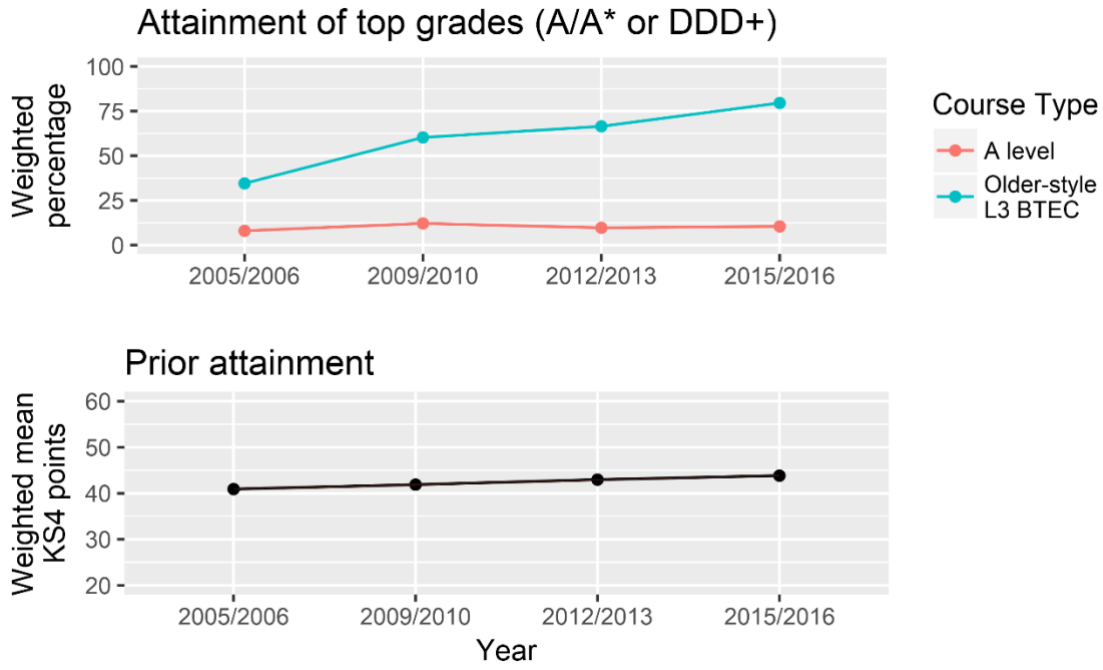


Figure 8. Top grades over time for candidates taking a mix of courses

Notes. Only 1 prior attainment line is shown in each graph, because this is a within candidates comparison.

4.2 Level 6 (undergraduate degree) outcomes

The findings thus far have all seemed consistent with the existence of grade inflation. However, one could still argue that there may be some unaccounted for factors which could support the legitimacy of these changes. So, assuming that the increases in 'older style' L3 BTEC outcomes do reflect genuine positive change in the ability of the cohorts, one might expect those changes to also be reflected in their performance at university and in employment. In this section we assess this, by comparing the performance of 'older style' L3 BTEC and A level students in their subsequent university degrees. Subsequent employment outcomes shall be considered in Section 4.3.

University graduates were identified who had only taken either 'older style' L3 BTECs or A levels at school. For this section, we focussed only on 'older style' L3 BTEC extended diploma students (equivalent in size to 3 A levels), as there were too few candidates who had only taken 3 subsidiary diplomas, and we wanted to be able to compare against the typical A level entry of 3 A levels. Similarly, there were no obvious student groups against which 'older style' L3 BTEC diplomas (equivalent in size to 2 A levels) could be compared. To make a fair comparison, we selected only A level candidates who had entered university with exactly 3 A level qualifications (although they may have held other level 3 qualifications such as AS levels). For

simplicity, the term ''older style' L3 BTEC student' will henceforth be used to describe university graduates that left school with an 'older style' L3 BTEC extended diploma, and 'A level student' to describe university graduates that left school with 3 A levels. Table 9 shows the size of the datasets used.

Table 9. Sample sizes for level 6 analyses

	Number of university graduates	
Academic year leaving school	'older style' L3 BTEC extended diploma students	A level students
2005/2006	5,350	52,285
2009/2010	11,910	68,465
2012/2013	10,655	51,300

Note. Values rounded to nearest 5.

We shall use UCAS tariff points throughout this section as a metric for making comparisons between 'older style' L3 BTEC and A level students with equivalent levels of achievement at level 3. Table 10 shows the mean tariff score for each type of graduate. As one can see, 'older style' L3 BTEC students had a higher mean tariff than A level students each year, with this difference expanding over time. This is simply a reflection of the grade distributions presented earlier (ie the fact that the number of top grades being awarded for 'older style' L3 BTECs has increased over time, whereas A level outcomes have remained fairly stable).

Table 10. Mean tariff entering university by cohort and course type

	Mean tariff	
Academic year leaving school	'older style' L3 BTEC extended diploma (closest equivalent grade)	A level (closest equivalent grade)
2005/2006	279 (DMM = 280)	272 (BBC = 280)
2009/2010	310 (DDM = 320)	282 (BBC = 280)
2012/2013	343 (DDD = 360)	286 (BBC = 280)

Figure 9 shows the cumulative percentage of graduates who achieved a 1st and at least a 2:1 or 2:2, according to their tariff upon entry, and level 3 course type. As we are only looking at those who completed their degree, 3rd class degrees are not shown here, because 100% of the dataset achieved at least a 3rd class degree.

This graph shows that 'older style' L3 BTEC students appear to consistently be at a comparative disadvantage to A level students, because fewer 'older style' L3 BTEC students achieved each degree classification, compared to A level students with the same tariff score. In other words, an 'older style' L3 BTEC student holding exactly the same number of tariff points as an A level student seemed to be less likely to achieve a 1st/2:1/2:2 degree.

Importantly, particularly for 1st class degree outcomes, this relative disadvantage appears to have increased over time. Figure 10 shows this more clearly – on average, 'older style' L3 BTEC students appear to have needed higher level 3 outcomes than A level students to achieve the same degree result, and this difference appears to have widened over time.

These findings call into question the idea that the abilities of 'older style' L3 BTEC students legitimately rose in the manner suggested by the increases in outcomes presented previously. Instead, they suggest that 'older style' L3 BTEC outcomes have offered a progressively lower reflection of students' preparation for university over time in comparison to equivalent A level outcomes. It seems hard to conclude, therefore, that 'older style' L3 BTEC and A level students with the same tariff scores are equally prepared for university studies. As before, however, there are other factors that may explain these findings, which need to be controlled for.

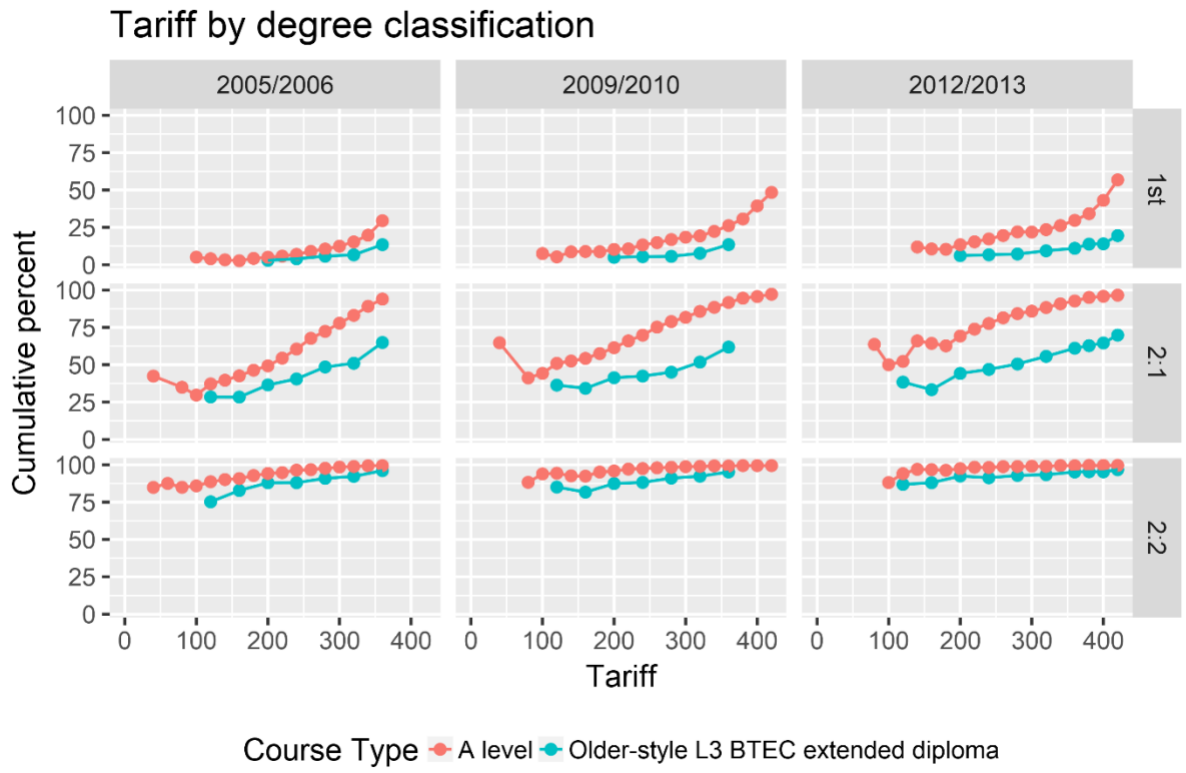


Figure 9. Cumulative percentage achieving each classification by tariff and year.

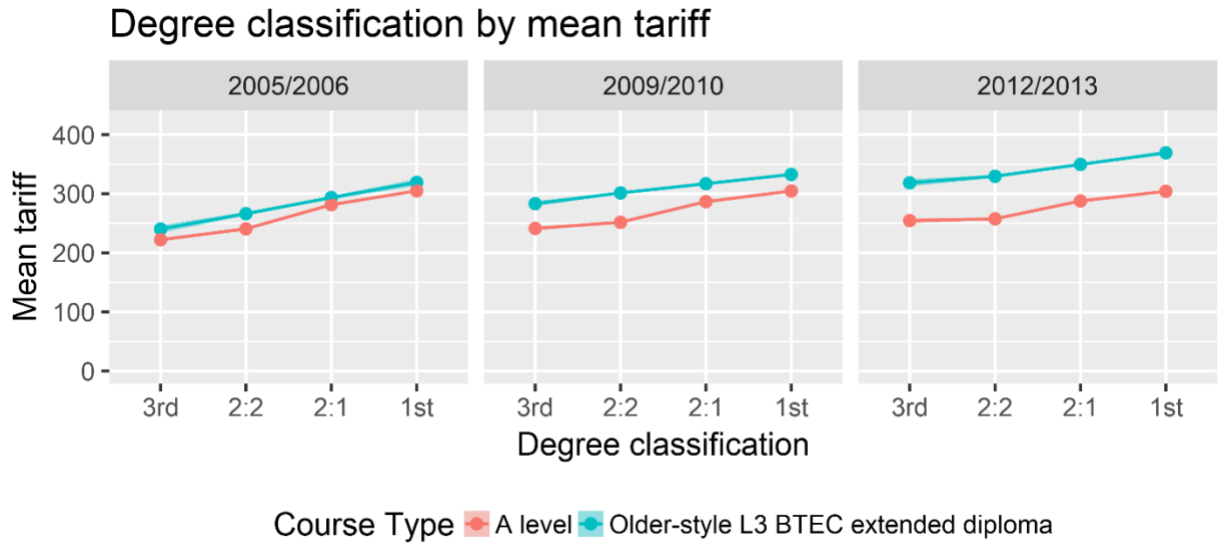


Figure 10. Mean tariff of candidates achieving each degree classification

Note. 95% confidence intervals are plotted, but are too small to be visible.

Using the same method as before (see Section 3.1), propensity score weights were calculated for each cohort, balancing the background characteristics of the 'older style' L3 BTEC and A level students. Using logistic regressions, the relationship between course type (BTECs vs. A levels) and the attainment of either a 1st or at least a 2:1 degree class was modelled for each year (applying the propensity score weightings). Standard errors were computed in a way to take into account the cross-classified structure/clustering of candidates both within schools and within universities. Tariff upon entry, higher education mission group, and degree subject area[14] were included as additional variables in the model. An interaction term was also included between level 3 course type and degree subject area, as any difference in performance at university between 'older style' L3 BTEC and A level students may be greater or lesser in some subject areas compared to others (ie given the more applied focus of 'older style' L3 BTECs). Table 11 shows the sample sizes for the regressions, following the removal of graduates with missing data.

Table 11. Sample sizes for level 6 regressions

	Number of university graduates	
Academic year leaving school	'older style' L3 BTEC extended diploma students	A level students
2005/2006	4,325	30,210
2009/2010	10,075	40,770
2012/2013	9,200	32,425

Note. Values rounded to nearest 5.

Table 12 shows the odds ratios relating to the 'course type' coefficient (for brevity, other coefficients are not shown, but are reflected in the graphs to follow). Values less than 1 indicate that the odds of 'older style' L3 BTEC students achieving a 2:1/1st are *lower* than the odds of A level students achieving a 2:1/1st (after accounting for the other variables included in the model). This was the case for all contrasts. All differences were also statistically significant, meaning that 'older style' L3 BTEC students were statistically significantly less likely to achieve a 2:1/1st, when all other variables are equal between the 2 groups. This relative disadvantage for 'older style' L3 BTEC students also increased over time (lower odds ratios indicate a greater disadvantage), suggesting that 'older style' L3 BTEC outcomes offered a progressively decreasing level of preparation for university, in comparison with equivalent A level outcomes.

[14] HESA assigns each degree subject to one of 19 subject areas. To further condense these into a more manageable number for the regressions, each HESA subject area was assigned to one of the following subject areas described by Bramley (2014): 'Applied', 'expressive', 'humanities', and 'STEM'. Bramley also described a 'languages' subject area, but as there were very few BTEC students studying language degrees, these students were excluded from the analysis.

These findings not only confirm the relative lower readiness of 'older style' L3 BTEC students for university, but also call into question the legitimacy of the increases in 'older style' L3 BTEC outcomes observed at level 3.

Table 12. Logistic regression: course type and the achievement of a top grades

Academic year leaving school	Odds ratio	95% confidence interval
Probability of achieving a 1st		
2005/2006	0.27***	0.18 – 0.42
2009/2010	0.25***	0.17 – 0.38
2012/2013	0.15***	0.11 – 0.21
Probability of achieving at least a 2:1		
2005/2006	0.32***	0.18 – 0.57
2009/2010	0.27***	0.21 – 0.35
2012/2013	0.19***	0.15 – 0.25

Note. Reference category = A levels. ****p* < .001.

As noted previously, odds ratios are not always easy to understand in real terms, and perhaps a more relatable metric is a difference in predicted probabilities. Figure 11 and Figure 12 show plots of the predicted probability of achieving each classification for each cohort, for 'typical graduates' from each course type, at each tariff score. 'Typical graduates' were defined as those who completed a degree in the most common subject area ('humanities'), in the most common university mission group (the 'other' group). As the graphs show, after controlling for background characteristics, those entering university with 'older style' L3 BTECs were less likely than A level students *with equivalent level 3 attainment* to achieve either a 1st or at least a 2:1 in their degree, with this comparative disadvantage becoming greater over time. Again, this calls into question what was observed at level 3, implying that outcomes had been subject to grade inflation. While this effect may appear somewhat more pronounced for higher ability candidates (probability of getting a 1st) than for more middle ability candidates (probability of getting at least a 2:1), this could be due to ceiling effects for A level students graduating with at least a 2:1.

Figure 13 and Figure 14 also show this effect for 'typical graduates' (mission group = 'other') in each degree subject area. The same pattern of results as above is again implied.

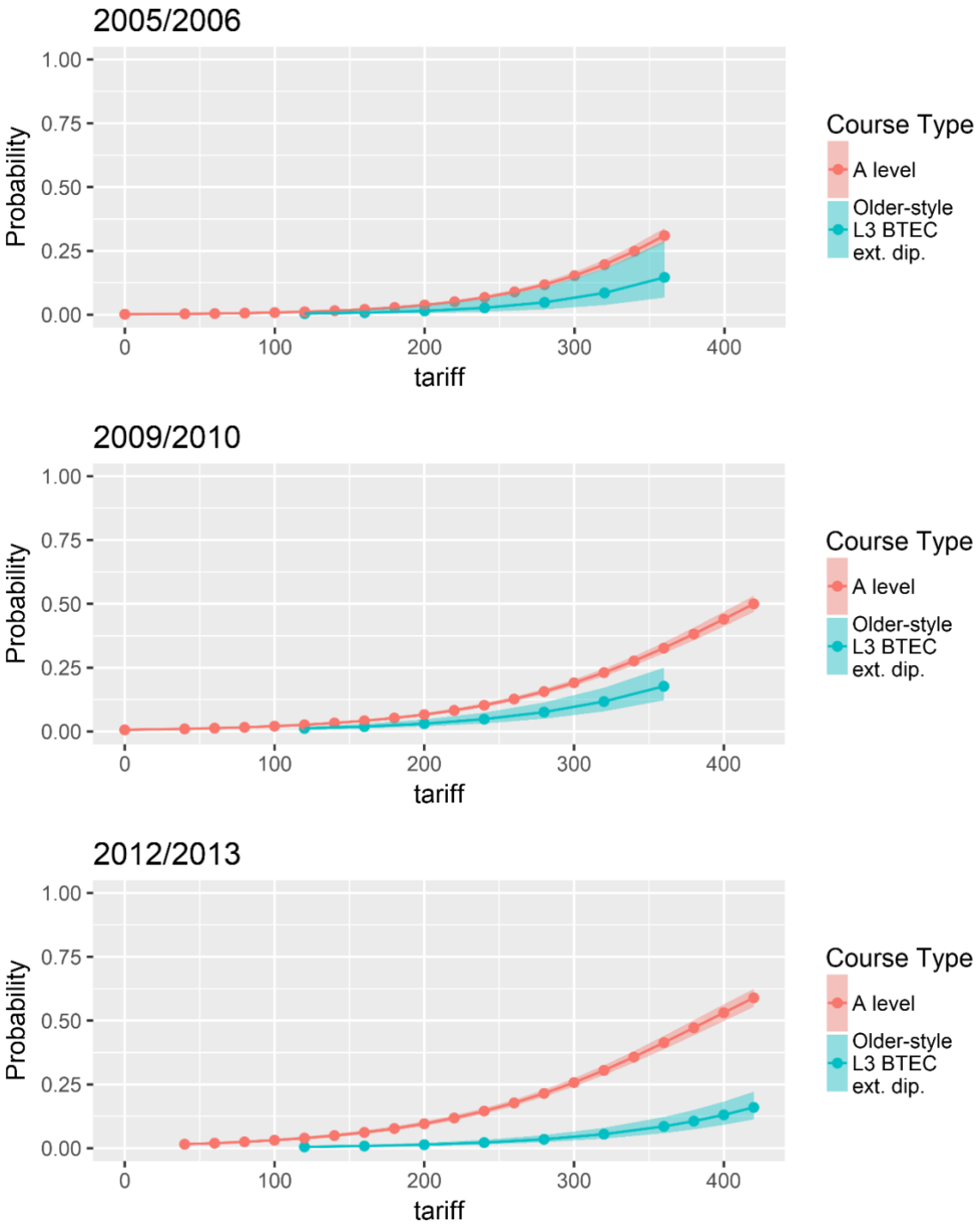


Figure 11. Probability of achieving a 1st by tariff for 'typical graduates'

Note. In all graphs, academic years refer to the school leaving cohort, not the year in which students graduated from university. "Ext. dip." = extended diploma.

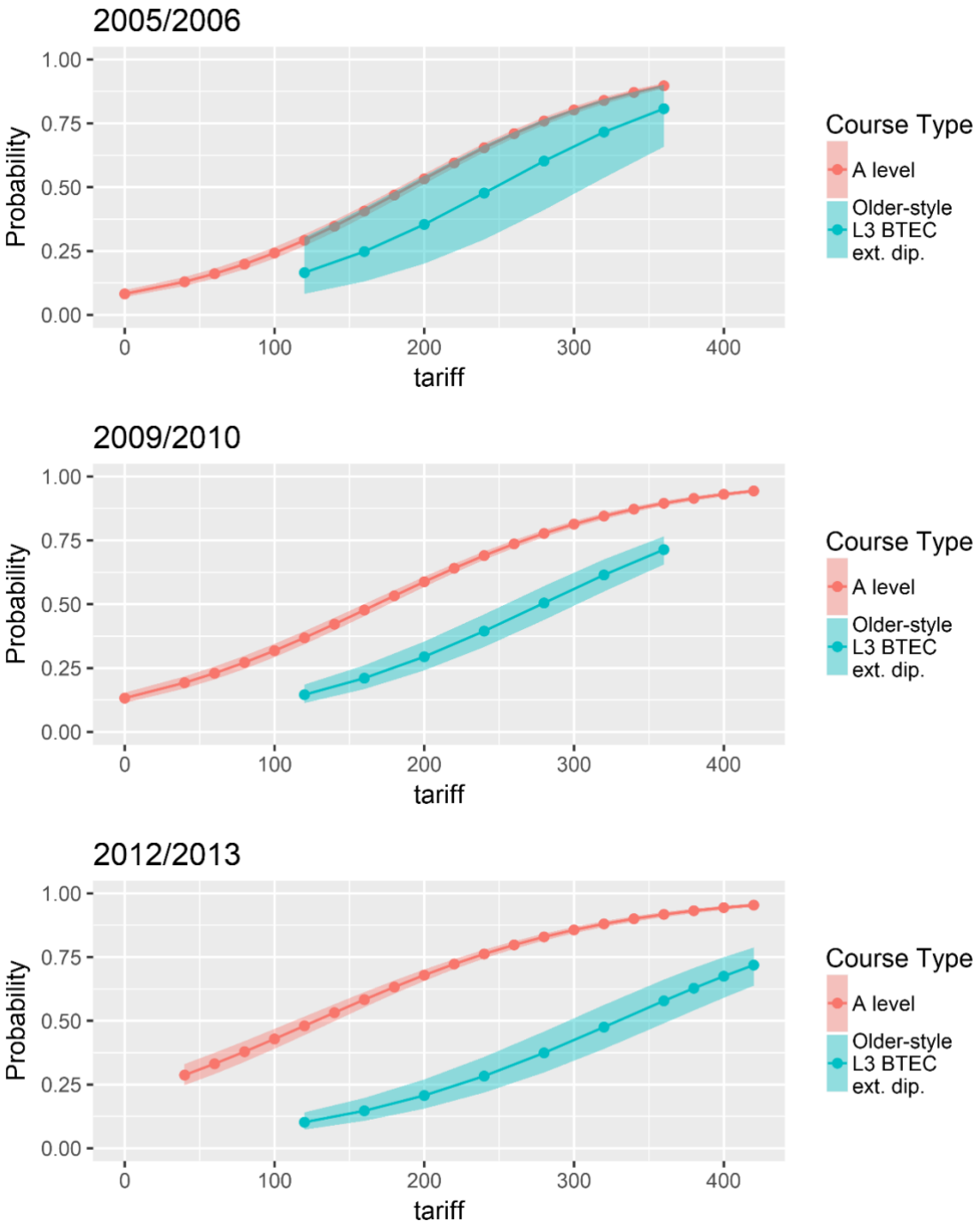


Figure 12. Probability of achieving at least a 2:1 by tariff for 'typical graduates'

Note. "Ext. dip." = extended diploma.

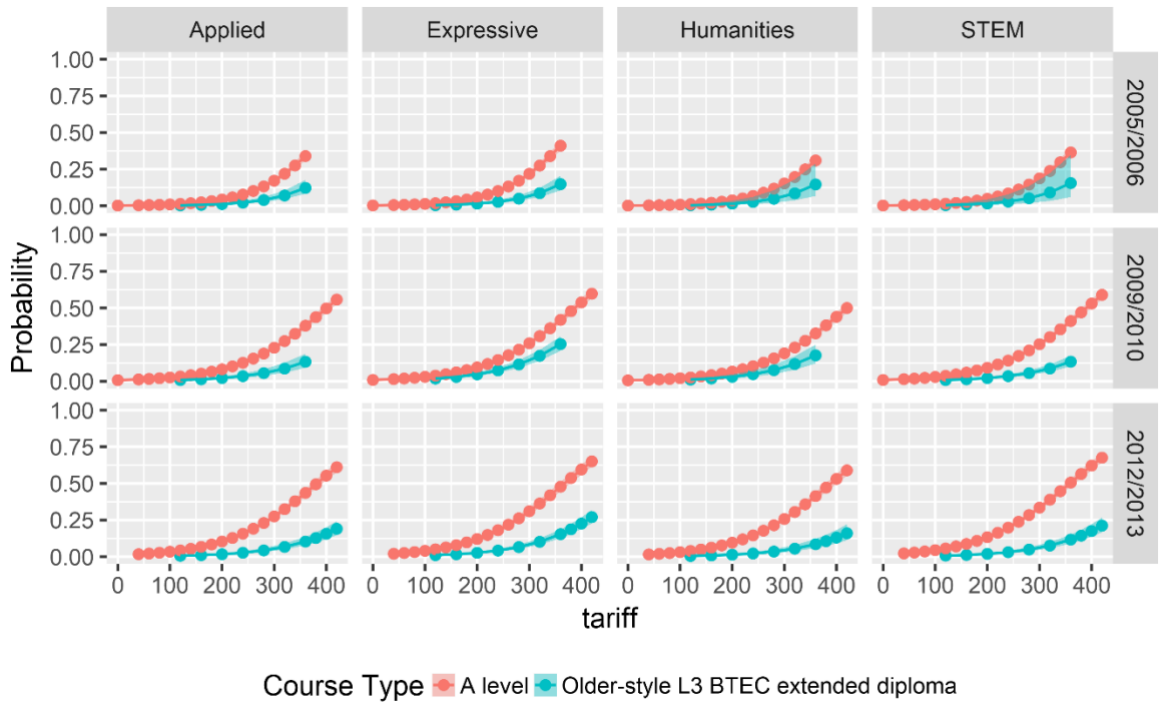


Figure 13. Probability of achieving a 1st by tariff for 'typical graduates' from each degree subject area

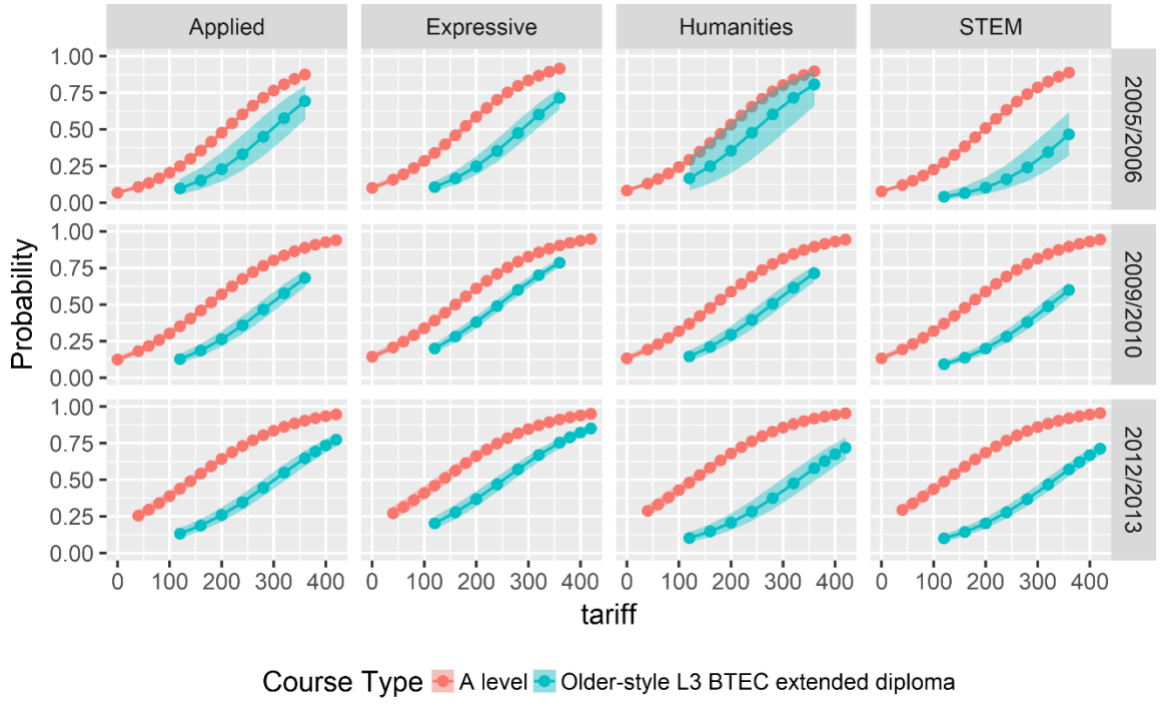


Figure 14. Probability of achieving at least a 2:1 by tariff for 'typical graduates' from each degree subject area

We also looked at university graduates who had completed their degree in the same subject area as their level 3 studies. As before, we made use of the subject areas described by Bramley (2014): 'applied', 'expressive', 'humanities', and 'STEM'. Again, Bramley also had a 'languages' category, but this was excluded as there are no 'older style' L3 language BTECs. Each level 3 qualification, as well as each degree, was assigned to one of these groups, and students were identified who had completed a degree in the same subject area as their level 3 studies. Because A level students could have completed 3 courses in different subject areas, students were included in this analysis if 2 out of 3 of their A levels were in the same subject area (this majority area was then considered their 'specialism'). Propensity scores were recalculated for this new 'same subject area' cohort. Sample sizes (after exclusions) are shown in Table 13.

Table 13. Sample sizes for level 6 regressions (same subject area)

	Number of university graduates	
Academic year leaving school	'older style' L3 BTEC extended diploma students	A level students
2005/2006	2,810	12,755
2009/2010	5,675	17,830
2012/2013	4,870	14,615

Note. Values rounded to nearest 5.

After running the same regressions as before (Table 14), probability graphs were replotted (Figure 15 and Figure 16). Again, A level students exhibited a greater probability of getting a 1st or at least a 2:1 than 'older style' L3 BTEC students at each level of the tariff, and this gap again widened over time. While alignment is better in some subject areas compared to others, this spreading over time is still apparent in most cases. The wide confidence intervals around the probability curves for STEM is due to small sample sizes in the 'older style' L3 BTEC group.

Table 14. Logistic regression: course type and the achievement of a top grades (same subject area)

Academic year leaving school	Odds ratio	95% confidence interval
Probability of achieving a 1st		
2005/2006	0.35***	0.20 – 0.62
2009/2010	0.26***	0.16 – 0.43
2012/2013	0.18***	0.11 – 0.28
Probability of achieving at least a 2:1		
2005/2006	0.29***	0.16 – 0.53
2009/2010	0.25***	0.18 – 0.35
2012/2013	0.17***	0.11 – 0.26

Note. Reference category = A levels. ***$p < .001$

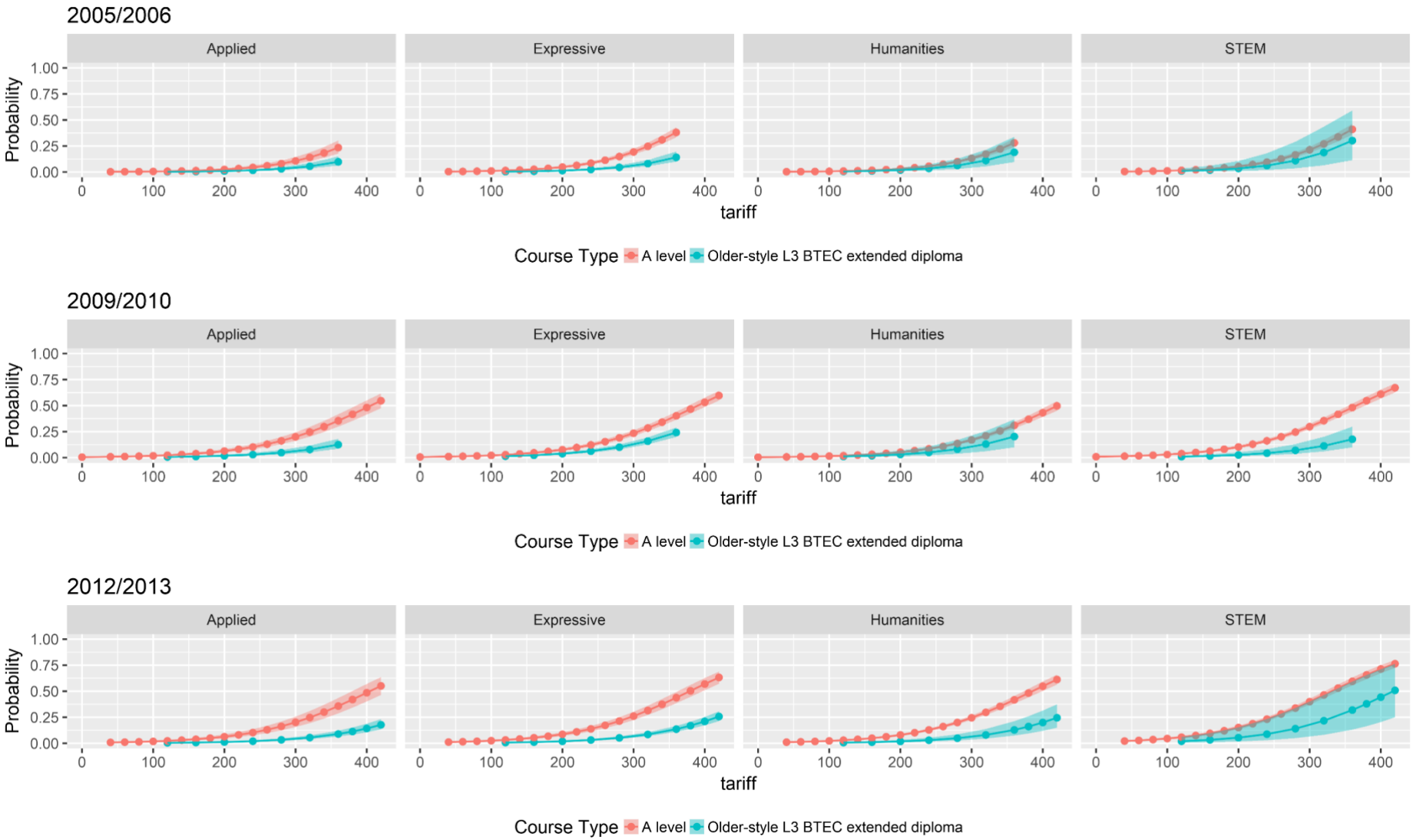


Figure 15. Probability of achieving a 1st by tariff for 'typical graduates' from each same subject area

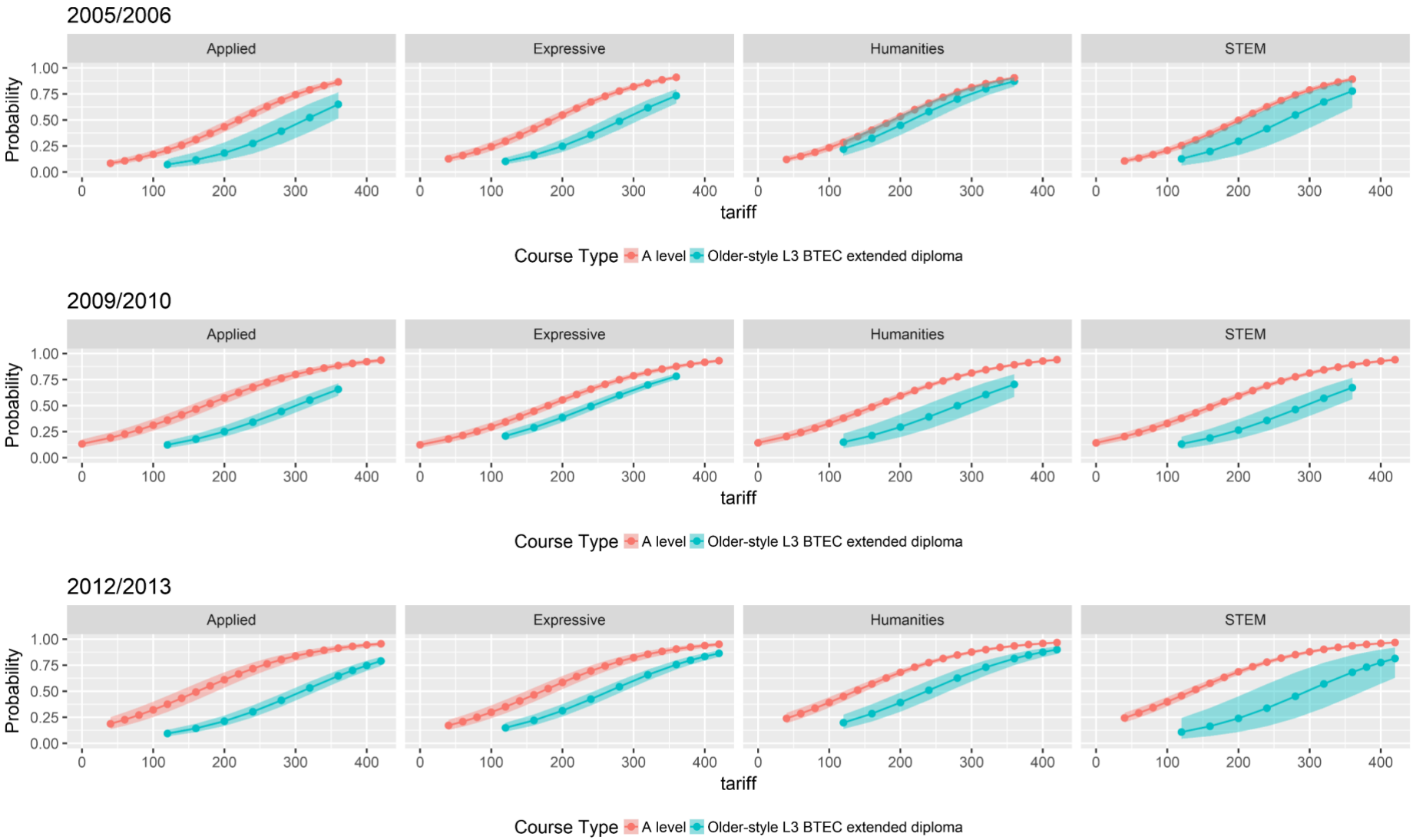


Figure 16. Probability of achieving at least a 2:1 by tariff for 'typical graduates' from each same subject area

To summarise this section on level 6 outcomes, the evidence relating to performance at university sheds further doubt on the legitimacy of the increases in level 3 attainment over time for 'older style' L3 BTEC students. Had those increases been a reflection of genuine improvements in cohort ability, one would expect 'older style' L3 BTEC and A level students with equivalent level 3 outcomes to perform equally well at university. However, we have demonstrated the opposite here. 'older style' L3 BTEC students seem to have a lower likelihood of achieving a 1st/2:1 than A level students with the same level 3 outcomes (tariff score), after also controlling for other potential confounds, such as background characteristics, the university attended, and subject studied at university. This comparative disadvantage has appeared to increase in nearly all cases, again suggesting that 'older style' L3 BTEC outcomes offered a progressively lower estimate of preparedness for university over time (comparative to A levels), which would imply grade inflation.

4.3 Employment outcomes

Our final set of analyses concerned the ability of 'older style' L3 BTEC students to gain access into 'higher quality' employment, after graduating from university[15]. The following analyses are based upon graduates' responses to the DLHE survey, 6 months after graduating. Propensity weights were again recalculated to balance the background characteristics of survey respondents who had left school with either an 'older style' L3 BTEC extended diploma or 3 A level qualifications. Table 15 outlines the sample sizes.

Table 15. Sample sizes for employment regressions

	Number of DLHE survey respondents	
Academic year leaving school	'older style' L3 BTEC extended diploma students	A level students
2005/2006	3,250	23,880
2009/2010	7,520	32,780
2012/2013	6,880	26,165

Note. Values rounded to nearest 5.

As nearly all graduates were in some form of employment or further study, we focussed the first analysis on those who were only in employment (not in combination with further study), and modelled the probability of being in full-time employment, as opposed to part-time employment. This gives some measure of the 'quality' of employment, as full-time work is usually more sought after. Essentially the same method as before was undertaken. Regressions included course type, degree class, tariff, university mission group, and degree subject as predictor variables.

[15] The DLHE dataset only targets university graduates. We did not have data on those who went straight into employment after school.

Table 16 shows the results of the regressions and Figure 17 shows the probability plots for 'typical' graduates (ie the most common university mission group, degree subject, and degree class). As one can see, while there was no significant difference between groups for the 2005/2006 school leaving cohort, 'older style' L3 BTEC students were at a small, but statistically significant, disadvantage compared to A level students with equivalent level 3 outcomes for 2009/2010 and 2012/2013, and this gap also widened to a small degree over time.

Table 16 and Figure 18 also show the probability of being in a 'highly skilled' occupational category (defined as being in a managerial or professional occupation[16]). 'older style' L3 BTEC students again appeared to be at a comparative disadvantage, with this disadvantage increasing to a small degree over time.

Figure 19 (and Table 16) shows that the same also appears to be true in terms of having a salary above £20,000 (the top 13% of graduates).

It is important to remember that these results all control for differences in degree classification (amongst other variables). In other words, 2 students with the same tariff and the *same degree classification* have different likelihoods of gaining higher quality employment, depending on whether they left school with 'older style' L3 BTEC or A level qualifications. In other words, these differences are to a larger extent due to the type of course taken at level 3, rather than the strength of their degrees. Even when controlling for university studies and background characteristics, those that took 'older style' L3 BTECs appear to be at a relative disadvantage compared to those that took A levels.

While these are not necessarily strong metrics for employment 'outcomes', these findings again seem inconsistent with the idea that 'older style' L3 BTEC students overall have increased in ability, as was implied by the large increases in level 3 outcomes over time. If those changes were a genuine reflection of changes in ability, and the standards required to meet those grades had remained the same, then one might expect those students in subsequent years with the same tariff as those in earlier years to have similar chances in employment, and in comparison to A level students. However, this does not seem to have occurred, which may again point towards those rises in level 3 outcomes being due to grade inflation.

[16] These categories come from the Standard Occupational Classification for the UK (SOC2010) as defined by the Office for National Statistics. Managerial occupations include corporate managers and directors, and other managers and proprietors. Professional occupations include science, research, engineering and technology professionals; health professionals; teaching and educational professionals; business, media and public service professionals. For more information, see https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2010

Table 16. Logistic regression: course type and the achievement of some measures of employment 'quality'

Academic year leaving school	Odds ratio	95% confidence interval
Probability of being in full-time employment, rather than part-time employment		
2005/2006	0.87	0.60 – 1.27
2009/2010	0.87†	0.75 – 1.00
2012/2013	0.68**	0.55 – 0.84
Probability of being in a 'highly skilled' occupational category		
2005/2006	0.89	0.56 – 1.42
2009/2010	0.77*	0.62 – 0.97
2012/2013	0.49***	0.41 – 0.60
Probability of having a 'good' salary		
2005/2006	0.72	0.35 – 1.47
2009/2010	0.63***	0.48 – 0.81
2012/2013	0.57***	0.45 – 0.71

Note. Reference category = A levels. †p = .05; *p < .05; **p < .01; ***p < .001

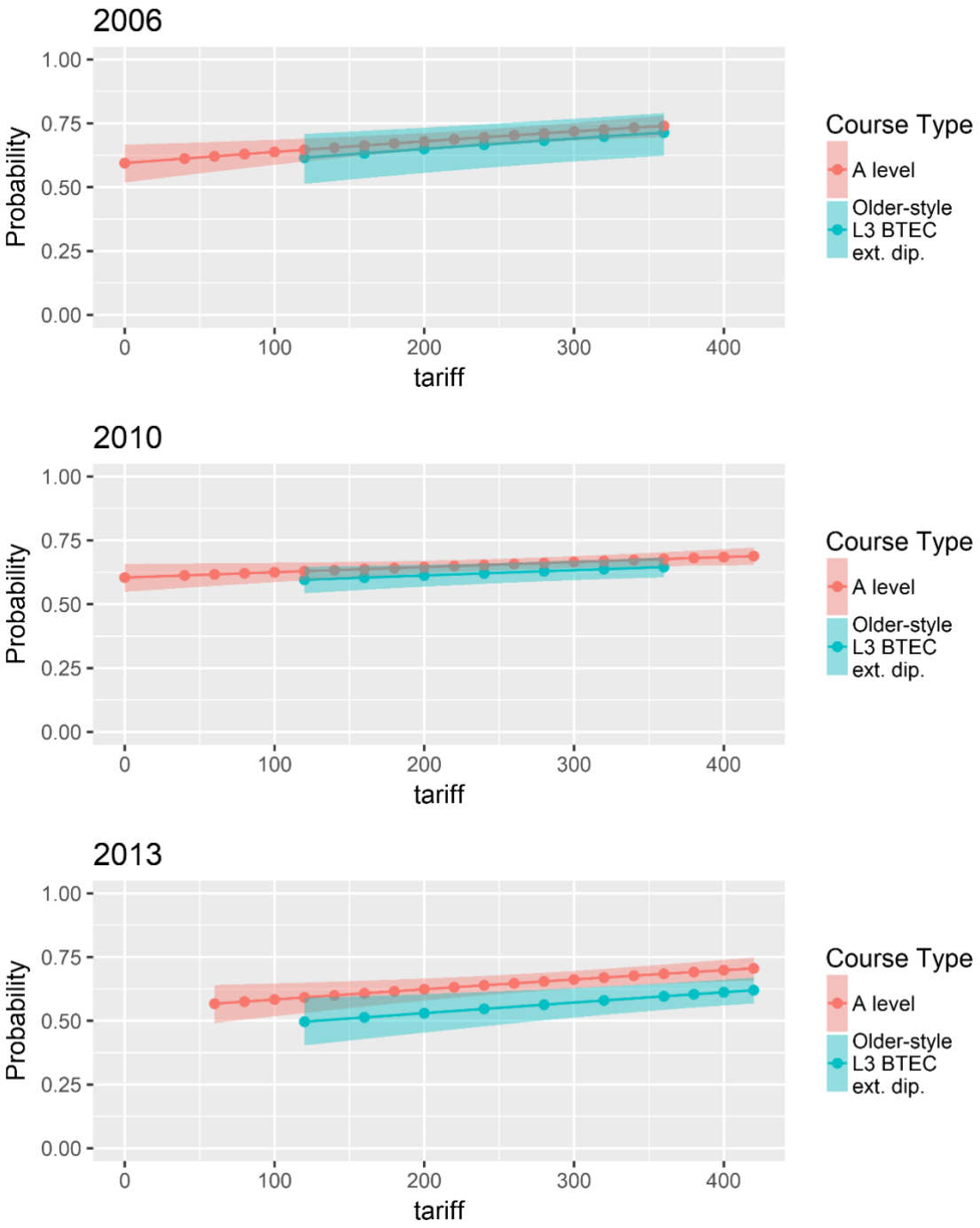


Figure 17. Probability of being in full-time employment, rather than part-time employment, for 'typical graduates'

Note. "Ext. dip." = extended diploma.

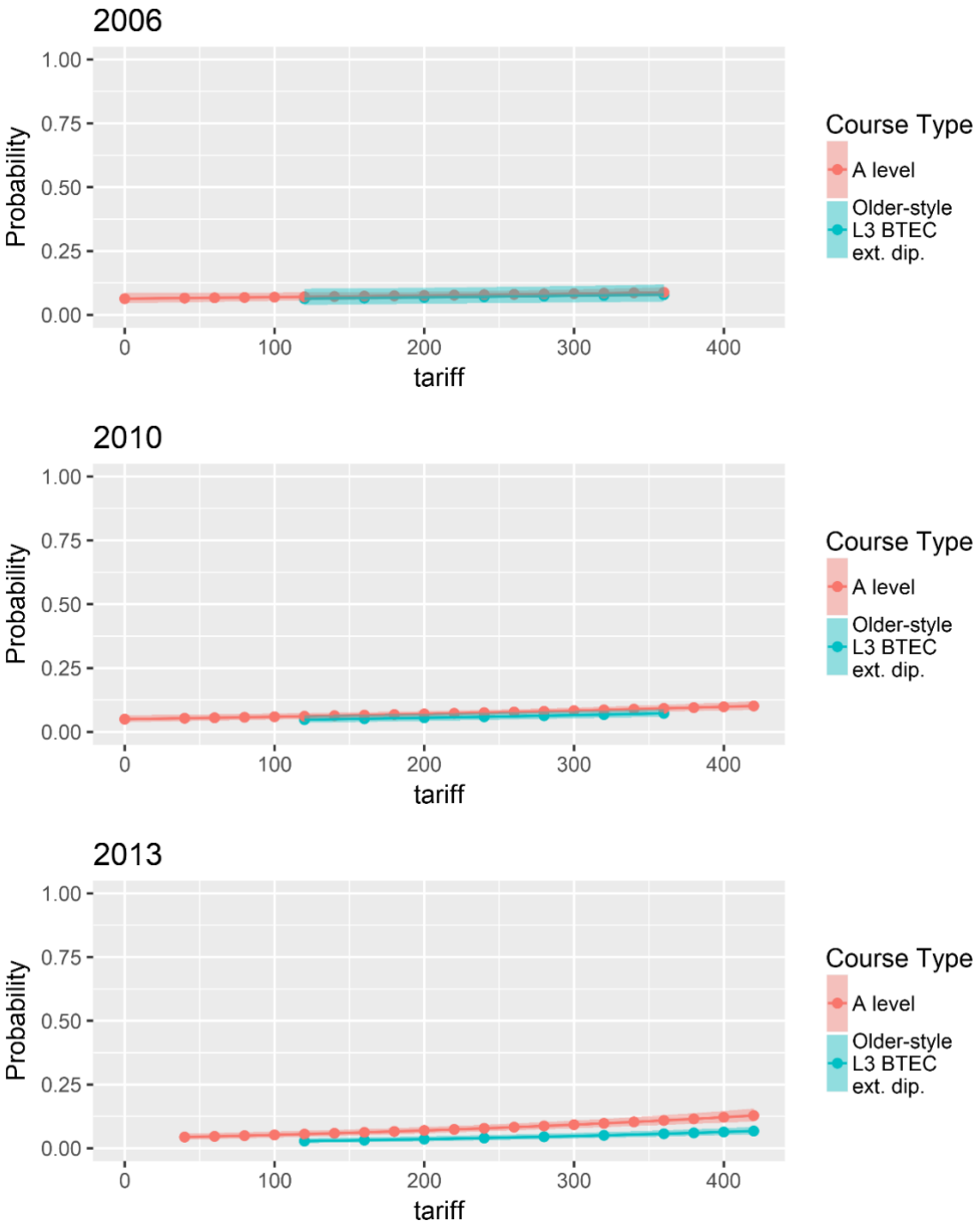


Figure 18. Probability of being in a 'highly skilled' occupation category for 'typical graduates'

Note. The 'highly skilled' occupational category includes managerial and professional occupations (see footnote 16). "Ext. dip." = extended diploma.

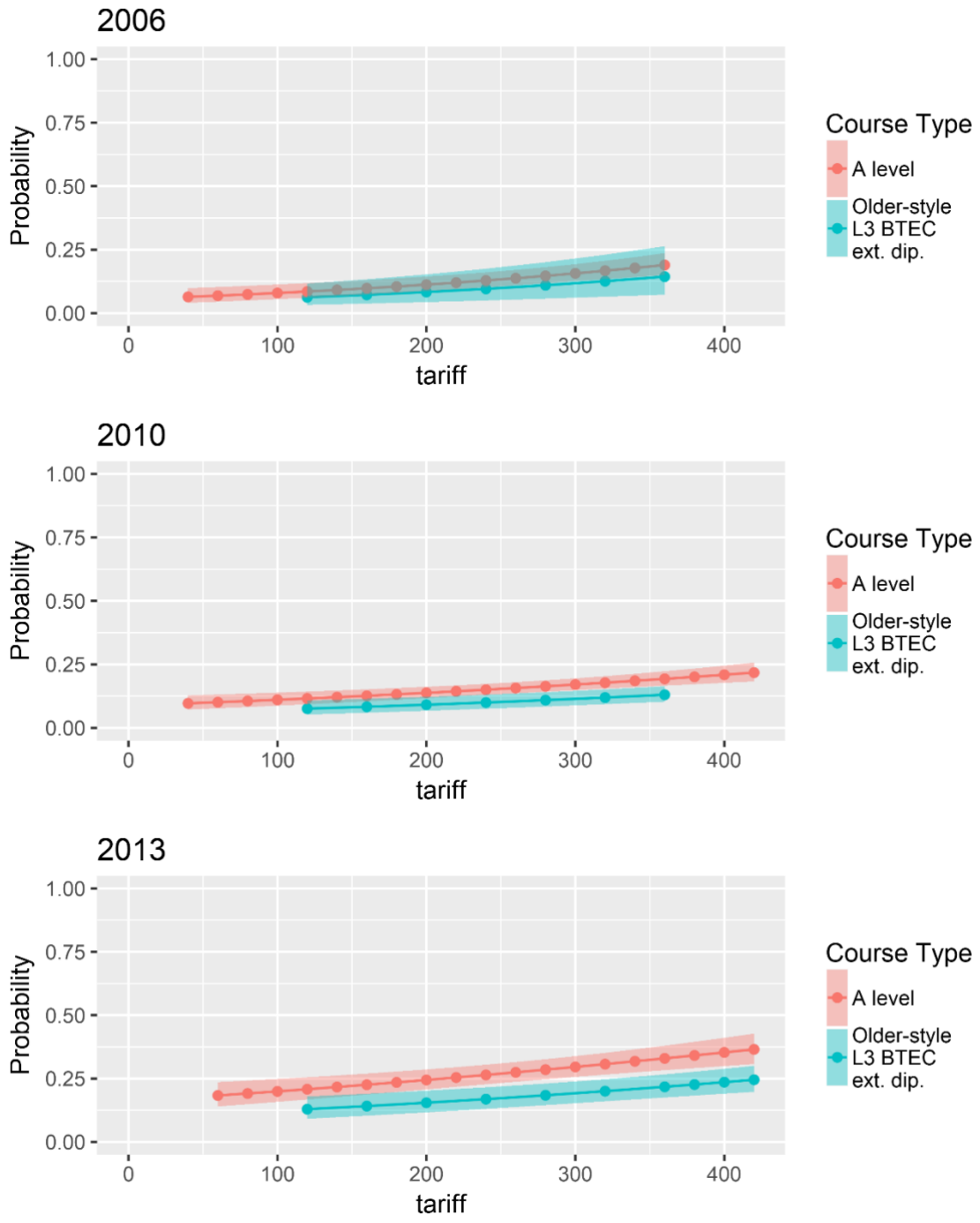


Figure 19. Probability of having a 'good' salary for 'typical graduates'

Note. "Ext. dip." = extended diploma. 'Good' salary is defined as above £20,000. Roughly 13% of the dataset exceeded this threshold. It is worth noting that we have not taken inflation into account here, and so this figure is not equitable over time. However, there is no reason to assume that inflation would affect the two groups differently, and so the comparisons between 'older style' L3 BTEC and A level students should be unaffected by this.

5 Discussion

From the multiple angles through which we have explored this issue, the findings that we have reported all point to the existence of grade inflation in 'older style' level 3 (L3) BTECs: 2006 - 2016. We have shown how the awarding of top grades has risen substantially over the years for these 'older style' L3 BTECs, and that these changes cannot be attributed to a change in the characteristics of the cohorts potentially affecting educational performance (including, but not restricted to, prior attainment). This was also true when we considered students who had taken a mixture of 'older style' L3 BTEC and A level courses. Given that preparation for university is one of the main purposes of 'older style' L3 BTECs, if these increases were legitimate, one might expect that this would be reflected in these students' performance at university. However, the findings of the level 6 analyses showed this not to be the case. Rather, it seems as though 'older style' L3 BTEC students and A level students with the same tariff score (ie apparently equivalent level 3 attainment[17]) are not equally prepared for university studies, and, importantly, this comparative disadvantage has become more pronounced over time. The same can also be said in terms of employment outcomes, as these 'older style' L3 BTEC students seem to be less likely to secure higher quality employment, with this comparative disadvantage increasing over time (although this was to a lesser degree than for the other analyses). Both the findings for the level 6 and employment analyses suggest that 'older style' L3 BTEC outcomes have become devalued over time, offering a progressively lower reflection of student ability at each level of attainment. This indicates there has been grade inflation in these qualifications over the period of analysis.

As with Gill (2015, 2016a, 2017), our results suggest that the UCAS tariff is misaligned between A levels and 'older style' L3 BTECs (ie equivalent achievement at level 3 does not suggest equivalent levels of preparation for university), and that this misalignment has become more pronounced over time. By implication, this means that grades at level 3 are not aligned in the manner implied by the UCAS tariff and described in Table 4 (ie 'older style' L3 BTEC D* with A level A*, 'older style' L3 BTEC D with A level A, etc.). This may not necessarily be problematic, so long as stakeholders do not treat these outcomes as being equivalent. Indeed, the Department for Education actually operates a slightly different system of alignment (DfE, 2018)[18]. However, the finding that alignment appears to be changing over time (due to grade inflation in 'older style' L3 BTECs) is more troublesome. This means that the same grades do not reflect the same level of ability from one year to the next, limiting the usefulness of these qualifications as indicators of student ability for relevant stakeholders (eg university admissions officers and employers).

Commentary does exist about the potential disadvantage 'older style' L3 BTEC students may face at university and in seeking employment. Potential causes include social factors, the academic focus of universities, a lack of information being available on vocational progression routes into university, and different levels of participation in activities to improve employability (eg Al Meselmani et al., 2018;

[17] It is also worth reiterating that Pearson promotes BTECs as having parity with A levels in the UCAS tariff (Pearson, 2018b) (other awarding organisations do the same for their similar qualifications).
[18] For example, in 2018, A levels receive 60 points for an A*, 50 points for an A, 30 points for a C, and 10 points for an E. Equivalent sized BTECs receive 50 points for a D*, 35 points for a D, 25 points for an M, and 15 points for a P.

Bathmaker, Ingram, & Waller, 2013; Budd, 2017; Dunbar-Goddet & Papageorgiou, 2008, cited in J. Carter, 2009). However, while these may explain some of the gap between 'older style' L3 BTEC and A level students in terms of university/employment outcomes, such explanations are less likely to account for the magnitude and duration of the changes over time that were observed. In other words, external factors would have needed to become progressively more disadvantaging to quite a large degree, in order to cause the changes over time that were observed. On balance, and in combination with other analyses presented here, grade inflation of 'older style' L3 BTECs is the more likely explanation for the changes over time that were observed.

As suggested previously, any grade inflation might be explained by the fact that 'older style' L3 BTECs were mostly assessed by non-examination assessments (NEAs) during the period of interest, without the use of a robust methodology to ensure that the qualification standard is maintained over time. While A level awarding is informed by the comparable outcomes approach, operationalised through statistical predictions (see Ofqual, 2015), 'older style' L3 BTEC awarding is not. Given that similar patterns of rising outcomes occurred for A levels prior to the introduction of tighter statistical controls in 2011 (eg see Ofqual, 2015, fig. 3 – at least part of which is likely to have been unwarranted inflation), the fact that grade inflation also appears to have occurred for 'older style' L3 BTECs is perhaps to be expected. The difference in the size of these changes in outcomes are notable, however: the percentage of A level students achieving a grade A rose by about 10% between 1996 and 2010 (Ofqual, 2015), while the percentage achieving a D/D* in 'older style' L3 BTEC subsidiary diplomas rose by 40% between 2006 and 2016 (Figure 2 of the current study). Outcomes for NEAs also tend to be higher compared to written exams, and less well spread out (Ofqual, 2013), meaning that NEA outcomes may give inflated estimates of student ability compared to exams. If school/college accountability/performance cultures lead to over-marking of NEAs (eg as suggested by Ofqual, 2012, 2013), and with few controls to manage over-marking[19], then an increase in pressures on teachers over time may have contributed to the trends observed for 'older style' L3 BTECs. This increase in pressure may have accompanied the increase in popularity of these qualifications over time. Although centres' marks for 'older style' L3 BTECs do undergo a process of verification, this process involves inspecting a sample of work, and relies upon expert judgments of grading accuracy (Pearson, 2015), which may be susceptible to confirmatory bias (ie verifiers might tend towards agreeing with a centre's marks) and 'benefit of the doubt' judgements, meaning that over marking may not always be challenged. Readers are reminded that 'older style' L3 BTECs are not the only qualifications which have these features – other similar qualifications offered by other awarding organisations do operate similar processes, and so may be subject to the same issues.

Nevertheless, while there may be issues associated with NEAs, any move towards further examination-based methods of assessment should be done with careful thought. Carter and Bathmaker (2017) in particular urged caution against ignoring historical concerns raised in the Haslegrave report (Haslegrave, 1969), that exams may not be an appropriate method of assessing vocational skills. It may be possible, for example, to have more robust mechanisms for marking/grading, quality assuring

[19] For these qualifications, centre marks are not numerical as such (eg 65 of 100), but rather just a grade – such as Pass or Merit or Distinction.

and standard maintaining in qualifications which contain a large proportion of internal assessment.

It is worth bearing in mind that recent changes to the vocational qualifications landscape could curtail these trends in future awards. Partly in response to concerns raised by the Wolf Report (Wolf, 2011), level 3 BTEC Nationals have been reformed (with first awards made in 2017). These new BTEC Nationals have been designed with more external assessments and mandatory content (see Pearson, 2017b), which may help to mitigate any potential grade inflation from internal assessments. However, the continued lack of a robust standard maintaining mechanism or awarding methodology (eg statistical controls), and the inability to adjust marks or grades or grade boundaries on the majority of the assessments comprising these qualifications, will continue to make the control of qualification standards challenging.

5.1 Limitations

Some limitations of our analyses should be kept in mind when considering the outcomes of this research.

Firstly, it should be noted that while we have made use of the relationship between 'older style' L3 BTEC outcomes and other variables as proxy measures of student achievement, we have not looked at actual examples of student work as part of this project. As it is not possible to directly attest to whether there has or has not been a change in student performance over time, these analyses do not necessarily provide definitive proof of grade inflation. Nevertheless, the findings presented here do cast doubt on the *plausibility* of the rises in outcomes observed, and whether they represent genuine increases in student achievement. In our view, the findings are consistent with what one would expect to observe under the presence of grade inflation.

Secondly, although we have tried to control for as many confounding variables as we can, it is still possible that some others could affect interpretation of these trends. For example, it is possible that although the prior attainment of 'older style' L3 BTEC cohorts has not changed over time, there might have been greater selectivity of students enrolling on 'older style' L3 BTEC courses, students who are better suited to do well in 'older style' L3 BTECs. However, it seems unlikely greater selectivity is at play here, since this would likely be seen in the context of a diminishing cohort, not an increasing cohort. It is also possible that teaching practices may have greatly improved as these qualifications increased in popularity. However, previous work (on A levels) has shown that improvements in student performance over time due to improved teaching and learning are much smaller, and reached their peak much sooner, than those observed in Figure 4 (see Ofqual, 2016a). In addition, improvements in teaching and learning might be expected to lead to increases in performance at university/employment, which we have not observed here. In fact, given the size of the changes observed in the level 3 analyses (eg a change from 21% in 2005/2006 to 61% in 2015/2016 in the proportion achieving top grades), any legitimate explanation of these rises would presumably have to be substantial enough to be readily recognisable.

Thirdly, while we have reported that 'older style' L3 BTEC students appear to be at a small comparative disadvantage in gaining access to higher quality employment, it is

possible that the changes observed reflect changes in public perceptions of these qualifications over time, rather than the effects of grade inflation. In other words, it may be that 'older style' L3 BTEC students are at a comparative disadvantage in the job market due to increasingly negative public perceptions of these qualifications (and those that hold them). Nevertheless, this would again be unlikely to affect the other analyses presented here (eg university outcomes are unlikely to be affected by public perceptions), and perceptions do often have at least some grounding in reality.

Finally, we were unable to explore university and employment outcomes for the 2015/2016 cohort, because data was not yet available. There is of course a possibility that this cohort will exhibit different patterns of outcomes when the data does become available. However, given the other results reported, and the fact that neither prior attainment nor quality assurance processes have changed during this time for 'older style' L3 BTECs[20], this seems unlikely.

5.2 Conclusions

Each of the analyses presented within this report appears to point to the existence of unwarranted and considerable grade inflation in 'older style' level 3 (L3) BTEC Nationals, between 2006 and 2016. Next steps should include a careful consideration of how such issues might be mitigated in the future, including for the older style L3 BTECs (which are still ongoing), and for other similar qualifications which show grading increases and which operate similar controls or models of assessment, standards maintenance and quality assurance. Continued monitoring of outcomes will be necessary to ensure that standards for these qualifications remain comparable over time, so that they offer valid and reliable indicators of student ability for all stakeholders that make use of them.

[20] It is important to note that steps have been taken to mitigate some of the risks to grade inflation in the newer style L3 BTECs. However, as some features are still shared with the 'older style' L3 BTECs (see Table 1), the potential for such risks should still be monitored.

References

Al Meselmani, M., Baker, Z., Basi, A., Broglia, E., Sutherland-Jarvest, A., & Towers, L. (2018). *Finding Potential: How a selective University can attract and retain high quality students with equivalent qualifications to A levels*. Sheffield, UK: The University of Sheffield. Retrieved from https://www.sheffield.ac.uk/polopoly_fs/1.779993!/file/Finding_Potential-Final_Report_Equivalent_Qualifications_Project.pdf

Allan, T. (2017). *Applied general user research*. (Ofqual Report 17/6217). Coventry, UK: Ofqual. Retrieved from https://www.gov.uk/government/publications/applied-general-user-research

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *46*, 399–424. http://doi.org/10.1080/00273171.2011.568786

Bailey, N., & Bekhradnia, B. (2008). *The Academic Experience and Outcomes of Students with Vocational Level 3 Qualification*. Oxford, UK: Higher Education Policy Institute. Retrieved from http://www.hepi.ac.uk/wp-content/uploads/2014/02/38AcademicexperienceofstudentswithvocationalL3s.pdf

Bathmaker, A.-M., Ingram, N., & Waller, R. (2013). Higher education, social class and the mobilisation of capitals: recognising and playing the game. *British Journal of Sociology of Education*, *34*, 723–743. http://doi.org/10.1080/01425692.2013.816041

Bramley, T. (2014). Multivariate representations of subject difficulty. *Research Matters: A Cambridge Assessment Publication*, *18*, 42–47. Retrieved from http://www.cambridgeassessment.org.uk/Images/174492-research-matters-18-summer-2014.pdf

Brookhart, M. A., Wyss, R., Layton, J. B., & Stürmer, T. (2013). Propensity Score Methods for Confounding Control in Nonexperimental Research. *Circulation: Cardiovascular Quality and Outcomes*, *6*, 604–611. http://doi.org/10.1161/CIRCOUTCOMES.113.000359

Budd, R. (2017). Disadvantaged by degrees? How widening participation students are not only hindered in accessing HE, but also during – and after – university. *Perspectives: Policy and Practice in Higher Education*, *21*, 111–116. http://doi.org/10.1080/13603108.2016.1169230

Cameron, A. C., & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, *50*, 317–372. http://doi.org/10.3368/jhr.50.2.317

Carter, A., & Bathmaker, A.-M. (2017). Prioritising progression over proficiency: limitations of teacher-based assessment within technician-level vocational education. *Journal of Further and Higher Education*, *41*, 460–474. http://doi.org/10.1080/0309877X.2015.1135881

Carter, J. (2009). *Progression from vocational and applied learning to higher education in England*. *Higher Education*. Bolton, UK: University Vocational Awards Council. Retrieved from http://www.uvac.ac.uk/wp-

content/uploads/2013/09/UVAC-BIS-Progression-Report-March-2010.pdf

DCLG. (2015). *The English Indices of Deprivation 2015: Research Report*. London, UK: Department for Communities and Local Government. Retrieved from https://www.gov.uk/government/publications/english-indices-of-deprivation-2015-research-report

DfE. (2018). Performance points for qualifications counting in the 16 to 18 performance tables 2018. Retrieved November 14, 2017, from https://www.gov.uk/government/publications/16-to-19-qualifications-discount-codes-and-point-scores

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*, 802–813. http://doi.org/10.1111/j.1365-2656.2008.01390.x

Gill, T. (2013). *Provision of level 3 qualifications in English schools 2008-2012: Statistics Report Series No. 58*. Cambridge, UK: Cambridge Assessment. Retrieved from http://www.cambridgeassessment.org.uk/Images/150206-provision-of-level-3-qualifications-in-english-schools-2008-2012.pdf

Gill, T. (2015). Using generalised boosting models to evaluate the UCAS tariff. *Research Matters: A Cambridge Assessment Publication*, *20*, 2–6. Retrieved from http://www.cambridgeassessment.org.uk/Images/255867-research-matters-20-summer-2015.pdf

Gill, T. (2016a). Assessing the equivalencies of the UCAS tariff for different qualifications. *Research Matters: A Cambridge Assessment Publication*, *21*, 16–23. Retrieved from http://www.cambridgeassessment.org.uk/Images/294145-research-matters-21-winter-2016.pdf

Gill, T. (2016b). *Provision of level 3 qualifications in English schools 2015: Statistics Report Series No. 104*. Cambridge, UK: Cambridge Assessment. Retrieved from http://www.cambridgeassessment.org.uk/our-research/all-published-resources/statistical-reports/

Gill, T. (2016c). *Uptake of level 3 qualifications in English schools: Statistics Report Series No. 105*. Cambridge, UK: Cambridge Assessment. Retrieved from http://www.cambridgeassessment.org.uk/our-research/all-published-resources/statistical-reports/

Gill, T. (2017). Preparing students for university study: a statistical comparison of different post-16 qualifications. *Research Papers in Education*, *1522*, 1–19. http://doi.org/10.1080/02671522.2017.1302498

Green, F., & Vignoles, A. (2012). An empirical method for deriving grade equivalence for university entrance qualifications: an application to A levels and the International Baccalaureate. *Oxford Review of Education*, *38*, 473–491. http://doi.org/10.1080/03054985.2012.713858

Haslegrave, H. L. (1969). *Report of the Committee on Technician Courses and Examinations*. London, UK: Department of Education and Science. Retrieved from https://archive.org/details/op1269366-1001

HEFCE. (2015). *Young participation in higher education*. Bristol, UK: Higher Education Council for England. Retrieved from http://www.hefce.ac.uk/pubs/year/2015/201503/

Lee, B. K., Lessler, J., & Stuart, E. A. (2009). Improving propensity score weighting using machine learning. *Statistics in Medicine*, *29*, 337–346. http://doi.org/10.1002/sim.3782

London Economics. (2013). *The outcomes associated with the BTEC route of degree level acquisition*. London, UK: London Economics. Retrieved from http://londoneconomics.co.uk/blog/publication/the-outcomes-associated-with-the-btec-route-of-degree-level-acquisition/

Macmillan, L., Tyler, C., & Vignoles, A. (2015). Who Gets the Top Jobs? The Role of Family Background and Networks in Recent Graduates' Access to High-status Professions. *Journal of Social Policy*, *44*, 487–515. http://doi.org/10.1017/S0047279414000634

Nelson, F. (2016a). If Theresa May was serious about helping those "just managing", she'd focus on Btecs, not grammar schools. Retrieved September 23, 2016, from http://www.telegraph.co.uk/news/2016/09/15/if-theresa-may-was-serious-about-helping-those-just-managing-she/

Nelson, F. (2016b). The problem with Btecs – a response to Pearson Plc. Retrieved September 23, 2016, from http://blogs.spectator.co.uk/2016/09/problem-btecs-response-pearson-plc/

Ofqual. (2012). *GCSE English 2012*. (Ofqual Report 12/5225). Coventry, UK: Office of Qualifications and Examinations Regulation. Retrieved from http://webarchive.nationalarchives.gov.uk/20141031163546/http://www.ofqual.gov.uk/files/2012-11-02-gcse-english-final-report-and-appendices.pdf

Ofqual. (2013). *Review of controlled assessment in GCSEs*. (Ofqual Report 13/5291). Coventry, UK: Office of Qualifications and Examinations Regulation. Retrieved from https://www.gov.uk/government/publications/review-of-controlled-assessment-in-gcses

Ofqual. (2015). *Setting GCSE, AS and A level grade standards in summer 2014 and 2015*. Coventry, UK: Office of Qualifications and Examinations Regulation. Retrieved from https://www.gov.uk/government/publications/setting-gcse-and-a-level-grade-standards-in-summer-2014-and-2015

Ofqual. (2016a). *An investigation into the "Sawtooth Effect" in GCSE and AS / A level assessments*. Coventry, UK: Office of Qualifications and Examinations Regulation. Retrieved from https://www.gov.uk/government/publications/investigation-into-the-sawtooth-effect-in-gcses-as-and-a-levels

Ofqual. (2016b). *Perceptions of A Levels, GCSEs and Other Qualifications in England – Wave 14*. (Ofqual Report 16/6040). Coventry, UK: Office of Qualifications and Examinations Regulation. Retrieved from https://www.gov.uk/government/statistics/perception-of-a-levels-and-gcses-wave-14

Ofqual. (2017). Regulating GCSEs, AS and A levels: guide for schools and colleges. Retrieved December 1, 2017, from https://www.gov.uk/guidance/regulating-gcses-as-and-a-levels-guide-for-schools-and-colleges/grading

Ofqual. (2018). *Annual Qualifications Market Report: 2016 to 2017 academic year*. (Ofqual Report 18/6350/1). Coventry, UK: Office of Qualifications and

Examinations Regulation. Retrieved from https://www.gov.uk/government/statistics/annual-qualifications-market-report-academic-year-2016-to-2017

Pearson. (2015). *BTEC international quality assurance handbook 2015-16*. London, UK: Pearson. Retrieved from http://qualifications.pearson.com/en/support/support-topics/quality-assurance/btec-quality-assurance-handbook.html#tab-InternationalcentresoutsidetheUK

Pearson. (2017a). BTEC Nationals. Retrieved November 13, 2017, from https://qualifications.pearson.com/en/qualifications/btec-nationals.html

Pearson. (2017b). New BTEC Nationals from 2016 to 2018. Retrieved December 1, 2017, from https://qualifications.pearson.com/en/qualifications/btec-nationals/btec-nationals-2016.html

Pearson. (2018a). BTEC assessment and verification tools. Retrieved January 10, 2018, from https://qualifications.pearson.com/en/support/support-topics/assessment-and-verification/btec-assessment-and-verification-tools.html

Pearson. (2018b). Comparing BTEC to other qualifications. Retrieved April 11, 2018, from https://qualifications.pearson.com/en/support/support-topics/understanding-our-qualifications/comparing-btec-to-other-qualifications/btec-ucas-points.html#tab-NewBTECNationalsRQF

Ridgeway, G., McCaffrey, D., Morral, A., Burgette, L., & Griffin, B. A. (2017). *Toolkit for Weighting and Analysis of Nonequivalent Groups: A tutorial for the twang package*. Retrieved from https://cran.r-project.org/web/packages/twang/vignettes/twang.pdf

Ridgeway, G., McCaffrey, D., Morral, A., Griffin, B. A., & Burgette, L. (2017). Package "twang." Retrieved from https://cran.r-project.org/web/packages/twang/twang.pdf

Robins, J. M., Hernán, M. Á., & Brumback, B. (2000). Marginal Structural Models and Causal Inference in Epidemiology. *Epidemiology*, *11*, 550–560. http://doi.org/10.1097/00001648-200009000-00011

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*, 41–55. http://doi.org/10.1093/biomet/70.1.41

Russell Group. (2017). *Informed Choices: A Russell Group guide to making decisions about post-16 education - Fifth edition*. Retrieved from http://russellgroup.ac.uk/for-students/school-and-college-in-the-uk/subject-choices-at-school-and-college/

Thomson, D. (2017). The equivalence of A-Levels and BTECs. Retrieved November 14, 2017, from https://educationdatalab.org.uk/2017/02/the-equivalence-of-a-levels-and-btecs/

Thomson, D. (2018). Long-term outcomes: How did life turn out for those who took Level 3 BTEC qualifications? Retrieved November 30, 2018, from https://ffteducationdatalab.org.uk/2018/11/long-term-outcomes-how-did-life-turn-out-for-those-who-took-level-3-btec-qualifications/

Thomson, D., & Keshwani, A. (2017). *Post-16 Qualification and Subject Mix- a Report for The Royal Society*. Retrieved from https://royalsociety.org/~/media/policy/topics/education-skills/Broadening the curriculum/education-datalab-report-post-16-qualifcations-and-subject-mix-05-09-2017.pdf

UCAS. (2003). *Expert group report for awards seeking admission to the UCAS tariff: BTEC national awards; BTEC national certificates; BTEC national diplomas*. Cheltenham, UK: UCAS. Retrieved from https://www.ucas.com/sites/default/files/btec-nats-report.pdf

UCAS. (2016). *UCAS Tariff tables: New Tariff points for entry to higher education from 2017*. Cheltenham, UK: UCAS. Retrieved from https://www.ucas.com/advisers/guides-and-resources/tariff-2017

UCAS. (2017). Tariff tables. Retrieved December 1, 2017, from https://www.ucas.com/advisers/guides-and-resources/information-new-ucas-tariff-advisers

Vidal Rodeiro, C. L., Sutch, T., & Zanini, N. (2013). *Prior learning of undergraduates in UK higher education institutions*. Cambridge, UK: Cambridge Assessment. Retrieved from http://www.cambridgeassessment.org.uk/Images/230030-prior-learning-of-undergraduates-in-uk-higher-education-institutions.pdf

Wolf, A. (2011). *Review of vocational education: The Wolf report*. London, UK: Department for Education. Retrieved from https://www.gov.uk/government/publications/review-of-vocational-education-the-wolf-report

Xu, S., Ross, C., Raebel, M. A., Shetterly, S., Blanchette, C., & Smith, D. (2010). Use of Stabilized Inverse Propensity Scores as Weights to Directly Estimate Relative Risk and Its Confidence Intervals. *Value in Health*, *13*, 273–277. http://doi.org/10.1111/j.1524-4733.2009.00671.x

Appendix A – Other similar qualifications, and their grade distributions over time

Level 3 qualifications are offered by other awarding organisations which are similar to level 3 'older style' BTECs in terms of structure and methods of standards maintaining. While this report has focussed on 'older style' L3 BTECs, which have the largest market share, it is important to note that any issues raised in this report may also apply to these other similar qualifications with similar controls over standard maintaining.

The figures on the following pages show the grade distributions over time for a number of these qualifications; the table shows the total number of grades that were awarded each year. Note that only qualifications types graded above a pass are included here.

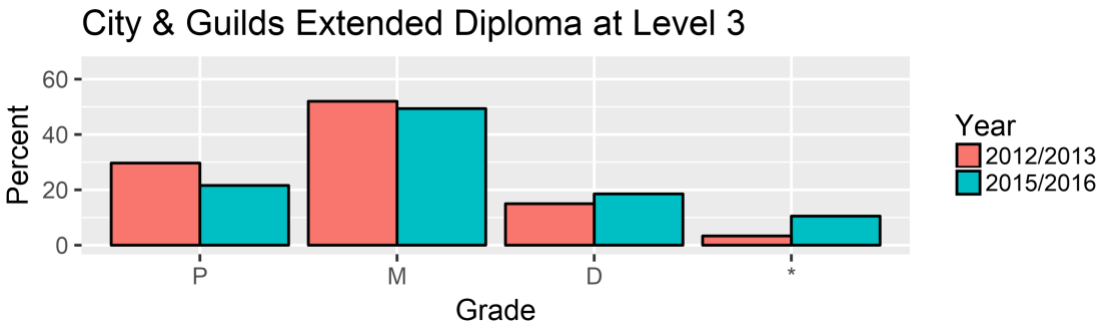
The patterns observed are mostly similar to those reported for 'older style' L3 BTECs, shown in Figure 2 in the main body of this report. So as to avoid showing large changes in distributions due to small entry sizes in early years, qualification/year combinations are only plotted where there were more than 100 entries in that qualification and year. NB data is sourced from the National Pupil Database and may not reflect outcomes for those students whose data is not recorded in this database.

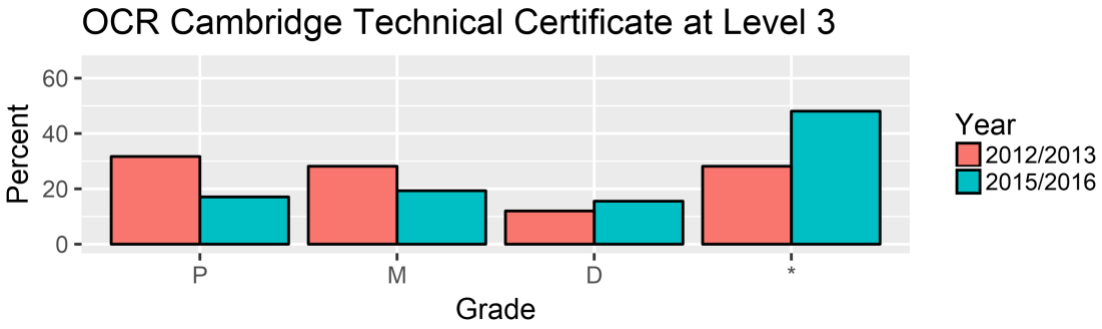
Table A.1. Number of grades awarded per year for the 'older style' L3 BTECs, and each of the other similar qualification types

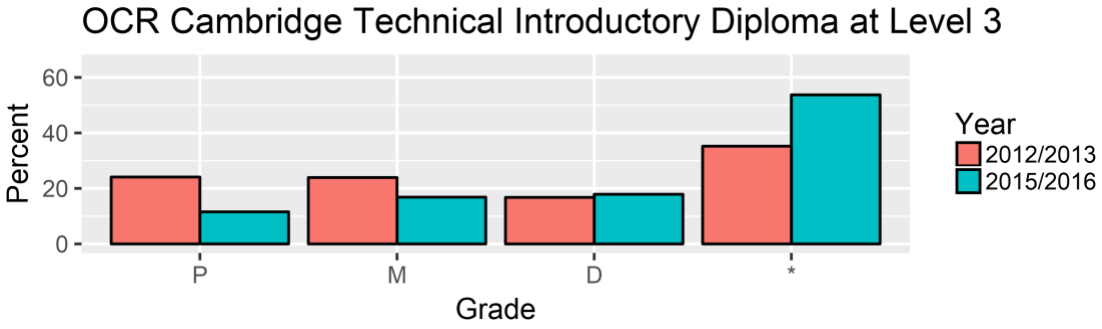
		Number of grades awarded			
Qualification type	Equivalence in size to 1 A level	2005/2006	2009/2010	2012/2013	2015/2016
'Older style' level 3 BTEC combined figures	-	53,490	141,935	200,985	124,720
- 'Older style' level 3 BTEC subsidiary diploma	1	9,600	50,765	88,125	63,410
- 'Older style' level 3 BTEC diploma	2	9,465	27,200	36,110	17,985
- 'Older style' level 3 BTEC extended diploma	3	34,425	63,970	76,750	43,325
City & Guilds Extended Diploma at Level 3	3	0	0	1,050	1,930
OCR Cambridge Technical combined figures	-	0	0	1,835	19,815
- OCR Cambridge Technical Certificate at Level 3	0.5	0	0	1,135	5,605
- OCR Cambridge Technical Diploma at Level 3	2	0	0	15	1,930
- OCR Cambridge Technical Extended Diploma at Level 3	3	0	0	0	1,010
- OCR Cambridge Technical Introductory Diploma at Level 3	1	0	0	530	9,290
- OCR Cambridge Technical Subsidiary Diploma at Level 3	1.5	0	0	160	1,975
OCR National combined figures	-	290	7,530	11,430	0
- OCR National Certificate at Level 3	1	165	4,605	8,375	0
- OCR National Diploma at Level 3	2	75	2,215	2,460	0
- OCR National Extended Diploma at Level 3	3	50	715	595	0
Rock School Limited (RSL) combined figures	-	9	50	475	1,575
- RSL VRQ Level 3	0.3	0	0	65	0

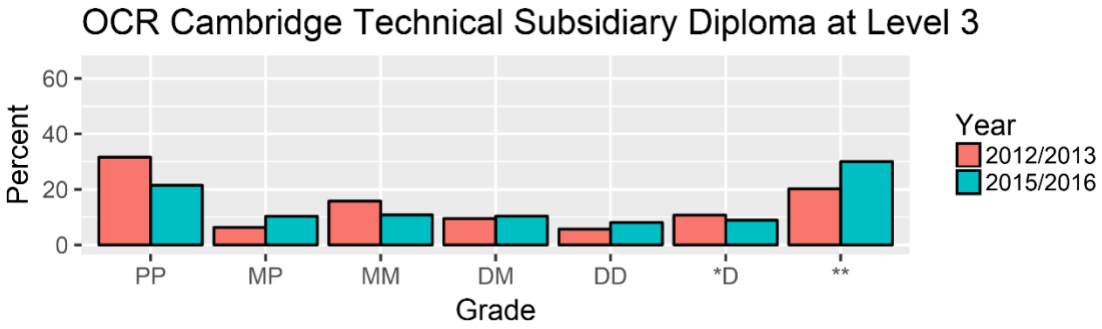
- RSL VRQ Level 3	0.5	0	0	45	0
- RSL VRQ Level 3	0.75	0	0	75	0
- RSL VRQ Level 3	1.25	10	0	155	0
- RSL VRQ Level 3	1.5	0	0	50	1,210
- RSL VRQ Level 3	2	0	0	15	10
- RSL VRQ Level 3	2.75	0	50	20	0
- RSL VRQ Level 3	3	0	0	50	355
University of the Arts London (UAL) combined figures	-	0	60	1,995	9,060
- UAL VRQ Level 3	1.5	0	0	210	4,530
- UAL VRQ Level 3	1.75	0	60	1,775	1,910
- UAL VRQ Level 3	3	0	0	15	2,620

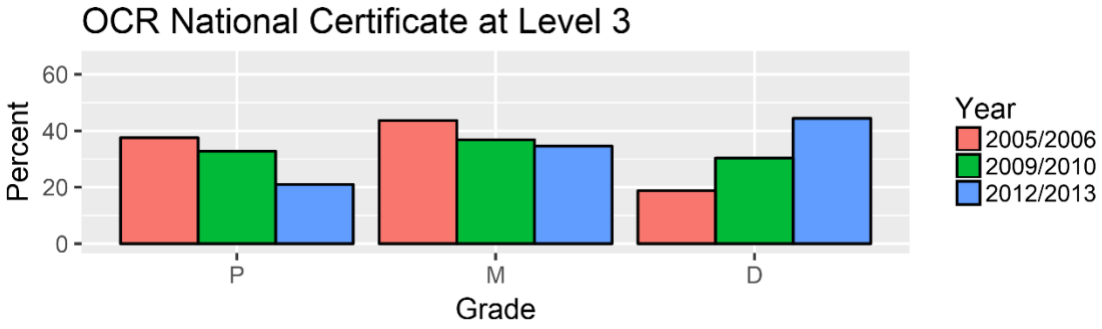
Note. Data only relates to entries present within the NPD dataset. Values have been rounded to the nearest 5.

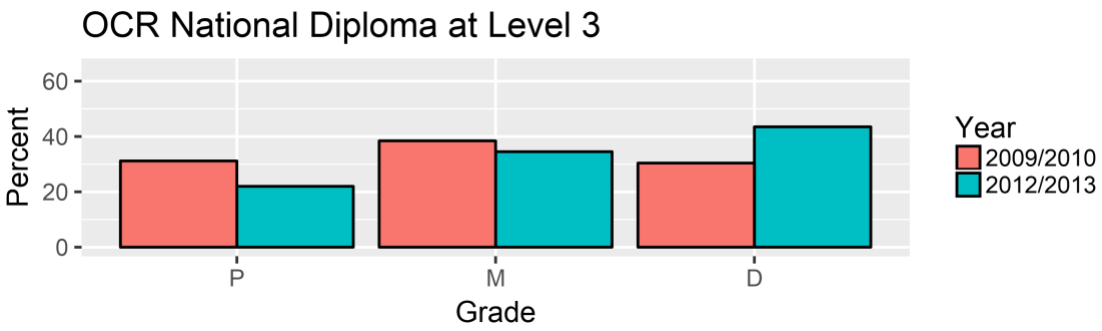

City & Guilds Extended Diploma at Level 3
Percent
60
40
20
0
P
M
D
*
Grade
Year
2012/2013
2015/2016

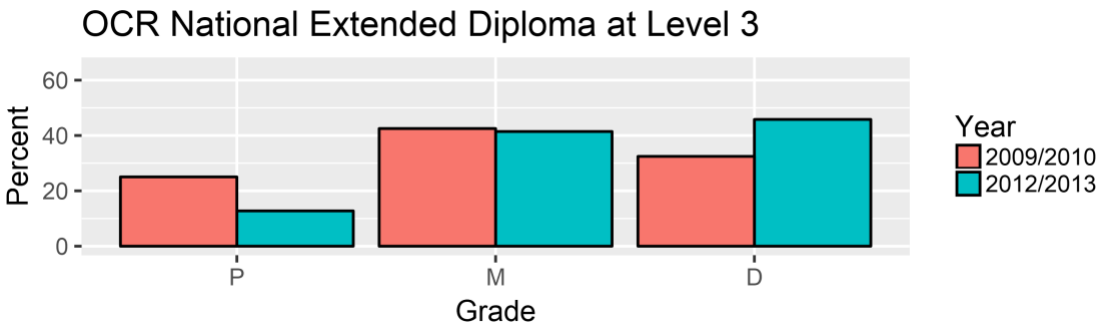

OCR Cambridge Technical Certificate at Level 3
Percent
60
40
20
0
P
M
D
*
Grade
Year
2012/2013
2015/2016

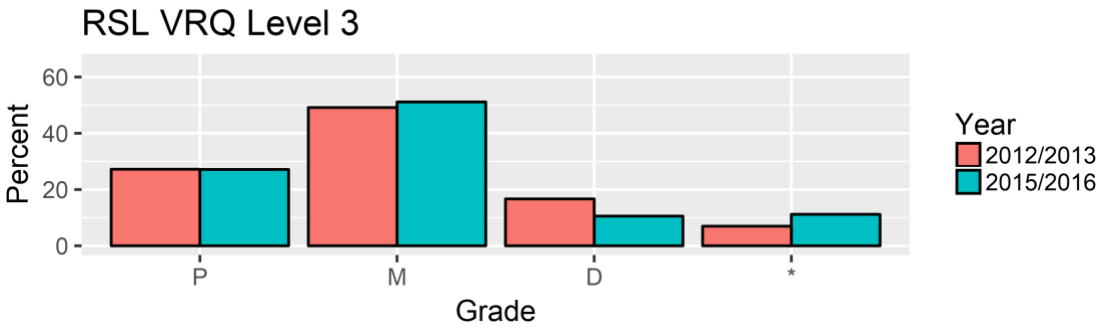

OCR Cambridge Technical Introductory Diploma at Level 3
Percent
60
40
20
0
P
M
D
*
Grade
Year
2012/2013
2015/2016

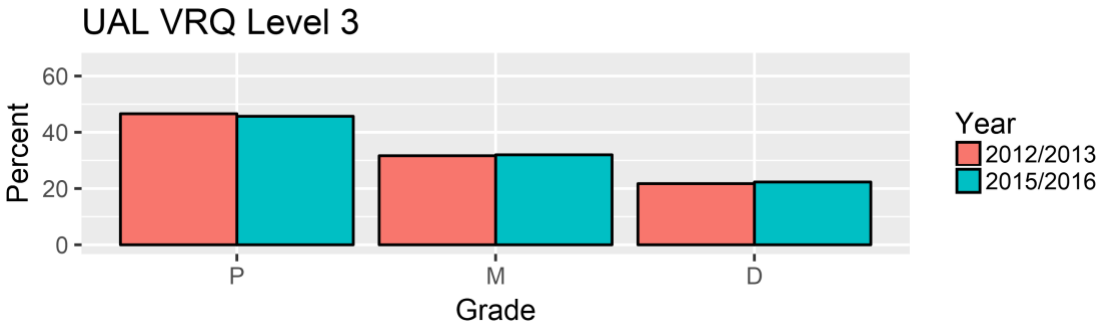

OCR Cambridge Technical Subsidiary Diploma at Level 3
Percent
60
40
20
0
PP
MP
MM
DM
DD
*D
**
Grade
Year
2012/2013
2015/2016


OCR National Certificate at Level 3
Percent
60
40
20
0
P
M
D
Grade
Year
2005/2006
2009/2010
2012/2013


OCR National Diploma at Level 3
Percent
60
40
20
0
P
M
D
Grade
Year
2009/2010
2012/2013


OCR National Extended Diploma at Level 3
Percent
60
40
20
0
P
M
D
Grade
Year
2009/2010
2012/2013


RSL VRQ Level 3
Percent
60
40
20
0
P
M
D
*
Grade
Year
2012/2013
2015/2016

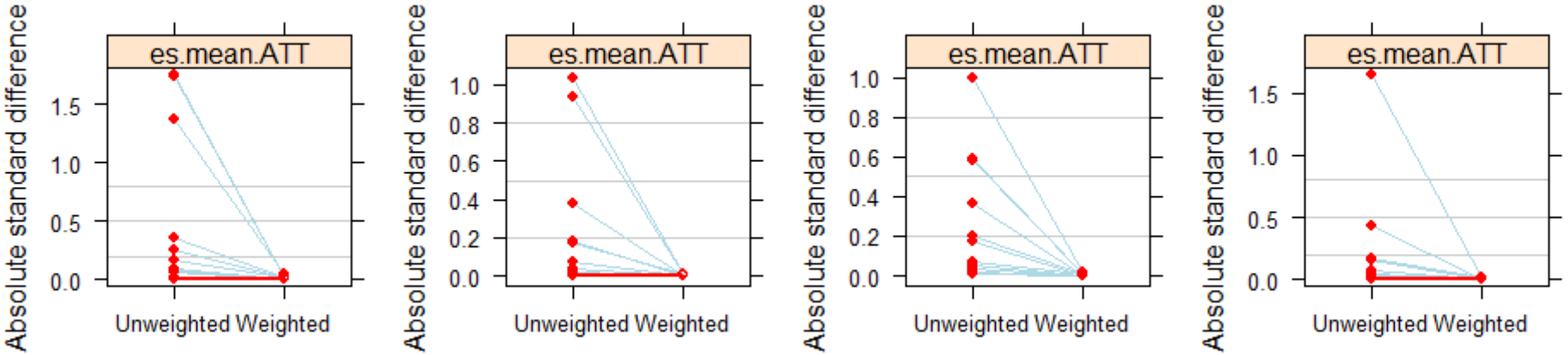

UAL VRQ Level 3
Percent
60
40
20
0
P
M
D
Grade
Year
2012/2013
2015/2016

Appendix B – Assessments of balance for propensity score weighting

The following graphs show the extent to which the propensity score weighting methodology was successful. Each point represents the size of the difference between ‘older style’ L3 BTEC and A level students on each variable included in the propensity score model. The points to the left of each graph represent differences before weighting, and the points to the right represent differences after weighting. The idea of propensity score weighting is that the differences after weighting should all be close to zero. Thin blue lines show where the difference has decreased after balancing; thicker red lines show where the difference has increased after balancing. While the latter is not desirable, all such changes here were negligible.

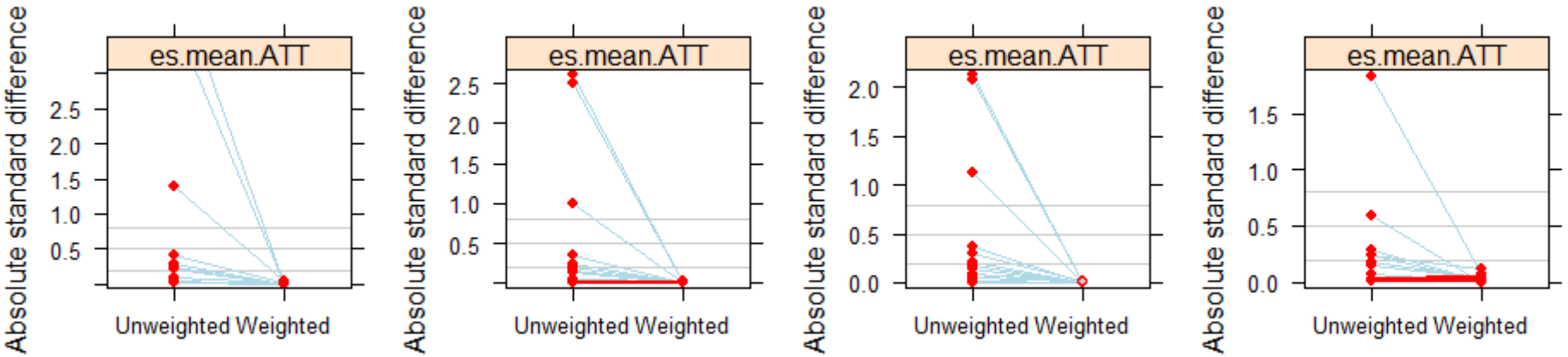
Level 3 – 'older style' L3 BTEC subsidiary diploma vs. A levels

2005/2006 2009/2010 2012/2013 2015/2016



Level 3 – 'older style' L3 BTEC extended diploma vs. A levels

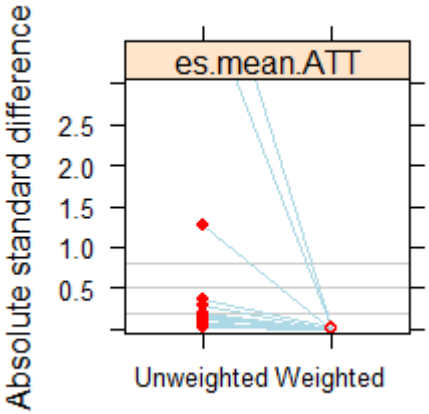
2005/2006 2009/2010 2012/2013 2015/2016



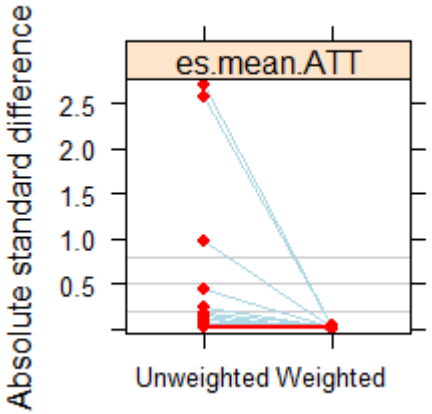
Note. Standardised differences of greater than 3 are not plotted (for unweighted differences only).

Level 6 – Main analyses

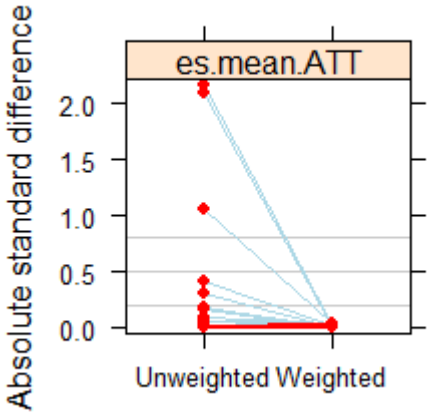
2005/2006



2009/2010



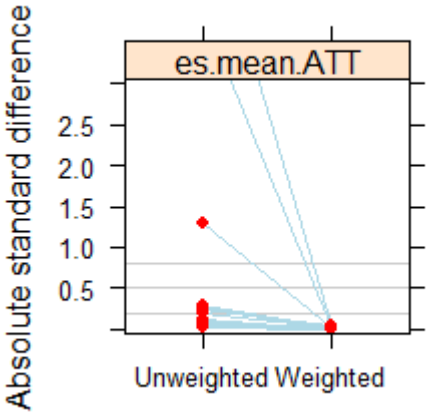
2012/2013



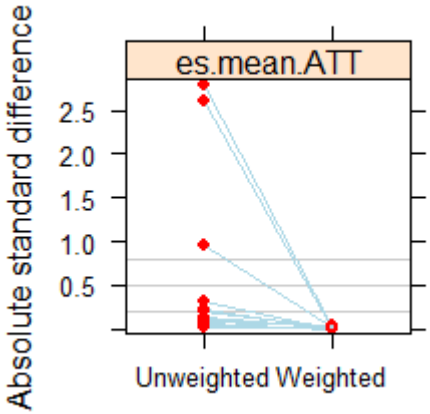
Note. Standardised differences of greater than 3 are not plotted (for unweighted differences only).

Level 6 – Same subject area analyses

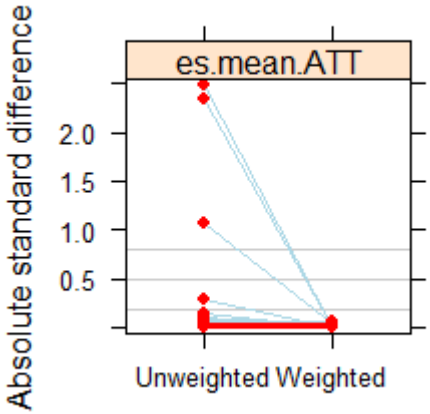
2005/2006



2009/2010



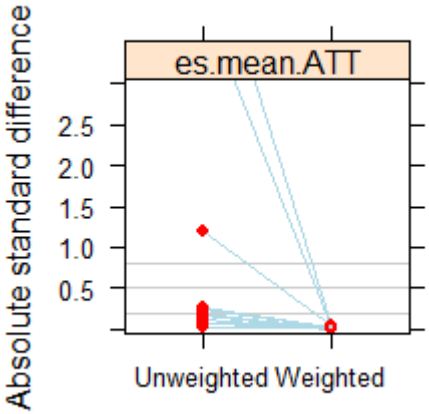
2012/2013



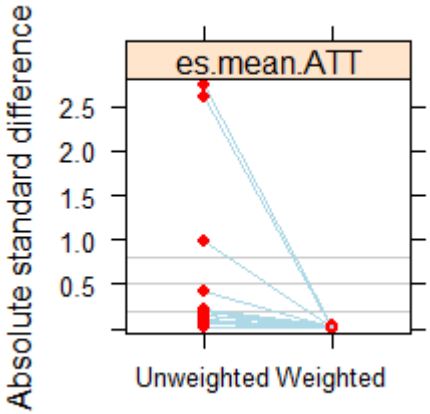
Note. Standardised differences of greater than 3 are not plotted (for unweighted differences only).

Employment – Employed full or part time

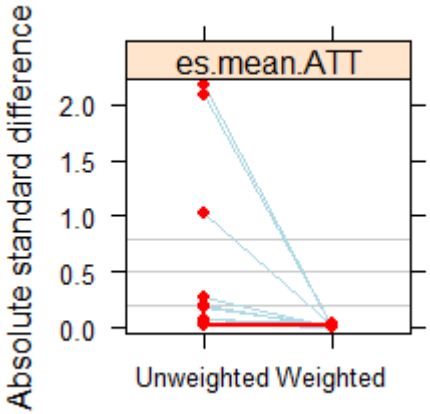
2005/2006



2009/2010



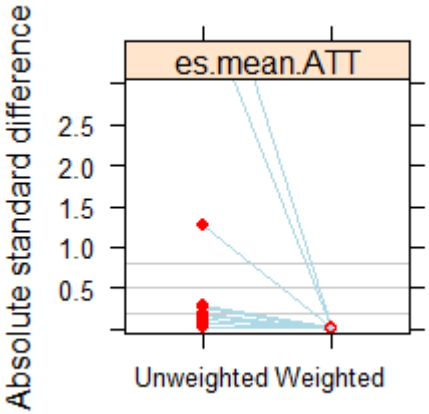
2012/2013



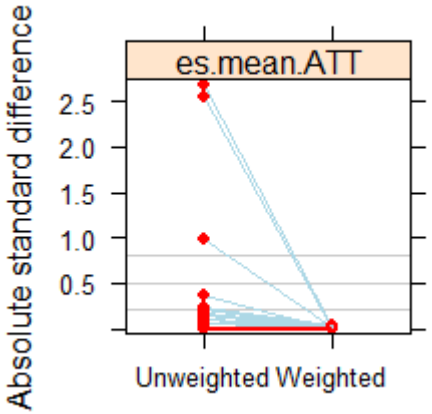
Note. Standardised differences of greater than 3 are not plotted (for unweighted differences only).

Employment – Occupational category

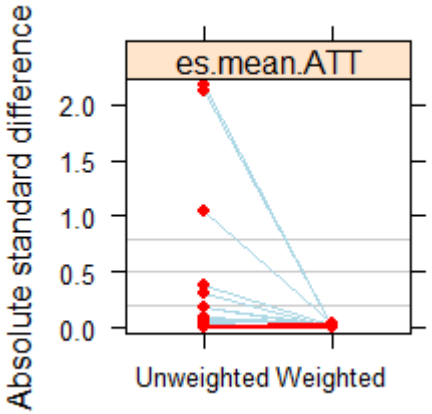
2005/2006



2009/2010



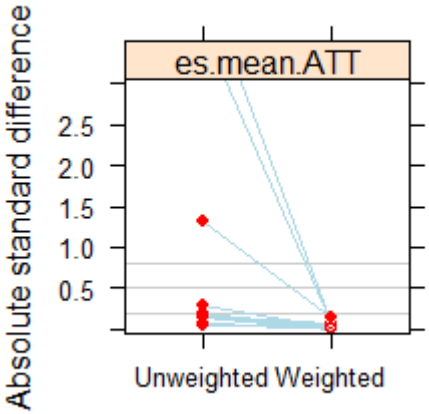
2012/2013



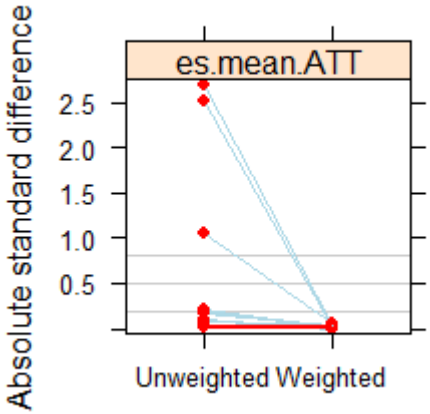
Note. Standardised differences of greater than 3 are not plotted (for unweighted differences only).

Employment – Salary

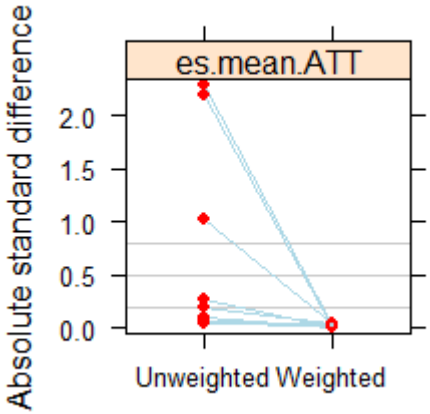
2005/2006



2009/2010



2012/2013



Note. Standardised differences of greater than 3 are not plotted (for unweighted differences only)

Published by:

ofqual

Earlsdon Park
53-55 Butts Road
Coventry
CV1 3BH

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

December 2018 **Ofqual/18/6459/2**