Department for Business Innovation & Skills

RESEARCH

BIS RESEARCH PAPER NUMBER 195

Estimation of the labour market returns to qualifications gained in English Further Education

DECEMBER 2014

David Bibby, Franz Buscha, Augusto Cerqua, Dave Thomson and Peter Urwin

Contents

List of Tables

List of Figures

Executive Summary

In this paper we update the findings from Buscha and Urwin (2013), estimating separately the (i) earnings, (ii) employment probability and (iii) probability of being on active benefits, for those who achieve their highest learning aim whilst studying at an English Further Education Institution (FEI), relative to those who have the same highest learning aim, but do not achieve. Estimates are obtained using the 2002-2012 ILR-WPLS administrative dataset, and the findings provide a robust estimate of the value added from each qualification, when held as an individual's highest.

For each learner we have 7 academic years (2004/2005 to 2010/2011) when they can possibly exit learning (as an achiever or non-achiever) and over these academic years, a learner can have multiple ILR learning spells. Across all of the learning spells for each individual we select the highest learning aim. The estimate of value added is obtained by comparing the returns of those who have a particular highest learning aim (for instance Full Level 2) and achieve; with the returns of those who have the same highest learning aim, but do not achieve.

We consider the question of how long any labour market returns persist beyond qualification, with a study of returns to different cohorts of learners up to six years after the end of learning. The findings influence our decision to report estimated earnings, employment and active benefit returns for the first, third, fourth and fifth tax year after learning. Returns in the first year are an important benchmark for comparison across studies and the 3 to 5 year average is chosen because it balances an implicit trade-off; as we require estimated premiums that (i) persist sufficiently far into the future, but (ii) are not over-reliant on a small number of cohorts that completed learning many years ago.

Updated Estimates of Value Added

Aggregated qualification categories

Earnings: *We find that all qualification categories provide positive and statistically significant earnings premiums, but returns are particularly high for 'Full Level' qualifications.* The 3 to 5 year average returns are 11% for Full Level 2; 9% for Full Level 3 and 8% for Level 4+. We observe smaller (but still statistically significant) earnings premiums of 2% for those achieving a highest qualification Below Level 2; 1% for Level 2 and 3% for Level 3. These 3 to 5 year averages are a good reflection of returns that remain relatively stable over the period of analysis; with the exception of Level 3 and Full Level 3. The possible confounding effects of unobservable moves to HE mean our estimates of 3% for L3 and 9% for FL3 understate the true return to these qualifications. To clarify:

- When considering estimated L3 and FL3 earnings returns, analysis that does not control for movements into HE will find low/negative earnings returns in the first 3 years after learning.
- The available data allow us to control for movements into HE to some extent in the present study.

- However, the previous report (Buscha and Urwin, 2013) contains a separate piece of analysis restricted to a subset of learners for whom we hold HE data, and we suggest that this is taken as the more accurate estimate of impact. In this case the premium for L3 achievers is estimated at 6% and for FL3 it is 14%.

Employment: ***All qualification categories from L2 and above provide positive and statistically significant employment premiums, but again returns are particularly high for 'Full Level' qualifications.*** The highest 3 to 5 year average return accrues to Full Level 3 achievers, who have employment rates that are 4 percentage points higher than non-achievers. Those achieving a FL2 qualification have employment rates 2 percentage points higher than the comparable non-achievers. These FL2 employment premiums seem higher for post-2007 cohorts of learners and this figure of 2 percentage points may be slightly higher for cohorts achieving in the present environment. Similarly, the employment premium for FL3 has become more pronounced for post-2007 cohorts and may now be closer to 5 percentage points for contemporaneous cohorts. The 3 to 5 year average for both L2 and L3 learners is 1 percentage point. The figure of 1 percentage point for L2 is a little better than it seems, because L2 achievers and non-achievers have lower absolute levels of employment (and therefore small percentage point differences amount to larger percentage differences). However, even low levels of employment amongst learners whose highest achievement is Below Level 2, does not alter the suggestion that the size of any employment effect is very small.

Job-search or 'Active' Benefits: ***In all qualification categories from L2 upwards, achievers have a significantly lower probability of being on active benefits than non-achievers, and the gap is more pronounced for 'Full Level' qualifications.*** Our category of Active Benefits associated with job-search activity includes Job Seekers Allowance (JSA), Job Training Allowance (JTA) and Employment and Support Allowance (ESA). Because the absolute proportion of individuals on benefits is small, our percentage point differences can be misleading. For instance, the 3 to 5 year average for Full Level 2 learners suggests that achievers are 2 percentage points less likely to be observed on active benefits, but because only about 10% of non-achievers and 5% of achievers are on benefits after learning, this 2 percentage point difference translates into an approximate 28% difference. The difference in active benefit probabilities for Level 2 learners is 1 percentage point (approximately 11%); 1 percentage point for Level 3 achievers over non-achievers (approximately 26%) and for Full Level 3 achievers we observe a 2 percentage point difference (approximately 40%). For those achieving a highest qualification Below Level 2 we once again observe poor returns, with our percentage point estimate only translating to a 5% lower probability of achievers being on active benefits compared to non-achievers following learning.

Fig. 1: Summary of three to five year averages for aggregated qualification categories

	Earnings Returns	Employment Probability Premiums	Benefit Probability Gaps
Below Level 2	2%	0 pp	0 pp
Level 2	1%	1 pp	-1 pp
Full Level 2	11%	2 pp	-2 pp
Level 3	3%	1 pp	-1 pp
Full Level 3	9%	4 pp	-2 pp
Level 4+	8%	1 pp	-1 pp

Variation by sex

Earnings: *Estimated earnings premiums for women are lower than those for men at all levels of learning, but the gap becomes much smaller at FL3.* For instance, the average three to five year earnings premium for female Full Level 2 achievers is 9%, compared to 13% for men. However, at FL3 the 3 to 5 year earnings premium is 9% for male achievers and 8% for female achievers. The gender differences in earnings premia at Level 4+ are also less pronounced, and at L3 we seem to suffer a pronounced dip in early earnings returns for women, which are actually negative and significant in the first and third years after learning (-3 and -2 per cent respectively). However, five years on from the end of learning the earnings premium for L3 women achievers is only 2 percentage points lower than the return for their male counterparts, and we do not observe the same early dip in returns for these men.

Our estimates for women at L3 likely understate true earnings returns and the same issue may also understate our estimates of earnings returns for women achievers at FL3. Unfortunately the poor returns for women at Level 2 and Below Level 2 seem to be more accurate. Women achievers secure no significant and systematic return that allows us to quote a valid three to five year average for these categories of learner. In contrast, their male counterparts secure a 5% and 3% average return three to five years from the end of learning, for Below Level 2 and Level 2 achievement respectively.

Employment: *Women secure more of a three to five year average employment premium from learning at all levels, when compared to men.* For instance, the highest 3 to 5 year average return accrues to women Full Level 3 achievers, who have employment rates that are 4 percentage points higher than non-achievers. Men achieving a FL2 qualification have employment rates 1 percentage point higher than non-achievers, whilst the comparable figure for women achievers is 3 percentage points. The lowest employment returns for both men and women are seen amongst those achieving a highest

qualification Below Level 2, where men and women seem to have a zero employment return.

These differences are confirmed when we check the raw employment figures, as the raw percentages of men and women in employment is roughly equivalent. Considering earnings results for women from the previous section and the findings here, female earnings returns may be understated because many work part-time and we are not able to fully identify this in our data – women working part-time will seem as if they are achieving a lower wage, because we are only able to capture daily earnings. When considering employment returns, women may do better partly because part-time and full-time employment are given an equal weight in the analysis.

Fig. 2: Summary of three to five year averages for women and men

	Women		Men	
	Earnings Returns	**Employment Probability Premiums**	**Earnings Returns**	**Employment Probability Premiums**
Below Level 2	0%	0 pp	5%	0 pp
Level 2	0%	1 pp	3%	1 pp
Full Level 2	9%	3 pp	13%	1 pp
Level 3	2%	2 pp	5%	1 pp
Full Level 3	8%	4 pp	9%	3 pp
Level 4+	7%	1 pp	10%	1 pp

Age Differences

Earnings: *Estimated earnings premiums for the 19-24 age group are higher than those for the 25+ age group, whichever qualification category we consider.* For instance, amongst learners aged 19 to 24 with a highest aim of Full Level 2, achievers secure a 3 to 5 year average earnings return of 10%; and at Full Level 3 the figure is also 10%. In contrast, achievers aged 25+ secure an 8% 3 to 5 year premium over similar aged non-achievers at Full Level 2 and 10% at Full Level 3. The majority of learners are located in the 25+ age group (between 4 and 2 millions) but this still leaves between 1 and 0.5 millions in our regressions for 19-24 year olds. It is a concern that 25+ achievers, amongst the 875 thousand whose highest aim is Level 2, secure a negligible premium on average.

Employment and Active Benefits: *Comparison of estimated employment probability premiums, and probabilities of being on active benefits, across our two age groups show some small differences.* For instance, the 25+ group of achievers have estimated employment probability premiums of 2 percentage points at Full Level 2 and Full Level 3, that are higher than those secured by achievers aged 19-24 (0 and 1 percentage points, respectively). When we consider the proportions in employment in the two groups and turn these percentage point premiums into percentages, these differences remain. Any differences in the percentage point probabilities of achievers being on active benefits, relative to non-achievers, following the completion of learning, disappear when we consider the raw figures and change percentage point into percentage figures.

Fig. 3: Summary of three to five year averages for 19 to 24, and 25+ year olds

	Age 19-24		Age 25+	
	Earnings Returns	**Employment Probability Premiums**	**Earnings Returns**	**Employment Probability Premiums**
Below Level 2	5%	0 pp	1%	0 pp
Level 2	5%	2 pp	1%	1 pp
Full Level 2	10%	0 pp	8%	2 pp
Level 3	6%	2 pp	3%	1 pp
Full Level 3	10%	1 pp	10%	2 pp
Level 4+	11%	1 pp	7%	1 pp

Variation by academic and vocational qualification categories

Earnings: *We find that vocational qualifications provide positive and statistically significant earnings premiums at all levels, with returns particularly high for work-based learning.* Earnings premiums are estimated separately for three categories of learner studying in settings that are (i) Academic (ii) Vocational Classroom Based (CBL) and (iii) Vocational Workplace-Based (WBL). We find that 96% and 91% of learners whose highest aim is Below Level 2, and L2 respectively, are in our Vocational CBL category. Therefore the Below Level 2 and L2 earnings premiums of 2% and 1% are very similar to the 2% and 1% three to five year averages previously seen for *aggregated qualification categories*. In contrast, the more favourable Below Level 2 earnings return of 7% in WBL contexts applies to very few learners. Learners in our Full Level 2 category are more evenly spread across the three categories - the average 3-5 year average return for Academic Full Level 2 is 12%; it is 10% for Vocational CBL and 10% for Vocational WBL. At both Level 3 and Full Level 3 we observe negative and significant returns to learning in

the first and third years for our Academic category. This reflects the impact of unobserved HE destinations, as it does not have as much impact for the earnings premium for Vocational CBL Level 3 achievers, which rises steadily from 2% to 4% (giving a 3 to 5 year average of 3%). Similarly, Vocational CBL Full Level 3 achievers secure a 3-5 year average premium of 6%, with the underlying estimated return rising from 2 to 9 over the five-year period. The return to Vocational WBL Full Level 3 Achievers averages 9% between 3 and 5 years.

Employment: ***Comparison of estimated employment probability premiums across our three qualification categories shows a lot of variability, with no clear overall pattern.*** The premiums secured by those achieving a highest qualification at Below Level 2, are positive and significant for both our WBL and CBL categories of Vocational learners, but in the CBL group this only equates to a 3 to 5 year average of zero, to the nearest whole number (compared to a figure of 2 in the WBL Group). In contrast, those achieving a highest aim of Level 2 in the WBL category secure no significant employment premium in the fourth and fifth years after learning (despite stronger performance in the 1st and 3rd years). The 3 to 5 year average return to achievers at Level 2 in the Vocational CBL category is 1 percentage point – not particularly high, but statistically significant in all years from the end of learning. Similarly, the employment premium for Level 3 Vocational CBL achievers averages 1 percentage point between 3 and 5 years, whilst the same level of learning secures a 2 percentage point premium in the Academic qualification category. At Full Level 3, achievers in the Academic category of qualifications seem to secure the highest return over non-achievers, with a 4 percentage point, 3 to 5 year average; compared to figures of 3 and 2 percentage points for Vocational CBL and Vocational WBL categories, respectively.

Fig. 4: Summary of three to five year averages for Vocational Learning

	Classroom-Based		Workplace-based	
	Earnings Returns	Employment Probability Premiums	Earnings Returns	Employment Probability Premiums
Below Level 2	2%	0 pp	7%	2 pp
Level 2	1%	1 pp	5%	0 pp
Full Level 2	10%	2 pp	10%	1 pp
Level 3	3%	1 pp	N/A	N/A
Full Level 3	6%	3 pp	9%	2 pp
Level 4+	8%	1 pp	9%	0 pp

Estimated returns for Apprenticeships

Earnings, Employment and Active Benefits: *At L2 and L3 we estimate significant and substantial earnings returns, together with significantly lower probabilities that achievers will be on benefits.* The 3 to 5 year average earnings premium for Level 2 apprenticeship achievers is estimated at 11% and for level 3 achievers the figure is 16%. The estimated employment returns are less pronounced, and by the third and fourth year after learning there is little difference in employment rates between achievers and non-achievers at both L2 and L3. In contrast, both L2 and L3 apprenticeship achievers have a significantly lower percentage point probability of being on active benefits (3 and 2 percentage points, respectively). Because the absolute proportions of apprenticeship achievers and non-achievers is relatively low, these translate into figures that suggest L2 achievers have an approximate 30% lesser probability of being on active benefits and for L3 achievers the figure is 48%. The lesser impact of apprenticeship achievement on employment is perhaps to be expected as learners are usually employed during the period of their apprenticeship, and even non-achievers will likely have a high probability of continuing in employment, but at a much lower wage.

Fig. 5: Summary of three to five year averages for Apprenticeship Learners

	Earnings Returns	Employment Probability Premiums	Benefit Probability Gaps
Level 2	11%	0 pp	-3 pp
Level 3	16%	0 pp	-2 pp

L1/L2 Maths and English qualifications

The L1 and L2 Maths and English qualifications we consider in this section will mostly be taken alongside more substantial qualifications – they are often taken as complements to more 'Full' or higher-level qualifications. As a result we adapt our approach to estimation. Elsewhere in the report we are capturing returns to qualifications gained as an individual's highest learning aim. Taking such an approach here would miss many learners, as we will only analyse individuals whose highest learning aim across their 2002 to 2012 FE spells is a L1 or L2 English or Maths qualification. This is a relatively unique group and as we have already seen, returns tend to be low for those who have Below Level 2 or L2 as their highest learning aim. In this section of the report, we find statistically significant returns to L1 and L2 Maths and English, when we adopt an approach to estimation that accommodates the complementary nature of these qualifications.

The 3 to 5 year average earnings premium for those achieving a L1 and/or L2 Maths/English qualification, relative to those who do not achieve their L1/L2 Maths or English qualification is 2% (with this estimate relevant for a population of all FL2 and FL3 achievers). More specifically, we estimate that Literacy/Numeracy Key Skills achievers secure a 3 to 5 year average earnings return of 6% over non-achievers (amongst a population of all FL2 achievers) and 4% (for a population of all FL3 achievers).

Fig. 6: Summary of three to five year average earnings return for L1 and L2 English and Maths Achievers

	Amongst FL2 achievers	Amongst FL3 achievers	Amongst FL2 or FL3 achievers
Return to L1 and L2 English and Maths Achievers (<u>when not held as highest qualification</u>)	4%	1%	2%

Variation in returns by sector subject area

Earnings: *We find that the majority of sectors provide positive and statistically significant earnings premiums to FL2 and FL3 learning.* In female-dominated sectors such as Adult Social Care, average 3 to 5 year earnings returns are just above 10 per cent at FL2 and FL3, and the same is true of achievers at FL3 in the area of Child Development and Wellbeing. Even where we see lower earnings returns, for instance in the areas of Hair and Beauty, and at FL2 for Child Development and Wellbeing; nowhere do these dip below 5%. In male-dominated areas such as Construction, Engineering & Manufacturing and Transportation we see earnings returns that are close to 10 percentage points higher than this. Across these Sectors we seldom see average three to five year returns dropping below 15% - only in the case of FL2 learners in Engineering and Manufacturing. In areas that are neither male nor female dominated such as ICT, Customer service and Hospitality and Catering, earnings returns at FL2 are around 10% and in Business Studies just below at 8%.

Employment: *In female dominated sectors we estimate positive and statistically significant employment returns for achievers, whilst in sectors dominated by men, and by neither sex, returns are more variable.* In Adult Social Care, Child Development and Wellbeing and Hair and Beauty average 3 to 5 year employment returns vary around 3 percentage points at FL2, and at FL3 they are 4, 6 and 2 percentage points respectively. There is some concern over employment returns for those achieving a highest FL3 qualification in Construction, FL2 in Engineering and Manufacturing and Full Level 3 in Transportation. Similarly, those achieving FL2 qualifications in ICT, Customer Service and Business Studies experience periods of negative returns. At FL3 we observe instances of negative and statistically significant employment probability returns in Customer Service. Those achieving Hospitality and Catering qualifications at FL2 and FL3 secure returns of 3 and 2 percentage points respectively; and FL3 Business Studies achievers secure a 6 percentage point average return between 3 and 5 years from the end of learning.

Fig. 7: Three to five year average earnings and employment returns, for female dominated sectors

	Adult Social Care		Child Development and Wellbeing		Hair and Beauty	
	Earn.	Emp.	Earn.	Emp.	Earn.	Emp.
Full Level 2	12%	3 pp	6%	3 pp	8%	3 pp
Full Level 3	14%	4 pp	11%	6 pp	7%	2 pp*

* These estimates should be considered with some care

Fig. 8: Three to five year average earnings and employment returns, for male dominated sectors

	Construction		Engineering and Manufacturing		Transportation	
	Earn.	Emp.	Earn.	Emp.	Earn.	Emp.
Full Level 2	16%	2 pp	12%	0 pp	16%	3 pp
Full Level 3	15%	0 pp	16%	3 pp	19%	0 pp

Fig. 9: Three to five year average earnings and employment returns, for sectors that are neither male nor female dominated

	ICT		Customer Service		Hospitality and catering		Business Studies	
	Earn.	Emp.	Earn.	Emp.	Earn.	Emp.	Earn.	Emp.
Full Level 2	12%	0 pp	9%	0 pp	11%	3 pp	8%	0 pp
Full Level 3	0%*	3 pp	8%	0 pp	7%	2 pp	10%*	6 pp

* These estimates are particularly impacted by the HE flag issue

How Robust are our Estimated Returns?

Whilst further enhancements are always going to be possible, the results produced in this report provide BIS with estimates of the value added of FE that are quality assured to the highest academic standards. We test the validity of our approach to estimation of the Value Added from qualifications in other areas of the report. We focus on NVQs, as they are often the focus of investigation in survey-based studies, and are also where concern has been raised over low (or even negative) estimated returns. The process we have gone through, to test the robustness of our own estimates, has been peer-reviewed by four academic experts and we are very grateful for their comments. Any mistakes remain our own and readers can consider the comments of reviewers and our responses at the end of this report. Our findings suggest that regression-based

techniques, which compare achievers and non-achievers in the ILR-WPLS, produce robust estimates of value added.

We also investigate the persistence of earnings and employment probability premiums over the years after the end of learning. For FL2 and L3 apprenticeship learners, the difference between a 3 to 5 year average earnings premium and a 4 to 6 year average, is less than 0.4 of a percentage point. The results for FL3 learners suggest that data issues may be particularly problematic in the pre-recessionary period when calculating earnings returns. We adopt a 3 to 5 year average in the remainder of the report, as it balances a number of competing issues – going with a 4 to 6 year average would leave us with only one cohort making up the estimate of year 6, which is as much of a concern as persistence.

Those achieving qualifications at Full Level 3 seem particularly well placed to secure employment in a recession than non-achievers, but for FL2 achievers there is an apparent one-off employment impact of recession for those in the 2007/2008 cohort – with the gap between achievers and non-achievers in this FL2 cohort, much greater than that seen for earlier or later cohorts. Considering the figures presented elsewhere in this report, this suggests that (i) for each cohort, percentage-point employment probability premiums exhibit persistence and (ii) premiums may be higher for recessionary cohorts. The implication is that, for FL3 our estimated 3 to 5 year average employment probability premiums will be dominated by the slightly lower pre-recession employment probability premiums, because we do not have a long time period to observe employment rates for post-2008 cohorts. This may be more appropriate looking forward to the post-recessionary era. For FL2 this may be less of an issue, as we only seem to have a one-off effect of recession.

Further Enhancements

In various sections of the report we highlight issues that arise because of the limited coverage of our HE flag, and our inability to accurately identify part-time working in the data. Work is needed to integrate the indicator of progression to HE learning for the full population of learners, and also to better understand any activity that isn't currently observed in the data. This also raises the wider issue of progress to further learning (whether HE or otherwise).

The HE-flag issue may be less of a problem when considering a three to five year average for returns, as it only acts to depress our estimated returns up to the third year after learning. Unfortunately, if our explanations of HE learning and part-time working are correct, when we see returns recover in the 4th and 5th years after learning, we may not be able to attribute all of the upturn to L3 and FL3 learning. However, it is possible that unobserved moves into HE from Level 3 learning is not the entire story here and there are various competing issues to consider. For the moment we refer readers to the estimated earnings returns of 6% for L3 and 14% for FL3 in the previous report, where this issue was tackled by focusing on the population for whom the HE flag was available.

Any future study would ideally capture progress to HE as an outcome in itself (as long as those moving from FE to HE, subsequently progress within HE). This would allow us to better capture returns to those who achieve at L3 and FL3 and then progress to the labour market. This touches on the wider issue of progression to further study, an issue we need

to keep in mind when considering other findings, particularly those at Below Level 2 and Level 2. Our estimates of the returns to these qualifications when held as a highest learning aim are for those who do not progress to further study. It may be that those achieving Below Level 2, who go on to further study, are achieving very good returns to this level of qualification, because it is an essential facilitator of their move to this higher level of learning.

Finally, our exploratory analysis of cohorts raises some very interesting issues. For FL2 achievers there is an apparent one-off employment impact of recession for those in the 2007/2008 cohort – with achievers much more likely to secure employment than non-achievers. FL3 achievers in cohorts after 2007 secure much higher percentage point employment premiums than those before the recession.

For the FL2 learners there will be a small impact of this in our estimates, but nothing substantial. In contrast, estimated 3 to 5 year average employment probability premiums for FL3 learners will be dominated by the slightly lower pre-recession employment probability premiums, and we flag where this is likely to occur. In contrast, for those who have secured jobs, the nominal earnings impacts of a recession seem minimal (though real earnings may of course have stagnated or fallen).

For policymakers, the question is whether the pre-recession returns or recessionary returns are more appropriate for the immediate future. We would suggest that the balance between the two that we achieve in the present study is highly appropriate in the present uncertain economic environment.

1. Introduction

In this paper we update the findings from Buscha and Urwin (2013) [from here, B&U]; estimating separately the (i) earnings, (ii) employment probability and (iii) probability of being on active benefits, for those who achieve their highest learning aim whilst studying at an English Further Education Institution (FEI), relative to those who have the same highest learning aim, but do not achieve. The findings from this analysis provide a robust estimate of the value added from each qualification, when held as an individual's highest.

This approach to estimation is made possible by the construction of a database linking administrative FE learner information, with benefit and PAYE employment histories. In the following *Data and Method* section we give some idea of the 'spells-based' approach adopted in constructing the ILR-WPLS dataset (also see Bibby et al., 2012). This draws together the Individualised learner Record (ILR), which contains over 54 million instances of individuals engaged in 114 million separate FE learning aims and the Work and Pensions Longitudinal Study (WPLS), which contains the benefit histories of over 9 million individuals, P45 records relating to over 15 million individuals and P14 earnings relating to approximately 14 million individuals[1].

Our categories of educational attainment relate to the National Qualifications Framework (NQF), running from Below Level 2 to a highest category of Level 4+; and estimates are created for those aged 19+ based upon,

- FE learning spells taking place between August 2002 and April 2011[2].
- Earnings data for financial years 2004/2005 to 2011/2012.
- Data on benefit and employment status from August 2002 to October 2012[3].

There are a number of enhancements to the approach used in B&U. We have (i) made improvements that impact the sample selected for analysis and (ii) introduced additional controls that make our regression estimates more accurate. These changes lead to a discernable reduction in estimated employment probability premiums, compared to B&U. In contrast, estimated earnings premiums and differences in the likelihood of being on active benefits are impacted only slightly. The estimated earnings premiums remain much higher than those obtained from survey-based studies, which often identify negligible (or even negative) returns to some qualifications, such as NVQs at Levels 1 and 2 (see for instance, Conlon, Patrignani and Chapman, 2011). The Data and Method section provides detail on the changes we have made, the motivation for each, and the reasons why they reduce estimated employment probabilities.

---

[1] We also incorporate an HE flag, created for a subset of learners (born after 31/08/1980 and who experience an FE learning spell between 2004/2005 and 2009/2010) who progress to Higher Education, by linking to the Higher Education Statistics Agency's (HESA) student record.

[2] Although we have information on learning spells that run to July 2012, the latest earnings and employment data we have is for 2011-2012 and so we are unable to use the latest year of learning data to estimate returns.

[3] For more detail see http://ofqual.gov.uk/qualifications-and-assessments/qualification-frameworks

These enhancements produce more robust estimates, but they still rely on a standard regression approach, which has potential limitations. Taking the example of wages, any study attempting to capture the returns to education needs to identify the difference between the earnings of those who have the qualification, and an otherwise identical group who do not have the qualification. We use regression analysis to control for any differences between the characteristics of achievers and non-achievers, so that we are more likely to be comparing 'like-with-like'. The assumption is that we control for all the important differences between achievers and non-achievers, and any remaining difference in earnings is a result of the achievers' additional learning.

The Data and Method section describes the regression approach used in this study, its potential limitations and how these might possibly bias estimated returns. Section 6 describes the results of an analysis that uses a combination of econometric methods (including matching and difference-in-differences) to investigate whether the achiever v non-achiever comparison using a standard regression approach produces estimates that suffer from bias as a result of these limitations.

The findings in Section 6 make an important contribution to the research on FE learning outcomes. The results suggest that estimated labour market returns gained from a comparison of achievers and non-achievers, using standard regression techniques, are very robust. The findings of Section 6 provide BIS with the necessary quality assurance to consider the achiever v non-achiever approach as a verified method of estimating the value added for FE.

Section 6 also considers the question of how long any labour market returns persist beyond qualification, with a study of returns to different cohorts of learners up to six years after the end of learning providing the necessary insight. The findings influence our decision to report estimated earnings, employment and active benefit returns for the first, third, fourth and fifth tax year after learning. Returns in the first year are an important benchmark for comparison across studies and results of the cohort analysis provide a rationale for adoption of a 3 to 5 year average as the appropriate policy metric[4].

Taken together, our enhancements to method and data selection; validation of estimates using difference-in-differences on matched samples; and investigation of the persistence of returns, provide a high level of confidence in the estimated labour market returns presented in Sections 3, 4 and 5.

More specifically, the comparison of achievers and non-achievers from administrative data,

- seems to overcome the problems of negative selection that may be the cause of low estimates in survey-based studies (for instance, Dearden et al., 2004; Greenwood et al., 2007; McIntosh, 2009; Dickerson and Vignoles, 2007; Garrett, Campbell and Mason, 2010; Conlon, G., Patrignani, P. and Chapman, J., 2011). *If individuals who would gain a lower wage (independent of their level of learning) are more likely to select into certain vocational qualifications, then estimated returns may be understated*

---

[4] The 3 to 5 year average is chosen because it best balances an implicit trade-off. We want estimated premiums that (i) persist sufficiently far into the future, after the end of learning, but (ii) are not over-reliant on a small number of older cohorts that completed learning many years ago (as we only observe more than five years of post-learning outcomes for the older cohorts of learners in our data).

if this group are compared to a control group who do not select into this qualification (and any analysis does not manage to control for the implied differences).

- but does not seem to suffer from the weakness that we might expect when comparing those who achieve with those who do not achieve. *If those who 'select' into achievement are more able/motivated in ways that we cannot observe, then regression estimates of the returns to achievers V non-achievers will be biased upwards.*

The analysis does not seem to suffer from the latter weakness because many individuals who select into the non-achieving group, 'drop-out' rather than 'fail'; and because this drop-out seems more a result of variation in course quality across FE institutions, rather than characteristics of the individuals themselves (Martinez and Munday, 1998; Hodkinson and Bloomer, 2000; Simm, Page and Miller, 2007; Hogarth, Gambin, Hasluck, de Hoyos and Owen, 2009). These findings have important implications for the literature on returns to FE and we consider these and other issues in the conclusion (Section 7).

Table 1 provides a simple overview of treatment and control groups when adopting the achiever v non-achiever framework, detailing all possible combinations of highest aim and qualification achievement. For instance, the cell marked 'X' contains all learners whose highest aim across all learning spells is Full Level 2, but whose highest level of achievement is recorded as Below Level 2. In contrast, there should be no observations in the cell marked 'Z' as one cannot have a highest aim of Full Level 2 and highest achievement of Full Level 3. For each treatment group represented by the grey diagonal, the yellow cell in the 'no achievement' column of the same row is the relevant control.

Table 1: Example of treatment (achievers) and control (non-achievers)

Highest Aim	Highest level of Achievement						
	Null	Below L2	Level 2	Full Level 2	Level 3	Full Level 3	Level 4+
Below Level 2							
Level 2							
Full Level 2		X				Z	
Level 3							
Full Level 3							
Level 4+							

In this report we estimate returns to qualifications, when held as an individual's highest. For each learner we have 7 academic years (2004/2005 to 2010/2011) when they can possibly exit learning (as an achiever or non-achiever) and over these academic years, a learner can have multiple ILR learning spells. Across all of the learning spells for each individual we select the highest learning aim. The estimate of value added is obtained by comparing the returns of those who have a particular highest learning aim (for instance Full Level 2) and achieve; with the returns of those who have the same highest learning aim, but do not achieve.

2. Data and Method

Section 2.1 reports on the creation of an ILR-WPLS spells-based dataset. Section 2.2 describes (i) the regression approach used to model earnings, employment and active benefit returns in Sections 3, 4 and 5; (ii) the limitations of this approach; and (iii) the difference-in-differences and matching methods we use in Section 6 to test the robustness of regression estimates produced using the achiever v non-achiever approach.

2.1 Creation of the ILR-WPLS dataset

We create the ILR-WPLS dataset linking FE learner information, benefit and PAYE employment histories for tens of millions of individuals. A pre-requisite for the construction of such a dataset is the creation of an over-arching Person-key to link records in the data sources (the Individualised learner Record (ILR) together with the Work and Pensions Longitudinal Study (WPLS)) reliably to the same individual. This Person-key identifies the same individual in the various data sources and is now used by BIS for in-house analysis. The creation of the ILR-WPLS dataset requires complex processes of imputation and merging. Details of the imputation processes and procedures to arbitrate between competing matching possibilities are described at length in B&U and Bibby et. al. (2012). Some issues with the data have implications for the analysis presented here, and these are set out in an accompanying Technical Annex.

Our categories of qualification are as follows:

- In the *Full Level 2* category we include all highest aims that are equivalent to 5 GCSEs at grade A* to C (including apprenticeship qualifications)
- In the *Full Level 3* category we include all highest aims that are equivalent to two A levels (including apprenticeship qualifications)
- *Level 2* and *Level 3* include learners with highest aims at these levels of study, but which fall below the equivalence required to be considered 'Full'.
- *Level 4+* and *Below Level 2* are self-explanatory.

These '*Level of study'* categories are based on <u>all</u> aims undertaken by a learner within a spell (not just their single highest aim). For instance, an individual undertaking 5 GCSEs (and nothing else) will be categorised as '*Full Level 2*' even though the single highest aim would only be worth 20% of the Level 2 threshold. NVQs will always figure in our Full categories, as will the majority of BTECs, whereas RSA and City & Guilds qualifications could fit into either *Full* or *Other* depending on the level of commitment required of the learner.

As suggested in the Introduction (page 8) we have, 'made improvements that impact the sample selected for analysis' and this impacts our estimated labour market returns (when compared to the results presented in B&U). Section 2.2 sets out the detail of these changes, the motivation for change and the impact on estimates.

2.2 Econometric approach and explaining enhancements since B&U

In observational (non-experimental) studies the treatment group usually has different characteristics to those of the control group. To create robust estimates of any treatment effects, we require estimators capable of controlling for such differences. Standard regression-based approaches, that do not utilise data discontinuities or instruments, simply control for differences in characteristics by adding regressors. Our estimates in Sections 3, 4 and 5 update the findings from B&U using a standard regression approach to arrive at premiums secured by Achievers over Non-achievers, with the same highest learning aim. The econometric models take the following form and are all estimated using Ordinary Least Squares (OLS):

$$y_i = \alpha + Qual_i'\beta + \mathbf{x}_i'\gamma_x + \varepsilon_i$$

where the dependant variable, y_i, takes one of the following forms:

- Log of deflated daily earnings (top and bottom 1% removed) in the whole financial tax year 1, 3, 4 and 5 years after the end of a learning spell.
- The probability of being employed (binary) exactly 1, 3, 4 and 5 years after the end of a learning spell[5].
- The probability of being on job-seeking ['Active'] benefits (binary) exactly 1, 3, 4 and 5 years after the end of a learning spell.

The qualification variables, $Qual_i$, are inserted as dummies where a value of 1 represents those who achieved their highest spell aim, whilst a value of 0 represents those who had the same highest spell aim, but failed to achieve any qualification within that spell (achieve=null in Table 1). The coefficient β then represents an estimate of the return (or premium) to that qualification level. The nature of our setup (achievers vs. non-achievers) requires each qualification level to be estimated in its own regression model.

It is worth emphasising that the earnings results only cover those achievers and non-achievers for whom we have earnings data in at least one year after the end of learning. For instance, if we observe earnings for an individual in the third tax year after the end of learning, then they will be included in the regression equation that estimates earnings returns for this third year. However, if we observe no earnings for this same person (whether achiever or non-achiever) in the first tax year, they will not feature in the regression equation estimating earnings for the first year after learning. Taking the three to five year average for returns, better ensures that our results are not confounded by the potential for individuals to be in one year's regression equation and not the next. Also, Section 6 considers a related issue, where we possibly have differential sample selection of achievers and non-achievers into employment (and therefore, earnings).

---

[5] Some regressions have binary variables on the left-hand side. These models are also estimated using OLS resulting in a linear probability model (LPM). Such models have the disadvantage that out-of-bound predictions can occur (probabilities greater than 1 or less than 0) and also suffer from heteroskedasticity. However, the sheer size of these data make logit/probit modelling with marginal effects highly computationally intensive; particularly as our identification strategy requires each qualification level to be estimated in its own regression equation. LPM models generally performed well and we report robust standard errors.

The Impact of, and motivation behind, Methodological Enhancements

As suggested in the introduction, whilst this report adopts the same standard regression approach as that used in B&U, the estimated labour market returns in Sections 3, 4 and 5 differ as a result of further enhancements. We now describe these changes and the reason why they have been implemented - with the discussions around Table 2 providing a detailed explanation of how changes impact our estimates.

One set of changes relates to the way we have selected our sample of achievers and non-achievers. Specifically,

a. The previous analysis in B&U focused on learning spells between August 2004 and July 2010. The analysis presented here includes information on learning spells between August 2002 and April 2011. This results from the addition of earnings data (2011-2012) that have become available, coupled with a desire to include as many spells as possible, even when these are undertaken much earlier in our period for analysis[6].

b. In the new analysis we have kept only the highest learning aim associated with the most recent learning spell, in situations where an individual has multiple spells with the same highest learning aim. In the previous analysis multiple observations were included for the same individual (in both treatment and control) where that individual had multiple spells with the same highest learning aim[7].

In addition, we have further refined the controls used in the regression equations to come up with the very best specification, given the data we have. Specifically,

c. Regression estimates in the previous report (*Basic Controls* in Table 2) were produced controlling for sex; age; an interaction of sex and age; ethnicity; disability; region; postcode; type of funding (none, LCS, ESF, both); mode of attending (FT/PT); offender; spell duration; number of previous FE learning spells; prior education level[8] and year dummies.

We have expanded this basic regression specification to include a number of additional controls. We have merged in the Index of Multiple Deprivation (IMD), replacing postcodes; we now include an indicator derived from Sector Subject Area (SSA)[9]; and

---

[6] Previously there was some concern over the learning data prior to 2004, because of changes in the funding arrangements for courses around this time – prior to 2004 funding was available for many more courses not formally accredited to an NQF level. However, our analysis does not include any courses that are not accredited to an NQF level and therefore we include these earlier learners. It may not seem immediately obvious why we would desire the introduction of earlier learners, but as the discussion in the second part of Section 6 underlines, by the time we get to an estimate of earnings in the sixth year after learning we are wholly reliant on our earliest cohorts of learners (see Section 6.2 for a more detailed discussion).

[7] We have opted for the new approach because the old approach increased effective sample size by including multiple observations for learners with more than one highest learning aim at the same level, and implicitly gave a greater weight to the experiences of these learners. This also avoids the situation where we attempt to estimate the returns to education of a learning spell, which is followed immediately by a learning spell at the same level.

[8] Prior education is constructed from two sources; (i) the records of those who undertake FE learning prior to the spell that includes their highest learning aim and (ii) self-reported qualification status on entry to FE.

[9] The derived Industry Sector classification we have used has the SSA variable in the ILR at its route, but also uses the title of the highest learning aim and the SSC footprint where SSA is missing or uninformative.

have also introduced controls for (i) the number of days an individual was on active benefits in the year before learning (ii) whether an individual has an inactive benefit spell in the year before learning and (iii) how many days in sustained (6 months) employment an individual has just before learning. The addition of these last three controls brings us into line with approaches in the literature focused on evaluating active labour market policy interventions (see for instance, Lechner et. al. 2011).

d. In the previous report the majority of estimates were produced without the benefit of information indicating whether an individual has moved on to HE after their FE learning spell (referred to as a 'flag' of HE Learning). This data are only available for a subset of learners, born after 31/08/1980 and who experience an FE learning spell between 2004/2005 and 2009/2010. For other individuals we have no indication of whether they attend HE at the end of FE learning. For the analysis here we include all available learners in the dataset, no matter when they are born and when they undertake learning; with those who have no HE flag because they fall outside of the subset of learners, being assigned to the category of 'missing', but retained in the analysis. The difference between the approach here and that of B&U is that here the HE flag is included as a control in the main analysis (where it is present), rather than only being used in analysing a subset of learners. It is not the scope of the population under analysis that has changed; rather, here we include the HE flag as a control across the whole analysis (again where it is present).

Before considering in detail how the changes under a., b. and c. impact estimates, it is important to note how the addition of an HE flag has impacted estimates of FE learning. In a previous study by Patrignani and Conlon (2011) using administrative data, we observe a large dip in returns up to the third year after FE learning, particularly when considering the earnings returns to Full Level 3 qualifications. However, these estimates were created without the benefit of an HE flag. The assumption has been that if we do not identify those who go on to HE from Full Level 3 FE learning, we have a significant number in the treatment (achievers) who will be taking poorly paid PT jobs (whilst HE students) and this will understate earnings relative to a control who exit to higher paid jobs. As returns tend to pick up 3 to 4 years from the end of learning (when the treatment will have finished HE) this seemed to fit with the facts.

The results presented in B&U were based on a more 'inclusive' approach to creation of the dataset and the dip in returns was less obvious, but still apparent (see Data and Method section of B&U for more detail); as the HE flag was still missing from the main body of analysis. As suggested above, the results presented here are from regression equations that include an HE flag, with those who have no HE flag being assigned to the category of 'missing', but retained in the analysis. As readers will note from the following results, there is still some slight dip in returns at Full Level 3, when we consider the trajectory of estimates from the 1st, to the 3rd and the 4th year, but this is much less pronounced.

It is possible that unobserved moves into HE from Level 3 learning is not the entire story here. There may be a 'foundation degree' issue or something about the wages received by those who complete a vocational qualification at Level 3 – they may simply not be earning as much as those who drop out and gain a return to their Level 2 qualifications, in the initial years after learning. Perhaps they go into firms on lower 'training-type' wages as they gain practical application of the skills learned at Full Level 3. After two or three years

doing this, their returns start to rise. We have carried out this analysis for specific sector subject areas where there are sufficient numbers and the same results still hold.

The HE flag issue is one for further research, but considering the results presented here it is (i) much less of a problem when we consider a three to five year average for returns and (ii) if anything it acts to depress our estimated returns, rather than falsely inflate them, especially when considering FE qualifications at Level 3.

Finally, Table 2 shows how the changes detailed under a., b. and c. impact estimates, taking the examples of Full Level 2 employment probability and earnings estimates, at three months and one year after learning[10].

First, **Column A** presents the old estimates using the standard regression specification (*Basic Controls*) and using the same approach to selection of treatment and control as described in B&U (that is, before the changes described under a. and b. above)[11]. **Column B** presents the results using the same regression equation, but with the new approaches to selecting treatment and control described under a. and b. The new approach to data selection alters the weight given to different cohorts in calculating our estimated returns. For instance, in Section 6.2 we shall see that the employment probability premium for those achieving a Full Level 2 qualification in 2008 (the 2007/2008 academic year cohort of achievers and non-achievers) is 4 percentage points or more for most of the years (2008/2009, 2009/2010, 2010/2011 and 2011/2012) where we observe returns. This is in stark contrast to the premiums secured by all other learner cohorts, which rarely rises above 2.5 percentage points, whether they complete learning before or after the 2007/2008 academic year. Our expansion of the dataset[12], to include more learner cohorts that secure employment premiums in line with those in (seemingly) more 'normal' years, reduces the weight given to the exceptionally high employment probabilities associated with the 2007/2008 cohort of learners. This results in a fall for the estimated employment returns by 3.5 percentage points in Table 2, but has only a 1 percentage point impact on earnings estimates. This is consistent with the analysis in Section 6.2 where we see no pronounced difference in earnings returns to FL2 across different cohorts whether they complete learning before, during or after 2008 – changing the cohort-composition of the sample has much less effect as we have no cohorts with 'exceptional' earnings returns. We return to consider the exceptional nature of employment probability premiums for the 2008 cohort in Section 6.2.

---

[10] The choice of 3 months and 1 year is simply due to the fact that these are the first periods for which we present estimated earnings (1st tax year) and employment (3 month employment probability premium) results in the following tables. The findings described in Table 2 hold for later periods where returns are estimated.
[11] All estimates presented include an HE flag in the specification.
[12] The changes under a. introduce some earlier learners and a 2010/2011 cohort, whilst the change under b. concentrates estimated returns on more recent cohorts.

Table 2: Impact of enhancements to basic regression approach: the example of employment probability and earnings premiums for FL2 achievers V non-achievers

	Column A: Old estimate (basic controls)	**Column B:** New estimate (basic controls)	New estimate (basic controls +IMD)	New estimate (basic controls +IMD+SSA)	New estimate (basic controls +IMD+SSA +Benefit Hist.)	**Column C**: New estimate (basic controls+IMD +SSA+Benefit and Employment Hist.)
Employment after 3 months	0.087	0.052	0.052	0.052	0.049	0.023
Earnings after one year	0.185	0.174	0.173	0.166	0.149	0.143

Comparing **Column B** with **Column C** for the employment probability estimate (and the intermediate columns which show how estimates change as we add controls one by one) we can see how the changes described under c. impact our estimates. The addition of IMD, our sector variable derived from SSA and previous Benefit status has little impact, but when we also control for employment status prior to learning the estimate falls from 0.049 to 0.023. The suggestion is that, in controlling for the previous employment status of individuals, we obtain a more accurate (lower) estimate of the employment impact of qualifications. The fact that the estimated probability premium falls, suggests that prior to controlling for previous employment status, achievers were more likely than non-achievers to be previously employed and this made them more likely to be employed after learning – something that we were ascribing to the learning itself, rather than their status prior to learning.

Change in the nature of our sample and introduction of new controls impacts employment probability estimates at FL2 across the board in the new report, whereas earnings and benefits are much less changed at all levels – the earnings estimate above moves from 17.4 to 14.3 with the introduction of the new estimates and in many of the tables in the report, the reduction is less than this. For the estimate of active benefit probabilities, there is similarly much less of a difference between previous and present estimates, even when considering FL2. To summarise,

- The difference between column A and B is mostly due to the fact that the new approach to data selection alters the weight given to different cohorts in calculating our aggregate estimated returns. The 2007/2008 cohort exhibit particularly pronounced employment probability returns, whilst the cohorts added to the data exhibit more typical (lower) premiums for Full Level 2 qualifications. The 2007/2008 cohort do not exhibit such exceptional earnings premiums. As a result, a reduction in the weight given to returns from the 2007/2008 cohort has little impact on earnings estimates, but has a marked impact on employment probability premiums.
- The difference between B and C is mostly due to the addition of 3 extra controls (a variable derived from SSA, previous benefit status, and previous employment status)

for the earnings equation, and only the addition of previous employment history for the employment probability equation makes a difference, as we are better controlling for the importance on post-learning employment of previously having a job.

These changes account for the slight reduction in estimated earnings returns between the previous estimates and those in the Tables across Sections 3, 4 and 5.

Difference-in-differences and Matching

As suggested in the introduction, whilst these changes improve the accuracy of our estimated labour market returns, they are still based on standard regression approaches. In Section 6 estimates obtained from an achiever v non-achiever comparison of earnings returns for NVQ2 and NVQ3 learners are subjected to a rigorous econometric investigation, to see if our findings contain a bias as a result of unobservable individual characteristics that are correlated with both achievement and earnings.

For instance, it is possible that unobservable motivation or commitment is important in differentiating achievers and non-achievers, so that achievers would potentially receive a higher return whatever their level of qualification. From a method perspective, there are a number of such 'unobservables' that can be assumed broadly constant over the period of our analysis; together with some, such as divorce and job loss which are time-varying in their potential (unobservable) impacts.

Section 6 sheds some light on the validity of our assumption that both time-varying and time-invariant unobservables are not biasing our regression estimates; and that even when we subject our results to more advanced econometric techniques that attempt to compensate for these possible problems, they remain substantially higher than estimates in the survey-based literature. Section 6 focuses analysis on the earnings returns of NVQ2 and NVQ3 learners because these are the qualification categories most commonly analysed in the existing academic literature (unfortunately the academic literature that analyses survey data does not adopt our 'Level' and 'Full Level' categorisations). Furthermore, the returns to NVQ2 and NVQ3 returns in the survey-based literature are those that have caused most concern (suggesting that they are qualifications suffering most from unobservable impacts). Results from Section 6 that support regression-based achiever v non-achiever approaches to estimation of returns to these qualifications, provide the required quality assurance for the results presented in Sections 3, 4 and 5 of this report.

The more advanced techniques we use are Coarsened Exact Matching (CEM) and difference-in-differences methods. First, matching methods account for any differences in characteristics between treated and control by matching each treated individual (achiever) to one or more control individual (non-achiever); who are as similar as possible with respect to a given set of pre-treatment variables. Matching methods mainly rely on two crucial assumptions. First, the conditional independence assumption (CIA), which assumes that all the relevant differences between achievers and non-achievers are captured in their observable attributes. Second, the common support assumption, assumes that every achiever has at least one counterpart in the control group. In recent years, a number of papers have highlighted the misapplication of matching methods by some researchers. As a result, a new class of matching methods has emerged - dubbed

"monotonic imbalance bounding (MIB)" (see Iacus et al., 2011) - that curtails the misuse of these techniques.

We implement one of these MIB methods, using coarsened exact matching (CEM). The idea of CEM is to temporarily 'coarsen' each conditioning variable into meaningful categories[13]; match exactly on these 'coarsened' data, and then retain only the original (uncoarsened) values of the matched data. If different numbers of treated and control units appear in different strata, the econometric model must weight or adjust for the different stratum sizes. This is why a weighted regression of the dependent variable on the covariates is adopted at the end of the matching procedure.[14] Iacus et al. (2011) show that the CEM dominates commonly used existing matching methods in its ability to reduce imbalance, model dependence, estimation error, bias, variance, mean square error, and other criteria.

It is important to remember that the inherent trade-off of matching remains. With the CEM approach, larger bins (more coarsening) will result in fewer strata; fewer strata will result in more diverse observations within the same strata and, thus, higher imbalance (Blackwell et al., 2009). As recognised by Ho et al. (2007), matching methods are data-preprocessing techniques and analysts must still apply statistical estimators to the data after matching.

In this paper we combine the CEM with the difference-in-differences estimator (DiD). By using a difference-in-differences approach we go a long way to accommodate unobservable impacts that are time invariant – for instance, if there are systematic differences in the level of motivation/commitment between achievers and non-achievers and these are constant through time, the diff-in-diffs estimates should not be impacted. However, the validity of this assertion is dependent on the assumption of parallel earnings trends[15] between achievers and non-achievers. If we find that earnings strands of achievers and non-achievers are not parallel prior to learning, the use of matching can 'force' the prior earnings of achievers and non-achievers to be parallel by matching on earnings in the two periods prior to learning.

The main matching hypothesis is now stated with respect to the before-after evolution instead of levels (Blundell and Costa Dias, 2009).

The CEM-DiD estimator is given by:

$$ATT \equiv \frac{1}{N_1} \sum_{i=1}^{N_1} ((Y_{1it} - Y_{0it'}) - \sum_{j=1}^{N_0} w_{ij} (Y_{0jt} - Y_{0jt'}))$$

---

[13] For instance, if we are matching on previous earning we may match on data that has been 'coarsened' by putting earnings into quartiles.

[14] Selecting matched samples reduces bias due to covariate differences, and regression analysis on those matched samples can adjust for small remaining differences and increase efficiency of our estimates (Stuart and Rubin, 2007).

[15] Diff-in-diffs compares the gain in earnings experienced by the treated, with the gain in earnings achieved by the control (before and after treatment). For this to provide a valid estimate of the counterfactual, we must assume that in the absence of the treatment, the gap between treatment and control earnings would be the same before and after treatment (hence, 'parallel trends'). Using two time periods of earnings previous to treatment we also adopt an extension of the diff-in-diffs estimator based on the weaker 'parallel growth rate' assumption.

where t and t' are time periods after and before the program start date, N_1 is the number of achievers, N_0 the number of non-achievers and w_{ij} the weight given to the jth non-achiever outcome difference.

This project to update and enhance estimates of the labour market returns to English Further Education has included a process of academic peer review. Four academics, expert in the area of education economics, have provided comment on the methods used, with a particular focus on the analysis carried out in Section 6 to check the robustness of our basic regression results using the example of NVQ2 and NVQ3 qualifications. We are very grateful for the comments provided and any remaining issues are the sole responsibility of the authors. For detail of our responses to reviewer comments, please see the final pages of this report.

3. Estimated returns for aggregated qualification categories

This section details the findings from analysis of the earnings, employment probability, and likelihood of being on active benefits, for achievers compared to non-achievers with the same learning aim, in each of the categories of learning discussed in Section 2.1. Section 3.1 details the estimated earnings premiums; Section 3.2 the employment probability premiums and Section 3.3 differences in the likelihood of being on active benefits. Finally, Section 3.4 sets out the possible variation to these headline findings when we consider returns separately for men and women, and for different age groups.

3.1 Earnings

Key Findings: The earnings premiums in Table 3 are obtained from 24 separate regressions run for each of our aggregated FE qualification categories. We find that all qualification categories provide positive and statistically significant earnings premiums, with 3 to 5 year average returns for Full Level 2 of 11.3%; Full Level 3 of 8.5% and Level 4+ of 8.4%. We observe smaller (but still statistically significant) earnings premiums of 1.9% for those achieving a qualification Below Level 2; 1.3% for Level 2 and 3.3% for Level 3. These 3 to 5 year averages are a good reflection of returns that remain relatively stable over the period of analysis; with the exception of Level 3 and Full Level 3. The possible confounding effects of unobservable moves to HE mean our estimates of 3.3% for L3 and 8.5% for FL3 slightly understate the true return to these qualifications[16].

These 3 to 5 year averages are constructed from regression estimates of returns across five years from the end of learning. For instance, the figure of 0.021 in the top left-hand corner of Table 3 is a coefficient obtained from a regression equation for individuals whose highest qualification aim, across all their learning spells, is Below Level 2. Amongst this 1,448,606 whose highest qualification aim is Below Level 2, there are 1,018,354 who achieve, and 430,252 who fail to achieve, this stated aim[17]. The coefficient of 0.021 is an estimate of the additional earnings that achievers receive in the first tax year after the learning spell ends, compared to the earnings of those who fail to achieve the aim, controlling for a variety of additional factors in our regression equation[18]. Those achieving

---

[16] For a more detailed explanation of why, see point d. in the section on *The Impact of, and motivation behind, Methodological Enhancements* (page 24).

[17] Readers will note that the numbers here are lower than those cited in the corresponding section of B&U, because we previously cited total numbers of learners in, for instance, the Below Level 2 category – here we cite the numbers within this category for whom we observe earnings returns. Similarly, in the remainder of this study, we report learner numbers that are linked to the specific regression equations estimated.

[18] As suggested in the Data and Method section, we control for sex; age; interaction sex-age; ethnicity; disability; region; type of funding (none, LCS, ESF, both); mode of attending (FT/PT); offender; spell duration; number of previous FE learning spells; prior education level; year dummies; HE flag; Index of Multiple Deprivation (IMD); Indicators derived from Sector Subject Area (SSA); the number of days an individual was on active benefits in the year before learning; whether an individual has an inactive benefit spell in the year before learning; and how many days an individual has spent in sustained (6 months) employment just before learning.

a qualification *Below Level 2* earn, on average, 2.1% more than those who do not achieve this stated highest aim in the first year after the learning spell ends.

The figure of 2.1% is therefore our estimate of the difference between (i) the post-qualification earnings of individuals who obtain Below Level 2, compared to (ii) the earnings of the same individuals if they had not taken this qualification (otherwise known as the 'counterfactual'[19]). Moving along this first row of Table 3 the results of three more regression equations, estimated for this group of achievers and non-achievers, suggests that this earnings premium stays relatively stable between the first and fifth year after the end of learning; with only a slight dip to 1.6% in the fourth year. As with all the tables in this report, the final column of Table 3 presents the earnings return averaged over the period three to five years on from learning. As suggested in Section 6.2, we consider this to be the most relevant policy metric and here there is an average 1.9% earnings return per annum for those who secure a qualification Below Level 2.

It is important to note that the return in the first year after learning will, necessarily, include all cohorts of learners (because we observe a first year of earnings for even those whose learning spell finishes in 2010-2011). However, by the fifth year after learning our estimate is based on learners who complete prior to the 2007/2008 academic year (as these are the only learners for whom we observe five or more years of earnings). This is an issue to which we return in Section 6.2 and the Conclusion, where we consider the 'currency' of our findings – for now it is sufficient to say that the findings do not suggest any systematic decline in earnings returns over the years. Similarly, on page 22 we have already described how the results in Table 3 cover those achievers and non-achievers for whom we have earnings data in at least one year after the end of learning[20].

Table 3: Daily earnings premium of achievers relative to non-achievers

| Achievement | Percentage Log Daily Earnings Premium in Tax Year after Spell End | | | | |
	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.021***	0.020***	0.016***	0.020***	0.019
Level 2	0.004*	0.007**	0.013***	0.020***	0.013
Full Level 2	0.143***	0.118***	0.114***	0.107***	0.113
Level 3	-0.005*	0.005	0.038***	0.057***	0.033
Full Level 3	0.063***	0.052***	0.086***	0.118***	0.085
Level 4+	0.089***	0.088***	0.086***	0.077***	0.084

*** significant at the 0.1% level; ** 1% and * 5%

In all tables included in this report, the Standard Errors associated with each parameter estimate have been removed to aid exposition. Readers can access the full Tables, including standard errors via the Technical Annex.

Nearly all the results in Table 3 are at least statistically significant at the 5% level and the majority at the 0.1% level (i.e. 99.9% level of confidence) and we are therefore confident of

---

[19] Called the 'Counterfactual', because it is 'counter' to the 'factual' state of the world. We never observe the counterfactual, but use various econometric estimation techniques to get as close an estimate as possible.
[20] The issue of selection into employment (or earnings) in each year is one that we consider in our response to reviewers comments, detailed at the end of this report.

a strong correlation between achievement of a qualification aim and higher earnings. Most of the coefficients in this report are significant, because of the exceptionally large numbers included in our regression equations. Returning to Table 3, we can see that the estimated earnings premium for achievers in our category of Level 2 qualification ranges from less than 1% in the first year after the end of learning, to only 2% by the fifth year. This is in stark contrast to the estimated premium for achievers v non-achievers whose highest aim is a Full Level 2 qualification, which goes from 14.3% in year 1 to 10.7% in year 5. This splitting of Level 2 aims into two categories leaves us with 906,363 people whose highest aim is Level 2 (260,224 were Non-Achievers) and 1,521,339 with Full Level 2 as their highest aim (336,929 of them did not achieve any qualification).

Within our category of Full Level 2 we have individuals who are aiming for BTEC, RSA, City and Guilds qualifications that are equivalent to 5 GCSEs at grades A* to C and it is worth noting that the existing literature suggests an amount of heterogeneity in returns across these different qualification types. Studies suggest a similar amount of heterogeneity in the estimated returns for qualifications gained at NQF Level 3 and we attempt to capture some of this by differentiating between 'Full' and other Level 3 aims. In the Full Level 3 category we have approximately 1.12 million achievers and 220 thousand non-achievers; compared to 600 thousand achievers and 200 thousand non-achievers in the Level 3 category. The estimated earnings premium for Full Level 3 qualifications varies from 6.3% in the first year after learning, to 11.8% in the fifth year. This compares to an earnings premium for Level 3 that starts at less than 1% in the first year and rises to 5.7% by the fifth year; and for Level 4+ there is an annual estimated return that remains relatively stable, averaging 8.4% between the 3rd and 5th years after learning.

Table 4 sets out the average daily earnings across our achiever and non-achiever groups from the 'raw' data[21]. Generally, we find that those attempting Full Level 2 or Full Level 3 qualifications have lower average post-learning wages (whether they are achievers or non-achievers) compared to those in our Level 2 and Level 3 qualification categories. We might speculate that a larger proportion of Level 2/3 learners are in full-time employment before, during and after learning, as they have highest learning aims that demand less of their time; in contrast to Full Level 2/3 learners who are perhaps taking qualifications at a point before they have entered the labour market for the first time. However, the picture seems more complicated, with Level 3 learners having a 58% employment rate one year before learning, compared to 38% amongst Full Level 3 learners; but the comparable figures for Level 2 and Full Level 2 are 57% and 56% respectively[22]. Also, as was the case in B&U, we see intermediate achievers (i.e. those achieving a qualification at a lower level than their highest aim) having raw average earnings that are almost everywhere lower than the corresponding control group (of non-achievers), something that we return to later in our discussions.

---

[21] Figures that are simply observed in the 'raw' data, with no control for potential differences in characteristics between achievers and non-achievers.
[22] Though it should be remembered that all figures here include both part-time and full-time working.

Table 4: Raw average daily earnings in year after the end of learning spell[23]

Spell Participation	Spell Achievement						
	No achieve	Below Level 2	Level 2	Full Level 2	Level 3	Full Level 3	Level 4+
Below Level 2	40.56	44.18					
Level 2	44.39	37.34	42.76				
Full Level 2	33.52	27.66	29.11	39.78			
Level 3	43.37	34.88	37.45	29.53	45.33		
Full Level 3	32.76	28.52	27.81	28.53	26.16	31.43	
Level 4+	50.10	45.43	44.01	42.44	43.43	35.39	52.27

3.2 Employment

Key Findings: Table 5 reports the employment probability premiums for our categories of educational achievement obtained from estimation of 30 separate regression equations. The highest 3 to 5 year average return accrues to Full Level 3 achievers, who have employment rates that are 3.8 percentage points higher than non-achievers. Those achieving a FL2 qualification have employment rates 1.5 percentage points higher than the comparable non-achievers.

As we shall see in Section 6.2, the FL2 employment premium seems higher for post-2007 cohorts of learners and this figure of 1.5 percentage points may more likely be 2 percentage points for cohorts achieving in the present environment. Similarly, the employment premium for FL3 has become more pronounced for post-2007 cohorts and may now be closer to 5 percentage points for contemporaneous cohorts. The figures for L2 and L3 are 0.9 and 1.4 percentage points respectively. The figure of 0.9 for L2 is a little better than it seems, because L2 achievers and non-achievers have lower absolute levels of employment (and therefore small percentage point differences amount to larger percentage differences). However, even the low levels of employment amongst Below Level 2 learners, does not alter the suggestion that the size of any employment effect (0.3 of a percentage point) is very small.

Table 5 presents findings from the individual regression results that are used to calculate these 3 to 5 year averages. For instance, Table 5 suggests that in the fifth year after learning those who achieve a FL3 qualification have an employment rate 3.3 percentage points higher than those who have the same highest qualification aim but do not achieve this. This estimated employment probability premium remains relatively stable, averaging 3.8 percentage points between the 3rd and 5th tax year after learning. Whilst this estimated premium is significant, it does not seem particularly large[24]. However, from Table 6 we can see that the 'raw' average employment rates of achievers and non-achievers within this

---

[23] These averages are lower than those presented in B&U as we have removed the top one per cent of outliers.

[24] A finding that is highly statistically significant can be small in magnitude. There may be a strong correlation, with a large proportion of those achieving a qualification *Below Level 2* being observed with higher employment probabilities. However, the difference in employment probabilities does not have to be particularly large.

qualification aim, one year after the end of a learning spell, are 69% and 61% respectively. From such a base, an average 3.8 percentage point difference implies a percentage premium of approximately 5.8%[25].

Table 5: Estimated employment probability premium for achievers over non-achievers

Achievement	Percentage Point Employment probability Premium in Time Period after Spell End					
	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.0007*	0.003***	0.003***	0.003***	0.002***	0.003
Level 2	0.006***	0.009***	0.010***	0.009***	0.009***	0.009
Full Level 2	0.023***	0.018***	0.015***	0.017***	0.014***	0.015
Level 3	0.011***	0.010***	0.011***	0.016***	0.015***	0.014
Full Level 3	0.048***	0.043***	0.038***	0.043***	0.033***	0.038
Level 4+	0.012***	0.011***	0.012***	0.011***	0.011***	0.011

*** significant at the 0.1% level; ** 1% and * 5%

Table 5 presents an average employment probability premium of less than 1 percentage point for those taking qualifications Below Level 2, and an approximate one percentage-point premium for Level 2 achievers. The raw employment probabilities (Table 6) amongst individuals whose highest-aim is Below Level 2, is 41% for achievers and 37% for non-achievers one year after the end of a learning spell. For those studying Level 2 qualifications, the one percentage point premium translates into an approximate 2% premium, as the raw employment figures suggest that only just over half of achievers (54%) are employed in the year after learning (compared to 50% of non-achievers).

Full Level 2 achievers are an average 1.5 percentage points more likely to be in employment in the 3rd to 5th years after completion of learning; from a raw base of 68% of achievers in employment one year after spell end and 59% for non-achievers (Table 6). This raw employment probability difference is quite pronounced and the fact that this reduces to only 1.5 percentage points in our regression outputs, suggests that the achievers at Full Level 2 are very different to non-achievers on a number of observable characteristics; most importantly the extent to which they are employed prior to learning – which is much higher amongst achievers.

The raw employment probabilities/proportions cited in Table 6 are those that we observe one year after spell end and strictly speaking we should consider how these change up to the period 5 years on from learning (as it changes the base from which to consider percentage point changes). However, the growth of proportions in employment between the first and fifth year of learning is very similar across our categories of learner. For instance, for both achievers and non-achievers in the categories of Below Level 2, Level 2 and Level 3 we observe a growth in raw employment between the first and the fifth year, of between 9 and 11 percentage points. Raw employment proportions for FL2 non-achievers

---

[25] A 3.8 percentage point premium from a base 65% employment probability (mid way between 61% and 69%) is a 5.8 percentage premium.

and FL3 achievers deviate only slightly from this, with growth of 13 percentage points between the first and fifth year; and for FL3 non-achievers, growth is 15 percentage points.

Table 6: Raw employment probabilities one year after spell end date

Spell Participation	Spell Achievement						
	No achieve	Below Level 2	Level 2	Full Level 2	Level 3	Full Level 3	Level 4+
Below Level 2	0.37	0.41					
Level 2	0.50	0.44	0.54				
Full Level 2	0.59	0.51	0.54	0.68			
Level 3	0.57	0.53	0.56	0.64	0.63		
Full Level 3	0.61	0.59	0.59	0.67	0.57	0.69	
Level 4+	0.71	0.72	0.75	0.78	0.74	0.77	0.76

Whilst we must be careful in comparing raw figures on employment in Table 6 and similarly 'raw' earnings figures in Table 4, comparison does raise some interesting questions. First, whilst the raw average earnings of Level 2 achievers (£42.76) and non-achievers (£44.39) are above those of Full Level 2 learners (£39.78 and £33.52 respectively); the proportions of Level 2 learners in employment (54% and 50% respectively for achievers and non-achievers) are much lower than the corresponding figures for Full Level 2 learners (68% and 59% respectively). Similarly, the earnings of Level 3 achievers and non-achievers are higher than those of Full Level 3 learners; but a higher proportion of FL3 achievers and non-achievers are in employment one year after learning ends.

We can gain some insight into this issue, by considering the *Descriptive Statistics* at the end of this report. The suggestion from these descriptive statistics is that Level 2 learners are on average 5 years older than Full Level 2 learners (36.8 compared to 31.8 years old, respectively) and there is an even more pronounced age difference between Level 3 and Full Level 3 learners (33.8 compared to 23.7 years, respectively). This likely goes a long way to explain these differences in earnings and employment probabilities.

Once again we see intermediate achievers (who achieve an outcome that is lower than their stated highest aim), having relatively poor outcomes (from the raw figures) when compared to both achievers and non-achievers in the relevant qualifications category. For instance, amongst those who aim for a Full Level 2 qualification, but only achieve Level 2 or Below Level 2, there are only 54% and 51% in employment one year after, respectively. This compares to 59% amongst those who achieve no outcome (non-achievers). The only exceptions to this pattern are those who aim for a Level 4+ qualification and achieve something at a lower level.

3.3 Job-search (or 'Active') Benefits

Key Findings: Table 7 reports the results of 30 regression equations that estimate the percentage probability of achievers being on active benefits, relative to non-achievers, following the completion of learning, for each category of educational achievement. Because the absolute proportion of individuals on benefits is small, the percentage point differences reported in Table 7 can be misleading. For instance, the 3 to 5 year average for Full Level 2 learners suggests that achievers are 2.1 percentage points less likely to be observed on active benefits, but because only about 10% of non-achievers and 5% of achievers are on benefits after learning (Table 8), this 2.1 percentage point difference translates into an approximate 28% difference. The difference in active benefit probabilities for Level 2 learners is 0.6 of a percentage point (approximately 11%); 0.9 of a percentage point for Level 3 achievers over non-achievers (approximately 26%) and for Full Level 3 achievers we observe a 1.8 percentage point difference (approximately 40%). For those studying Below Level 2 we once again observe poor returns, with our 0.3 percentage point estimate only translating to a 5% lower probability of achievers being on active benefits compared to non-achievers following learning. Whilst there is some variation in the first year after learning, the coefficient estimates are all negative and significant in the three to five year period after the end of learning.

The benefits considered as 'active' in Table 7 (i.e. they are associated with job-search activity) include Job Seekers Allowance (JSA), Job Training Allowance (JTA) and Employment and Support Allowance (ESA)). The coefficients identify the percentage point differences in active benefit probability between achievers and non-achievers. Those achieving a Full Level 2 or Full Level 3 qualification have a much lower likelihood of being on active benefits, relative to comparable non-achievers. In contrast, Level 2 and Level 3 achievers are only 0.6 and 0.9 percentage points less likely to be on active benefits than non-achievers after learning.

Table 7: Estimated probability of Achievers being on *Active Benefits*, compared to non-achievers

	Percentage Point Probability of Achievers V Non-achievers being on Active Benefits					
Achievement	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.009***	0.0003	-0.003***	-0.002***	-0.003***	-0.003
Level 2	0.005***	-0.003***	-0.005***	-0.006***	-0.006***	-0.006
Full Level 2	-0.025***	-0.027***	-0.022***	-0.022***	-0.020***	-0.021
Level 3	-0.003***	-0.010***	-0.010***	-0.008***	-0.008***	-0.009
Full Level 3	-0.026***	-0.032***	-0.021***	-0.016***	-0.016***	-0.018
Level 4+	-0.005***	-0.010***	-0.008***	-0.008***	-0.009***	-0.008

*** significant at the 0.1% level; ** 1% and * 5%

In the previous section we saw how Full Level 2 and Full Level 3 achievers secure more of an employment probability premium compared to non-achievers; even though average earnings of achievers and non-achievers at Level 2 and Level 3 are higher (for those who manage to secure employment). Here the suggestion is that securing a Full Level 2 and Full Level 3 qualification leads to a lower probability of claiming an active benefit following achievement, relative to the appropriate non-achievers.

This finding remains even when we turn <u>percentage point</u> into <u>percentage</u> differences, but the difference is not as pronounced as one might think simply considering the percentage point premiums presented in Table 7. The premium for Level 2 learners translates into an approximate 11% difference and for Level 3 achievers over non-achievers we observe an approximate 26% difference[26]. This contrasts with figures of 28% and 40% for Full Level 2 and Full Level 3 learners respectively, having turned *percentage point* into *percentage* differences.

Table 8: Raw proportions on *Active Benefits* one year after the end of learning spell

Spell Participation	Spell Achievement						
	No achieve	Below Level 2	Level 2	Full Level 2	Level 3	Full Level 3	Level 4+
Below Level 2	0.07	0.05					
Level 2	0.06	0.09	0.05				
Full Level 2	0.10	0.17	0.13	0.05			
Level 3	0.04	0.08	0.06	0.12	0.03		
Full Level 3	0.06	0.08	0.08	0.11	0.07	0.03	
Level 4+	0.04	0.05	0.04	0.05	0.04	0.05	0.02

The raw proportion of non-achievers on *Active Benefits* is highest for those who attempt, but do not achieve a Full Level 2 qualification (10%), whilst those who aim for either a Level 3 or Level 4+ qualification, but do not achieve, have the lowest proportions amongst non-achievers (4%). However, it is amongst the group who achieve an intermediate qualification, having aimed for something higher, that we find particularly high benefit rates. Seventeen per cent of those who aim for a Full Level 2, but achieve Below Level 2, are on benefits in the first year after the end of learning; whilst 12% of those who aim for Level 3 but achieve Full Level 2 are on benefits one year on from learning.

Whilst these are all raw figures, Tables 4,6 and 8 suggest once again that intermediate achievers perform particularly badly when compared to either achievers or non-achievers. This is a finding that is particularly important for our 'identification strategy', as it is the first piece of evidence that suggests our non-achievers are not necessarily those who 'fail' but rather those who drop-out – an issue we return to in Section 6 and the concluding section. For now, it is worth noting that the intermediate achievers who enrolled on a FL2 course are on average much younger than FL2 Achievers (24 years old, compared to 32.8 respectively). Also the intermediate achievers who enrolled on a FL3 course are on average much younger than the FL3 Achievers (20 years old, compared to 23.5 respectively). The median duration for intermediate achievers who enrolled on a FL2 course is substantially longer than the median duration for FL2 Achievers (366 days, compared to 252 days, respectively),whilst the opposite is true for FL3 courses (where the median duration is 583 days, compared to 660).

From this discussion, it seems quite possible that the characteristics of these intermediate achievers are so different from our categories of non-achievers that in a regression

---

[26] If we take a midpoint of 3.5 from the raw figures, a 0.9 percentage point difference amounts to an approximate 26 per cent difference.

framework, controlling for these differences we might still identify a significant return to learning, compared to a situation where they secure nothing from their studies. However, there would seem to be some potential for selection effects here, with less 'able' individuals following this route of aiming high and achieving low.

3.4   Variation by sex and age

In this section of the report we see whether the estimated premiums presented above differ for men and women; and whether they vary across different age groups. The regressions estimated to this point contain controls for gender and age, but this does not account for the possibility that the same learning experience impacts differently for men and women, and across different age groups. For instance, labour market discrimination may mean that men and women engaged in exactly the same learning may secure very different labour market returns. Only by running separate regressions for men and women (or using interaction terms) will we capture any such differences. Similarly, there is clear evidence[27] that older and younger learners undertaking the same learning secure different labour market returns. Running separate equations for men and women, and for different age groups, accommodates this potential for differential impacts across our main covariates of interest[28].

Sex differences in earnings returns

Key Findings: Table 9 presents estimated earnings premiums for women that are lower than those in Table 3 at all levels of learning. For instance, the average three to five year earnings premium for female Full Level 2 achievers is 9.4% in Table 9, compared to 11.3% in Table 3. The estimated returns presented in Table 3 are underpinned by lower returns for women (Table 9) and higher returns for men (Table 10) at all levels of learning. However, these gender differences almost disappear at FL3 (where the 3 to 5 year earnings premium is 8.7% for male achievers and 8.4% for female achievers). The gender differences in earnings premia at Level 4+ are also less pronounced, and at L3 we seem to suffer a pronounced dip in early earnings returns for women, which are actually negative and significant in the first and third years after learning (-3.3 and -1.5 per cent respectively). However, five years on from the end of learning the earnings premium for L3 women achievers is only 1.5 percentage points lower than the return for their male counterparts, and we do not observe the same early dip in returns for these men.

Our estimates for women at L3 likely understate true earnings returns and the same issue may also understate our estimates of earnings returns for women achievers at FL3. Unfortunately the poor returns for women at Level 2 and Below Level 2 seem to be more accurate. Women achievers secure no significant and systematic return that allows us to quote a valid three to five year average for these categories of learner. In contrast, their male counterparts secure a 4.6% and 3.2% average return over three to five years from the end of learning, for Below Level 2 and Level 2 achievement respectively.

---

[27] Blanden, J., Buscha, F., Sturgis, P. and Urwin, P. (2012) "Measuring the Returns to Lifelong Learning", *Economics of Education Review*, Volume 31, No. 4; pp. 501-514.
[28] In the more technical terminology of the academic literature, we are attempting to control for heterogeneous treatment effects.

Our findings in Table 9 for women taking Level 3 qualifications would warrant an entire strand of research of themselves (there are 310,223 women achievers in this category of Level 3). Why do the returns for women and men at L3 behave so differently? A brief analysis of the differences between men and women taking Level 3 qualifications suggests no significant age differences, but quite substantial differences in sector subject area. Whilst 20% of men studying at Level 3 are taking qualifications related to Engineering, this is true of only 0.5% of the women; 7% of men are on courses related to Urban Planning, compared to only 0.1% of women and whilst only 6% of men are studying courses associated with 'Care', this is true of 24% of women. The differing nature of these Level 3 qualifications may mean that we miss more subsequent HE learning amongst female achievers, compared to male achievers. However, the HE flag issue is likely one of [hidden] part-time working, and it is still the case in the UK that women are much more likely to be working part-time.

It could be that many women achieving a Level 3 qualification are much more likely to move on to part-time working (whether related to HE learning or otherwise), compared to non-achievers. It may be that the apparent lower earnings return secured by women studying at Level 3 in the first years after learning, when compared to men, simply reflects the greater propensity for young women achievers to continue on to HE and work part-time, when compared to men achieving at this level.

This may also understate our initial estimates of earnings returns for women achievers at FL3. Whilst our estimates suggest that women secure much lower returns to L3 and Full Level 3 in the early years after FE Learning, they close this gap by the fifth year. By this time the premium for women at Level 3 is 4.8%, compared to 6.3% for men and for Full Level 3 female achievers the premium of 12.2% has become larger than the corresponding figure for men (11.3). For a more detailed explanation of the HE flag issue, see point d. in Section: *The Impact of, and motivation behind, Methodological Enhancements* (page 24).

We return to discuss this issue in the conclusion, when we comment on the robustness of our estimated returns in the report as a whole. Unfortunately, this issue is unlikely to explain the poor returns for women at Level 2 and Below Level 2, where the insignificant returns seem to be an accurate representation. This is of great concern, given that we have between 615 and 382 thousand women in our *Level 2* equations[29] and Table 9 suggests they receive no earnings benefit compared to non-achievers.

---

[29] These figures relate to the numbers included in each of our four separate equations estimated for each year following learning aim, with year 1 always having more observations, relative to year 5, because of the nature of the data (we have fewer individuals with learning and earnings records over the full 5 years that we estimate earnings returns).

Table 9: Daily earnings premium of female achievers relative to non-achievers

| Achievement | Percentage Log Daily Earnings Premium in Tax Year after Spell End | | | | |
	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	-0.0002	0.0008	-0.005	0.002	-0.001
Level 2	-0.004	-0.006*	0.003	0.010**	-0.004
Full Level 2	0.120***	0.097***	0.097***	0.087***	0.094
Level 3	-0.033***	-0.015***	0.025***	0.048***	0.019
Full Level 3	0.045***	0.040***	0.090***	0.122***	0.084
Level 4+	0.084***	0.079***	0.075***	0.063***	0.072

*** significant at the 0.1% level; ** 1% and * 5%

Table 10 identifies statistically significant positive returns to men across all categories of learning, which are also substantial in magnitude[30]. Even for those studying at Level 2, where we observe negative average returns for women, estimated earnings returns for men average 3.2% between the third and fifth year after learning. We once again have a potential impact from our HE flag. Whilst we observe returns Below Level 2 remaining pretty stable across the five year period; Level 2 rising slightly and Full Level 2 falling from 16.1 to 12.5; returns to Level 3 and Full Level 3 experience a dip between the first and second year after learning and then recover up to the fifth year after learning. However, as we have already mentioned, this is much less pronounced than the dip we observe in the returns for women achievers.

Table 10: Daily earnings premium of male achievers relative to non-achievers

| Achievement | Percentage Log Daily Earnings Premium in Tax Year after Spell End | | | | |
	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.048***	0.046***	0.046***	0.046***	0.046
Level 2	0.022***	0.028***	0.030***	0.038***	0.032
Full Level 2	0.161***	0.135***	0.129***	0.125***	0.130
Level 3	0.028***	0.026***	0.050***	0.063***	0.046
Full Level 3	0.088***	0.068***	0.080***	0.113***	0.087
Level 4+	0.092***	0.098***	0.100***	0.095***	0.098

*** significant at the 0.1% level; ** 1% and * 5%

As was the case in Table 4, Table 11 suggests that the earnings of female achievers and non-achievers in the Level 2 and Level 3 categories of learning aim are higher than those for Full Level 2 and Full Level 3. It would seem that our control group in the former categories have 'relatively' high daily earnings and it may be harder to secure a further return from these additional qualifications. However, this still does not explain the relatively poor returns of women, as the same situation is apparent for men in table 12, where

---

[30] It is of some interest that the male equations are a much better fit (in terms of R-square) than the female equations [R_square is 0.26 for men and only 0.17 for women].

earnings of achievers and non-achievers are lower in our 'Full' categories. The low estimated returns to women at Level 2 and Below Level 2 in a framework where we expect (if anything) estimates to be biased upwards, is a real cause for concern.

Table 11: Raw average daily earnings one year after the end of learning spell for women

Spell Participation	Spell Achievement						
	No achieve	Below Level 2	Level 2	Full Level 2	Level 3	Full Level 3	Level 4+
Below Level 2	35.60	36.33					
Level 2	38.40	33.58	36.84				
Full Level 2	27.83	24.19	25.69	30.60			
Level 3	37.69	31.82	33.97	25.36	37.87		
Full Level 3	29.49	26.75	26.20	25.85	24.54	27.26	
Level 4+	46.04	41.87	40.93	39.23	40.16	34.95	47.74

Table 12: Raw average daily earnings one year after the end of learning spell for men

Spell Participation	Spell Achievement						
	No achieve	Below Level 2	Level 2	Full Level 2	Level 3	Full Level 3	Level 4+
Below Level 2	46.69	53.17					
Level 2	52.53	42.86	52.21				
Full Level 2	38.85	30.39	32.70	47.21			
Level 3	50.42	39.64	42.81	32.82	54.86		
Full Level 3	37.08	31.01	29.91	31.64	28.03	37.72	
Level 4+	55.00	50.93	49.44	46.60	48.45	36.01	59.53

Sex differences in employment probability premiums

Key findings: Tables 13 and 14 suggest that women secure more of a three to five year average employment premium from learning at all levels, when compared to men. For instance, the highest 3 to 5 year average return accrues to women Full Level 3 achievers (Table 13), who have employment rates that are 4.3 percentage points higher than non-achievers. Men achieving a FL2 qualification have employment rates 0.5 of a percentage point higher than non-achievers (Table 14), whilst the comparable figure for women achievers is 2.6 percentage points. The lowest employment returns for both men and women are seen at Below Level 2, where men seem to have no significant employment return; and women secure a 0.4 percentage point premium.

These differences are confirmed when we check the raw employment figures, as the raw percentages of men and women in employment is roughly equivalent in both Tables 15 and 16. Considering earnings results for women from the previous section and the findings here, female earnings returns may be understated because many work part-time and we are not able to fully identify this in our data – women working part-time will seem as if they

are achieving a lower wage. In this section of the report women do better, partly because part-time and full-time employment are given an equal weight in our analysis.

The fact that women's 3-5 year average employment probability premiums in Table 13 are all higher than those for men in Table 14 (apart from the returns estimated for men and women at Level 4+), seems driven by a divergence in employment premiums later on in our period of analysis, when considering *Full Level* qualifications. For instance, at Full Level 2 employment returns in the first year after learning are 2.5 for women and 1.1 percentage points for men, but by the fifth year whilst this is still 2.5 for women, the figure has fallen to 0.2 of a percentage point for men. Similarly, first year returns at Full Level 3 are 4.6 for women and 4.2 for men in the first year after learning, but by the fifth year these are 4 for women and only 2.5 percentage points for men.

Table 13: Estimated employment probability premium for female achievers over non-achievers

	Percentage Point Employment probability Premium in Time Period after Spell End					
Achievement	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.001**	0.003***	0.004***	0.004***	0.003***	0.004
Level 2	0.007***	0.009***	0.011***	0.010***	0.010***	0.010
Full Level 2	0.027***	0.025***	0.026***	0.026***	0.025***	0.026
Level 3	0.010***	0.009***	0.012***	0.018***	0.017***	0.016
Full Level 3	0.048***	0.046***	0.041***	0.049***	0.040***	0.043
Level 4+	0.011***	0.012***	0.013***	0.011***	0.011***	0.012

*** significant at the 0.1% level; ** 1% and * 5%

When comparing earnings and employment returns between men and women there seems to be a continuing issue with the data, in terms of (i) the coverage of our HE flag and (ii) the (sometimes associated) issue of part-time working, which we are not able to explicitly identify in the data. This may explain the lower earnings returns to women at Full Level 2 and above in the previous section, which are compensated in this section by improved employment premiums. In contrast, the pattern of falling returns for men studying *Full* qualifications, between the first and fifth year in Table 14 could be due to a number of factors. We have already mentioned the possibility of cohort effects and it could be that such cohort effects are much more pronounced for men – if male achievers in more recent cohorts have better employment performance than non-achievers, this would explain the drop-off in returns seen in Table 14[31]. Section 6.2 gives more detail on this cohort analysis, but without a specific dedicated cohort study, we are not able to dig further into the detail of this particular gender difference.

---

[31] Remember that all cohorts contribute to the estimate of year 1 returns, but by the fifth year, only cohorts prior to 2007/2008 have enough post-learning observations to contribute to this estimate.

Table 14: Estimated employment probability premium for male achievers over non-achievers

	Percentage Point Employment probability Premium in Time Period after Spell End					
Achievement	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.0003	0.003***	0.002**	0.001	0.0002	0.001
Level 2	0.006***	0.009***	0.008***	0.007***	0.008***	0.008
Full Level 2	0.020***	0.011***	0.005***	0.007***	0.002*	0.005
Level 3	0.013***	0.011***	0.009***	0.011***	0.010***	0.010
Full Level 3	0.050***	0.042***	0.035***	0.036***	0.025***	0.032
Level 4+	0.015***	0.011***	0.011***	0.012***	0.011***	0.011

*** significant at the 0.1% level; ** 1% and * 5%

The comparison of differences in percentage point premiums between men and women in Tables 13 and 14 remain valid as the raw employment proportions in Tables 15 and 16 suggest that they translate into similar percentage differences. For instance, at Level 3 the three to five year average premium for women is 1.6 percentage points, compared to 1 percentage point for men. Considering the relevant percentage employment rates from Tables 15 and 16, these percentage point premiums translate into approximately 2.7 per cent for women and 1.6 per cent for men.

Table 15: Raw employment probabilities one year after spell end date for women

Spell Participation	Spell Achievement						
	No achieve	Below Level 2	Level 2	Full Level 2	Level 3	Full Level 3	Level 4+
Below Level 2	0.38	0.40					
Level 2	0.51	0.47	0.55				
Full Level 2	0.59	0.50	0.55	0.70			
Level 3	0.56	0.55	0.57	0.65	0.62		
Full Level 3	0.62	0.59	0.61	0.68	0.59	0.71	
Level 4+	0.71	0.73	0.76	0.79	0.74	0.78	0.77

Table 16: Raw employment probabilities one year after spell end date for men

Spell Participation	Spell Achievement						
	No achieve	Below Level 2	Level 2	Full Level 2	Level 3	Full Level 3	Level 4+
Below Level 2	0.36	0.42					
Level 2	0.49	0.41	0.51				
Full Level 2	0.59	0.52	0.53	0.66			
Level 3	0.58	0.51	0.54	0.63	0.64		
Full Level 3	0.60	0.58	0.57	0.67	0.56	0.67	
Level 4+	0.70	0.72	0.75	0.78	0.74	0.76	0.75

Age differences in earnings premiums

Key Findings: Tables 17 and 18 present estimated earnings premiums separately for 19 to 24 year olds, and those aged 25 and above. Estimated earnings premiums for the 19-24 age group (Table 17) are higher than those for the 25+ age group (Table 18), whichever qualification category we consider. For instance, amongst learners aged 19 to 24 with a highest aim of Full Level 2, achievers secure a 3 to 5 year average earnings return of 10.4%; and at Full Level 3 the figure is 10.2%. In contrast, achievers aged 25+ secure an 8.1% 3 to 5 year premium over similar aged non-achievers at Full Level 2 and 9.6% at Full Level 3. The majority of learners are located in the 25+ age group (between 4.22 and 2.06 millions) but this still leaves between 1 and 0.47 millions in our regressions for 19-24 year olds. It is a concern that achievers, amongst the 875 thousand Level 2 learners aged 25+, secure a negligible premium on average.

Considering the trajectory of these premiums from the first to the fifth year after learning, for the 19-24 year old group we observe premiums steadily increasing for Level 2, Level 3 and Full Level 3. In contrast, returns for Full Level 2 fall from 16.1 to 8.8 between the first and fifth years after learning. At Level 3 the estimated returns for 19 to 24 year olds actually move from negative and significant (-3.5%) in the first year, through insignificance in the third year and end with a significant positive return (9.9%) in the fifth year after completion of learning. A similar pattern is apparent for the 25+ age group, where we see rising returns for Level 2, Level 3 and Full Level 3 qualifications as we move from the first to the fifth year after learning. It is of some concern that achievers amongst the 875 and 532 thousand *Level 2* learners aged 25+ (in our first and last regression of Table 18, respectively) secure a negligible premium on average. We need to remember that amongst this group there is potential for this average to hide an amount of heterogeneity, with some securing a more positive return. But the suggestion is that a large proportion will be securing no return.

Table 17: Daily earnings premium of 19-24 year old achievers relative to non-achievers

	Percentage Log Daily Earnings Premium in Tax Year after Spell End				
Achievement	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.042***	0.053***	0.047***	0.056***	0.052
Level 2	0.026***	0.043***	0.049***	0.064***	0.052
Full Level 2	0.161***	0.126***	0.098***	0.088***	0.104
Level 3	-0.035***	0.010	0.071***	0.099***	0.060
Full Level 3	0.068***	0.083***	0.100***	0.122***	0.102
Level 4+	0.100***	0.113***	0.124***	0.106***	0.114

*** significant at the 0.1% level; ** 1% and * 5%

Table 18: Daily earnings premium of 25+ year old achievers relative to non-achievers

Achievement	Percentage Log Daily Earnings Premium in Tax Year after Spell End				
	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.017***	0.017***	0.012***	0.014***	0.014
Level 2	0.005*	0.005*	0.008***	0.011***	0.008
Full Level 2	0.107***	0.085***	0.083***	0.076***	0.081
Level 3	0.007***	0.016***	0.032***	0.040***	0.029
Full Level 3	0.085***	0.087***	0.098***	0.102***	0.096
Level 4+	0.075***	0.071***	0.066***	0.060***	0.066

*** significant at the 0.1% level; ** 1% and * 5%

Finally, we can see from Table 17 that amongst the 19 to 24 age group achievers at Level 4+ secure a particularly high premium, ranging from 10% to 12.4% over the five years after learning. This is a relatively small group of learners [64,000 to 32,000], when compared to the 227,000 to 106,000 *Level 4+* learners in our equation for the 25+ year age group; where we see achievers securing much lower (but still respectable) returns of between 6% and 7.5% at various points during the five years after the end of learning.

Age differences in employment and active benefit outcomes

Key Findings: Tables 19 and 20 set out the percentage point employment probability premiums across our two age groups. Comparison of estimated returns in the two tables shows some small differences with, for instance, the 25+ group of achievers having estimated employment probability premiums (of 1.8 percentage points) at Full Level 2 and Full Level 3 that are higher than those secured by achievers aged 19-24 (0.2 and 1.3 percentage points, respectively). When we consider the proportions in employment in the two groups and turn these percentage point premiums into percentages, these differences remain. For our two age groups Tables 21 and 22 report the estimated probability of achievers being on active benefits, relative to non-achievers, following the completion of learning. Any differences disappear when we consider the raw figures and change *percentage point* into *percentage* figures.

Other than the employment probability differences in our *Full Level* learning categories, there is only one other difference that stands out in Tables 19 and 20 when comparing returns - that is the very small estimated employment premium for achievers aged 19 to 24 who secure a Full Level 2 qualification, which in the third and fourth years after learning is insignificant. Whilst the equivalent premium for those aged 25+ is only one percentage point (3 to 5 year average), it is pretty stable from year one to five and always significant.

Table 19: Estimated employment probability premium of 19-24 year old achievers relative to non-achievers

Achievement	Percentage Point Employment probability Premium in Time Period after Spell End					
	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.0002	0.003**	0.003**	0.003**	0.003	0.003
Level 2	0.012***	0.014***	0.017***	0.016***	0.016***	0.016
Full Level 2	0.016***	0.005***	0.0003	0.002	0.004*	0.002
Level 3	0.008***	0.005**	0.015***	0.021***	0.025***	0.020
Full Level 3	0.027***	0.011***	0.009***	0.013***	0.017***	0.013
Level 4+	0.010***	0.006*	0.008**	0.010**	0.012***	0.010

*** significant at the 0.1% level; ** 1% and * 5%

Table 20: Estimated employment probability premium of 25+ year old achievers relative to non-achievers

Achievement	Percentage Point Employment probability Premium in Time Period after Spell End					
	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	-0.0003	0.002***	0.003***	0.004***	0.003***	0.003
Level 2	0.005***	0.009***	0.010***	0.010***	0.010***	0.010
Full Level 2	0.014***	0.010***	0.010***	0.014***	0.019***	0.018
Level 3	0.002**	0.003***	0.005***	0.010***	0.012***	0.009
Full Level 3	0.017***	0.009***	0.011***	0.020***	0.022***	0.018
Level 4+	0.008***	0.009***	0.010***	0.009***	0.008***	0.009

*** significant at the 0.1% level; ** 1% and * 5%

Tables 21 and 22 report the estimated probability of achievers being on active benefits, relative to non-achievers, across our two age groups. At all levels of learning, apart from Below Level 2, 19 to 24 year old achievers have a lower percentage point probability of being on active benefits, compared to non-achievers of the same age; when compared to our 25+ estimated achiever v non-achievers where the gap is much less pronounced. However, this gap disappears when we consider percentage differences. For instance, at Full Level 2 the three to five year average for those aged 19 to 24 of 2.1 percentage points translates into an approximate 25% difference if we consider the base proportions on benefits. For those aged 25+ who achieve a Full Level 2 qualification the 1.2 percentage point premium over similar non-achievers of the same age also translates into a 25% premium.

Table 21: Estimated probability of 19-24 year old Achievers being on *Active Benefits*, compared to non-achievers

| Achievement | Percentage Point Probability of Achievers V Non-achievers being on Active Benefits | | | | | |
	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.008***	-0.003**	-0.002*	-0.001	0.0001	-0.001
Level 2	0.003**	-0.003**	-0.006***	-0.004***	-0.007***	-0.006
Full Level 2	-0.041***	-0.035***	-0.023***	-0.022***	-0.017***	-0.021
Level 3	-0.013***	-0.017***	-0.011***	-0.008***	-0.008***	-0.009
Full Level 3	-0.039***	-0.037***	-0.018***	-0.013***	-0.015***	-0.015
Level 4+	-0.005**	-0.011***	-0.011***	-0.009***	-0.011***	-0.010

*** significant at the 0.1% level; ** 1% and * 5%

The trajectory of returns is also similar for both age groups in Tables 21 and 22. At all points in both tables, achievers are less likely to be on active benefits than non-achievers. However, this gap gets smaller with time for Full Level 2 and Full Level 3 achievers, relative to non-achievers, between the first and fifth year after learning. For instance, in the first year after learning, Full Level 2 achievers aged 19 to 24 have a probability of being on benefits that is 3.5 percentage points lower than non-achievers – by the fifth year there is only a 1.7 percentage point difference. In contrast, at Level 2 there is an increase in the gap between achievers and non-achievers, over the years for both age groups. The only difference between age groups is the tendency for the gap between achievers and non-achievers to narrow over the years for 19 to 24 year olds achieving a Level 3 qualification; compared to the situation for 25+ year olds, where size of the gap remains constant.

Table 22: Estimated probability of 25+ year old Achievers being on *Active Benefits*, compared to non-achievers

| Achievement | Percentage Point Probability of Achievers V Non-achievers being on Active Benefits | | | | | |
	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.009***	0.002***	-0.0008**	-0.0008**	-0.002***	-0.001
Level 2	0.006***	-0.001**	-0.003***	-0.004***	-0.004***	-0.004
Full Level 2	-0.021***	-0.021***	-0.014***	-0.012***	-0.009***	-0.012
Level 3	-0.002***	-0.006***	-0.007***	-0.005***	-0.006***	-0.006
Full Level 3	-0.020***	-0.020***	-0.011***	-0.009***	-0.009***	-0.010
Level 4+	-0.005***	-0.009***	-0.007***	-0.007***	-0.008***	-0.007

*** significant at the 0.1% level; ** 1% and * 5%

4. Estimated returns for disaggregated qualification categories

In this section of the report we set out the variation in estimated premiums for more disaggregated categories of learning aim. In Section 4.1 we begin by setting out the results for categories of qualification, differentiated according to whether they can be considered as (i) *Academic*, (ii) Vocational and [mainly] Classroom-Based Learning (*CBL*) or (iii) Vocational and delivered via Workplace-Based Learning (*WBL*). This final category excludes Apprenticeship learners, as Section 4.2 focuses specifically on the premiums secured by achievers amongst *Apprentices*. In Section 4.3 we describe an analysis that helps identify the relevant premiums for those who achieve *L2 Maths and English* qualifications[32].

4.1 Premiums estimated separately for academic and vocational qualification categories

Differences in earnings premiums

Key Findings: Tables 23, 24 and 25 present estimated earnings premiums separately for our three categories of learner studying in settings that are (i) Academic (ii) Vocational Classroom Based (CBL) and (iii) Vocational Workplace-Based (WBL). We find that 96% and 91% of learners whose highest aim is Below Level 2, and L2 respectively, are in our Vocational CBL category. Therefore the Below Level 2 and L2 earnings premiums of 1.5% and 1% in Table 24 are very similar to the 1.9% and 1.3% three to five year averages previously seen in Table 3. Learners in our Full Level 2 category are more evenly spread across the three categories - in Table 23 the average 3-5 year average return for Academic Full Level 2 is 11.6%, in Table 24 it is 10.3% for Vocational CBL and in Table 25 9.6% for Vocational WBL. At both Level 3 and Full Level 3 we observe negative and significant returns to learning in the first and third years for our Academic category. This re-enforces continuing concern over unobserved moves to HE, as it does not seem to have such an impact in Table 24 where the earnings premium for Vocational CBL Level 3 achievers rises steadily from 2% to 3.6% (giving a 3 to 5 year average of 2.9%). Similarly, Full Level 3 achievers in Table 24 secure a 3-5 year average premium of 5.8%, with the underlying estimated return rising from 1.9 to 9.1 over the five-year period. The return to Vocational WBL Full Level 3 Achievers averages 9% between 3 and 5 years (Table 25).

The figures in Table 23 do not include Academic Below Level 2 (or Level 4+) because numbers are too small. We have approximately 1.2 to 0.9 million observations[33] in our Below Level 2 equations in the Vocational CBL category (Table 24), and this constitutes 96% of all learners whose highest aim is Below Level 2. The remaining 4 per cent of learners whose highest aim is Below Level 2 are in the Vocational WBL category (between

---

[32] In the previous report of B&U these were referred to as *Skills for Life* qualifications.
[33] We provide a range for total population numbers that reflect the difference between numbers in our regressions estimated for the first and fifth years after learning.

50,000 and 24,000), and in all years after learning this group has much higher estimated returns than those studying within an FE institution – culminating in a three to five year average of 7.4% in Table 25 for Below Level 2 learners. This is encouraging but the numbers are relatively small and in this situation non-achievers, who make up approximately 10% of learners, may not be such a good match to achievers.

Within our category of Level 2 qualifications, 91% of learners are in our Vocational CBL category (between approx. 0.9 and 0.6 million); compared to approximately 74,000 to 51,000 learners who have highest aims related to Academic Level 2 qualifications. The estimated premium for Level 2 achievers who fall within our category of Vocational CBL moves from 0.4 of a per cent to 1.5 per cent over the five years following the end of learning (resulting in an average 1% effect in Table 24). Whilst we observe no statistically significant premium for those who achieve a highest aim of Level 2 in our Vocational WBL category, this is possibly due to small sample sizes (of approximately 17,000 to 3,000 observations as we move from the first to fifth year).

Table 23: Daily earnings premium for Achievers v Non-achievers in the *Academic* category of learners

	Percentage Log Daily Earnings Premium in Tax Year after Spell End				
Achievement	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Level 2	-0.009	-0.005	0.027	0.035*	0.019
Full Level 2	0.113***	0.083***	0.154***	0.112***	0.116
Level 3	-0.097***	-0.046***	0.063***	0.119***	0.045
Full Level 3	-0.296***	-0.147***	0.023*	0.135***	0.003

*** significant at the 0.1% level; ** 1% and * 5%

Learners in our Full Level 2 category are more evenly spread across the three categories that differentiate types of vocational and academic learning, and across all categories we estimate significant and substantial returns. In Table 23 the average 3-5 year average return for Full Level 2 is 11.6%, in Table 24 it is 10.3% and in Table 25 9.6%. Whilst premiums for the Vocational and WBL category in Table 25 fall from the first to the fifth year after learning, these averages tend to reflect relative stability in estimated premiums between the third and fifth year after learning.

Table 24: Daily earnings premium for achievers v non-achievers engaged in *Vocational Classroom Based Learning*

	Percentage Log Daily Earnings Premium in Tax Year after Spell End				
Achievement	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.017***	0.016***	0.013***	0.017***	0.015
Level 2	0.004*	0.006**	0.009***	0.015***	0.010
Full Level 2	0.109***	0.099***	0.114***	0.097***	0.103
Level 3	0.020***	0.022***	0.030***	0.036***	0.029
Full Level 3	0.019***	0.020***	0.063***	0.091***	0.058
Level 4+	0.084***	0.086***	0.086***	0.075***	0.082

Achievement	Percentage Log Daily Earnings Premium in Tax Year after Spell End				
	1st Year	3rd Year	4th Year	5th Year	3-5 year average

*** significant at the 0.1% level; ** 1% and * 5%

Considering the findings for Level 3 and Full Level 3 within our Academic category of learning, we have the strongest evidence yet that negative returns are related to our inability to observe all progression to HE from learning at Level 3. At both Level 3 and Full Level 3 in Table 23 we observe negative and significant returns to learning in the first and third years after learning – for Full Level 3 the suggestion is that achievers earn approximately 30% less in the first year after learning, compared to non-achievers. In the third year after learning, achievers are earning 15% less than non-achievers at Full Level 3, but by the fifth year this has reversed to provide a 13.5% positive premium.

Unfortunately the positive premium seen in the fifth year after learning, for those achieving a Full Level 3 Academic qualification, is likely dominated by those who have now left HE and are receiving a premium to achievement at Level 6 (the final year of a degree). In contrast, Table 24 suggests that the earnings premium for Level 3 achievers in our Vocational CBL category rises steadily from 2% to 3.6% (giving a 3 to 5 year average of 2.9%). Full Level 3 achievers in Table 24 secure a 3-5 year average premium of 5.8%, with the underlying estimated return rising from 1.9 to 9.1 over the five-year period. In contrast the return to Full Level 3 in Table 25 (which averages 9% between 3 and 5 years) experiences some decline from 11.4% in the first year to 8.9% in the fifth.

Table 25: Daily earnings premium for achievers v non-achievers engaged in *Vocational Workplace Based Learning* (excluding apprenticeships)

Achievement	Percentage Log Daily Earnings Premium in Tax Year after Spell End				
	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.058***	0.079***	0.080***	0.062***	0.074
Level 2	0.030	0.056*	0.042	0.037	0.045
Full Level 2	0.130***	0.103***	0.094***	0.092***	0.096
Full Level 3	0.114***	0.097***	0.085***	0.089***	0.090
Level 4+	0.094***	0.093***	0.087***	0.082***	0.087

*** significant at the 0.1% level; ** 1% and * 5%

At Level 4+ the large majority of learners (between 303,000 and 144,000) fall within our category of Vocational CBL (Table 24), and between approximately 34,000 and 17,000 are included in the regressions focusing on vocational qualifications in a WBL context (Table 25). Tables 24 and 25 suggest that estimated earnings returns are similar for these learners in our two vocational categories, with 3 to 5 year averages of 8.2% and 8.7% respectively.

It is important to raise a note of caution when interpreting these findings. For instance, when comparing the 9% return secured by Vocational learners in a WBL context with the 5.8% for Vocational learners in a CBL context, we should not take these findings as evidence that learners should be moved to a WBL approach from an FE classroom based

model. It is quite possible that these two categories (of WBL and FE classroom based) contain very different types of learner. Vocational achievers in the CBL environments are securing statistically significant returns, relative to those who do not achieve in this environment - a move to a WBL environment would not necessarily raise this return further, as the environment may be less appropriate for them. This is an issue to which we return, but it is worth noting here that survey-based studies would utilise the same comparison group when calculating returns in WBL and CBL environments (i.e. all those at the qualification level below).

In Figure 1 we take the regression results from Tables 24 and 25 for FL2 learners and present them in a way that gives some indication of what the premiums mean for the daily earnings of achievers [on average], compared to non-achievers who have the same FL2 highest learning aim. For instance, the upper shaded area in Figure 1 is bounded below by the predicted daily earnings that Full Level 2 non-achievers studying in a WBL context can expect, between one [approx. £43] and five years [approx. £43.5] after the end of learning, calculated from our regression equations. The upper bound of this shaded area is calculated by applying the percentage premium for achievers from Table 25[34]. For instance, those who achieve a Full Level 2 WBL qualification can expect a premium of 13 per cent in the year directly after the end of learning (approximately £5.50-a-day more than the approximate £43 gained by non-achievers), and 9.2 per cent in the fifth year after learning.

These are predictions gained from regression analysis with a full set of controls, estimated for data between 2002 and 2012. We can therefore be more confident that the differences in these daily earnings are a result of the gaining of the relevant qualification, rather than other confounding factors. However, we must also remember that the actual estimated level of daily earnings are averaged across many years and may have been subject to change over the last few years; and this is in addition to the technical difficulties mentioned in footnote 34. They are presented only to provide some indication of how our regression results might translate into real world gains for those achieving these qualifications.

---

[34] This is not a wholly orthodox approach, but it is the only way to create graphs that are anywhere near consistent with the regression estimates. The problems associated with re-transformation of log daily earnings leave us with directly transformed earnings estimates from the regression equations that are inconsistent with the premiums estimated in the regression equations themselves.

Figure 1: Average Daily Earnings Returns for FL2 Achievers v Non-achievers in (i) *Vocational CBL* **and (ii)** *Vocational WBL* **contexts**
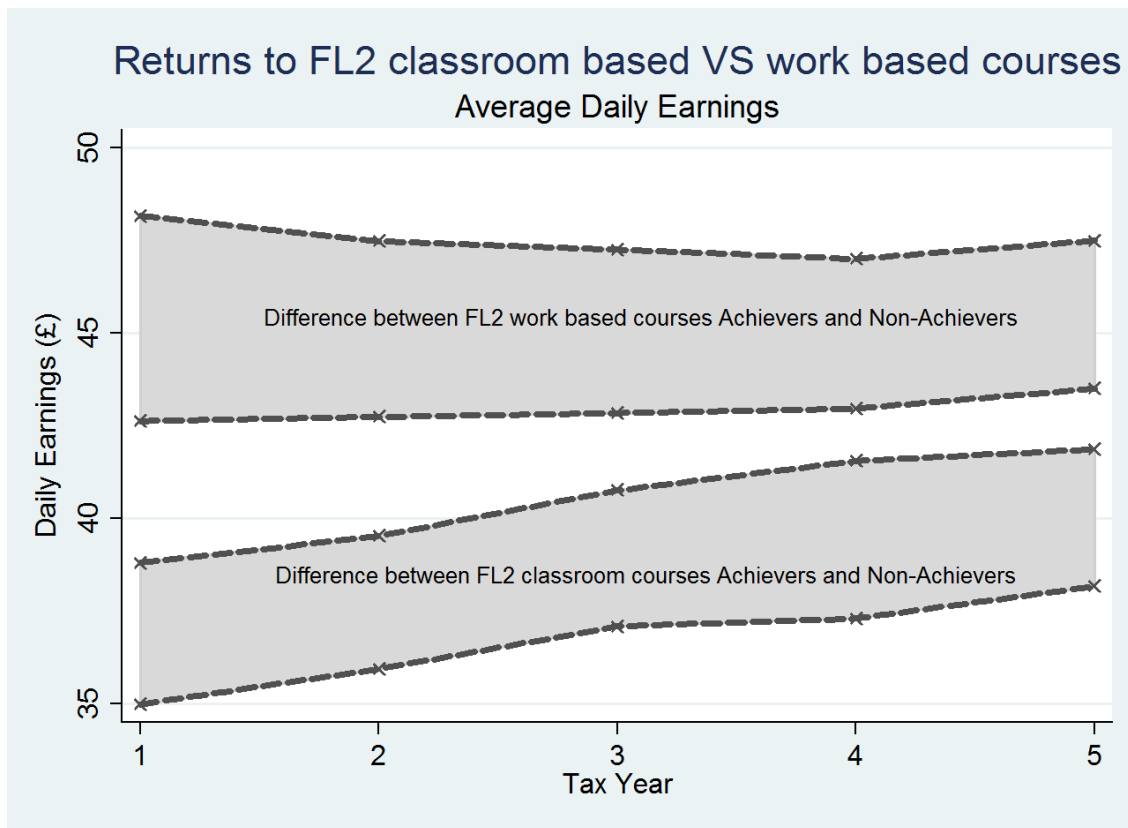
Similarly, Figure 2 presents the regression results from Tables 24 and 25 for FL3 learners in a way that gives some indication of what the premiums mean for the daily earnings of achievers [on average], compared to non-achievers who have the same FL3 highest learning aim. As we can see, the returns to classroom-based FL3 achievers behave in a very unusual way, with both achievers and non-achievers experiencing rapidly rising wages over the period that we observe them – this suggests there may be other issues to consider, alongside any problems that stem from the limited coverage of our HE flag.

Figure 2: Average Daily Earnings Returns for FL3 Achievers v Non-achievers in (i) *Vocational CBL* **and (ii)** *Vocational WBL* **contexts**
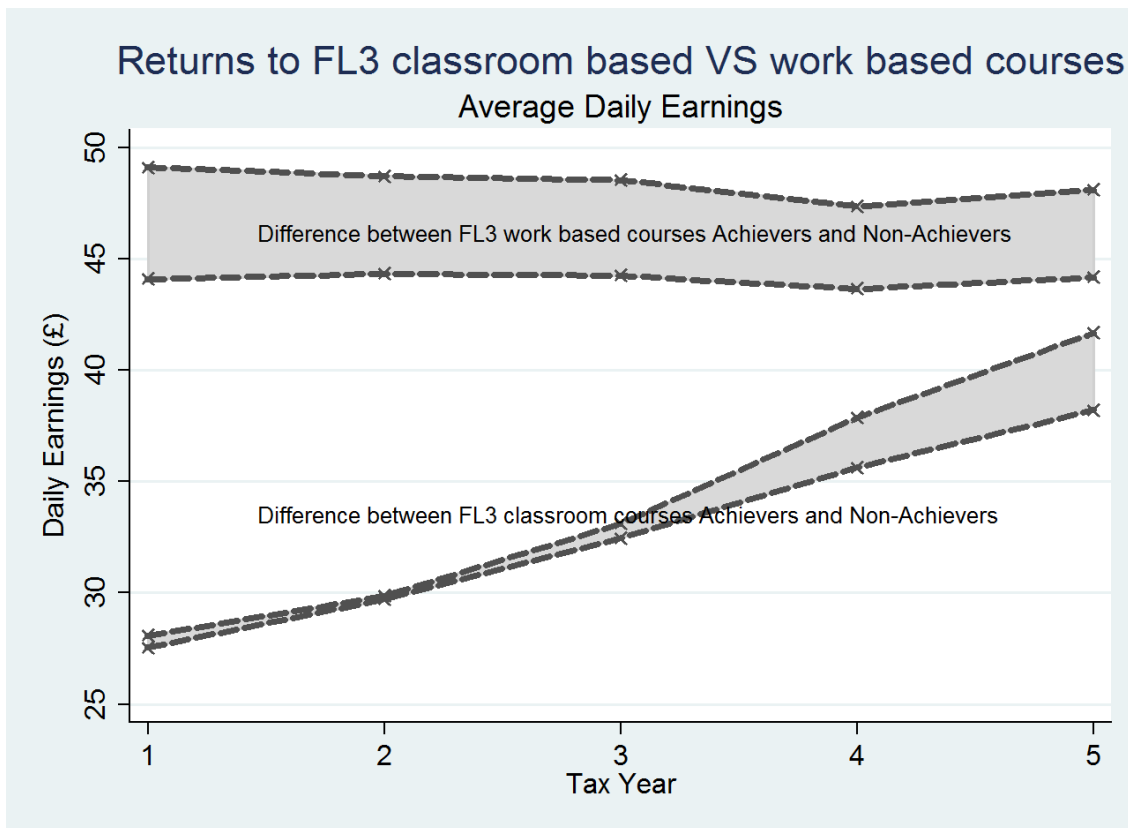
Employment Probability Premiums

Key Findings: Tables 26 to 28 set out the estimated employment probability premiums for achievers over non-achievers in each of our categories of Academic, Vocational CBL and Vocational WBL. The premiums secured by those achieving a highest qualification at Below Level 2, are positive and significant for both our WBL and CBL categories of Vocational learners, but in the CBL group this only equates to a 3 to 5 year average of 0.3 of a percentage point (compared to a figure of 1.7 in the WBL Group). In contrast at Level 2, achievers in the WBL category secure no significant employment premium in the fourth and fifth years after learning (despite stronger performance in the 1st and 3rd years); whilst the 3 to 5 year average return to achievers at Level 2 in the Vocational CBL category is 0.9 of a percentage point – not particularly high, but statistically significant in all years from the end of learning. Similarly, the employment premium for Level 3 Vocational CBL achievers averages 0.9 of a percentage point between 3 and 5 years, whilst the same level of learning secures a 2.3 percentage point premium in the Academic qualification category. At Full Level 3, achievers in the Academic category of qualifications seem to secure the highest return over non-achievers, with a 4.1 percentage point, 3 to 5 year average; compared to figures of 3.3 and 1.9 for Vocational CBL and Vocational WBL categories, respectively.

In Table 26 we are forced to drop the category of Academic Below Level 2, as there are less than 3,000 learners and therefore the non-achieving control group becomes rather small and less reliable. The results highlighted in our Key Findings section are confirmed when we change percentage point differences into percentages. For instance, when

considering Below Level 2 qualifications, Tables 30 and 31 show a much lower raw employment rate for the Vocational BCL category (40% and 37% for achievers and non-achievers respectively), compared to the Vocational WBL category (52% and 46% respectively), but the percentage differences remain as stark (0.8% relative to 3.5% respectively) as the percentage point differences highlighted in Key Findings.

Table 26: Estimated employment probability premium for Achievers v Non-achievers in the *Academic* category of learners

| | Percentage Point Employment probability Premium in Time Period after Spell End | | | | | |
Achievement	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Level 2	0.010***	0.006*	0.008*	0.010**	0.011**	0.010
Full Level 2	-0.012**	-0.015***	-0.025***	-0.021**	-0.026***	-0.024
Level 3	0.005**	0.006**	0.013***	0.027***	0.030***	0.023
Full Level 3	0.037***	0.032***	0.026***	0.049***	0.048***	0.041

*** significant at the 0.1% level; ** 1% and * 5%

Considering premiums at Full Level 2, where we are able to make comparisons across all three categories of qualification, the findings for Academic qualifications are particularly unusual as we observe negative and significant returns across all five years after the end of learning. This group is dominated by very young learners, with a median age of 17 amongst both achievers and non-achievers (even the last decile is only 19 years old). It is likely that most of the achievers continue to some form of continued learning that we are not picking up, even though the ILR data tell us that FL2 is their highest aim across the 10 years we observe them (78% of this group achieved after or during 2010).

At Full Level 3 achievers in the Academic category of qualifications seem to secure the highest return over non-achievers, with a 4.1 percentage point, 3 to 5 year average; compared to figures of 3.3 and 1.9 for Vocational CBL and Vocational WBL categories respectively. When we take into account the raw employment proportions in Tables 29 to 31 these differences become even more pronounced, resulting in a figure of 7.5% for the Academic category, 5.5% for the *Classroom Based* category and 2.4% for achievers over non-achievers in the WBL category.

Table 27: Estimated employment probability premium for achievers v non-achievers engaged in *Vocational Classroom Based Learning*

| | Percentage Point Employment probability Premium in Time Period after Spell End | | | | | |
Achievement	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.0003	0.002***	0.003***	0.003***	0.002***	0.003
Level 2	0.005***	0.008***	0.009***	0.009***	0.008***	0.009
Full Level 2	0.013***	0.016***	0.016***	0.020***	0.020***	0.019
Level 3	0.009***	0.009***	0.008***	0.010***	0.009***	0.009
Full Level 3	0.052***	0.050***	0.038***	0.040***	0.022***	0.033
Level 4+	0.012***	0.011***	0.012***	0.012***	0.011***	0.012

*** significant at the 0.1% level; ** 1% and * 5%

Table 28: Estimated employment probability premium for achievers v non-achievers engaged in _Vocational Workplace Based Learning_ (excluding apprenticeships)

Achievement	Percentage Point Employment probability Premium in Time Period after Spell End					
	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Below Level 2	0.010***	0.017***	0.022***	0.014***	0.015***	0.017
Level 2	0.024***	0.031***	0.041***	0.011	0.013	0.022
Full Level 2	0.016***	0.007***	0.005***	0.009***	0.011***	0.008
Level 3	n/a	n/a	n/a	n/a	n/a	n/a
Full Level 3	0.021***	0.015***	0.017***	0.020***	0.021***	0.019
Level 4+	0.007*	0.004	0.007*	0.008*	0.006	0.007

*** significant at the 0.1% level; ** 1% and * 5%

Finally, at _Level 4+_ the majority of learners (approximate 467,000) fall within the Vocational CBL category, where employment probability premiums vary only slightly from the 3 to 5 year average of 1.2 percentage points. For the approximate 52,000 to 30,000 observations in the Vocational WBL category, returns are low and insignificant.

Tables 29 to 31 provide baseline raw employment percentages to contextualise the discussions of percentage point differences from Tables 26 to 28. Generally, for the Academic and Vocational CBL categories we observe a similar pattern to that seen in previous tables, with Full Level 2 and Full Level 3 learners having lower employment rates than Level 2 and Level 3 learners. However, in the category of Vocational WBL qualifications it is the Full Level 2 and Full Level 3 qualifications where (amongst both achievers and non-achievers) we see the highest employment probability premiums. This seems in line with expectations, as those studying for Full Level 2 and Full Level 3 qualifications in a WBL context are more likely to have employers who are fully engaged with the training programme (and for it to be a formal part of the job they are doing).

Table 29: Raw employment probabilities one year after spell end date for _Academic_ qualification category

Spell Participation	Spell Achievement					
	No achieve	Below Level 2	Level 2	Full Level 2	Level 3	Full Level 3
Level 2	0.51	0.49	0.56			
Full Level 2	0.35	0.38	0.38	0.33		
Level 3	0.51	0.49	0.50	0.61	0.59	
Full Level 3	0.49	0.49	0.53	0.63	0.53	0.60

Table 30: Raw employment probabilities one year after spell end date for *Vocational Classroom Based Learning*

Spell Participation	Spell Achievement						
	No achieve	Below Level 2	Level 2	Full Level 2	Level 3	Full Level 3	Level 4+
Below Level 2	0.37	0.40					
Level 2	0.50	0.44	0.53				
Full Level 2	0.47	0.45	0.49	0.52			
Level 3	0.60	0.55	0.59	0.64	0.64		
Full Level 3	0.53	0.55	0.57	0.63	0.62	0.66	
Level 4+	0.70	0.72	0.74	0.78	0.74	0.77	0.76

Table 31: Raw employment probabilities one year after spell end date for *Vocational Classroom Based Learning* (excluding apprenticeships)

Spell Participation	Spell Achievement						
	No achieve	Below Level 2	Level 2	Full Level 2	Level 3	Full Level 3	Level 4+
Below Level 2	0.46	0.52					
Level 2	0.52	0.49	0.61				
Full Level 2	0.70	0.71	0.75	0.74			
Level 3	n/a	n/a	n/a	n/a	n/a		
Full Level 3	0.75	0.78	0.83	0.77	0.78	0.81	
Level 4+	0.75	0.81	0.85	0.81	0.82	0.81	0.80

4.2 Estimated returns for Apprenticeships

Key Findings: Tables 32, 34 and 36 consider the earnings, employment probability and probability of being on active benefits, for those who achieve an apprenticeship qualification as their highest learning aim, relative to those who have the same highest learning aim, but do not achieve. The findings suggest significant and substantial earnings returns, with a 3 to 5 year average earnings premium for Level 2 apprenticeship achievers estimated at 11.1% and for level 3 achievers the figure is 15.6%. The estimated employment returns are less pronounced, and by the third and fourth year after learning there is little difference in employment rates between achievers and non-achievers at both L2 and L3. In contrast, both L2 and L3 apprenticeship achievers have a significantly lower percentage point probability of being on active benefits (2.5 and 1.9 percentage points, respectively). Because the absolute proportions of apprenticeship achievers and non-achievers is relatively low, these translate into figures that suggest L2 achievers have an approximate 30% lesser probability of being on active benefits and for L3 achievers the figure is 48%. The lesser impact of apprenticeship achievement on employment is perhaps to be expected as learners are usually employed during the period of their apprenticeship, and even non-achievers will likely have a high probability of continuing in employment, but at a much lower wage.

Table 32 sets out the earnings premium estimated from regressions containing between 181,000 and 85,000 learners who study for a Level 2 apprenticeship and between 154,000

and 54,000 who study for a Level 3 apprenticeship. The premium for those who achieve at Full Level 3 (a.k.a. Advanced apprenticeships), compared to those who do not achieve, falls slightly from 17.1 per cent in the year directly after the end of learning, to 15.9 per cent in the fifth year after learning – leaving us with a 3 to 5 year average of 15.6%. For achievers of Level 2 apprenticeships (a.k.a. Intermediate or Foundation apprenticeships), there is a similar fall in returns from the first year, but by years 4 and 5 estimated returns are very similar (10.9% and 10.4%). Table 33 provides figures on average daily earnings from the raw data.

Table 32: Estimated daily earnings premium for apprenticeship achievers v non-achievers

Achievement	Percentage Log Daily Earnings Premium in Tax Year after Spell End				
	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Level 2 apprenticeship	0.145***	0.120***	0.109***	0.104***	0.111
Level 3 apprenticeship	0.171***	0.162***	0.147***	0.159***	0.156

*** significant at the 0.1% level; ** 1% and * 5%

Table 33: Raw average daily earnings in year after the end of apprenticeship learning spell

Spell Participation	Spell Achievement		
	No achieve	Level 2	Level 3
L2 apprenticeship	28.90	34.16	
L3 apprenticeship	34.95	34.97	45.54

Figure 3 presents the results from Table 32 in a way that gives some indication of what the premiums mean for the daily earnings of achievers [on average], compared to non-achievers who have the same apprenticeship aim. The upper shaded area in Figure 3 is bounded below by the predicted daily earnings that Full Level 3 non-achievers can expect, between one [approx. £35] and five years [approx. £42] after the end of learning, calculated from our regression equations. The upper bound of this shaded area is calculated by applying the percentage premium for achievers from Table 32. For instance, those who achieve a Level 3 apprenticeship qualification can expect a premium of 17.1 per cent in the year directly after the end of learning (approximately £6-a-day more than the approximate £35 gained by non-achievers), and 16 per cent in the fifth year after learning.

Figure 3: Average Daily Earnings Returns to Apprenticeship L2 and L3 qualifications
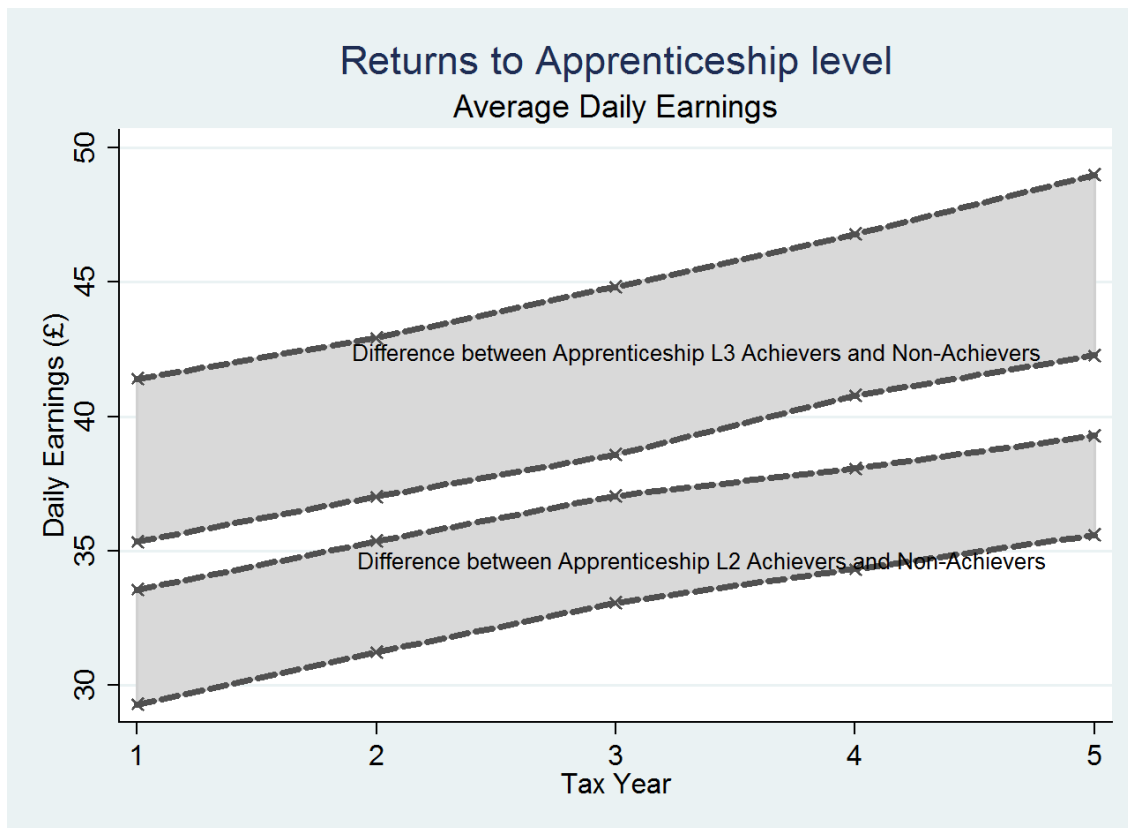
Readers should refer to the discussion around Figure 1 on page 50 for understanding of the limitations of Figure 3. There are technical limitations placed on our ability to present the results of regression equations, not least the re-transformation of log daily earnings (mentioned in footnote 34). The Figures in this report are presented only to provide some indication of how our regression results might translate into real world gains for those achieving these qualifications and should be taken as approximations.

As suggested in B&U, we must be careful comparing these results to those obtained from investigation of datasets such as the Labour Force Survey. However, it is interesting to note that the range of estimates here are in line with those obtained in such studies. For instance, McIntosh (2007) arrives at an estimate of 17.7% for Level 3 apprenticeships using weekly earnings and 15.6% for those with Level 2 apprenticeship qualifications. The National Audit Office (2012) replicate the approach of McIntosh with more up-to-date LFS data and arrive at estimates of 13.1% for Level 2 apprentices and 21.5% for *Level 3*. In their BIS-commissioned study Conlon et. al. (2011) arrive at estimates that are very close to those of McIntosh, with returns of 22% for Level 3 apprenticeships and 12% for Level 2.

Readers of these studies will note the differences in approaches to estimation, the exact definition of treatment and control, together with a variety of other factors. However, the ultimate aim is the same. We wish to compare the average earnings of those who secure an apprenticeship qualification with the earnings of a group who do not have the qualification, but are otherwise identical (in each case we are ultimately pursuing the same 'counterfactual').

61

Moving on, Table 34 presents the estimated employment probability premium for apprenticeship achievers over non-achievers; with Table 35 providing raw employment rates as a context for discussions. As we would expect, Table 35 suggests that both achievers and non-achievers amongst our Level 3 learners have higher employment rates than those whose highest aim is a Level 2 apprenticeship.

We must treat the results in Table 34 with care as there are points later on in the period of analysis where employment premiums are statistically insignificant. One year after learning, L2 and L3 achievers secure almost identical percentage point employment premiums, of 2.7 and 2.8 respectively. However, these percentage point premiums have become insignificant by the end of our period of analysis. This finding is consistent with the suggestion that apprenticeship training provides skills that benefit even non-achievers, in terms of eventually securing employment, but achievement is required to secure higher earnings returns.

The employment probability estimates presented here are much lower than those reported in B&U. Using the same approach detailed in Section 2.2 for apprenticeship learners[35], we find the drop in estimated employment premiums is mainly due to the new approach to data selection and to the addition of previous employment history, i.e. achievers were more likely than non-achievers to be employed prior to the commencement of learning. This means that achievers have longer employment durations before the start of learning and this makes them more likely to be employed after learning (something that in B&U was ascribed to the learning itself, rather than their status prior to learning).

Table 34: Estimated employment probability premium for achievers v non-achievers: apprenticeships

| Achievement | Percentage Point Employment probability Premium in Time Period after Spell End | | | | | |
	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
L2 apprentice.	0.027***	0.019***	0.011***	0.008***	-0.001	0.006
L3 apprentice.	0.028***	0.011***	0.005*	0.001	0.001	0.002

*** significant at the 0.1% level; ** 1% and * 5%

Table 35: Raw employment probabilities one year after spell end date for apprenticeships

| Spell Participation | Spell Achievement | | |
	No achieve	Level 2	Level 3
L2 apprenticeship	0.60	0.73	
L3 apprenticeship	0.75	0.81	0.83

Table 36 presents the estimated percentage point probability that those who achieve an apprenticeship qualification as their highest learning aim, will be on active benefits, relative to non-achievers. Table 37 provides context to the discussion of these results with raw

[35] The results of which are available from the authors, upon request.

benefit proportions. The suggestion from Table 36 is that L2 achievers have a probability of being on benefits that is much lower than non-achievers, and that this gap is much more pronounced than that for L3 achievers and non-achievers. However, the raw proportions in Table 37 are such that this initial finding is not born out when we calculate the percentage differences. The average 3 to 5 year percentage point gap in active benefit proportions between L2 apprenticeship achievers and non-achievers is -2.5 percentage points and this translates into an approximate 30% lesser probability of being on active benefits. However, for L3 achievers the -1.9 percentage point gap translates into an approximate 48% difference. This turnaround occurs because the base for our L2 calculations is approximately 8.5 per cent (half way between 0.06 and 0.11), whilst for FL3 the base is only approximately 4 per cent (half way between 0.03 and 0.05).

Table 36: Estimated probability of achievers being on *Active Benefits* (job-seeking), compared to non-achievers: apprenticeships

| Achievement | Percentage Point Probability of Achievers V Non-achievers being on Active Benefits | | | | | |
	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
L2 apprentice.	-0.032***	-0.031***	-0.026***	-0.025***	-0.023***	-0.025
L3 apprentice.	-0.026***	-0.022***	-0.017***	-0.020***	-0.019***	-0.019

*** significant at the 0.1% level; ** 1% and * 5%

Table 37: Proportion of apprentices on *Active Benefits* (job-seeking) one year after learning spell ends

| Spell Participation | Spell Achievement | | |
	No achieve	Level 2	Level 3
L2 apprenticeship	0.11	0.06	
L3 apprenticeship	0.05	0.07	0.03

Finally, there has been an expansion of Higher Apprenticeships since 2009, with the focus of such expansion initially across the engineering and IT sectors. As we can see from Table 38, even with this expansion we have too few observations to carry out a regression analysis – even in the first tax year after the end of learning, Table 38 suggests we only have 82 non-achievers who would constitute our control group, and we cannot observe raw earnings figures beyond 3 years. Table 38 presents the raw average earnings estimates for this group of learners, with the population size underpinning each estimate in brackets.

Table 38: All those with Higher Apprenticeship aim, raw average daily earnings in each year after the end of learning (population size in brackets)

| Tax year after learning | Spell Achievement | | |
	No achieve	Full Level 3	Level 4+
1st Year	£43.67	£38.18	£45.01
	(82)	(93)	(716)
2nd Year	£46.07	£38.71	£53.09
	(32)	(14)	(112)

Tax year after learning	Spell Achievement		
	No achieve	Full Level 3	Level 4+
3rd Year	£43.73	£67.75	£58.76
	(8)	(2)	(12)

4.3 Returns to L1/L2 Maths, English and Preparation for Work

Key Finding 1: The L1 and L2 Maths and English qualifications we consider in this section will mostly be taken alongside more substantial qualifications – they are often taken as complements to more '*Full'* or higher-level qualifications. As a result we need to adapt our approach to estimation. Elsewhere in the report we are capturing returns to qualifications gained as an individual's highest learning aim. Taking such an approach here would miss many learners, as we will only analyse individuals whose highest learning aim across their 2002 to 2012 FE spells is a L1 or L2 English or Maths qualification. This is a relatively unique group and as we have already seen, returns are low for those who have Below Level 2 or L2 as their highest learning aim[36]. In this section of the report, we find statistically significant returns to L1 and L2 Maths and English, when we adopt an approach to estimation that accommodates the complementary nature of these qualifications. Readers should consider Figure 4 and the accompanying explanation below before reading *Key Findings 2*.

Key Findings 2: The 3 to 5 year average earnings premium for those achieving a L1 and/or L2 Maths/English qualification, relative to those who do not achieve their L1/L2 Maths or English qualification is 2.4% (with this estimate relevant for a population of all FL2 and FL3 achievers). More specifically, we estimate that Literacy/Numeracy Key Skills achievers secure a 3 to 5 year average earnings return of 6% over non-achievers (amongst a population of all FL2 achievers) and 3.7% (for a population of all FL3 achievers).

In the previous B&U study, the Maths and English qualifications considered in this Section were referred to as *Skills for Life* and included ESOL (English for Speakers of Other Languages). Our approach to evaluation of the returns to these qualifications remains the same as in B&U, with some modification of the achiever v non-achiever approach needed, as Maths and English qualifications taken at L1 and L2 are often forms of 'complementary learning'. In Figure 4 we describe an analysis that better suits this complementarity and helps us obtain a truer estimate of the value added of these qualifications.

Figure 4 shows how we select our population for analysis, starting from the top, at {a}.

- First we select a group of achievers from our existing categories. In the example of Figure 4, we select all those who, between 2002 and 2012, Achieve their highest learning aim of a Full Level 2 Qualification {a}.

---

[36] Though, obviously not all of those in our Below Level 2 and L2 categories in other sections of the report will be studying Maths and English.

- Within this group of Full Level 2 Achievers, we then identify all learners who also have a L1 and/or L2 English or Maths aim (this will not be held as a highest learning aim) {b}.
- Within group {b} we can then differentiate those who {c} Achieve their L1 &/or L2 English/Maths aims and {d} those who Do not Achieve L1 &/or L2 English/ Maths aim.

Figure 4: Example approach to capture returns to L1 and/or L2 English or Maths

{a} Learners who Achieve their highest learning aim of a Full Level 2 Qualification

{b} L1 and/or L2 English or Maths aim   No L1 and/or L2 English or Maths aim

{c} Achieve L1 &/or L2 English/Maths   {d} Do not Achieve L1 &/or L2 English/ Maths
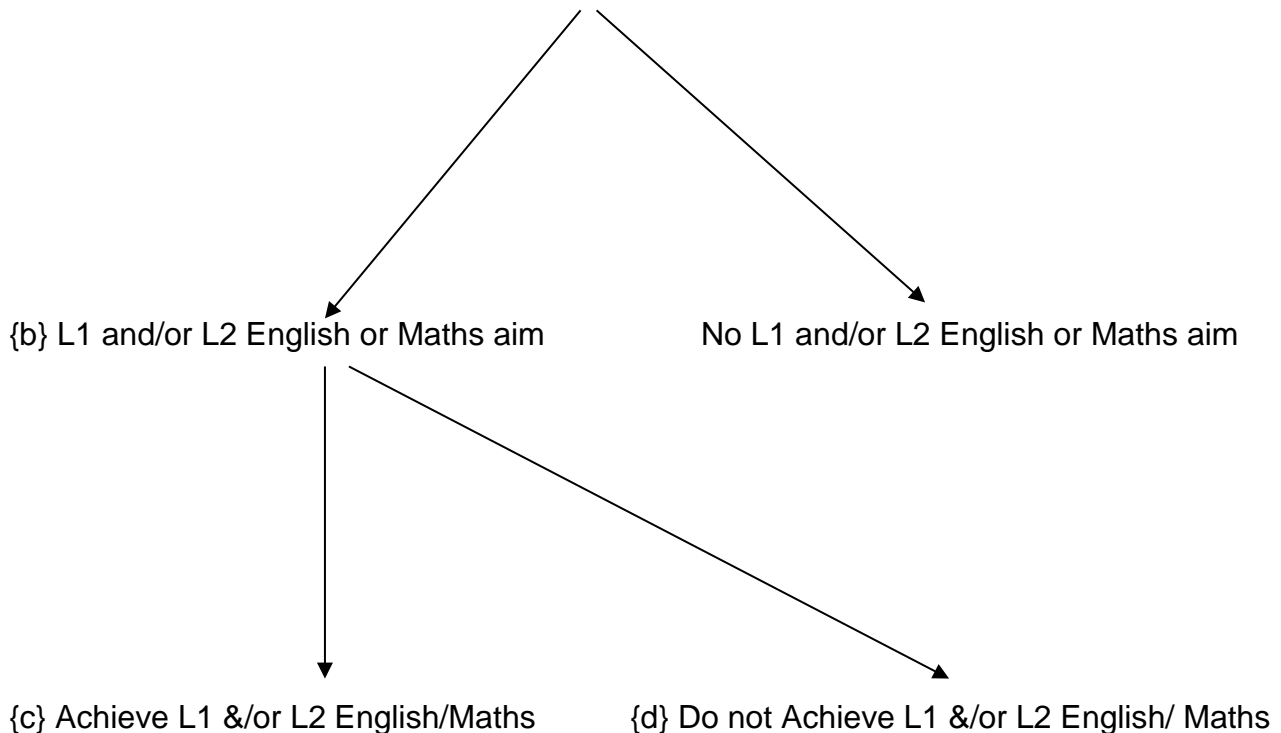
In the example of Figure 4, a comparison of earnings between achievers {c} and non-achievers {d}, provides us with a robust estimate of the value added of L1 or L2 English or Maths qualifications. The only caveat is that this estimate is obtained by only considering the population of learners who achieve a Full-Level 2 qualification. In the following tables we also carry out a comparison of achievers and non-achievers who have a L1 &/or L2 English/Maths aim, for all those who achieve Full Level 3. The analysis works in the same way, with the only difference being that, in Figure 4, we would first select [at {a}] those Learners who Achieve their highest learning aim of a Full Level 3 Qualification.

The first row of Table 39 presents the estimated returns to L1 and/or L2 Maths/English qualifications gained from the comparison of achiever and non-achievers as suggested in Figure 4 (i.e. for a population of FL2 achievers); the second row carries out the same analysis for a population of FL3 achievers and the third row combines these populations. More specifically:

- The first row of Table 39 estimates the value added of L1 and/or L2 Maths/English qualifications, with the population of individuals included in the regression equations

restricted to those who have achieved a highest aim of Full Level 2[37]. The 3 to 5 year average earnings return of 4.2% is therefore the earnings premium that those achieving a L1 and/or L2 Maths/English qualification secure, relative to those who do not achieve their L1/L2 Maths or English qualification (with this estimate relevant for the population of FL2 achievers).

- The second row of Table 39 estimates the value added of L1 and/or L2 Maths/English qualifications, with the population of individuals included in the regression equations restricted to those who have achieved a highest aim of Full Level 3[38]. The 3 to 5 years average earnings return of 1.2% is therefore the earnings premium that those achieving a L1 and/or L2 Maths/English qualification secure, relative to those who do not achieve their L1/L2 Maths or English qualification (with this estimate relevant for the population of FL3 achievers). This is likely an underestimate, as we again have a potential problem from the HE flag issue in the first three years after the end of FE learning – something we return to discuss in the conclusion.

- The third row of Table 39 estimates the value added of L1 and/or L2 Maths/English qualifications, with the population of individuals included in the regression equations restricted to those who have achieved a highest aim of either FL2 or FL3 - amalgamating the populations used to produce the results in the first two rows. The 3 to 5 years average earnings return of 2.4% is therefore the earnings premium that those achieving a L1 and/or L2 Maths/English qualification secure, relative to those who do not achieve their L1/L2 Maths or English qualification (with this estimate relevant for a population of FL2 and FL3 achievers).

Table 39: Returns to daily earnings for L1 and L2 English and Maths achievers [for population of FL 2 and/or FL3 achievers]

Achievement	Percentage Log Daily Earnings Premium in Tax Year after Spell End				
	1st Year	3rd Year	4th Year	5th Year	3-5 year average
FL2 Achievers	0.060***	0.035***	0.048***	0.042***	0.042
FL3 Achievers	-0.002	0.001	0.018***	0.017***	0.012
FL2 or FL3 Achievers	0.021***	0.015***	0.031***	0.027***	0.024

*** significant at the 0.1% level; ** 1% and * 5%

Essentially, Table 39 estimates the earnings return for an amalgamated category of learners who achieve L1 Maths, and/or L1 English, and/or L2 Maths, and/or L2 Maths. Within the ILR data we are able to further differentiate these learners by the following four categories:

i) Key Skills

ii) Certificate Skills

---

[37] They are the Full Level 2 Achievers in Table 3.
[38] They are the Full Level 3 Achievers in Table 3.

iii) Basic Skills

iv) Functional Skills

Table 40 uses the same approach to estimation as detailed in Figure 4 and Table 39, but with, for instance, the first row comparing the earnings of those who achieve 'only literacy key skills' with those who have 'only a literacy key skill' aim, but do not achieve (with this estimate, and all other estimates in Table 40, relevant for a population of FL2 achievers)[39]. When we get down to this level of disaggregation, even the ILR starts to suffer from low numbers – in this case we have few failures on which to construct our control group of non-achievers. Having said this, in row three we see Literacy/Numeracy Key Skills achievers secure a 3 to 5 year average earnings return of 6%.

Table 40: Returns to daily earnings for subgroups of L1 and L2 English and Maths achievers [for population of FL 2 achievers]

| Achievement | Percentage Log Daily Earnings Premium in Tax Year after Spell End | | | | |
	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Only Literacy Key Skills	0.029***	0.027**	0.040***	0.039***	0.035
Only Numeracy Key Skills	0.018	0.016	0.016	0.018	0.017
Literacy/Numeracy Key Skills	0.060***	0.047***	0.066***	0.066***	0.060
Only Literacy Certificate	0.035***	0.021	0.039	0.001	0.020
Only Numeracy Certificate	0.030*	0.024	0.027	0.004	0.018
Literacy/Numeracy Cert.	0.023	0.068*	0.094*	0.090	0.084

*** significant at the 0.1% level; ** 1% and * 5%

Table 41 uses the same approach to estimation as detailed in Table 40, but this time our earnings comparison in the first row is between 'only literacy key skills' achievers and non-achievers for a population of FL3 achievers. Again, we face limitations in terms of the number of non-achievers we can identify in each category and this likely drives our lack of statistical significance in some areas. However, in row three we once again see Literacy/Numeracy Key Skills achievers securing a consistent statistically significant 3 to 5 year average earnings return, in this case of 3.7%. We can also see that in many other areas the HE flag issue results in a number of negative estimated returns.

---

[39] Returns for Basic Skills are not reported because of the small population size, while returns for Functional Skills Maths and English qualifications are not reported because of their very recent introduction, which does not allow a proper evaluation of the earnings returns.

Table 41: Returns to daily earnings for subgroups of L1 and L2 English and Maths achievers [for population of FL 3 achievers]

Achievement	Percentage Log Daily Earnings Premium in Tax Year after Spell End				
	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Only Literacy Key Skills	-0.010*	0.012*	0.020**	0.014*	0.015
Only Numeracy Key Skills	-0.006	0.009	0.026***	0.011	0.015
Literacy/Numeracy Key Skills	0.037***	0.033***	0.044***	0.033***	0.037
Only Literacy Certificate	-0.011	-0.007	0.015	0.022	0.010
Only Numeracy Certificate	-0.049***	0.007	0.055**	0.073**	0.045
Literacy/Numeracy Cert.	-0.051*	-0.004	-0.005	0.035	0.009

*** significant at the 0.1% level; ** 1% and * 5%

5. Variation in returns by sector subject area

Key Findings: In female-dominated sectors such as *Adult Social Care*, average 3 to 5 year earnings returns are just above 10 per cent at FL2 and FL3, and the same is true of achievers at FL3 in the area of *Child Development and Wellbeing*. Even where we see lower earnings returns, for instance in the areas of *Hair and Beauty*, and at FL2 for *Child Development and Wellbeing*; nowhere do these dip below 5%. In male-dominated areas such as *Construction*, *Engineering & Manufacturing* and *Transportation* we see earnings returns that are close to 10 percentage points higher than this. Across these Sectors we seldom see average three to five year returns dropping below 15% - only in the case of FL2 learners in *Engineering and Manufacturing.* In areas that are neither male nor female dominated such as *ICT*, *Customer service* and *Hospitality and Catering*, earnings returns at FL2 are around 10% and in Business Studies just below at 7.5%.

In *Adult Social Care*, *Child Development and Wellbeing* and *Hair and Beauty* average 3 to 5 year employment returns vary between 2.5 and 3.3 percentage points at FL2 and at FL3 they are 4.2, 5.9 and 1.8 percentage points respectively. There is some concern over employment returns for those achieving FL3 *Construction*, FL2 *Engineering and Manufacturing* and Full Level 3 *Transportation qualifications.* Similarly, those achieving FL2 qualifications in ICT, *Customer Service* and *Business Studies* experience periods of negative returns. At FL3 we observe instances of negative and statistically significant employment probability returns in *Customer Service*. Those achieving *Hospitality and Catering* qualifications at FL2 and FL3 secure returns of 2.5 and 1.8 percentage points respectively; and FL3 *Business Studies* achievers secure a 5.7 percentage point average return between 3 and 5 years from the end of learning.

One of the changes to our econometric approach (when compared to B&U) is the introduction of detailed sector subject areas as controls in the regression equations underpinning results in Section 3, 4 and 5. As the discussion in Section 2.1 suggests, the addition of controls derived from Sector Subject Area variables (see footnote 9) has a limited impact on estimated premiums for the broader qualification categories[40]. However, it is still possible that, whilst sector subject area does not substantially alter our results in these regression equations, we may observe a lot of variation in returns across specific sectors.

In the academic literature, survey-based studies identify such variability in returns when disaggregating by Sector of qualification. For instance, the analysis of Greenwood et. al. (2007) suggests that males in construction with an NVQ2 achieve a return of 11%; much higher than their estimated average return for NVQ2s as a whole. In this section we see if our results exhibit a similar level of variability in returns, adopting the same approach seen in the academic literature, where returns for men and women are considered separately.

---

[40] The addition of sector controls, without substantially altering our estimated premiums, suggests that there is less of a correlation between estimated returns and drop-out rates across sector subject areas than we might have expected a-priori.

Many industry sectors are dominated by men (e.g. Construction), and some by women (e.g. Child Development and Wellbeing), to such an extent that consideration of returns for all achievers in a certain sector, is equivalent to consideration of returns to a particular sex. In the following tables we present the earnings and employment premiums for learners within each sector subject area as a whole (i.e. not calculating results separately for men and women), but Table 42 flags those sectors that are dominated by one sex. Similarly, we present earnings and employment returns for each sector as a whole, but Table 43 focuses on female-dominated sectors; Table 44 on male-dominated sectors and Table 45 on sectors that are neither male nor female dominated.

As we can see from Table 42, in *Adult Social Care*, *Child Development & Wellbeing* and *Hair & Beauty*, women dominate; whereas men make up the vast majority of learners in *Construction* (for instance, 98.5% at FL2); *Engineering & Manufacturing* (particularly at FL3) and *Transportation*.

Table 42: Concentrations of Men and Women across Sector Subject Areas, for Full Level 2 (FL2) and Full Level 3 (FL3) qualifications

Sector (derived from SSA)	Proportion of Women FL2	Proportion of Women FL3
Adult Social Care	86%	86%
Construction	1.50%	3.60%
Engineer. & manufacturing	14.40%	4.80%
Inf. & Comm. Technology	48%	19%
Child devel. & wellbeing	97%	96%
Customer Service	62%	70%
Hair & beauty	94%	97%
Hospitality and Catering	58%	62%
Transportation	7%	4%
Business studies	34%	45%

Table 43 sets out estimated earnings returns for those sectors where women make up approximately 90% or more of learners. For learners in these sectors there are substantial returns to achievers, compared to those with the same highest learning aim who do not achieve. Some Sector Subject areas, for instance *Hair and Beauty*, are derided as less valuable (or 'Mickey Mouse') areas for study by the popular press. However, for many young people with more limited labour-market options, these qualifications can help secure a more stable position in the labour market. It is often hard to identify accurate estimated returns in qualification areas dominated by learners with lower levels of academic achievement, especially in survey-based studies. It is very difficult to find an accurate control or comparison group (to give us an indication of what these young people would achieve in the absence of learning). The ILR-WPLS data allow us to identify a group who, we would argue, provide a more robust estimate of the labour market outcomes that might be expected if these specific learners had not secured a *Hair and Beauty* qualification.

Table 43: Estimated daily earnings premiums for female-dominated Sector Subject Areas[41]

Adult social care

Achievement	Log Daily Earnings in Tax Year after Spell End				
	1st Year	3rd Year	4th Year	5th Year	Average 3-5
Full Level 2	0.185***	0.132***	0.117***	0.099***	0.116
Full Level 3	0.130***	0.101***	0.142***	0.173***	0.139

Child development and wellbeing

Achievement	Log Daily Earnings in Tax Year after Spell End				
	1st Year	3rd Year	4th Year	5th Year	Average 3-5
Full Level 2	0.059***	0.060***	0.069***	0.058**	0.062
Full Level 3	0.135***	0.097***	0.109***	0.115***	0.107

Hair and beauty

Achievement	Log Daily Earnings in Tax Year after Spell End				
	1st Year	3rd Year	4th Year	5th Year	Average 3-5
Full Level 2	0.096***	0.098***	0.076***	0.058***	0.077
Full Level 3	0.099***	0.087***	0.068**	0.056*	0.070

*** significant at the 0.1% level; ** 1% and * 5%

Table 44 presents estimated earnings returns for Sector Subject areas dominated by men. As we can see, when compared to the returns estimated for female dominated sectors, the returns for men in *Construction*, *Engineering & Manufacturing* and *Transportation* are close to 10 percentage points higher. There are many factors driving these differences, not least the fact that in the male dominated sectors of Table 43 there is a longer history of vocational training and in such sectors returns tend to be higher (see Greenwood et. al. 2007 and Blanden et. al. 2012). Across the levels of qualification in Sectors included in Table 44 we seldom see average three to five year returns dropping below 15% - only in the case of FL2 learners in *Engineering and Manufacturing.*

Table 44: Estimated daily earnings premiums for male-dominated Sector Subject Areas

Construction

Achievement	Log Daily Earnings in Tax Year after Spell End				
	1st Year	3rd Year	4th Year	5th Year	Average 3-5
Full Level 2	0.212***	0.178***	0.160***	0.131***	0.156
Full Level 3	0.224***	0.165***	0.155***	0.121***	0.147

Engineering and manufacturing

---

[41] The approach taken to estimation of premiums using the achiever v non-achiever comparison is identical to that adopted in the rest of the report. First we select only the latest, highest learning aim for each individual and then within these latest, highest learning aims we select the relevant sector subject area.

| Achievement | Log Daily Earnings in Tax Year after Spell End | | | | |
	1st Year	3rd Year	4th Year	5th Year	Average 3-5
Full Level 2	0.194***	0.127***	0.117***	0.107***	0.117
Full Level 3	0.229***	0.196***	0.148***	0.136***	0.160

Transportation

| Achievement | Log Daily Earnings in Tax Year after Spell End | | | | |
	1st Year	3rd Year	4th Year	5th Year	Average 3-5
Full Level 2	0.150***	0.166***	0.163***	0.169***	0.163
Full Level 3	0.239***	0.203***	0.172***	0.202***	0.192

*** significant at the 0.1% level; ** 1% and * 5%

Table 45 sets out the earnings returns for men and women in sectors where there is more of a gender balance. Here we again get some indication that our HE flag is not capturing the full extent of movement into HE from FE learning, as there are initial negative returns to FL3 qualifications in areas such as ICT and Business Studies, which are more often oriented towards subsequent HE learning. It is interesting to note that in the areas of *Customer Service* and *Hospitality and Catering*, women achievers secure higher returns over non-achievers than male achievers secure over non-achievers; at both FL2 and FL3 levels (these supplementary results are available on request to the authors).

Table 45: Estimated daily earnings premiums for Sector Subject Areas that are neither male nor female dominated

Information & Communication Technology

| Achievement | Percentage Log Daily Earnings Premium in Tax Year after Spell End | | | | |
	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Full Level 2	0.077***	0.116***	0.124***	0.125***	0.122
Full Level 3	0.056***	-0.037**	-0.013	0.069***	0.006

Customer service

| Achievement | Percentage Log Daily Earnings Premium in Tax Year after Spell End | | | | |
	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Full Level 2	0.110***	0.093***	0.090***	0.096***	0.093
Full Level 3	0.088***	0.085***	0.083***	0.081***	0.083

Hospitality and Catering

| Achievement | Percentage Log Daily Earnings Premium in Tax Year after Spell End | | | | |
	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Full Level 2	0.121***	0.098***	0.120***	0.108***	0.109
Full Level 3	0.093***	0.061***	0.078***	0.069***	0.069

Business studies

	Percentage Log Daily Earnings Premium in Tax Year after Spell End

Achievement	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Full Level 2	0.114***	0.086***	0.079***	0.061**	0.075
Full Level 3	0.012	-0.015	0.104***	0.199***	0.096

*** significant at the 0.1% level; ** 1% and * 5%

Table 46 provides evidence of statistically significant employment probability premiums for achievers over non-achievers in all subject sector areas where women make up approximately 90 per cent or more of learners (though by the fifth year of learning, returns become insignificant for FL3 learners in the area of *Hair and Beauty*, due to low numbers). These estimated returns are relatively stable in the first to fifth years for most of the subject areas, and therefore in each area estimated returns deviate very little from the 3 to 5 year average.

Table 46: Estimated employment probability premiums for Sectors Subject Areas dominated by women

Adult social care

	Percentage Point Employment probability Premium in Time Period after Spell End					
Achievement	3mths	1yr	3yr	4yr	5yr	Average 3-5
Full Level 2	0.045***	0.037***	0.034***	0.035***	0.032***	0.033
Full Level 3	0.046***	0.038***	0.038***	0.048***	0.040***	0.042

Child development and wellbeing

	Percentage Point Employment probability Premium in Time Period after Spell End					
Achievement	3mths	1yr	3yr	4yr	5yr	Average 3-5
Full Level 2	0.026***	0.026***	0.026***	0.027***	0.037***	0.030
Full Level 3	0.078***	0.073***	0.063***	0.061***	0.053***	0.059

Hair and beauty

	Percentage Point Employment probability Premium in Time Period after Spell End					
Achievement	3mths	1yr	3yr	4yr	5yr	Average 3-5
Full Level 2	0.020***	0.021***	0.027***	0.029***	0.020***	0.025
Full Level 3	0.021***	0.024***	0.021**	0.020*	0.012	0.018

*** significant at the 0.1% level; ** 1% and * 5%

In Table 47 we observe estimated employment probability premiums in sectors dominated by men that are much less consistent than those in Table 46. At FL2 learners in *Engineering and Manufacturing* experience negative employment returns from the third year onwards; and at FL3 men in construction and Transportation experience a similarly negative return across all years after the first tax year after learning. It is not wholly clear what is driving these unusual findings and we need to be cautious in our interpretation, but

there are questions over employment rates for those achieving FL3 *Construction*, FL2 *Engineering and Manufacturing* and Full Level 3 *Transportation* qualifications.

Table 47: Estimated employment probability premiums for male-dominated Sector Subject Areas

Construction

Achievement	Percentage Point Employment probability Premium in Time Period after Spell End					
	3mths	1yr	3yr	4yr	5yr	Average 3-5
Full Level 2	0.032***	0.027***	0.022***	0.017***	0.008*	0.016
Full Level 3	0.026***	0.006	-0.007	-0.014*	-0.016*	-0.012

Engineering and manufacturing

Achievement	Percentage Point Employment probability Premium in Time Period after Spell End					
	3mths	1yr	3yr	4yr	5yr	Average 3-5
Full Level 2	0.030***	0.015***	-0.001	-0.004	-0.012***	-0.006
Full Level 3	0.062***	0.046***	0.032***	0.030***	0.021***	0.028

Transportation

Achievement	Percentage Point Employment probability Premium in Time Period after Spell End					
	3mths	1yr	3yr	4yr	5yr	Average 3-5
Full Level 2	0.018***	0.010***	0.019***	0.039***	0.029***	0.029
Full Level 3	0.029***	-0.001	-0.005	-0.003	-0.016*	-0.008

*** significant at the 0.1% level; ** 1% and * 5%

In Table 48 we observe those achieving FL2 qualifications in ICT, *Customer Service* and *Business Studies* experiencing periods of negative returns. At FL3 we observe instances of negative and statistically significant employment probability returns in *Customer Service.* In contrast to the findings in other sections of the report, negative returns in these areas do not look to be driven by HE-flag data issues. We find evidence of significant employment returns for those securing Hospitality and Catering qualifications at FL2 and FL3 (2.5 and 1.8 percentage points respectively); and FL3 Business Studies achievers secure a 5.7 percentage point average return between 3 and 5 years from the end of learning. To shed more light on why some areas are securing negative employment probability returns and some positive, it would be useful to expand the cohort analysis in Section 6.2 – it is possible that certain cohorts are particularly important in driving these findings.

Table 48: Estimated employment probability premiums for Sector Subject Areas that are neither male nor female dominated

Information & Communication Technology

Achievement	Percentage Point Employment probability Premium in Time Period after Spell End					
	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Full Level 2	0.003	0.010***	-0.011**	-0.023***	-0.029***	-0.021
Full Level 3	0.052***	0.053***	0.031***	0.033***	0.023***	0.029

Customer service

Achievement	Percentage Point Employment probability Premium in Time Period after Spell End					
	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Full Level 2	0.003**	-0.005**	-0.007**	-0.002	-0.001	-0.003
Full Level 3	0.0002	-0.008**	-0.011**	-0.008	-0.005	-0.008

Hospitality and Catering

Achievement	Percentage Point Employment probability Premium in Time Period after Spell End					
	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Full Level 2	0.033***	0.026***	0.024***	0.025***	0.027***	0.025
Full Level 3	0.038***	0.028***	0.025***	0.023***	0.007	0.018

Business studies

Achievement	Percentage Point Employment probability Premium in Time Period after Spell End					
	3mths	1st Year	3rd Year	4th Year	5th Year	3-5 year average
Full Level 2	0.023***	0.012***	-0.005	-0.010*	-0.017*	-0.011
Full Level 3	0.057***	0.056***	0.053***	0.064***	0.055***	0.057

*** significant at the 0.1% level; ** 1% and * 5%

6. How robust are our estimated returns?

Key Findings 1: Section 6.1 tests the validity of our approach to estimation of the Value Added from qualifications in other areas of the report. Here we focus on NVQs, as they are often the focus of investigation in survey-based studies, and are also where concern has been raised over low (or even negative) estimated returns. The process we have gone through, to test the robustness of our own estimates, has been peer-reviewed by four academic experts and we are very grateful for their comments. Any mistakes remain our own and readers can consider the comments of reviewers and our responses at the end of this report. Our findings suggest that regression-based techniques, which compare achievers and non-achievers in the ILR-WPLS, produce robust estimates of value added. Whilst further enhancements are always going to be possible, the results produced in this report provide BIS with estimates of the value added of FE that are quality assured to the highest academic standards.

Key Findings 2: Section 6.2 Section 6.2 investigates the persistence of earnings and employment probability premiums over the years after the end of learning. We also investigate the persistence of earnings and employment probability premiums over the years after the end of learning. For FL2 and L3 apprenticeship learners, the difference between a 3 to 5 year average earnings premium and a 4 to 6 year average, is less than 0.4 of a percentage point. The results for FL3 learners suggest that data issues may be particularly problematic in the pre-recessionary period when calculating earnings returns. We adopt a 3 to 5 year average in the remainder of the report, as it balances a number of competing issues – going with a 4 to 6 year average would leave us with only one cohort making up the estimate of year 6, which is as much of a concern as persistence.

We have evidence that those achieving qualifications at Full Level 3 are better placed to secure employment in a recession than non-achievers, but for FL2 achievers there is an apparent one-off employment impact of recession for those in the 2007/2008 cohort – with achievers much more likely to secure employment than non-achievers. For each cohort, percentage-point employment probability premiums exhibit persistence, but for FL3 our estimated 3 to 5 year average employment probability premiums will be dominated by the slightly lower pre-recession employment probability premiums.

Validity of the results in Sections 3, 4 and 5 relies heavily on the assumption that important differences between achievers and non-achievers are accounted for in our regression framework. In Section 6.1 we see whether this assumption is justified, by analysing the returns to NVQ2 and NVQ3 qualifications. In this section we see if returns estimated using a regression-based, achiever v non-achiever comparison, are substantially altered when applying difference-in-differences on matched achiever v non-achiever data. Having considered the findings from this first part of the investigation, we then vary the specific treatment and control groups used to shed further light on the validity of our returns.

Section 6.2 considers the findings from a cohort analysis that investigates the persistence of earnings and employment probability premiums over the years after the end of learning; ultimately informing our decision to opt for a 3 to 5 year average return as the relevant

policy metric in the remainder of the report.

6.1   Testing the validity of estimates from achiever v non-achiever comparisons

The more advanced econometric approaches used in this section of the report have the potential to overcome many of our concerns when adopting a standard regression approach on achiever v non-achiever data (see Smith and Todd, 2005). More specifically, difference-in-differences attempts to rectify any bias arising from <u>time-invariant</u> unobservables (such as motivation), which are correlated with both achievement and earnings. For instance, it is possible that motivation or commitment is important in differentiating achievers and non-achievers, so that achievers would potentially receive a higher return whatever their level of qualification – if we do not control for this difference, then estimated returns will be biased upwards. From a method perspective, there are a number of such 'unobservables' that can be assumed broadly constant over the period of our analysis and if these are confounding our regression estimates, then estimates produced using difference-in-differences would be much lower.

Together with these <u>time-invariant</u> unobservables, we would ideally accommodate more time-varying impacts - such as divorce. When difference-in-differences is combined with matching (particularly on prior earnings histories) there is some potential to overcome bias arising from time-varying unobservables, if these are correlated with observables, such as prior earnings. One can think of the control group as perhaps having family relationships that are, on average, more likely to deteriorate over time and this has an impact that sets them on a different earnings trajectory. Earnings prior to learning act as a proxy for this unobservable time-varying difference between treatment and control, and matching on prior earnings has the potential to remove much of this[42].

Table 49 presents the results from a number of different approaches to estimation that attempt to remove bias arising from unobservable impacts (correlated with both achievement and earnings). The first two columns of Table 49 set out a starting point for investigation, with estimates obtained using the same regression approaches adopted in Sections 3, 4 and 5, but on a smaller sample used for matching and difference-in-differences[43]. The suggestion in these first two columns is that estimated returns are 14.86% and 14.67% for men achieving at NVQ2 and NVQ3 level, when we consider the outcome of a fully specified regression model; and for women the figures are 11.73% and 13.29% respectively.

---

[42] Matching allows us some possibility of 'forcing' our treated and control samples to have parallel trends within a difference-in-differences framework, if this is not already the case (see for instance the discussion of synthetic control methods in Abadie, Diamond and Hainmueller, 2010). Where time-varying unobservables are correlated with observable characteristics on which we can match (such as prior earnings), we have an opportunity to overcome the problem of non-parallel trends.
[43] To estimate using difference-in-differences we have to limit ourselves to a sample that has earnings observations both before and after achievement at NVQ2 or NVQ3 level. To allow comparison we therefore estimate earnings returns in the first two columns using our standard regression approach (with the same controls as those specified on page 16 of the Data and Method Section), but restricting the sample to one used in the subsequent diff-in-diff analysis.

Table 49: Subjecting our Regression Estimates to Coarsened Exact Matching and Difference-in-differences

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS (basic)	OLS (fully specified)	DiD (PT)	DiD (PGR)	CEM (no earnings)	CEM-DiD PT (no earnings)	CEM-DiD PGR (no earnings)	CEM (earnings)	CEM-DiD PT (earnings)	CEM-DiD PGR (earnings)
NVQ2 Men	22.00	14.86	9.31	6.99	14.81	9.43	7.69	10.93	11.59	11.33
N	349309	349309	349309	349309	349309	349309	349309	349309	349309	349309
r2_a	0.014	0.121	0.041	0.008	0.107	0.040	0.007	0.107	0.042	0.008
NVQ2 Women	11.97	11.73	8.85	7.51	12.44	9.18	7.89	9.84	11.36	11.89
N	274114	274114	274114	274114	274114	274114	274114	274114	274114	274114
r2_a	0.005	0.109	0.028	0.009	0.106	0.025	0.006	0.106	0.027	0.007
NVQ3 Men	20.97	14.67	8.23	6.83	16.83	9.34	8.46	13.87	11.32	11.04
N	87035	87035	87035	87035	87035	87035	87035	87035	87035	87035
r2_a	0.016	0.156	0.138	0.031	0.144	0.135	0.027	0.147	0.138	0.029
NVQ3 Women	11.98	13.29	10.39	9.38	13.30	10.69	9.44	12.4	11.68	12.48
N	172305	172305	172305	172305	172305	172305	172305	172305	172305	172305
r2_a	0.005	0.008	0.045	0.012	0.093	0.043	0.009	0.096	0.043	0.010

NB: Estimates in columns 3 and 4 are based on the same OLS specification of column 2. The difference-in-differences estimator based on the parallel trends assumption is denoted as 'DiD PT', while its extension based on the parallel growth rate assumption is denoted as 'DiD PGR'. There are 97.73% of observations on the common support for NVQ2 men; 97.31% for NVQ2 women; 92.25% for NVQ3 men and 96.83% for NVQ3 women.

Columns 3 to 7 in Table 49 present estimates from stages in the development of results least likely to be impacted by the bias arising from unobservables (presented in Columns 8 to 10). Firstly, columns 3 and 4 simply estimate the same OLS specification, but with (respectively) the dependent variable as the difference between earnings in the period prior to the relevant learning spell and immediately after ($Y = y_{t+1} - y_{t-1}$) and the same difference but in first differences ($Y = (y_{t+1} - y_{t-1}) - (y_{t-1} - y_{t-2})$). Whilst this approach controls for average differences between treatment and control in a regression framework, it does not ensure that we are comparing 'like with like'. In contrast, column 5 presents estimates arising from the use of Coarsened Exact Matching (CEM)[44], which is only possible when we have such large sample sizes[45]. Columns 3 to 7 do not match on, or

---

[44] The matching procedure has been carried out on the most relevant covariates. We have exactly matched on the variables region, subject sector area, ethnicity and disability. The pre-treatment growth rate of earnings is coarsened at the median, while the earnings two periods before the treatment are coarsened in quartiles. The more strata the larger the loss in treated observations. This is why we do not add more coarsening variables or we do not coarsen the latter variables for a larger number of intervals.
[45] The reason we often use Propensity Score Matching (PS Match) is that Exact Matching approaches suffer from the curse of dimensionality. PS Matching allows us to reduce the problem of matching to one dimension. We match on an individual's overall propensity to be observed in the treatment, given their various characteristics; with this propensity obtained from a Probit or Logit model that predicts the probability that each individual will be observed in the treatment. CEM matches on individual characteristics, rather than on an overall propensity score (See Iacus et al., 2011).

control for, previous earnings - they do not ensure that we are comparing the earnings gain of achievers, relative to non-achievers with the same prior earnings.

Similarly, whilst columns 6 and 7 present the results of difference-in-differences estimation[46] on [CEM] matched samples of treated and control, they do not force (through matching) the earnings of achievers to be the same as comparable non-achievers prior to learning. As we can see, the general finding here is that, whenever we use an approach that includes an element of difference-in-differences, estimated premiums fall quite substantially, until we combine this with a process of matching that includes prior earnings. This suggests that achievers have slightly higher earnings (on average) prior to learning and only when we account for this by matching on prior earnings do we uncover the true difference in earnings differences (or 'gains') between achiever and non-achievers.

More specifically, our results suggest that the (unmatched) earnings trends of achievers and non-achievers are not parallel prior to learning; the earnings of achievers prior to learning are, on average, higher and on a slightly steeper trajectory than the earnings of non-achievers. Any of the difference-in-differences approaches that do not match on prior earnings (columns 3, 4, 6 and 7) capture percentage increases in earnings for achievers that are understated, because they are at a higher average absolute level prior to learning[47]. In contrast, the matching approaches without difference-in-differences that do not match on prior earnings (columns 5 and 8) produce over-estimates, because they do not take account of the higher prior earnings of achievers and simply reflect their higher post-learning earnings. Matching on prior earnings trends (two periods prior to learning), coupled with difference-in-differences better ensures that we are comparing like with like in columns 9 and 10.

Dehejia and Wahba (2002) and Heckman et al. (1998) show that controlling for past wages goes a long way to reduce the bias of matching estimators. Columns 9 and 10 therefore provide our final most robust estimates (with column 9 being our preferred estimate), because matching is also carried out on the two periods of earnings prior to learning. It is important to note the strengths and limitations of this final estimate. By using a difference-in-differences approach we go a long way to accommodate unobservable impacts that are time invariant – for instance, if there are systematic differences in the level of motivation/commitment between achievers and non-achievers and these are constant through time, the diff-in-diffs estimates should not be impacted. However, the validity of this assertion is dependent on the assumption of parallel earnings trends[48] between achievers and non-achievers. In our matching framework we 'force' the prior

---

[46] Column 7 reports the estimates using the DiD extension based on the parallel growth rate assumption instead of the classical parallel trend assumption. This extension relies on the assumption that the treated would have experienced the same trend growth rate as the controls during the treatment period for the outcome variable (see Mora and Reggio, 2013).

[47] Even though they are on a slightly steeper trend, the higher absolute level means that percentage increases are lower, relative to non-achievers.

[48] Difference-in-differences compares the gain in earnings experienced by the treated, with the gain in earnings achieved by the control (before and after treatment). For this to provide a valid estimate of the counterfactual, we must assume that in the absence of the treatment, the gap between treatment and control earnings would be the same before and after treatment (hence, 'parallel trends'). Using two time periods of earnings prior to treatment we also adopt an extension of the diff-in-diffs estimator based on the weaker 'parallel growth rate' assumption.

earnings of achievers and non-achievers to be parallel by matching on earnings in the two periods prior to learning.

As we move from the basic achiever v non-achiever specification in column 2, to estimates from diff-in-diffs on matched samples in column 9, the nature of our effective sample changes. Even though we start with OLS estimates for the diff-in-diffs sample, the various forms of matching change the nature of the control as some observations fall off the common support[49]. However, even taking this into account and the substantial set of econometric manipulations we have carried out, all estimated returns remain above 11%. The estimated returns for NVQ2 qualifications for men fall from 14.86 to 11.59% (3.27 percentage points) and for women from 11.73% to 11.36% (0.37 percentage points). For NVQ3 qualifications the fall for men is from 14.67% to 11.32% (3.35 percentage points) and for women 13.29% to 11.68% (1.61 percentage points). The large numbers involved in our estimations mean that these are almost all statistically significant falls but [especially for women], the change in results is limited and estimated premiums remain much higher than those seen in any of the survey based literature.

These findings suggest initial support for the much higher estimated earnings returns obtained from regression approaches using the achiever v non-achiever comparisons in Tables 3 to 6. There is some limited upward bias in the achiever v non-achiever regression approach, but this is much less than might be expected, and these first set of results suggest that it may be much less of a problem than the apparent downward bias in survey-based studies.

It is interesting to consider the issue of Ashenfelter's dip in our achiever v non-achiever context (Ashenfelter, 1978). This is the general finding that earnings of individuals tend to fall in the period immediately before training is undertaken, and if we do not take account of this we can over-estimate the returns to training in a diff-in-diffs framework. In our context, both achievers and non-achievers select into treatment and therefore we might expect the same pre-treatment earnings patterns for both groups. However, we have seen that the (unmatched) earnings of achievers prior to learning are, on average, higher and on a slightly steeper trajectory than the earnings of non-achievers. This is why with the CEM-DiD we control for this source of bias matching on pre-treatment earnings both in level and in growth rates.

Investigating the nature of our control groups

Whilst our use of a variety of econometric approaches has the potential to provide more robust findings, our focus has still been on comparison of achievers and non-achievers. We now utilise more of our administrative data to pursue a detailed dissection of the non-achiever control groups. First, Figure 5 presents an overview of spell length distributions for our achievers and non-achievers taking NVQ2 and NVQ3 qualifications. As one would expect, the peak in spell-length distributions for non-achievers is to the left of all those for achievers within each qualification category, but other than this there is a close correspondence in the general shape of the distributions.

---

[49] During the process of matching, some observations have to be left out because of the lack of a match – they 'fall off the common support'.

These spell duration distributions identify a group of non-achievers who drop out early from their course. This group are represented by a spike in [short] spell durations for non-achievers which is to the left of the achiever distributions. This group of non-achievers are particularly interesting as we may consider them to be more likely 'non-achieving' because of dissatisfaction with the course on which they have enrolled, rather than because of 'failure'. The main reasons individuals drop out are (i) finding that the course on which they are enrolled turns out to be a bad match to their expectations, with (ii) poor quality teaching and (iii) the course not pitched at the correct level being close seconds (Martinez and Munday, 1998; Simm, Page and Miller, 2007). Together these account for about 25-30% of the reasons for drop-out (most respondents offer only one reason for drop-out). It would seem likely that individuals that secure a 'bad learning match' drop out early and therefore it seems sensible to use this group of non-achievers with a shorter duration as a specific control group[50].

What we find is that estimated returns are actually higher when comparing NVQ achievers with non-achievers who drop-out early. Running the same regression equations for NVQ2 learners, but with the control group reduced from all relevant non-achievers to only those in the first decile of spell durations; we observe an increase in the estimated earnings premium by 1.2 percentage points. Limiting the control group to the first quartile of spell durations raises the estimated earnings premium by another 4.2 percentage points.

When we limit the control group of NVQ3 non-achievers to those who have spell durations in the first decile, the percentage returns rise by one percentage point and limiting the control the first quartile raises this by a further 3.1 percentage points. The restriction of our control group to early drop-outs does not provide unqualified support for our approach, but the fact that this results in higher estimated returns provides yet another piece of evidence in support of the achiever v non-achiever approach.

---

[50] Clearly there are individuals in this group with shorter durations who have dropped-out for other reasons (such as divorce and negative one-off financial impacts), but we may expect the bulk of those who drop-out because they are a poor match to the course to be concentrated amongst those with shorter durations. As we argue in the conclusion to this study, those who drop-out because of course quality are likely to provide the best estimate of the counterfactual outcomes for our achievers because their selection into the non-achieving group contains more of a random element.

Figure 5: Distribution of spell lengths for treated and control
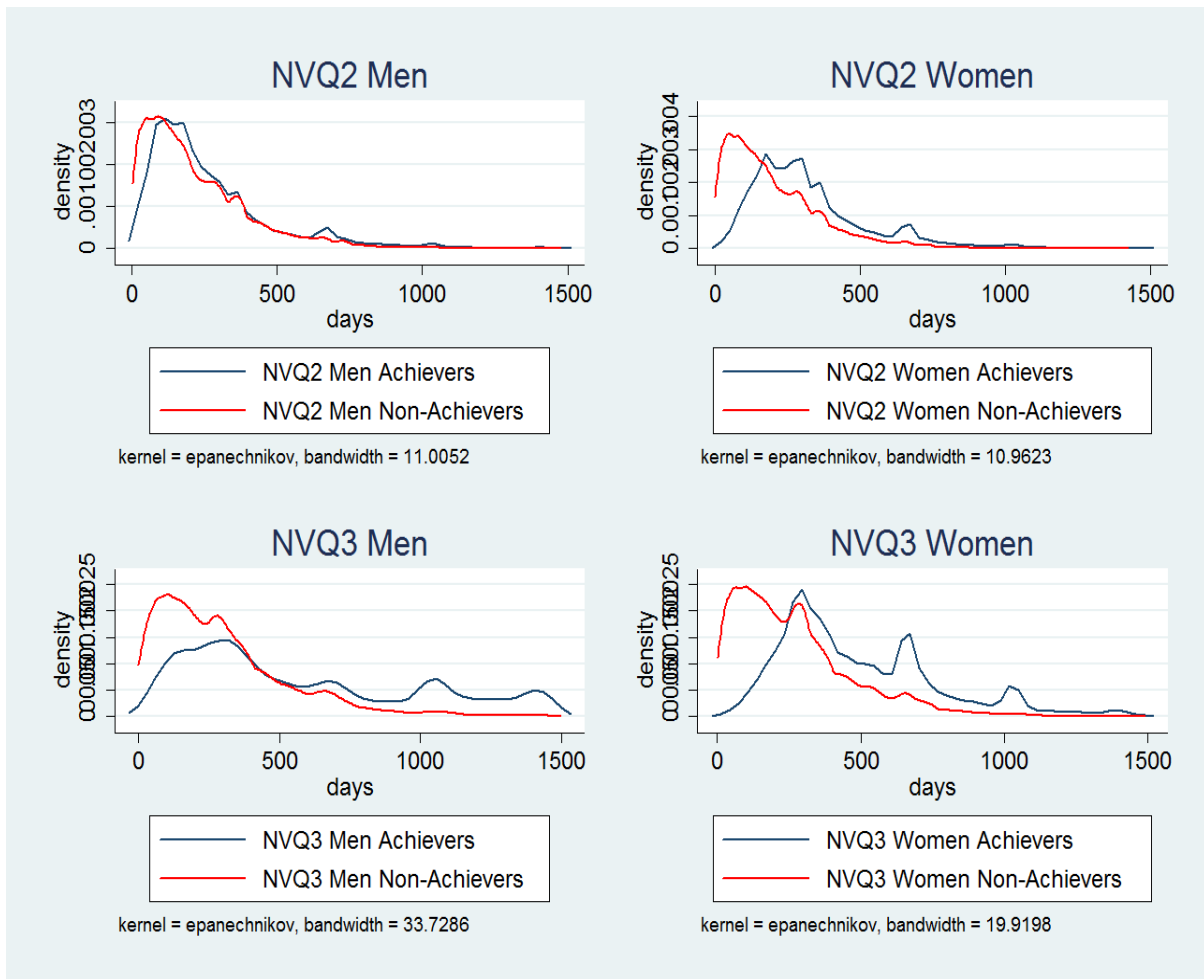
Finally, it is particularly interesting to compare the returns to non-achievers at Level 3 with the returns to Non-achievers at Level 2; returns to non-achievers at Level 3 with achievers at Level 2; non-achievers at Level 2 with achievers at Level 1 and so on. This allows some insight into the extent to which non-achievers at levels 2 and 3 are different to their 'peers' in observable and unobservable ways. For instance, the majority of non-achievers at Level 3 are a subset of achievers at Level 2[51]. Comparing their returns to other achievers at Level 2 will give us an additional insight into the validity of using those who aim, but do not achieve at Level 3 as a control group for achievers at Level 3. It is hard to speculate on exactly what we would expect to see from these comparisons. However, taking non-achievers at Level 3 as an example we might expect,

- a slightly higher return to non-achievers at Level 3 compared to achievers at Level 2 (perhaps due to any skills learnt whilst attending Level 3, even without the signal that a Level 3 qualification gives);
- a higher return to non-achievers at Level 3 compared to non-achievers at Level 2.

---

[51] FE learning journeys are often not as 'linear' as those in HE (for instance, we often find individuals attempting an NVQ2 in one area, when they already have an NVQ3 in another), but for the majority we do observe linear patterns of qualification achievement (see McIntosh, 2009).

82

Table 50 presents findings from a process of comparison between these various combinations of achievers and non-achievers, using the standard regression approaches set out above[52]. The findings suggest that the non-achieving groups conform to exactly this pattern of estimated returns. For instance, when comparing the returns of non-achievers at NVQ2 (NVQ2 NA) with non-achievers at NVQ1 (NVQ1 NA), the former group secure a substantial premium. However, when comparing NVQ3 non-achievers with NVQ2 achievers, there is a small (but significant) premium for the former of 2.4%. A similarly small, but statistically significant, premium is apparent when comparing NVQ2 Non-achievers to NVQ1 achievers.

Table 50: Estimated Earnings Premiums Returns for various Achiever (A) and Non-achiever (NA) comparisons

	NVQ1 NA	NVQ1 A	NVQ2 NA	NVQ2 A	NVQ3 NA	NVQ3 A	NVQ4+ NA	NVQ4+ A
NVQ1 NA	/	/	/	/	/	/	/	/
NVQ1 A	Basic: +12.28% All contr +9.01%	/	/	/	/	/	/	/
NVQ2 NA	Basic: +15.63% All contr +11.13%	Basic: +3.35% All contr +2.51%	/	/	/	/	/	/
NVQ2 A	/	Basic: +28.08% All contr +18.09%	Basic: +24.73% All contr +15.97%	/	/	/	/	/
NVQ3 NA	/	/	Basic: +15.77% All contr +17.23%	Basic: -8.96% All contr +2.40%	/	/	/	/
NVQ3 A	/	/	/	Basic: +4.42% All contr +18.79%	Basic: +13.38% All contr +14.65%	/	/	/
NVQ4+ NA	/	/	/	/	Basic: +40.75% All contr +27.69%	Basic: +27.37% All contr +17.23%	/	/
NVQ4+ A	/	/	/	/	/	Basic: +34.40% All contr +29.19%	Basic: +7.04% All contr +9.27%	/

If we had found that non-achievers at Level 3 did nearly as badly as non-achievers at Level 2 and much worse than achievers at Level 2 (even having controlled for all observable differences), this would raise questions over the validity of non-achievers at Level 3 as a control group for achievers at Level 3. This is clearly not the case and, taken together with the econometric results in Table 48, we have strong evidence that achiever v non-achiever returns estimated using standard regression approaches are very robust.

6.2　Gauging the persistence of returns

In this section of the report we present estimated earnings and employment returns for Full Level 2, Full Level 3 and Level 3 Apprenticeship learners, to gain some insight into the persistence of returns beyond the end of learning. To effectively analyse the persistence of returns we have to perform the analysis for separate yearly cohorts. For instance, asking the question: for those achieving their highest learning aim of a FL2 qualification in the 2004/2005 academic year (i.e. the '2005 cohort'), how do returns persist up to the sixth year after learning? For the 2006 cohort how do returns persist up to the fifth year after learning? Continuing this process up to the 2011 cohort of learners, where we only observe one post-learning year of earnings.

Adopting a cohort approach to analysis allows a clearer picture of any change in earnings returns in later years, because our standard approach to estimation (in other Sections of

---

[52] The 'Basic' premium is the estimate secured from the naïve regression, whilst 'All contr' is the estimate obtained from the fully specified model with all controls.

the report) results in each year's estimated earnings premium being composed of a slightly different sample of cohorts (see below for more detail). The findings from this section of the report determine our adoption of a 3 to 5 year average as the relevant policy metric in Sections 3, 4 and 5.

Before setting out the findings, it is important to note the definition of 'cohort' that we adopt; taking as an example the 2005 cohort for Full Level 2 learners. This contains all those who either achieve their highest learning aim of a Full Level 2 qualification in the 2004/2005 academic year (treatment) or drop-out from their highest Full Level 2 learning aim in the 2004/2005 academic year (control). Both groups are selected so that they then do not engage in a learning aim for the rest of the time that we observe them in the ILR (in this case until 6 years on from the end of learning)[53].

Table 51 presents the results of our cohort analysis for Full Level 2 achievers. The approach to estimation is the same as in other sections of the report, with a fully specified regression equation attempting to control for additional differences between achievers and non-achievers, to arrive at estimated earnings premiums. The difference is that, whilst the regression analysis in Sections 3, 4 and 5 includes individuals across all 10 years, the first row of Table 51 includes only individuals who have their latest highest learning aim in 2005. The findings from Table 51 are encouraging, in that they are relatively stable through time. By the fifth and sixth years after learning, any decline in returns over time for the 2005 and 2006 cohort seems to have flattened out. A quick calculation suggests that the difference between a 3 to 5 year average and a 4 to 6 year average, is less than 0.4 of a percentage point. Given that we only have year 6 observations for two cohorts, this inclines us towards a 3 to 5 year average, and therefore we use this elsewhere in the report.

Table 51: Estimated daily earnings premiums for different cohorts of learners whose highest learning aim is Full Level 2 in the relevant year

	1 year	2 year	3 year	4 year	5 year	6 year		
FL2 (cohort 2005)	15.23	13.7	11.86	11.44	11.69	11.04	Non-Ach.	63310
N obs	88719	85498	80723	80914	81935	84076	Achievers	93334
FL2 (cohort 2006)	16.84	14.4	12.62	13.68	12.51	11.62	Non-Ach.	57852
N obs	100690	94617	94194	95109	97497	95633	Achievers	118240
FL2 (cohort 2007)	16.77	13.4	12.9	11.79	10.39	/	Non-Ach.	58074
N obs	102638	101393	101165	103441	100908	/	Achievers	136253
FL2 (cohort 2008)	15.09	12.67	12.36	11.79	/	/	Non-Ach.	80649
N obs	155688	153376	155516	150949	/	/	Achievers	209731
FL2 (cohort 2009)	16.34	14.53	12.68	/	/	/	Non-Ach.	103371
N obs	235092	237628	229170	/	/	/	Achievers	319300

---

[53] This does have the potential to slightly alter the nature of our treatment and control groups, by bringing in an element of differential selection for any one cohort. This is not likely to alter our findings substantially, but it is an issue that is worth considering in any future work on cohorts.

	1 year	2 year	3 year	4 year	5 year	6 year		
FL2 (cohort 2010)	13.9	11.16	/	/	/	/	Non-Ach.	107870
N obs	337672	323569	/	/	/	/	Achievers	437347
FL2 (cohort 2011)	11.72	/	/	/	/	/	Non-Ach.	108496
N obs	355474	/	/	/	/	/	Achievers	486706

Table 51 gives a clear indication of which cohorts make up our estimates for the 1st year of earnings, 2nd year etc. in the remainder of the report – for instance we can see that seven cohorts contribute the estimate of 1st year earnings returns in Table 3. In contrast, Figure 6 presents the results of Table 51 in a way that allows us to gauge whether particular calendar years are having an impact on our estimated earnings returns – for instance, we can see whether the start of the recession in 2007/2008 is associated with a slight dip in returns for a earlier cohorts (independent of whether this is their 1st, 2nd, 3rd or other year after learning). There does not seem to be anything particularly pronounced in this respect from Figure 6.

Figure 6: Daily earnings premiums for cohorts of learners whose highest learning aim is Full Level 2 in the relevant year
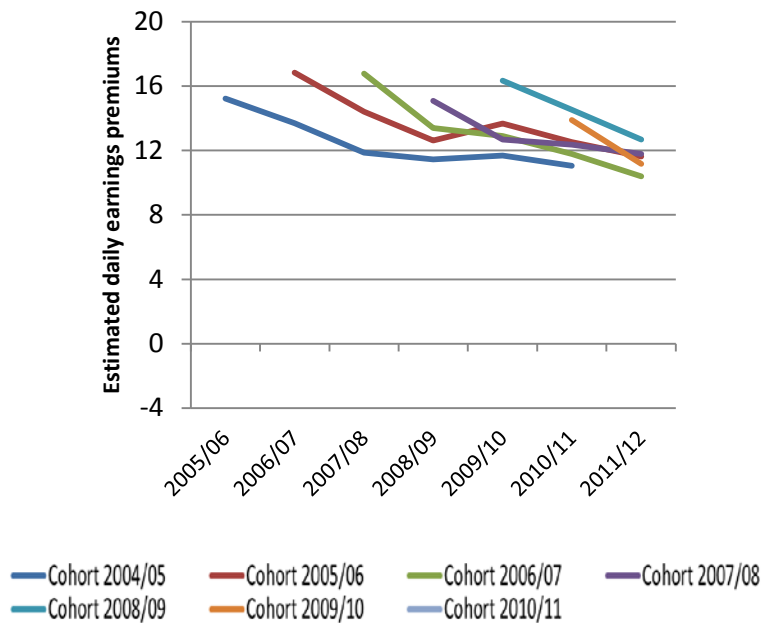
The results in Table 52 for Full Level 3 learners (which includes apprenticeships in the definition) are problematic. We observe a large dip in returns for the earlier cohorts in the second year after learning. This seems driven by the HE flag issue, once again, as returns tend to pick up 3 or 4 years from the end of learning (when achievers will have finished HE). To see how we investigate this further, please refer to the discussion in Section on *The Impact of, and motivation behind, Methodological Enhancements* on page 25.

Table 52: Estimated daily earnings premiums for different cohorts of learners whose highest learning aim is Full Level 3 in the relevant year

	1 year	2 year	3 year	4 year	5 year	6 year		
FL3 (cohort 2005)	1.13	-2.9	3.11	6.66	11.48	12.25	Non-Ach.	46032
N obs	102636	100410	101503	104880	109317	113268	Achievers	126245
FL3 (cohort 2006)	3.71	0.35	3.23	8.84	12.97	12.8	Non-Ach.	45737
N obs	125506	121852	127612	135186	140430	138589	Achievers	165950
FL3 (cohort 2007)	2.8	-1.31	4.7	10.94	13.72	/	Non-Ach.	42503
N obs	141437	142321	152473	161197	159407	/	Achievers	206804
FL3 (cohort 2008)	4.17	2.27	7.09	11.8	/	/	Non-Ach.	43956
N obs	148366	149392	158950	157117	/	/	Achievers	231700
FL3 (cohort 2009)	14.12	9.64	12.97	/	/	/	Non-Ach.	50781
N obs	190403	195797	192290	/	/	/	Achievers	271825
FL3 (cohort 2010)	11.79	7.99	/	/	/	/	Non-Ach.	57804
N obs	251856	240828	/	/	/	/	Achievers	342865
FL3 (cohort 2011)	6.25	/	/	/	/	/	Non-Ach.	65186
N obs	261912	/	/	/	/	/	Achievers	379744

Figure 7 presents the results of Table 52. There is a lot more work to be done on this to identify exactly why we see so much less of an under-estimation of returns after the 2008 cohort, but there is one explanation that seems particularly intuitively appealing. From 2008 the UK experienced a recession that raised levels of unemployment, but as of May 2014, 73.1 per cent of those aged 16-64 were in employment – a high that has not been seen since 1974 and 2004-2005 (ONS). UK unemployment has been relatively subdued in the present recession, not rising much above 8% even at its peak in 2011/2012 and falling below 7% in 2014. There are many reasons for this, but one manifestation has been a high proportion of individuals working part-time who report that they would prefer a full-time job. This may have limited opportunities for students in HE to engage in the same levels of part-time working and therefore reduced the confounding impacts of HE on our estimates. This is one of many possible explanations, and we cannot rule out other data issues, but even without further research the pronounced differences between pre-, and post-2008, cohorts suggests a labour-market explanation.

Figure 7: Daily earnings premiums for cohorts of learners whose highest learning aim is Full Level 3 in the relevant year
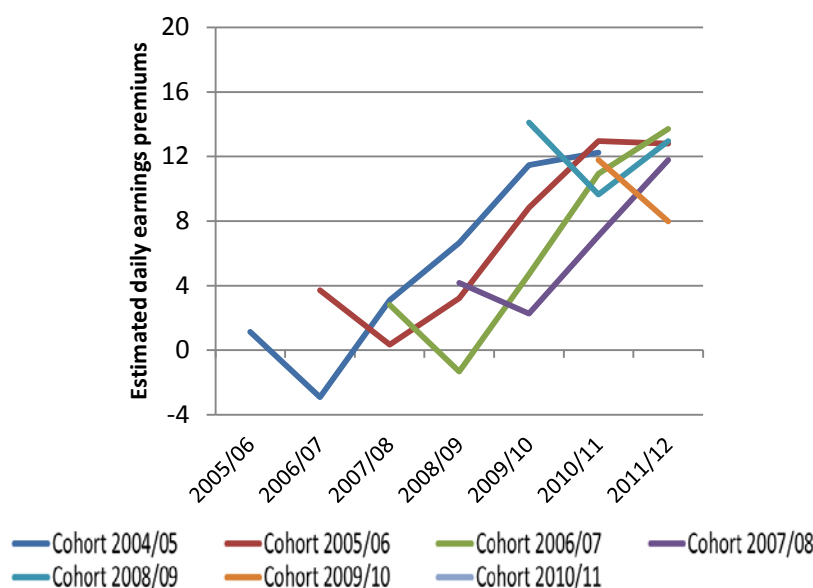

Table 53 and Figure 8 consider Apprenticeship learners at Level 3 and show some minor dip in returns for certain cohorts in the second year after the end of learning, but overall the table shows stable and solid returns, with no indication that returns are falling over the years. A quick calculation suggests that a 4 to 6 year average would produce higher estimated returns than a 3 to 5 year equivalent, but again with a very small difference (at most 0.4 of a percentage point). Clearly, from Figure 7 and Table 52 we have more of an issue with our estimated returns for FL3 when taking a 3 to 5 year, or other, average. The 3 to 5 year average seems best placed to counter some of the problems with our data, but as we suggest in the conclusion, if there are substantial numbers of FL3 learners going on to HE, then some of the returns we capture will be to degree-level study – an important issue for further research.

Table 53: Estimated daily earnings premiums for different cohorts of learners whose highest learning aim is Level 3 Apprenticeship in the relevant year

	1 year	2 year	3 year	4 year	5 year	6 year		
FL3 Appr. (cohort 2005)	14.78	10.34	14.86	13.94	15.59	16.13	Non-Ach.	46032
N obs	12833	12319	11798	11881	12218	12664	Achievers	126245
FL3 Appr. (cohort 2006)	17.35	14.3	15.62	14.73	15.42	15.89	Non-Ach.	45737
N obs	19596	18679	18658	19324	20005	19823	Achievers	165950
FL3 Appr. (cohort 2007)	16.45	16.97	16.98	17.99	15.05	/	Non-Ach.	42503
N obs	24623	24444	25043	26033	25608	/	Achievers	206804
FL3 Appr. (cohort 2008)	13.24	16.45	14.42	12.91	/	/	Non-Ach.	43956
N obs	29716	29910	31046	30487	/	/	Achievers	231700
FL3 Appr. (cohort 2009)	19.74	15.89	17.76	/	/	/	Non-Ach.	50781

	1 year	2 year	3 year	4 year	5 year	6 year		
N obs	35559	36639	35735	/	/	/	Achievers	271825
FL3 Appr. (cohort 2010)	16.13	12.59	/	/	/	/	Non-Ach.	57804
N obs	50501	48604	/	/	/	/	Achievers	342865
FL3 Appr. (cohort 2011)	13.74	/	/	/	/	/	Non-Ach.	65186
N obs	57934	/	/	/	/	/	Achievers	379744

Figure 8: Daily earnings premiums for cohorts of learners whose highest learning aim is Level 3 Apprenticeship in the relevant year
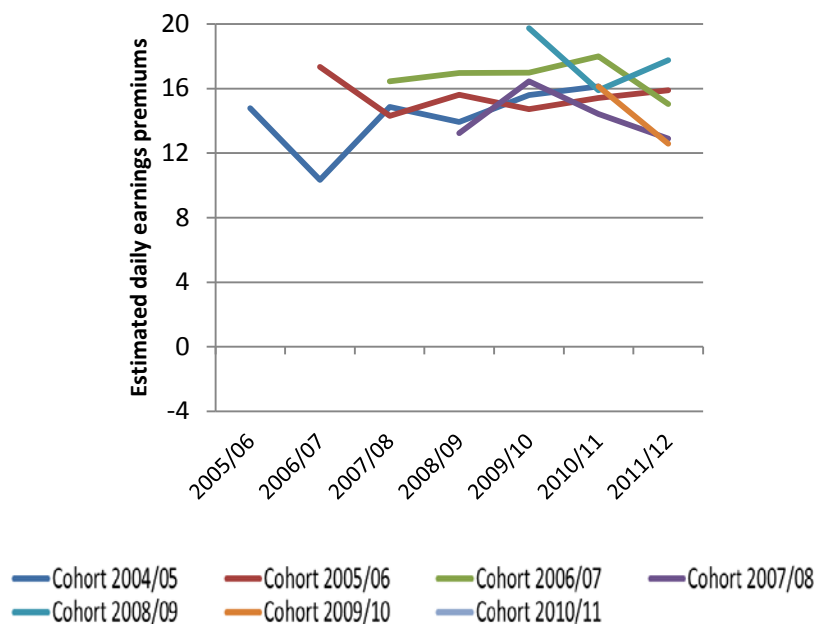
Figures 9 and 10 present the estimated percentage point employment probability premiums for different cohorts of learners. We have already mentioned the interesting pattern of employment probability premiums across cohorts (on page 27 when considering the change in our results compared to B&U and also in discussions where employment premiums fall in later years, for instance Table 14). Figure 9 suggests that the 2008 cohort of FL2 learners have sustained particularly high employment probability premiums compared to non-achievers of the same cohort, when compared to cohorts before and after. Figure 10 suggests that cohorts from 2007/2008 onwards have secured higher returns than those before the recession.

Figure 9: Employment probability premiums for cohorts of learners whose highest learning aim is Full Level 2 in the relevant year
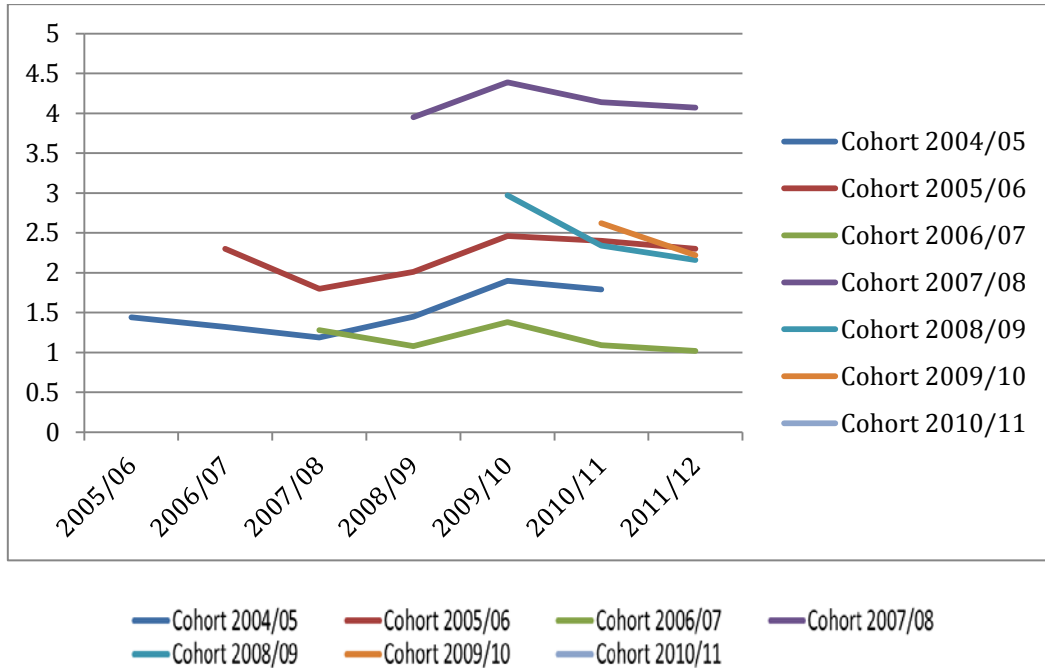
There are a number of factors possibly driving the patterns in Figure 9 and Figure 10 and further research is needed. We seem to have evidence that those achieving qualifications at Full Level 3 are better placed to secure employment in a recession than non-achievers, but for FL2 achievers there is an apparent one-off impact of recession for those in the 2007/2008 cohort. These are very interesting initial findings and more work on cohorts would be extremely valuable to inform a variety of policy debates. However, for the figures presented elsewhere in this report, Figures 9 and 10 suggest that (i) for each cohort, percentage-point employment probability premiums exhibit persistence and (ii) premiums may be higher for recessionary cohorts. The implication is that, for FL3 our estimated 3 to 5 year average employment probability premiums will be dominated by the slightly lower pre-recession employment probability premiums, because we do not have a long enough time period to observe earnings for post-2008 cohorts. This may be more appropriate looking forward to the post-recessionary era. For FL2 this may be less of an issue, as we only seem to have a one-off effect of recession.

Figure 10: Employment probability premiums for cohorts of learners whose highest learning aim is Full Level 3 in the relevant year
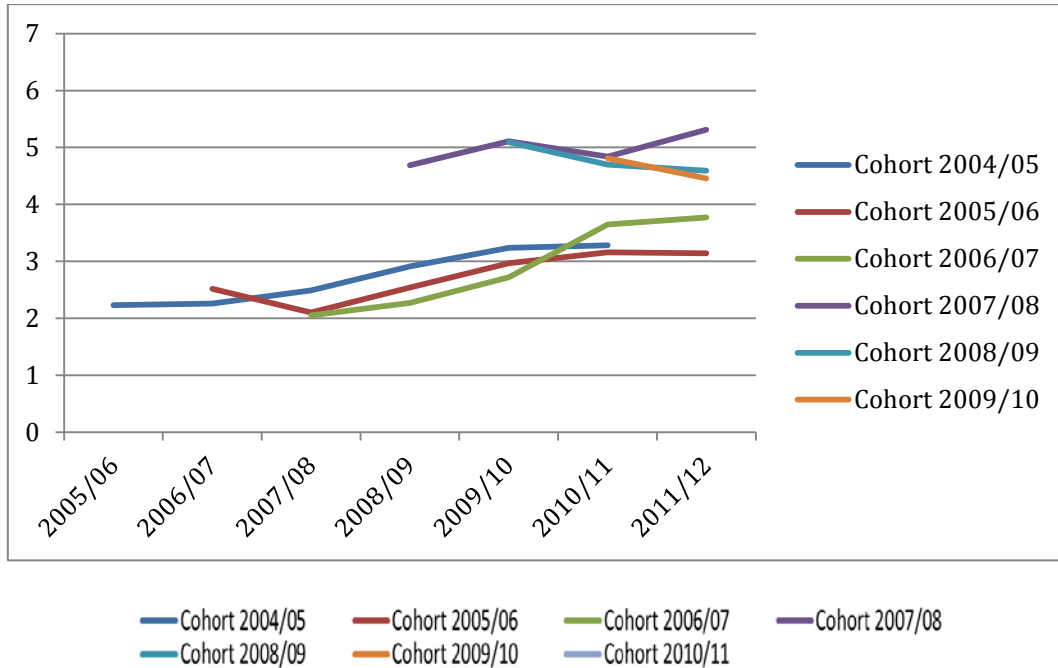
The results of this analysis inform our decision to present returns for the 1st, 3rd, 4th, and 5th tax year after learning in other sections of this report, and to suggest that the relevant policy metric is the 3 to 5 year average. To be clear on the motivation for our decision, it is important to consider the following implicit trade-off, which is in addition to the specific data issue mentioned already.

We need a measure (averaged over a number of years from the end of learning) that is justified because (on average) returns persist beyond this point. However, the proof of such persistence is always going to be based on the evidence of returns from cohorts at the very start of our period for analysis. Even if we were able to include additional years of earnings in Tables 51 to 53, any estimated returns in these additional years would be based on returns for very early cohorts – whilst the earnings returns would be occurring in the current period (up to 2012), they would be returns to those who finished learning in 2005. There is a question over whether one would want to base policy on an estimate for all learners that is essentially the return accruing to older cohorts, who may have received a slightly different learning package.

We have suggested a 3 to 5 year average as it balances these competing issues – going with a 4 to 6 year average would leave us with only one cohort making up the estimate of year 6, which is as much of a concern as persistence. Also, it makes very little difference whether one adopts the 3-5 or 4-6 year average in most of the instances discussed in this section.

7. Conclusions

In Sections 3,4 and 5 of this report we have updated the findings from Buscha and Urwin (2013), estimating separately the (i) earnings, (ii) employment probability and (iii) probability of being on active benefits, for those who achieve their highest learning aim whilst studying at an English Further Education Institution (FEI), relative to those who have the same highest learning aim, but do not achieve. Estimated returns to qualifications, when held as an individual's highest, are obtained using the 2002-2012 ILR-WPLS administrative dataset.

For each learner we have 7 academic years (2004/2005 to 2010/2011) when they can possibly exit learning (as an achiever or non-achiever) and over these academic years, a learner can have multiple ILR learning spells. Across all learning spells for each individual we select the highest learning aim. The estimate of value added is obtained by comparing the returns of those who have a particular highest learning aim (for instance Full Level 2) and achieve; with the returns of those who have the same highest learning aim, but do not achieve. The one exception to this approach is in Section 4.3, where we attempt to capture returns to L1 or L2 Maths/English qualifications when taken alongside other qualifications.

In Section 6 we have carried out a more advanced econometric investigation of the findings and this provides important quality assurance that these estimates (obtained using standard regression approaches with ILR-WPLS data) are robust. In the following discussion we highlight areas for further improvement (most notably, tackling the HE flag issue and the problems of identifying part-time working in the data). Whilst these enhancements would further improve the accuracy of estimates, in various sections of the report we have underlined where, at present, they may lead to slightly under-estimated returns. This report provides BIS with estimates of the value added of FE that are quality assured to the highest academic standards.

From the evidence presented in this study, what type of learning works best and for whom?

Full Level 2, held as highest FE qualification aim:

Earnings: When held as a highest learning aim, the 3 to 5 year average earnings premium for FL2 achievers over non-achievers is 11.3%, made up of a figure for women that is slightly lower (9.4%), and one for men, that is higher (13%). The estimate for women may be understated because many work part-time and we are not able to fully account for this in our data. Amongst learners aged 19 to 24 with a highest aim of Full Level 2, achievers secure a 3 to 5 year average earnings return of 10.4%, compared to 8.1% for those aged 25+. Learners in our Academic Full Level 2 category secure an average 3-5 year return of 11.6%; 10.3% for FL2 Vocational CBL and 9.6% for Vocational WBL.

In female-dominated sectors such as Adult Social Care, average 3 to 5 year earnings returns are just above 10 per cent at FL2. Even where we see lower earnings returns, for instance at FL2 for Child Development and Wellbeing; nowhere do these dip below 5%. In male-dominated areas such as Construction, Engineering & Manufacturing and Transportation we see earnings returns that are close to 10 percentage points higher than

this. Across these Sectors we seldom see average three to five year returns dropping below 15% - only in the case of FL2 learners in Engineering and Manufacturing. In areas that are neither male nor female dominated such as ICT, Customer service and Hospitality and Catering, earnings returns at FL2 are around 10% and in Business Studies just below at 7.5%.

Employment: Those achieving a FL2 qualification as their highest learning aim, have employment rates 1.5 percentage points higher than comparable non-achievers. These FL2 employment premiums seem higher for post-2007 cohorts of learners and this figure of 1.5 percentage points may therefore be closer to 2 percentage points for cohorts achieving in the present environment. Men achieving a FL2 qualification have employment rates 0.5 of a percentage point higher than non-achievers, whilst the comparable figure for women achievers is 2.6 percentage points. The 25+ group of FL2 achievers secure an estimated employment probability premium of 1.8 percentage points over non-achievers, whilst for achievers aged 19-24 the figure is only 0.2 percentage points.

In Adult Social Care, Child Development and Wellbeing and Hair and Beauty average 3 to 5 year employment returns vary between 2.5 and 3.3 percentage points at FL2. There is some concern over employment returns for those achieving a highest FL2 qualification in Engineering and Manufacturing. Similarly, those achieving FL2 qualifications in ICT, Customer Service and Business Studies experience periods of negative returns. Those achieving Hospitality and Catering qualifications at FL2 secure returns of 2.5 and 1.8 percentage points respectively.

Active Benefits: The 3 to 5 year average for Full Level 2 learners suggests that achievers are 2.1 percentage points less likely to be observed on active benefits, but because only about 10% of non-achievers and 5% of achievers are on benefits after learning, this 2.1 percentage point difference translates into an approximate 28% difference.

Full Level 3, held as highest FE qualification aim:

Earnings: The 3 to 5 year average earnings premium for FL3 achievers over non-achievers is 8.5%, made up of a figure of 8.4% for women and 8.7% for men. These estimates may understate true earnings returns (due to the related issues of subsequent HE learning and part-time working), especially for women achievers. For those aged 19 to 24 we estimate a return of 10.2% for Full Level 3 achievers compared to a figure of 9.6% for those aged 25+.

At both Level 3 and Full Level 3 we observe negative and significant returns to learning in the first and third years for our Academic category of learners. This re-enforces continuing concern over unobserved moves to HE, which do not seem to have as much of an impact on the earnings premium for Vocational CBL Level 3 achievers, which rise steadily from 2% to 3.6% (giving a 3 to 5 year average of 2.9%). Similarly, Vocational CBL Full Level 3 achievers secure a 3-5 year average premium of 5.8%, although the underlying estimated return rises from 1.9 to 9.1 over the five-year period. The return to Vocational WBL Full Level 3 Achievers averages 9% between 3 and 5 years.

In female-dominated sectors such as Adult Social Care, average 3 to 5 year earnings returns are just above 10 per cent at FL3, and the same is true of achievers at FL3 in the area of Child Development and Wellbeing. In male-dominated areas such as Construction,

Engineering & Manufacturing and Transportation we see earnings returns that are close to 10 percentage points higher than this. Across these Sectors we seldom see average three to five year returns dropping below 15%.

Employment and Active Benefits: The highest 3 to 5 year average employment return accrues to Full Level 3 achievers, who have employment rates that are 3.8 percentage points higher than non-achievers. This employment premium has become more pronounced for post-2007 cohorts and may now be closer to 5 percentage points for contemporaneous cohorts. Women Full Level 3 achievers have employment rates that are 4.3 percentage points higher than non-achievers, whilst for men the figure is 3.2 percentage points.

The 25+ group of achievers have an estimated employment probability premium of 1.8 percentage points at Full Level 3 that is higher than that secured by achievers aged 19-24 (1.3 percentage points). At Full Level 3, achievers in the Academic category of qualifications secure the highest estimated return over non-achievers, with a 4.1 percentage point premium; compared to figures of 3.3 and 1.9 for Vocational CBL and Vocational WBL categories, respectively.

In Adult Social Care, Child Development and Wellbeing and Hair and Beauty average 3 to 5 year employment returns are 4.2, 5.9 and 1.8 percentage points respectively at FL3. There is some concern over employment returns for those achieving a highest FL3 qualification in Construction and in Transportation. At FL3 we observe instances of negative and statistically significant employment probability returns in Customer Service. Those achieving Hospitality and Catering qualifications at FL3 secure a return of 1.8 percentage points; and FL3 Business Studies achievers secure a 5.7 percentage point average return between 3 and 5 years from the end of learning.

For Full Level 3, achievers have a probability of being on active benefits that is 1.8 percentage points lower than the probability for non-achievers and this translates into an approximate 40% difference.

Below Level 2, L2 and L3 held as highest FE qualification aim

Earnings: We observe smaller (but still statistically significant) 3 to 5 year earnings premiums of 1.9% for those achieving a highest qualification Below Level 2; 1.3% for Level 2 and 3.3% for Level 3. These 3 to 5 year averages are a good reflection of returns that remain relatively stable over the period of analysis, with the exception of Level 3, which suffers from the possible confounding effects of unobservable moves to HE. However, the returns at Below Level 2 and L2 are entirely driven by the experiences of men, with women achievers securing no significant and systematic earnings return that allows us to quote a valid three to five year average. In contrast, their male counterparts secure a 4.6% and 3.2% average return three to five years from the end of learning. The female earnings returns may be understated because many work part-time and we are not able to fully identify this in our data. It is a concern that 25+ achievers, amongst the 875 thousand whose highest aim is Level 2, secure a negligible premium on average.

Employment: The 3 to 5 year average for L2 and L3 learners is 0.9 and 1.4 percentage points respectively. The figure of 0.9 for L2 is a little better than it seems, because L2 achievers and non-achievers have lower absolute levels of employment (and therefore

small percentage point differences amount to larger percentage differences). However, even low levels of employment amongst learners whose highest achievement is Below Level 2, does not alter the suggestion that the size of any employment effect (0.3 of a percentage point) is very small - mainly driven by experiences in vocational work-based [WBL] and classoom-based [CBL] courses. Those achieving a highest aim of Level 2 in the WBL category secure no significant employment premium in the fourth and fifth years after learning (despite stronger performance in the 1st and 3rd years); whilst the 3 to 5 year average return to achievers at Level 2 in the Vocational CBL category is 0.9 of a percentage point.

Active Benefits: The difference in active benefit probabilities between achievers and non-achievers for Level 2 learners is 0.6 of a percentage point (approximately 11%), and 0.9 of a percentage point for Level 3 learners (approximately 26%). For those achieving a highest qualification Below Level 2 we observe poor returns, with our 0.3 percentage point estimate only translating to a 5% lower probability of achievers being on active benefits compared to non-achievers following learning.

L2 and L3 Apprenticeship returns

The 3 to 5 year average earnings premium for Level 2 apprenticeship achievers is estimated at 11.1% and for level 3 achievers the figure is 15.6%. Estimated employment returns are less pronounced and this is perhaps to be expected as learners are usually employed during the period of their apprenticeship, and even non-achievers will likely have a high probability of continuing in employment. For instance those completing, but not achieving, their apprenticeship qualification are likely to exhibit skills that mean they are more likely to be retained in the workplace. In contrast, both L2 and L3 apprenticeship achievers have a significantly lower percentage point probability of being on active benefits (2.5 and 1.9 percentage points, respectively). Because the absolute proportions of apprenticeship achievers and non-achievers is relatively low, these translate into figures that suggest L2 achievers have an approximate 30% lesser probability of being on active benefits and for L3 achievers the figure is 48%.

L1/L2 Maths and English qualifications, not held as highest FE qualification aim

Elsewhere in the report we are capturing returns to qualifications gained as an individual's highest learning aim. When taking such an approach at L2 and below, this misses many learners, as we will only analyse the relatively unique group of individuals whose highest learning aim across the 2004/2005 to 2010/2011 cohorts of learners is a L1 or L2 English or Maths qualification. As we have seen in the previous discussion, returns tend to be low for this group of learners. In contrast we find statistically significant returns to L1 and L2 Maths and English, when we adopt an approach to estimation that accommodates the complementary nature of these qualifications.

The 3 to 5 year average earnings premium for those achieving a L1 and/or L2 Maths/English qualification, relative to those who do not achieve their L1/L2 Maths or English qualification is 2.4% (with this estimate relevant for a population of all FL2 and FL3 achievers). More specifically, we estimate that Literacy/Numeracy Key Skills achievers secure a 3 to 5 year average earnings return of 6% over non-achievers (amongst a population of all FL2 achievers) and 3.7% (for a population of all FL3 achievers).

Our findings on earnings returns that consider those taking L1/L2 English/Maths qualifications as forms of complementary learning refer to around 450,000 achievers (amongst our population of FL2 and FL3 achievers) between 2002 and 2012. In contrast, our less favourable findings on earnings for those studying qualifications at L2 and Below L2 as a highest learning aim apply to around 750 and 900 thousand achievers respectively over the same time period. This underlines the diversity of experiences that we are likely to observe at these levels of learning. Even the less favourable average returns for those studying Below Level 2 and L2 have the potential to hide positive experiences amongst sub-groups of the population[54], such as those who are unemployed when they engage in learning; who are the focus of a current BIS study.

How Robust are our Estimated Returns?

The process we have gone through to test the robustness of our estimates, has been peer-reviewed by four academic experts and we are very grateful for their comments. Any mistakes remain our own and readers can consider the comments of reviewers and our responses at the end of this report. Our findings suggest that regression-based techniques, which compare achievers and non-achievers in the ILR-WPLS, produce robust estimates of value added.

We identify the following continuing limitations and potential enhancements for future research:

First, it is important to note that we are forced to incorporate the information on HE learning in a less-than-perfect way, and as a result the HE flag in our regression equations may be endogenous. In such a situation we would expect a falsely inflated level of significance for the variable (see the online Technical Annex for more details). In various sections of the report we have highlighted where this is an issue. For instance,

- At various points in our analysis we observe returns that are clearly understated in the years prior to the fourth tax year after the end of learning for L3 and FL3 achievers who are more likely to move on to Higher Education and take lesser-paid part-time jobs, compared to non-achievers.
- We observe negative returns to those achieving L1/L2 Maths and English for a population of FL3 achievers. This is likely driven by the same phenomenon, with those achieving L2 Maths or English qualifications that they need to get into university (alongside the FL3) working in less well-paid part-time jobs while at university.
- In areas where we expect to see much less progression to HE, such as in FL3 work-based qualifications, returns average around 9% over 3 to 5 years from the end of learning and we see much less of a dip in returns before this.
- The less-than-full coverage of our HE flag, for those studying L3 and FL3 as their highest learning aim, seems particularly acute for women. It may be that the apparent lower earnings return secured by women studying at Level 3 and Full Level 3 in the first years after learning, when compared to men, simply reflects the greater propensity for

---

[54] Readers should refer to a recent BIS-commissioned study that identifies potential diversity in the impact of Below L2 learning https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/253585/bis-13-1261-evaluation-of-the-impact-of-learning-below-level-2.pdf

young women achievers to work part-time when progressing to HE, compared to men achieving at this level.

- The problematic dip in returns for FL3 learners is much less pronounced after the start of the recession in 2008. There has been a much higher proportion of individuals working part-time who report that they would prefer a full-time job since this time, and this may have limited opportunities for students in HE to engage in the same levels of part-time working (reducing the confounding impacts of HE on our estimates).
- Our inability to explicitly identify part-time working in the data is an important part of this HE flag issue. This may explain the lower earnings returns to women at Full Level 2 and above, which are compensated by improved employment premiums.

This suggests a need for more work to integrate a more robust indicator of HE learning into the data and also raises the wider issue of progress to further learning (whether HE or otherwise).

First, we may suggest that the HE-flag issue is much less of a problem when considering a three to five year average for returns; as it only acts to depress our estimated returns up to the third year after learning. However, this is not always necessarily the case. Unfortunately, if the above explanations of HE learning and part-time working are correct, when we see returns recover in the 4th and 5th years after learning, we may not be able to attribute all of the upturn to L3 and FL3 learning. By this point many of our hidden HE learners will have obtained their degree and we will actually be capturing the return to qualification at Level 6 (the level at which individuals complete a three year degree). As we suggest, Buscha and Urwin (2013) have a separate piece of analysis restricted to a subset of learners for whom we hold HE data, and their estimates of 6% for L3 achievers and 14% for FL3 achievers should be considered as the most appropriate.

Any future study would ideally introduce more data from HESA identifying (i) the specific destination of FE learners who move on to HE and (ii) the drop-out rate of these learners. In this report we are concentrating on earnings, employment and active benefit outcomes, but for many FE learners, HE is an important and valuable outcome – but only if these learners progress within HE.

This would allow us to better capture returns to those who achieve at L3 and FL3 and then progress to the labour market. This touches on the wider issue of progression to further study, and it is worth emphasising again that we need to keep this in mind when considering our findings at Below Level 2 and Level 2. Our estimates of the returns to these qualifications when held as a highest learning aim are for those who do not progress to further study. It may be that those achieving Below Level 2, who go on to further study, are achieving very good returns to this level of qualification, because it is an essential facilitator of their move to this higher level of learning.

Finally, our exploratory analysis of cohorts raises some very interesting issues. For FL2 achievers there is an apparent one-off employment impact of recession for those in the 2007/2008 cohort – with achievers much more likely to secure employment than non-achievers. FL3 achievers in cohorts after 2007 secure much higher percentage point employment premiums than those before the recession.

For the FL2 learners there will be a small impact of this in our estimates, but nothing substantial. In contrast, estimated 3 to 5 year average employment probability premiums

for FL3 learners will be dominated by the slightly lower pre-recession employment probability premiums, and we flag where this is likely to occur. In contrast, for those who have secured jobs, the nominal earnings impacts of a recession are minimal (though real earnings may of course have stagnated or fallen).

This is an important issue for further research and it is the first time that such findings have been uncovered, as the data have previously not been available to carry out such an analysis. For policymakers, the question is whether the pre-recession returns or recessionary returns are more appropriate for the immediate future. We would suggest that the balance between the two that we achieve in the present study is highly appropriate in the present uncertain economic environment.

As a final related point, there is some interest amongst policy colleagues of how changes in the funding arrangements for courses over time may impact results. However, our analysis excludes courses that are not accredited to an NQF level, and whilst we have added some learners to the earliest cohorts, our analysis of the returns to different cohorts of FL2, FL3 or Apprentice FL3 learners shows little systematic difference between earnings and employment returns in our very earliest 2004/2005 cohort (where many of these additional learners will appear) and other pre-recession cohorts. This is not unequivocal, if many of our additional early-cohort learners studied at L2 or Below Level 2. However, if their learning were associated with strongly negative atypical returns this would show up as a fall-off in estimated returns in the 4th and 5th years after learning. As we can see from most of the Tables from 3 onwards, earnings and employment premiums change very little between the first and fifth years after the end of learning and if anything there is a slight increase in the third, fourth and fifth years.

References

Bibby, D., Knight, T., Speckesser, S. and Thomson, D. (2012), "Measuring Further Education Outcomes Using Matched Administrative Data: Constructing a Research Database", RM Data Solutions and Institute for Employment Studies

Blundell, R. and Costa Dias, M. (2009), "Alternative Approaches to Evaluation in Empirical Microeconomics", *Journal of Human Resources*, 44 (3): 565-640.

Buscha, F. and Urwin, P. (2013), "Estimating the labour market returns from qualifications gained in English further education using the Individualised Learner Record (ILR)", BIS research and analysis report.

Conlon, G., Patrignani, P. and Chapman, J. (2011), "Returns to Intermediate and Low Level Vocational Qualifications", BIS Research Paper No. 53.

Dearden, L., McGranahan, L. and Sianesi, B. (2004), "An in-depth analysis of the returns to national Vocational Qualifications obtained at Level 2", Centre for the Economics of Education, Discussion PaperNo. 46.

Dickerson, A. and Vignoles, A. (2007), The Distribution and Returns to Qualifications in the Sector Skills Councils, SSDA Research Report No. 21.

Garrett, R., Campbell, M. and Mason, G. (2010), "The Value of Skills: An Evidence Review", the UK Commision for Employment and Skills and the National Institute for Economic and Social Research.

Greenwood, C., Jenkins, A. and Vignoles, A. (2007), "The returns to qualifications in England: Updating the evidence base on Level 2 and Level 3 vocational qualifications", Centre for the Economics of Education, Discussion PaperNo. 89.

Hodkinson, P. and Bloomer, M. (2000), "Accountability, audit and exclusion in Further and Higher Education", Paper presented at SCUTREA, 30th Annual Conference, 3-5 July 2000, University of Nottingham.

Hogarth, T., Gambin L., Hasluck, C., de Hoyos, M. and Owen, D. (2009), "Maximising Apprenticeship Completion Rates", *Learning and Skills Council*.

Lechner, M., Miquel, R. and Wunsch, C. (2011), "Long Run Effects of Public Sector Sponsored Training in West Germany", *Journal of the European Economic Association*, 9: 742–784.

Martinez, P. and Munday, F. (1998), "9,000 voices: student persistence and drop-out in further education", *Further Education Development Agency*.

McIntosh, S. (2007), "A cost benefit analysis of apprenticeships and other vocational qualifications", Department for Education and Skills Research ReportRR834.

McIntosh, S., (2009), "The economic value of intermediate vocational education and qualifications", UK Commission for Employment and Skills, Evidence Report 11.

National Audit Office (2012), Estimating economic benefits from apprenticeships – technical paper, February.

Patrignani, P. and Conlon, G. (2011), The long-term effect of vocational qualifications on labour market outcomes, Department for Business, Innovation and Skills, Research Paper No. 47

Simm, C., Page, R. and Miller, L. (2007), Reasons for Early Leaving from FE and Work-Based Learning Courses, *Department for Education and Skills*, Research Report RR849.

Descriptive statistics

	Below Level 2		Level 2		Full Level 2		Level 3		Full Level 3		Level 4+	
	A	NA	A	NA	A	NA	A	NA	A	NA	A	NA
Proportion of women (%)	53.55	52.47	58.25	56.19	43.57	48.88	55.59	55.54	58.40	56.86	60.31	54.05
Average age	36.05	35.78	37.06	35.95	32.82	28.46	34.57	31.91	23.47	24.95	33.59	33.25
Proportion of learners with an offender history (%)	3.30	5.66	4.40	3.61	0.27	0.20	0.43	0.46	0.03	0.04	0.04	0.03
Proportion of learners with white ethnicity	72.24	69.93	80.95	77.60	83.45	83.53	82.86	78.55	82.08	81.47	84.04	78.54
Proportion of learners with no LSC Funding	5.47	5.14	5.99	5.93	1.30	2.02	10.73	8.99	1.54	2.04	39.93	46.29
Proportion of full-time attendants	7.79	5.83	7.31	5.25	5.32	7.64	17.52	13.61	32.51	18.34	14.78	13.68
Proportion of learners with learning difficulty	5.10	4.47	2.79	2.74	3.84	4.05	2.70	2.68	4.44	3.69	2.67	2.56
Proportion of non-UK domicile	5.94	5.39	5.41	3.85	4.83	3.97	3.72	2.76	3.71	3.45	5.48	4.83
Prob. of living in Greater London	10.85	13.39	9.40	12.52	7.27	9.07	9.84	10.47	9.34	8.59	8.08	8.78
Prob. of being employed 1 year before learning spell	46.65	43.69	58.32	54.83	57.48	49.97	58.58	54.67	36.09	50.26	67.19	67.42
Prob. of having spent at least 1 day on active benefits in the year before learning spell	27.71	32.76	22.78	27.72	16.31	25.62	16.22	22.36	8.85	18.93	10.98	16.07
Average highest spell duration (in months)	5.88	3.41	6.27	4.26	9.91	5.95	9.75	5.40	20.93	7.43	19.49	9.85
Average number of learning spell before the highest learning spell	0.222	0.186	0.350	0.278	0.311	0.288	0.394	0.294	0.339	0.411	0.566	0.483

Responses to Reviewer Comments

In this section we set out our response to four sets of reviewer comments, grouped according to common themes identified across these comments (a complete set of reviewers' comments is included in the final Appendix).

i) It would be interesting to have more detail about the pre-learning situation. Do they have to be in employment before learning and so have a prior wage in order to calculate the diff-in-diffs? Doesn't this rule out a lot of young learners still in full-time education, who have progressed straight to FE from school? Does this mean the results are not representative for all learners?

The study of NVQ learners in Section 6 of the report focuses the matching diff-in-diffs analysis only on achievers with positive earnings in the two tax years before starting the course (to better ensure parallel trends). In the main body of the report our OLS equations that estimate earnings returns include learners without earnings prior to learning. We mention this as an issue in Section 6, as the application of more advanced matching diff-in-diffs estimation necessarily requires restriction of the sample for analysis.

In response to the reviewers' comment we have re-estimated the matching equations to also include individuals who did not work before commencement of learning. Matching on the whole population, instead of only selecting those who are employed before the learning spell, helps us partially offset this issue, whilst still retaining a form of analysis that allows investigation of factors that could be impacting our OLS estimates. By matching on the entire population of learners, we compare earnings returns for achievers and non-achievers, who are matched so that they are similar, whether or not they were employed or not prior to learning. This provides us with estimated returns that are more representative of the entire population and we have been able to check if these alter our main findings. The results of this analysis are available on request (and will feature in a forthcoming academic paper), but the upshot is that they do not differ substantially from the estimates obtained when matching only on individuals with a previous employment history.

A similar issue has been raised by a reviewer, whose concern is over the potential for differential selection from the population of all achievers and non-achievers; into achievers and non-achievers for whom we have earnings observations.

The matching equations we use in Section 6 are designed to overcome the problems of differential selection (on observables) into the achiever and non-achiever samples. However, as the first reviewer's comment implies, in the matching diff-in-diffs analysis we are actually matching achievers who previously had an earnings observation, with non-achievers who also had a previous earnings observation. If this group are different to the entire population of achievers and non-achievers, our analysis will be less representative. Similarly, each time we estimate an earnings return in, for instance, the third year after completion of learning, we have the potential for selection into the population of employed/earning achievers, and non-achievers. In each year after the end of learning we estimate a separate matching equation. Each year we have a different sample of

achievers and non-achievers for whom we observe earnings and these may not be representative of all achievers and non-achievers in each of these years.

If selection of achievers and non-achievers into our population of interest (i.e. those with an earnings observation in a particular year) is *random*, it would not be a cause for concern.

However, there is a possibility that selection is *non-random on observables* (that is, the achiever/employed population looks very different to the achiever/not employed population, and the same for non-achievers, in each year). The approach described above essentially checks this issue - in each year after learning we use the entire population of achievers and non-achievers in our matching equations. The matching equations for the first, third and fifth year of earnings are therefore based on a balancing of the entire relevant population (because we include individuals with zero earnings observations in the matching equation). As described above, the results suggest that (on observable characteristics) this issue of selection does not substantially alter our findings.

If selection into the employed/achiever and employed/non-achiever populations is systematically driven by unobservable characteristics, then we are less able to accommodate this. In this instance, we would hope that observable characteristics proxy for any unobservables driving selection.

ii) Why do you not use a falsification test pretending that the period of learning occurred earlier? The matching diff-in-diffs analysis would be further validated if you do not observe a positive difference in the wage changes of the two groups around this earlier time period.

The present format of the analytical[55] dataset does not allow us to fully investigate this issue, without changing the econometric specification of our matching diff-in-diffs analysis. At present we do not have observations on the number of days worked in a particular year, earlier than two years from the start of learning. In one of the data tables we do have earnings data for the full period of analysis, but not in a format that allows us to retrieve the daily earnings of an individual three and four years prior to learning (and a similar issue applies to our benefit variables).

In the forthcoming academic paper it will be important to fully implement the suggested falsification test using our preferred econometric specification. However, whilst the data do not presently allow us to estimate our preferred matching diff-in-diffs specification for an earlier period, we are still able to carry out a falsification test with a more limited specification. We estimate our matching diff-in-diffs specification for individuals aged between 25 and 59 who started attending their highest qualification after 5th April 2008 (with a continued focus on NVQ2 and NVQ3 qualifications).

Using as our dependent variable the log of deflated daily earnings (top and bottom 1% removed) in the financial year prior to the start of learning and matching on prior annual earnings trends (two and three periods before the beginning of the course), we find that returns from the "fake treatment" for men are positive and statistically significant both for

---

[55] The Analytical Dataset is an extract from the original ILR-WPLS Production Dataset and the aim is to update the extract to allow (i) inclusion of 2012/2013 earnings data and (ii) a falsification test.

NVQ2 (+2.55%) and NVQ3 (+1.49%). Such returns are positive also for Women but statistically significant only for NVQ2 (+2.31%). Although these "fake" returns suggest a difficulty in completely controlling for differences in characteristics between the treatment and the control group, they are small in size and obtained with a more limited econometric specification. More specifically, these small, statistically significant findings will likely disappear when we run our preferred specification that overcomes (i) the incompleteness of prior employment and benefit controls and (ii) uses daily, rather than yearly earnings, for the years before the start of learning.

iii) The HE flag might be seen as an outcome variable, while the learning spell variable is probably endogeneous to the outcome variable. Why do you control for such variables?

Considering the HE flag issue, we are very aware that our treatment of learning that leads to individuals attending HE institutions is imperfect. As the extensive discussion of our HE flag in the main body of the report suggests, it will be important in the next stages of any ILR-WPLS analysis to better capture these transitions as an outcome in themselves. What we have at present is an indicator (HE flag) of whether individuals born after 31/08/1980 experienced an HE learning spell between 2004/2005 and 2009/2010. This is flagged as an issue in our discussion of returns, particularly at Level 3 and Full Level 3, in the main body of the report.

Considering our treatment of learning spell length, the thrust of this reviewer's comment is analogous to the issue we face when including years of education as an explanatory variable in the classic Mincer Earnings equation. We may have a situation where individuals, who are more able in ways that we cannot observe, select into additional years of education (or, in the analogy here, they select into longer learning spells). This would make the learning spell length variable endogenous. The difference here is that we have both the achievers and non-achievers selecting into spells of different lengths for the same qualification level; and therefore we can vary the comparison of treatment and control groups according to spell length. Tables 54 to 63 in the accompanying Technical Annex present estimated results for different combinations of achievers and non-achievers with differing spell lengths.

As we highlight in the main body of the report, we are in the process of carrying out a number of variations to the standard achiever V non-achiever approach, based on learning spell length, as this is a way of investigating the validity of our identification strategy. More specifically, we would expect those with shorter learning spells to have a higher proportion of non-achievers who drop out because they are a poor match to the course. If this poor match is driven by the characteristics of the course (the quality of which varies across regions of England), rather than the individual, then they may be a better control for achievers. In contrast those with a longer duration, who then drop out, are likely to sacrifice their learning to date for very different reasons (and we can split this group into those who exit to employment, those who exit to unemployment and those who exit to an unknown destination).

iv) Is there any other data that can show what the average duration "should be" for each course? Some individuals seem to take a very long time and it would be good to know what the expected time for full and part time students to complete the course would be.

In the dataset there is no variable reporting the "expected duration" of a course. However, in response to this comment we have created a proxy for expected duration. First, considering only those who achieve their course, we create an expected duration variable that is conditional on the level of the course, the sector subject area and the type of attendance (part-time or full-time). We use the findings on achievers to impute the expected duration for non-achievers and we can then split the non-achievers into different groups according to the point in time they dropped out. In particular, we are now able to pin down "early dropouts" – which is obviously a relative term.

Tables 54 to 63 of the Technical Annex respond to both questions iii. and iv., as we have different parameter estimates constructed from different sub-populations of learners according to their expected duration of learning, and within these expected duration categories we compare returns of achievers V non-achievers with differing observed learning spell durations. For instance, Table 54 considers all men whose highest learning aim is an NVQ2 with an expected duration of less than 6 months. Within this sub-population we compare the returns of achievers to those of non-achievers who have (i) a learning spell of less than one months duration; (ii) < six months duration; (iii) > six months duration and (iv) all non-achievers.

Each one of the tables considers a different sub-population of NVQ2 and NVQ3 learners according to the expected duration of their learning spell, and then within this we vary the non-achieving control group according to their actual observed spell length. The results suggest that across almost all of our equations, those non-achievers who drop out from their course later, perform better, when compared to earlier dropouts. This is consistent with non-achievers increasing their human capital to some extent and gaining a reward for this, even if they do not obtain the same return as those who secure the 'signal' that achievement provides. Also, across the board we see OLS estimates that are close to those achieved using the matching diff-in-diffs approach, particularly when comparison is made between achievers and non-achievers dropping out earlier – though readers are advised to treat with some care any results based on less than 5,000 learners, as the controls group of non-achievers will start to drop below 1,000.

v) The proposed diff-in-diffs analysis should be conducted for a particular subpopulation within the broader framework (e.g. those starting FE at age 30+ who are likely to have an earnings observation prior to treatment).

We have estimated the matching diff-in-diffs analysis for a sub-population of NVQ learners who start FE at age 30+ and compare the estimated returns with those obtained from a standard OLS approach. As we would expect, the estimated returns do differ from those estimated for a wider sub-population, but the gaps between our matching diff-in-diffs estimates and those from OLS are very similar to those set out in Section 6 of the report.

vi) You would ideally need to control for a) total months of work experience b) part time working prior to the qualification, and c) whether they changed employer during the period. This is because if an employer is assisting them to take the qualification alongside work, the earnings trajectory might look quite different from someone taking it because they lost their job or because they were a student working part time throughout their studies. Do you have all this data?

One of the main shortcomings of the ILR-WPLS matched data is the lack of detailed information on specific jobs. We are not able to differentiate PT and FT workers, though we proxy for this in the forthcoming academic paper by carrying out estimation for a subset of learners who are recorded as being employed for most days of the year[56]. At present we are not able to identify total months of work experience with a specific employer or job changes between employers. This is perhaps something to consider for future developments with the data, but at present in the basic OLS specification we control for (i) the number of days an individual was on active benefits in the year before learning (ii) whether an individual has an inactive benefit spell in the year before learning, (iii) how many days in sustained (6 months) employment an individual has just before learning, (iv) number of previous FE learning spells, and (v) in the matching diff-in-diffs analysis we additionally match on earnings two years before learning. Similarly, whilst we are not able to identify the occupation or industry of specific jobs, we are able to control for the sector subject area of the qualification being taken. Ideally we would introduce further controls as suggested, but we would suggest that the controls listed go a long way to counter the concerns raised here.

vii) It would be very useful to compare the estimated returns obtained from various combinations of achievers and non-achievers – for instance, NVQ3 non-achievers V NVQ2 non-achievers, NVQ3 non-achievers and NVQ2 achievers. Similarly, it would be very helpful to have a specification that allows comparison between results in this study and those using survey data.

The first one of these suggestions has been included in Section 6 and the results are presented in Table 50. They are an important part of the evidence that supports the use of an achiever V non-achiever comparison. The second suggestion is one that we are pursuing, as we are able to recover from the data all those who only attended a Level 0 qualification. This group could be included as part of the control to move our population closer to that in survey-based studies. However, this is something for the future and has not proved possible within the current time frame for this project.

---

[56] Whilst it is obviously possible for individuals to work half days and to do this for a full year, the suggestion here is that we would leave out all those whose part-time working means that they work a lesser number of days per year. This would also leave out many other individuals, but if the results for this subset of learners do not differ enormously from the main body of findings we can be more confident that an inability to accurately identify part-time working in the data is not impacting estimates.

Appendix: reviewer comments

The following comments were obtained from reviewers at the start of the project, after distribution of a brief explanation of our proposed approach and some initial findings (one reviewer has provided verbal feedback, with issues raised by them included in the previous section). As a result a number of the requests are for clarification, and these have been kept in mind when writing up the Data and Method section in the accompanying report.

Reviewer 1

The inclusion of all individuals in your sample, with or without a HE flag, seems sensible to me.

You state that you keep only the highest learning aim associated with the latest spell and that this only matters where an individual has multiple spells with achievement/aims at the same level. So if someone has done 3 different NVQ2s sequentially then you will use only the last one achieved/aimed for? If someone did GCSEs at school, went on and did a NVQ2 and then finished off with a NVQ3, they would not appear in the sample since you only take those who achieve NVQ2 as their highest qualification? In the regressions you control for prior qualifications and the number of FE spells so to be precise you are estimating the return to NVQ2 for: a sample of individuals for whom the NVQ2 is their latest but not necessarily their highest qualification. This is quite a distinct group, it does not include those who go on to take higher NVQs for example, but equally it can include someone who has a degree. Is my understanding correct? *The discussion in the main report is now hopefully clearer on exactly what we are capturing when estimating returns.*

There may be age versus cohort effects as you suggest. You control for age. You control for year that the wage is observed I think (the year dummies). You sometimes control for the year the qualification was acquired by focusing only on one year after qualification but in the regressions which include earnings up to 7 years after completing the qualification, how do you control for time since acquired? Perhaps I missed it. *Hopefully the discussion now clarifies this issue. Both achievers and non-achievers are included in each regression because they have the same time since the end of learning.*

You control for duration of spell. Is there any other data that can show what the average duration "should be" for each course? Some individuals seem to take a very long time and it would be good to know what the expected time for full and part time students to complete the course would be. *See the previous section of main report.*

The achiever versus non achiever approach is, as you say, problematic. You state that it becomes less credible as the qualification level increases, since qualifications increasingly act as a filter of ability so that non achievers will be less able than achievers. I agree however that the reverse may be true. Non achievers that fail to secure a qualification when the qualification can be obtained simply by turning up may have a serious lack of motivation or lack of ability leading to bias. An alternative explanation is that they may have simply got a job and that is why they did not complete their qualification, then

perhaps the bias may go the other way. Either way the unobservables problem might be quite serious. *See the previous section of main report, discussion of unobservables.*

The difference in difference strategy would produce an estimate of the value of NVQs for a specific sub group i.e. slightly older workers with an employment history, clearly this is a select group. For example this group is likely to be different from younger workers who did not work before taking an NVQ2. The problem is that you might end up comparing a worker who has done a number of years of work then acquired an NVQ2 and carried on with their career, perhaps even with the same company, with a younger or less experienced worker who just did some part time work while studying before progressing on to their career. You would ideally need to control for a) total months of work experience b) part time working prior to the qualification, and c) whether they changed employer during the period. This is because if an employer is assisting them to take the qualification alongside work, the earnings trajectory might look quite different from someone taking it because they lost their job or because they were a student working part time throughout their studies. Do you have all this data? *See the previous section of main report for a detailed response.*

More generally the difference in difference does not really get around the criticism of the previous work. Basically by definition those who complete and those who do not must have experienced different shocks – positive or negative, a job offer or financial difficulties, unless you really believe that completing or not is random. So the two groups have (tautologically) time varying unobservables. Not to say it is not worth doing but you will continue to get the criticism that you are comparing failures with successful people (sorry about the language!). Perhaps an alternative might be to have treatment and control groups of individuals who complete their NVQ2 but comparing different individual qualifications (ie by area/subject)? That would give some insight with individuals who are arguably very similar. That said comparing non achievers (the control) with achievers and non achievers in the level below makes sense and I think is a good way to go. *This issue is tackled in a number of ways in the previous section of main report.*

In any case, prior trends on earnings do need to be similar if you use the difference in difference. This will restrict your sample still further as you need a few years of data to check the prior trends of the treatment and control groups. *See the previous section of main report.*

Matching methods also do not overcome the time varying unobservables problem. I realise you know this but I am just not sure your matching approach is really any better than what has been done before. I would emphasise the difference in difference approach. *We have carried out matching with diff-in-diffs.*

How much do we know about the level 2 qualifications in the data base. Is there any way of checking? Could it include GCSEs at all or individuals taking GCSEs alongside? *The tests we have carried out in Section 6 are for NVQ level 2 and Level 3 qualifications taken as a highest learning aim, but the main body of the report has analysis of categories that include GCSEs.*

Reviewer 2

This research topic is very interesting and policy relevant. I have had a look at the original (April 2013) report in preparing these comments too as I have only heard a presentation on one occasion (at the small meeting in BIS). The amount of work you have undertaken in bringing these data sets together and trying to estimate returns is very impressive.

I have a long list of comments and suggestions. Of course, I understand that you may not agree with some of the below but I hope it is of some use.

1. You are careful to identify the potential problems of this approach – i.e. primarily the difficulty of comparing achievers to non-achievers. Although the latter may be a good control group in the sense that at some point, they planned to take a similar path to the achievers, they may not be a good control group for other reasons. For example: (a) they may be less able (in dimensions of cognitive and non-cognitive skills) or less motivated; (b) they may receive a lower quality of FE education than achievers; (c) they may suffer a shock (such as a health problem) that interferes with their completion of their studies or their performance. It is not obvious that such a control group is any better that the group typically chosen by researchers using survey data (e.g. the group with no qualifications). In fact, the nature of selection in this sample may make things worse (depending on why the control group are non-achievers). In my view, the results from this research should be presented alongside those in the literature – and not written up as though this methodology were superior. *We are aware of the need to be realistic and measured in our conclusions. However, we would argue that the analysis presented here allows us to go a little further in our assertions.*

2. It would be very helpful to have a specification that allows comparison between results in this study and those in the literature (e.g. where qualifications are similarly defined and there are similar controls). I agree that it would be interesting to also try to find a control group which is similar (which you suggest). However, the limitations need to be clearly spelt out. As I understand it, this data set only consists of people who entered FE at some stage. *The focus of analysis in Section 6 is on NVQ2 and NVQ3 qualifications and this allows some comparison. However, as we suggest in the previous section, further work on creating a control group closer to that seen in survey-based analysis is something to consider for the future.*

3. It would be useful to have an appendix with the most common qualifications listed under main categories reported in Table 1 (i.e. what are the most common qualifications for people with 'below level 2', 'level 2' etc...). *A description was previously included in the 2013 report and has now been incorporated into the Technical Annex (page 9).*

4. Why the distinction between 'level 2' and 'full level 2'? Is this really a distinction that is meaningful in practice? *This is very much a distinction driven by the data, funding and policy considerations (see page 9 of Technical Annex for a brief explanation).*

5. A great advantage of this data is its huge size and the fact that linked data is available for a number of years. I think it would be useful to break down the analysis (and data) according to some sensible and policy relevant criteria. For example, you could analyse

the data in groups according to age brackets in which they started their time in FE (such as 16-18; 19- 21; 21-30; 30+). This would have the advantage of making the subsamples a little more homogeneous. *We have split the data and carried out separate analyses for various sub-groups, both in the main report and in response to reviewers' comments.*

6. I think there is a need to describe the treatment and control groups more fully. For each treatment and control group, we should see summary statistics, which describes their characteristics. We should always know the sample size (in every table) and the percentage of the sample that is classified as being part of the treatment or control group. *We now include summary statistics for each of the main treatment and control groups in the appendix to the report and we provide some indication of sample sizes around a number of tables. However, the main body of the report now includes over 50 Tables and runs to over 100 pages, so additional information is only added where absolutely necessary. We flag where low numbers make our findings less reliable.*

7. It would be useful to see how the treatment and control group compare based on some prior achievement characteristic. For example, suppose you look at a particular age cohort of individuals. Then estimate regressions where KS4 points score is the outcome (but controlling for KS3 and/or KS2 scores). If there is no significant difference between the two groups with regard to their KS4 points score, this would give more credibility to 'post treatment' results (i.e. FE outcome). *We do have prior qualification information, but not to this level of detail.*

8. It would also be useful to give most emphasis to earnings differences several years after entering FE. This is for two reasons: (a) because it allows the 'non-achievers' to catch up – particularly important if the reason they did not achieve is because of a temporary shock; (b) it allows the earnings trajectory to widen a bit before estimating returns (which we expect to happen). *We now estimate earnings returns five years on from the end of learning.*

9. Whether a person goes on to HE could be seen as an outcome variable. I don't think it makes much sense to control for it. The possibility that people might go on to HE just illustrates another reason for why it is important to estimate earnings differences several years after people enter HE. *See discussion of HE flag in previous section.*

10. Also, I'm not sure it makes sense to control for learning spell. This is another variable that is probably endogenous to the outcome variable. I would approach the problem by dividing the population into categories (e.g. by age at start of FE and gender) and then estimating outcome equations X number of years later for the sub-population. *See the previous section and the cohort analysis in Section 6.2.*

11. Another big selection issue arises because not everyone gets a job after leaving FE. I think investigation of employment probabilities needs to be undertaken before getting on to earnings differentials. We should know how treatment and control groups compare based on the probability of getting a job X years after entering FE etc (with and without controls). As is well known, estimating earnings equations while giving no attention to this problem can be very misleading. So it is important that the issue is discussed fully – including on the direction of expected bias in the earnings equations. Another reason for analysing the

sector in a more disaggregate way (e.g. by age category and gender) is that this source of selection is probably more important in some categories than in others. *See the discussion of selection issues in the previous section and also the following table of employment probabilities.*

Table 54: OLS and Coarsened Exact Matching estimates of Employment probabilities, comparing achievers and non-achievers

	Year 1			Year 3			Year 5		
	(1)	(2)	(9)	(1)	(2)	(9)	(1)	(2)	(9)
	OLS (basic)	OLS (fully specified)	CEM	OLS (basic)	OLS (fully specified)	CEM	OLS (basic)	OLS (fully specified)	CEM
NVQ2 Men	4.51	1.12	1.15	1.81	0.43	0.56	0.97	0.55	0.73
N	1121806	1121806	1121806	697137	697137	697137	285021	285021	285021
r2_a	0.002	0.530	0.538	0.001	0.423	0.432	0.001	0.336	0.346
NVQ2 Women	10.63	2.80	2.66	8.53	2.66	2.49	6.57	2.39	2.34
N	940523	940523	940523	627529	627529	627529	314940	314940	314940
r2_a	0.011	0.462	0.466	0.009	0.382	0.388	0.006	0.333	0.345
NVQ3 Men	5.35	1.88	2.21	3.55	1.29	1.51	2.10	0.75	0.88
N	336647	336647	336647	201600	201600	201600	100324	100324	100324
r2_a	0.003	0.317	0.312	0.002	0.265	0.259	0.001	0.247	0.237
NVQ3 Women	5.73	1.77	1.87	5.31	1.83	1.89	5.92	2.20	2.25
N	573890	573890	573890	366233	366233	366233	200722	200722	200722
r2_a	0.003	0.398	0.397	0.004	0.336	0.334	0.005	0.317	0.315

12. I agree that it is particularly interesting to compare the returns to non-achievers at level 3 with the returns to non-achievers at level 2; returns to non-achievers at level 3 with achievers at level 2 etc... (In addition, it may be interesting to pool the data such that the same control group is used as the baseline for all levels of achievement). *See the previous section and Table 50.*

13. I think the proposed diff-in-diff analysis needs to be conducted for a particular sub-population within the broader framework (e.g. those starting FE at age 30+ ...who are likely to have an earnings observation prior to treatment). You would need to describe carefully what this sub-population looks like and say something about the % dropped because they do not have a pre-treatment earnings observation. I think this approach is worth a try but may be difficult in practice. My concern is that the sample may be too heterogenous for the approach (e.g. stopping work at all different times in their career for different reasons; and then waiting for varying lengths of time before they start FE). *See the previous section and results of CEM and diff-in-diff analysis.*

14. If using the whole sample together, I think it is important to control for year dummies and birth cohort dummies in all specifications. *See list of control variables.*

15. Although you can use matching approaches, I think it is important to also show results from OLS regressions with controls (showing sensitivity of results to including different sets of controls). In my view, matching probably adds very little to this. *See the findings presented in Section 6 and also the discussion in previous section.*

Reviewer 3 (who additionally saw a copy of the academic paper prior to comment)

I've had a read of your report. I think you've done a good job in terms of constructing the dataset, and then using it for careful analysis.

You have clearly taken great care when writing the report not oversell the results - you continually remind the reader of the risk of using drop-outs/failures as the control group. The key question is how well your methods control for the obvious differences between those who complete a course and those who don't.

I would say that you have done about as well as you could do - matching (I must admit I am unfamiliar with the CEM technique that you used), matching on both prior wages and the change in prior wages, and then finally diff-in-diffs to control for unobserved characteristics between the two groups, are all good things to do, and I can't think what else you could have done to control for selection effects on those who complete. Two sections that particularly strengthened the results, I thought, were (i) the analysis comparing the non-achieving at level 3 (2) versus the achieving at level 2 (1) analysis (page 17) which found a small positive effect, suggesting that the non-achievers are not a group with low labour market returns for unobserved reasons, and (ii) the discussion of the reasons for dropping out (page 18), where failure to understand/cope with the course is not mentioned, with the reasons more likely to be things beyond the individual's control such as a bad match, poor teaching etc.

I can think of at least a couple of reasons why your results could even under-estimate the effects:

(i) the reason for drop-out/non-completion could be important. If individuals are not completing because they have been offered a good job, so that the earlier leavers are those being offered the best jobs, then the estimated gap between the wages of the two groups would underestimate the return to the qualification.

(ii) you are essentially assuming that the achievers-non-achievers wage gap measures the differences in wages between having the qualification and not. But the non-achievers will have spent some time on the course (maybe only slightly less than those who complete in some cases). In a strict human capital theory world, as opposed to say, signalling, then the non-completers will have increased their human capital to at least some extent, and so will have a higher wage than those who never took the qualification. *The results in Tables 54 to 63 support this assertion.*

The extensions you propose for future work look sensible, particularly the time spent on the course by the non-achievers. A couple of other things that could be considered:

(i) a 'falsification' test - pretend that the period of learning occurred earlier and see whether you still observe a positive difference-in-differences in the wage changes of the two groups

around this earlier time period. Hopefully, the difference won't be significantly positive in this period when a qualification wasn't actually taken. *See previous section.*

(ii) I would be interested in more detail about the pre-learning situation. Do they have to be in employment before learning and so have a prior wage in order to calculate the diff-in-diffs?   Doesn't this rule out a lot of young learners still in full-time education, who have progressed straight to FE from school? Does this mean the results are not representative for all learners? *See previous section.*

(iii) are you going to do anything with the subject of qualification. You say in footnote 8 that you control for subject, and suggest that results are available in a detailed appendix. *The report now provides detailed analysis of returns across sectors (using OLS) and the following table reports equivalent CEM diff-in-diffs estimates for one male dominated sector (Construction) and one female-dominated (Child Development).*

Table 55: Coarsened Exact Matching and Difference-in-differences for two sectors

	Year 1			Year 3			Year 5		
	(1)	(2)	(9)	(1)	(2)	(9)	(1)	(2)	(9)
			CEM-DiD			CEM-DiD			CEM-DiD
	OLS (basic)	OLS (fully specified)	PT (earnings)	OLS (basic)	OLS (fully specified)	PT (earnings)	OLS (basic)	OLS (fully specified)	PT (earnings)
Construction Men (NVQ2)	22.78	17.22	10.18	20.63	16.66	11.36	20.72	14.71	13.57
N	58323	58323	58323	26417	26417	26417	2578	2578	2578
r2_a	0.011	0.073	0.047	0.009	0.069	0.057	0.012	0.092	0.125
Construction Men (NVQ3)	17.04	18.18	13.33	23.84	24.50	17.89	35.52	33.87	30.92
N	11642	11642	11642	3727	3727	3727	384	384	384
r2_a	0.006	0.090	0.207	0.014	0.101	0.234	0.061	0.242	0.264
Child Dev. Women (NVQ2)	23.12	18.67	16.43	17.57	13.80	12.74	12.50	9.04	13.23
N	65782	65782	65782	29003	29003	29003	5529	5529	5529
r2_a	0.016	0.054	0.024	0.010	0.052	0.036	0.006	0.062	0.061
Child Dev. Women (NVQ3)	18.99	17.49	16.54	15.41	14.16	11.89	18.77	13.72	11.14
N	78537	78537	78537	28520	28520	28520	5135	5135	5135
r2_a	0.011	0.056	0.039	0.009	0.054	0.046	0.017	0.080	0.060

BIS Evaluation Summary and Peer Review

The BIS Expert Peer Review Group for Evaluation reviews all BIS impact evaluation publications, and provides an independent assessment of the methodological quality of the evaluation.

Click here to read the peer review for this report.