National Institute of Economic and Social Research

A COMPARISON OF EARNINGS RELATED TO HIGHER LEVEL VOCATIONAL/TECHNICAL AND ACADEMIC EDUCATION

Héctor Espinoza[1]
Stefan Speckesser[1]

[1]National Institute of Economic and Social Research

CENTRE FOR VOCATIONAL EDUCATION RESEARCH

NIESR Discussion Paper No.502

Date: 02 April 2019

About the National Institute of Economic and Social Research

The National Institute of Economic and Social Research is Britain's longest established independent research institute, founded in 1938. The vision of our founders was to carry out research to improve understanding of the economic and social forces that affect people's lives, and the ways in which policy can bring about change. Over eighty years later, this remains central to NIESR's ethos. We continue to apply our expertise in both quantitative and qualitative methods and our understanding of economic and social issues to current debates and to influence policy. The Institute is independent of all party political interests.

National Institute of Economic and Social Research
2 Dean Trench St
London SW1P 3HE
T: +44 (0)20 7222 7665
E: enquiries@niesr.ac.uk
niesr.ac.uk
Registered charity no. 306083

This paper was first published in April 2019

A COMPARISON OF EARNINGS RELATED TO HIGHER LEVEL VOCATIONAL/TECHNICAL AND ACADEMIC EDUCATION

Héctor Espinoza* and Stefan Speckesser*

Discussion Paper

March 2019

Abstract:

Using rich administrative data for a full cohort of English secondary school leavers (2002/03 academic year), we compare earnings of people with higher vocational/technical qualifications to the earnings of degree holders at the age of 30, while controlling for prior attainment and background characteristics.

We find that by the age of 30 the early earnings differential associated with high-level vocational/technical education tends to disappear and degree holders earn more on average. However, there is strong heterogeneity by gender and subject area. There are especially high returns related to higher vocational/technical education in STEM subjects, which remain significantly above those of many degree holders several years after graduation.

Keywords: Returns to Education, Tertiary Education, High-Level Technical Education, Vocational Education, Administrative Data.

JEL codes: I21, I24, I26, J64

Acknowledgements: We would like to thank Sandra McNally, Jeffrey Smith, and Bernd Fitzenberger for comments and suggestions. We would also like to thank the Department for Education (DfE) for the data and financial support under the CVER work programme.

Corresponding author: Héctor Espinoza, h.espinoza@niesr.ac.uk

Affiliations:

* National Institute of Economic and Social Research and Centre for Vocational Education Research (CVER)

Executive summary

Background

Education investment is one of the major sources driving productivity and innovation in the economy. Successful higher education is particularly important as it also represents one of society's key mechanisms to create social mobility and prosperity, especially for those from deprived families. It does this because of the significant earnings returns associated with high-level qualifications (e.g. see Belfield et al. (2018), and because of many other benefits to individuals (such as non-cognitive skills) and to the wider community.

Most of the micro-econometric research on the earnings effects and social mobility created by higher education focuses on honours degrees, i.e. Bachelor of Arts (BA) and Bachelor of Science (BSc) degrees, which represent the vast majority of higher education. However, there are also programmes of higher vocational and technical education, which are e.g. important for people with university-entry qualifications ("Level 3") obtained in vocational programmes.

To date, there are no econometric studies of earnings benefits focused specifically on such "higher vocational and technical programmes" of tertiary education and how they compare to e.g. earnings of degree holders. This paper provides estimates for these programmes, using the new Longitudinal Education Outcomes (LEO) data, which link earnings from administrative data at census-level to individual records from England's central education register covering education from primary schools all the way up to university.

Aim

We take a deep look into the value of qualifications of tertiary education in England, exploring differences in earnings profiles between higher vocational/technical and academic programmes. Higher vocational programmes last either for one year in tertiary education ("Level 4" qualifications) like Higher National Certificates (HNCs) and Level 4 National Vocational Qualifications (NVQs), or two years ("Level 5"), which include Higher National Diplomas (HNDs), Foundation Degrees, NVQs Level 5 and other qualifications. In contrast, academic programmes are predominantly bachelor's degrees and a few other three years programmes like graduate certificates, graduate diplomas, etc. ("Level 6").

We analyse rich administrative data for a cohort of English secondary school leavers, who are old enough for us to be able to estimate some medium-term differences in labour market outcomes when choosing between high-level vocational/technical or academic education. We focus on the cohort finishing compulsory education (age 16) in the summer of 2003 and look into their earnings by age 30, at time when most of them will have been in

the labour market for several years. However, it may well be that differential returns of these qualifications change as the cohort ages.

Empirical research design

We estimate empirical earnings functions in the tradition of the Mincer (1974) human capital model using Ordinary Least Squares (OLS) and Inverse Probability Weighting Regression Adjustment (IPWRA), combined with the Least Absolute Shrinkage and Selection Operator (LASSO) that refines the most exhaustive specification. The rich set of covariates chosen by LASSO includes gender, work experience, ethnicity, Free School Meal (FSM) eligibility, region, Index of Multiple Deprivation at the Lower Layer Super Output Area level, GCSEs results, broad subject area, and school type.

Findings

In an initial descriptive analysis, we explore education progression and describe the highest level of education of individuals over time. This shows that higher level vocational/technical qualifications (Level 4-5) are the highest education achievement of very few people in the cohort (around 2%). While tertiary education attainment increases over time, Level 4-5 vocational qualifications tend to be acquired relatively late compared to degrees, which are mostly achieved by age 22/23 (i.e. until 2009 for this cohort). We also observe that students with Level 4-5 vocational qualifications have very diverse education backgrounds, ranging from Entry level to Level 3. This is very different to students aiming for degrees, who almost exclusively take A-Levels. Male and female students also make very different subject choices. Looking into their earnings over time, we find comparatively similar earnings trajectories for Level 4-5 students, which look very different from the earnings of those with Level 6 academic qualifications.

In the econometric analysis, we estimate whether earnings of achievers of Level 4-5 vocational/technical qualifications in 2017 differ from those having acquired degrees. We find that earnings for male degree holders are similar to higher vocational/technical education if they studied in non-Russell group universities, and higher for those from Russell group universities. Earnings for female degree holders are found to be higher regardless of the university type compared to those who achieved higher vocational/technical education. Within these overall findings, there is strong heterogeneity of effects by subject area and also by gender. When looking into results by subject, we find that earnings of males with Level 4-5 Science, Technology, Engineering and Mathematics ("STEM") qualifications are comparable or higher than earnings of STEM degree holders. Results for male students with qualifications in construction are similar, showing high returns for both academic and vocational qualifications. These empirical findings remain valid after controlling for prior attainment, and estimating human capital models with more sophisticated econometric techniques such as the IPWRA combined with LASSO method.

These findings are produced by carefully specified econometric models and estimated using rich observational data. However, estimates can only be interpreted as reflecting the causal effect of qualifications on earnings if all relevant characteristics influencing both are captured within the model. This will not be true if omitted variables (such as non-cognitive skills) are important for influencing both earnings and choice of qualification path – and if they are not adequately captured by included controls (such as prior attainment). This is an important caveat to bear in mind when interpreting the results of this paper and all research using similar methodologies. The size of the data set has no bearing on this issue. Notwithstanding this important caveat, it is informative to compare the earnings differential to higher level vocational versus academic education using the biggest data set available to consider this issue in England.

Further research

In the next stage of this work, we will use several more recent cohorts of secondary school leavers in order to provide a comprehensive set of estimates of earnings differentials associated with the full range of higher education options (vocational compared to academic) compared to counterfactual Level 3 attainment. The use of multiple cohorts, which – given the small number of individuals studying for Level 4 and 5 qualifications – will be important in boosting the sample and improving the precision of our estimates. This will be particularly useful for a range of sub-group analyses. Also, we will be considering the implications of differential drop-out between individuals who pursue different routes.[1]

[1] This research is in the process of development with colleagues at the Institute for Fiscal Studies and the University of Cambridge.

1. Introduction

Education investment is one of the major sources driving productivity and innovation in the economy[2]. Good education attainment at each step of the learning journey from primary to secondary school and beyond results in long-term economic benefits compared to low achievement. Successful higher education is even more important as it represents one of society's key mechanisms to create social mobility and prosperity, especially for those coming from deprived families. It does this because of the significant earnings returns associated with high-level qualifications, see e.g. the recent IFS study (Belfield et al., 2018), and because of many other benefits such as non-cognitive skills and benefits to the wider community. As in most countries, higher education in England offers a wide range of academic and vocational/technical qualifications. In the English education system, these qualifications are classified as between "Level 4", i.e. a successful first year of tertiary education and "Level 8", i.e. the completion of a doctorate.

Research on the earnings effects and further social benefits of higher education has mainly focused on honours degrees, i.e. Bachelor of Arts (BA) and Bachelor of Science (BSc) degrees, which represent the vast majority of higher education outcomes. However, there are also programmes of higher vocational/technical education, which are of particular importance for people with university-entry qualifications ("Level 3") obtained in vocational programmes, for example "Business and Technology Education Council" qualifications (BTECs). To date, there are no econometric studies of earnings differentials focused specifically on the comparison between "higher vocational/technical programmes" of tertiary education and "academic programmes" in England.

In this paper, we summarise findings from original empirical research comparing earnings differentials of vocational and academic qualifications gained in tertiary education. We benefit from the new Longitudinal Education Outcomes (LEO) data, which link individual earnings from the tax register at census-level to all individual education records from England's central register covering education from primary schools all the way up to university. This allows us to understand the type of tertiary education followed, but also allows us to control for many important characteristics observed in individual learning trajectories, which can explain why people made particular choices to follow vocational/technical or academic programmes.

Our analysis focuses on a cohort of school leavers, whose education and labour market trajectories we can follow until adulthood (those finishing their compulsory education at age 16 in the summer of 2003). The qualifications of interest in this study at Level 4 are Certificates of Higher Education, Higher National Certificates (HNCs), National

[2] See e.g. LSE Growth Commission (2013) "Investing for Prosperity"

Vocational Qualifications (NVQs), and various professional diplomas and certificates. Level 5 corresponds to a successful second year at university and attainment with a Diploma of Higher Education or other outcomes like Higher National Diplomas (HNDs), Foundation Degrees and NVQs. Level 6 qualifications, i.e. successful completion of three years full-time higher education, are primarily BA/BSc degrees, but there are various Certificates, Diplomas and NVQs also at this level.

Throughout this paper, we distinguish between academic and vocational/technical qualifications. Academic qualifications offer comprehensive subject knowledge and generic skills at different levels, ranging from the General Certificate of Secondary Education (GCSEs) via A-Levels to degrees. In contrast, vocational/technical qualifications emphasise technical and procedural knowledge and skills with focus on practical application, combined with abilities, and competences relevant to executing occupational roles, for example National Vocational Qualifications (NVQ) or Higher National Certificates (HNCs) and Higher National Diplomas (HNDs).

The key question this research paper seeks to answer is: What are the medium-term differences in earnings and employment trajectories between groups with vocational/technical and academic achievement in tertiary education? Similar to Brunello and Rocco (2017), we estimate returns to vocational and general education, but focusing on tertiary education only. Implementing a quantitative research design, we estimate empirical earnings functions in the tradition of the Mincer (1974) human capital model using Ordinary Least Squares (OLS) and Inverse Probability Weighting Regression Adjustment (IPWRA), combined with the Least Absolute Shrinkage and Selection Operator (LASSO) that refines the most exhaustive specification. By implementing LASSO, we aim to make best use of our rich universe-level data set, which includes attainment in primary and secondary education, measures of disadvantage and local areas, and many other important drivers of education participation and success.

Our findings show similar earnings trajectories of Level 4 and 5 achievers, which differ greatly from those with Level 6 academic qualifications. While data descriptions show an early earnings advantage for students with vocational qualifications, the average earnings of people achieving degrees are the same or higher by age 30. However, we also observe strong heterogeneity by subject area and gender. When looking into results by subjects, we find that earnings from Level 4-5 Science, Technology, Engineering and Mathematics ("STEM") subjects are comparable or higher for men than earnings of STEM degree holders. This finding remains valid when controlling for measures of prior educational attainment, in particular GCSEs, using a variety of specifications in OLS and IPWRA models.

The rest of the paper is organised as follows: Section two provides a brief summary of the literature relevant to this study. In Section three we describe the different sources of

administrative data, their information on education participation and labour market outcomes. We also discuss the empirical approach in this section. In section four, we present findings of the empirical analysis contextualising Level 4-5 vocational and Level 6 academic qualifications. Section five offers some conclusions.

2. A brief literature review

The question about the returns to vocational compared to general/academic education has been explored with both cross-sectional and longitudinal (individual life-course) data. Ryan (2001), focusing on cross-sectional studies, concludes that vocational programmes have a positive effect, particularly apprenticeships, as they increase success in early working life. Similar, Hanushek, et al. (2017) using a life-cycle approach, also found that vocational education has an early advantage compared to general education. For the UK, using the two rich cohort studies of children born in 1958 and 1970, Brunello and Rocco (2017) found – though only in the most recent cohort – that having a vocational education can lead to short-term real wage advantages, but that these advantages may evolve over time and become long-term disadvantages with respect to having an academic education (Brunello and Rocco, 2017)[3]. Other studies based on data for the United Kingdom have found lower returns to vocational education in the long-run. For instance, Dearden, McIntosh, Myck and Vignoles (2002) find that academic education leads to higher returns, but also document that the majority of vocational education programmes increase earnings relative to non-vocational qualifications, especially for low achieving school leavers.

In contrast to our paper, previous studies in the UK focus mainly on lower levels of vocational education. One could argue that the key finding of short-term benefits from vocational education, in particular from apprenticeships and for those less academically inclined, and related longer-term disadvantages, might not apply to those making choices in favour of higher level vocational education, who hold high levels of pre-exiting qualifications (usually Level 3) and could equally opt for higher academic education. However, previous research has dealt more generally with the issue that vocational skills depreciate relatively more quickly (Goldstein and Stenberg, 2017), which might well affect higher level technical skills in the same way. Moreover, some authors emphasise that general education enhances the ability to learn new skills, so that people can more easily cope with long-term changes in labour demand, for instance caused by fundamental technological changes (e.g. Krueger and Kumar (2004), or Autor, Dorn and Hansen (2015). Therefore, there might still likely be long-term benefits from general academic programmes compared to more focused high level vocational/technical education. There are very few studies exploiting exogenous policy changes to address self-selection (into different tracks). In general, they find no statistically

[3] Brunello and Rocco (2017) also did not find employment trade-offs affecting these cohorts.

significant differences in labour market outcomes associated with type of education (e.g. Oosterbeek and Webbink, 2007; Pischke and von Wachter, 2008; Hall, 2012).

We contribute to this literature by focusing specifically on high level vocational versus academic education with the use of richer administrative data that covers the universe of English school leavers for a particular cohort.

3. Data and empirical approach

Data

Our empirical analysis is based on one full cohort of English pupils completing their compulsory education at the end of the 2002/03 academic year. At this point, everyone has completed the last of the four compulsory "Key Stages" of education, Key Stage 4 (KS4), which coincides with the completion of the General Certificates of Secondary Education (GCSEs) taken in a variety of subjects including core subjects such as English, mathematics, science, a modern foreign language and a humanity. Data for these students, including their previous attainment at earlier Key Stages, are collected in the National Pupil Data (NPD), which can be linked at the individual level to further education registers, such as the Individualised Learner Records data (ILR) covering post-16 education outside schools and universities, the Higher Education Students Data (HESA) provided by universities and earnings from Her Majesty's Revenue and Customs (HMRC).

First, we make use of the National Pupil Data (NPD). The NPD is a census of all students attending English schools (in the state system), from which we select the full cohort of students completing secondary school at age 16 in 2002/03. These data were linked to further National Pupil Data covering the time period when they were sixteen to eighteen years old ("Key Stage 5", KS5), which is when most people acquire the entry qualifications to higher education ("Level 3"). Students can do this either in an academic track (A-Levels) or a vocational track, which can prepare for both employment and tertiary education. NPD data are used both to understand routes into higher education and to obtain important characteristics on previous learning trajectories and achievement during secondary education.

Second, we use linked records from the Individualised Learner Record (ILR) register, which collects information on each of the aims the student is enrolled in post-16 education outside of schools and universities, i.e. mainly further education colleges. These data capture further participation and achievement in education, including low level vocational or general education, but also tertiary education taken outside of universities.

Third, we link the data to the register of students in universities. These data are supplied by the Higher Education Statistics Agency (HESA) and contain full records of higher

education participation and outcomes. They are our primary source of data on progression to higher education and related achievement.

Finally, we use linked information about annual earnings from income tax records provided by Her Majesty's Revenue and Customs (HMRC). These data include both earnings from employment and self-employment. Linked to education register data at individual level, these "Longitudinal Education Outcomes" data (LEO) have only become available recently and provide detailed data on all taxable earnings for people living in the UK from the age of sixteen onwards. We use these data to create measures of annual earnings by aggregating all individual-level earnings records within a tax year (running April to March).

With this unique dataset, we track all education and labour market activity for all 622,000 English pupils in the fourteen years after leaving secondary education, i.e. until individuals are aged 29/30, when most individuals are well established in the labour market. Thus, this analysis exposes how well individuals have progressed within the labour market after opting for different educational paths, e.g. vocational or academic, and how their observed earnings evolve in subsequent years.

Empirical approach

As the main purpose of this study is to estimate medium-term earnings effects associated with different types of tertiary qualifications, we make use of different specifications of empirical earnings functions estimated in linear regression models, which aim to account for important differences amongst individuals driving the selection into academic or vocational/technical trajectories. We apply Ordinary Least Squares (OLS) models, which we augment using Least Absolute Shrinkage and Selection Operator models (LASSO) and Inverse Probability Weighting Regression Adjustment (IPRA). We briefly describe this approach here.

Linear Regression model using OLS

The key explanatory variable in the empirical model is the choice of individuals to follow academic (Level 6) or higher vocational/technical tertiary (Levels 4-5) education. We estimate the effect of this variable on annual earnings, controlling for a variety of covariates. Our approach follows the standard empirical human capital model, a version of the Mincer equation (Mincer 1974), which explores the relationship between investment in human capital, work experience and the income distribution.

The key equation is defined as follows:

$$\text{Ln}(\textit{Income}) = \beta_0 + \beta_1 \textit{Education} + \beta_2 \textit{Exp}_{2004} + \beta_2 \textit{Exp}_{2005\ldots} + \beta_2 \textit{Exp}_{2016+} X + u$$

[Equation 1]

where 'Ln (*Income*)' is a continuous variable that represents the natural logarithm of the annual income in 2017, '*Education*' is a categorical variable that takes value 1 if the

highest level of education is Level 6 Academic[4] or 0 for Level 4-5 vocational/technical education. *Exp* variables control for individual work experience made since the end of secondary education included by a set of yearly dummy variables[5] while X represents a set of covariates. The X include gender (1 if male), ethnicity (1 if white, 0 otherwise), eligibility for Free School Meals as an indicator of family disadvantage (1 if eligible, 0 otherwise), GCSEs results (number of "good" GCSEs graded at A*-C), region and the Index of Multiple Deprivation (IMD) related to small areas where pupils are resident in the final year of secondary education. Further control variables include subject areas and the type of school attended. '*u*' represents the error term of the empirical model.

Even when controlling for a number of characteristics, 'Education' representing the principal choice between academic and vocational/technical tertiary education is likely to suffer from endogeneity due to omitted variables bias. Unobserved personal attributes, such as ability or motivation, might be correlated with both the particular educational choice and the labour market outcomes observed (i.e. the earnings). As a consequence, the statistical association between education choice and labour market outcomes could be 'contaminated' by the effect of the unobserved attributes and differences in earnings patterns may be the result of self-selection between academic and vocational education.

The fact that we control for GCSEs results helps to attenuate this problem. This can be thought of as a measure of ability at a time earlier in the education biography, before the decision to select into particular programmes of tertiary education.[6] In addition, we improve how to account for observable characteristics by implementing an Inverse Probability Weighting Regression Adjustment (IPWRA) combined with the selection of the variables based on LASSO, which makes the best use of the available information from the linked administrative data.

Least Absolute Shrinkage and Selection Operator (LASSO)

The LASSO (Tibshirani, 1996) is a machine learning technique recently incorporated into econometrics to assist with model selection where many independent variables (or predictors) are available. It helps to find a parsimonious specification, e.g. for a linear regression model, in high dimensional settings as they exist in rich administrative data offering many potential predictors[7]. We use the LASSO method to remove variables that are redundant. This means that we do not add unnecessary information to the Mincer model[8], thereby reducing the risk of "overfitting" the data.

[4] Whenever possible we further distinguish according to Level 6 academic group, i.e. Russell and Non-Russell.
[5] Flexible specification also explored by Heckman et al., 2008. In practice, the variable work experience takes value 1 in any given year if we observe positive income in that particular year.
[6] We also tried UCAS tariff points and the results did not change. We opted for GCSE scores because there is a consistent measure for all the individuals in the study, and they do not reduce the Level 4/5 sample (as some Level 4/5 achievers do not have Level 3 qualifications).
[7] Considering categorical variables, we have more than fifty explanatory variables in our version of the Mincer equation.
[8] Also, it will support the first step in the IPWRA estimation.

The LASSO method estimates the regression parameters minimising the Residual Sum of Squares (RSS) with a penalty term to achieve sparse solutions. During this regularisation process some coefficients will be penalised (set to zero), and those that still have a non-zero coefficient after the shrinking process are selected to be part of the model. Formally, the LASSO estimate is defined as follows:

$$\hat{\beta}^{\text{LASSO}} = \arg\min_{\beta \in \mathbb{R}_p} || y - X\beta ||_2^2 + \lambda || \beta ||_1$$

[Equation 2]

The first term corresponds to the Residual Sum of Squares (RSS), λ is the tuning parameter that controls the strength of the penalty, and $|| \beta ||_1$ is the L_1 norm[9].

Essentially, in the empirical section, we apply the standard LASSO method mainly as a first step of model selection, i.e. supporting and improving our high-dimensional OLS and IPWRA estimations. In the OLS case, for instance, we take the following steps (following Belloni et al., 2014):

Step 1: Use the LASSO technique to estimate the income equation excluding the independent variable *Education* (academic versus vocational):

$$\text{Ln}(Income) = \beta_0 + \beta_1 Exp_{2004} + \beta_2 Exp_{2005\ldots} + \beta_3 Exp_{2016} + X + u$$

[Equation 3]

After estimating the first model by equation 3, we denote the set of LASSO-selected variables by (a).

Step 2: Use the LASSO to estimate the education equation, where *Education* (Academic vs Vocational) is now the dependent variable, maintaining the same set of covariates:

$$Education = \beta_0 + \beta_1 Exp_{2004} + \beta_2 Exp_{2005\ldots} + \beta_3 Exp_{2016} + X + u$$

[Equation 4]

After estimating the second model by equation 4, we denote the set of LASSO-selected variables by (b).

[9] L_1 Norm is the sum of the magnitudes of the vectors in a space.

Step 3: Estimate an augmented model via OLS:

$$\text{Ln}\,(Income) = \beta_0 + \beta_1\, Education + W + u$$

[Equation 5]

where 'W' is simply the union of selected controls (a + b) from steps 1 and 2.

Inverse Probability Weighting[10]: Average effect of "treatment-on-the-treated"

A second approach to improve the estimates involves using an Inverse Probability Weighting Regression Adjustment (IPWRA) combined with the LASSO method, where the different qualification choices are modelled as "multiple treatments", i.e. selection into particular routes, depending on variables used in the first and second stage of the IPWRA, which are selected via LASSO.

The procedure is as follows: First, after implementing the LASSO method, we estimate a multinomial logit model to obtain the probability of each individual studying for a particular highest qualification, depending upon all the relevant observable characteristics resulting from LASSO. In the second step, individuals are then weighted with the inverse of this probability to create weighted samples of treatment groups, which are similar in terms of the observable characteristics included in the logit model.

Formally, the two stages are defined as follows:

1. First stage (Multinomial Logit): Estimate the probability that an individual studies at a particular qualification level (l), i.e. Level 4-5 vocational or Level 6 academic or, within Level 6, to do this at a Russell group university or elsewhere) as a function of observable characteristics X:

$$Pr\,(\text{Level} = l_j) = \frac{exp(X'\beta_j^l)}{\sum_k exp(X\prime\beta_k^l)} \qquad \forall\, k \neq j,\ k \in l$$

[Equation 6]

2. Second Stage: Derive the IPWRA estimator by adding the IPW obtained from the model to the linear regression to obtain an average "treatment effect on the treated" (ATET):

$$ATET = E\left[Y^L - Ylev^{voc} \middle| D = L\right] \qquad \text{[Equation 7]}$$

In our study, the *ATET* represents multiple options of different treatments: That is, *L* has two different outcomes representing treatment, namely achieving Level 6 at Russell Group or Level 6 outside of Russell Group universities, while lev^{voc} represents the treatment

[10] Following Wooldridge (2007).

control outcome (i.e. Level 4/5 vocational). $D = L$ restricts the expected value to include only those individuals who actually receive the specific treatment L[11].

4. Empirical findings

Descriptive statistics

As discussed, we analyse the full cohort (622,000 students) leaving secondary education in 2002/03 and follow them through all the education registers. Similar to Brunello and Rocco (2017), we classify individuals to vocational or academic education using their highest qualification acquired during the period 2003-2015[12].

Table 1 shows the highest level of education achieved by this group up until 2015. First, it is worth noting the high shares of people with very low or very high attainment: 22.8% of the cohort only reach Level 1 or below (e.g. Entry Level) by their end twenties. In contrast, 31.8% of this cohort acquires Level 6 or above academic qualifications. Second, a considerable number have acquired both academic and vocational qualifications during their education (see for instance Level 3 qualifications). And third, for approximately 1.5%, Level 4-5 vocational qualifications are the highest level attained. While this is a small percentage, it still represents a sizeable number of people (around 9,000 observations). Importantly, this description shows that education in England emphasises general and academic education, although some of these academic qualifications may have an important vocational content (see Makepeace and Dolton, 2001).

[11] For two reasons we did not bring more groups (e.g. Level 3) into the analysis. First, restricting the sample to Level 4 plus effectively focuses this study in the upper-end spectrum of qualifications, where we can obtain a direct comparison between high-level technical and academic qualifications (that is the main research question). Second, and perhaps a more important reason, adding more groups to the analysis made in several cases the IPWRA computation unfeasible.

[12] Since we focus on education attainment, dropouts from a given qualification are classified according to the highest level of qualification actually achieved. In future work we will analyse differential drop-out from academic versus vocational qualifications.

Table 1: Highest level of education achieved by 2015

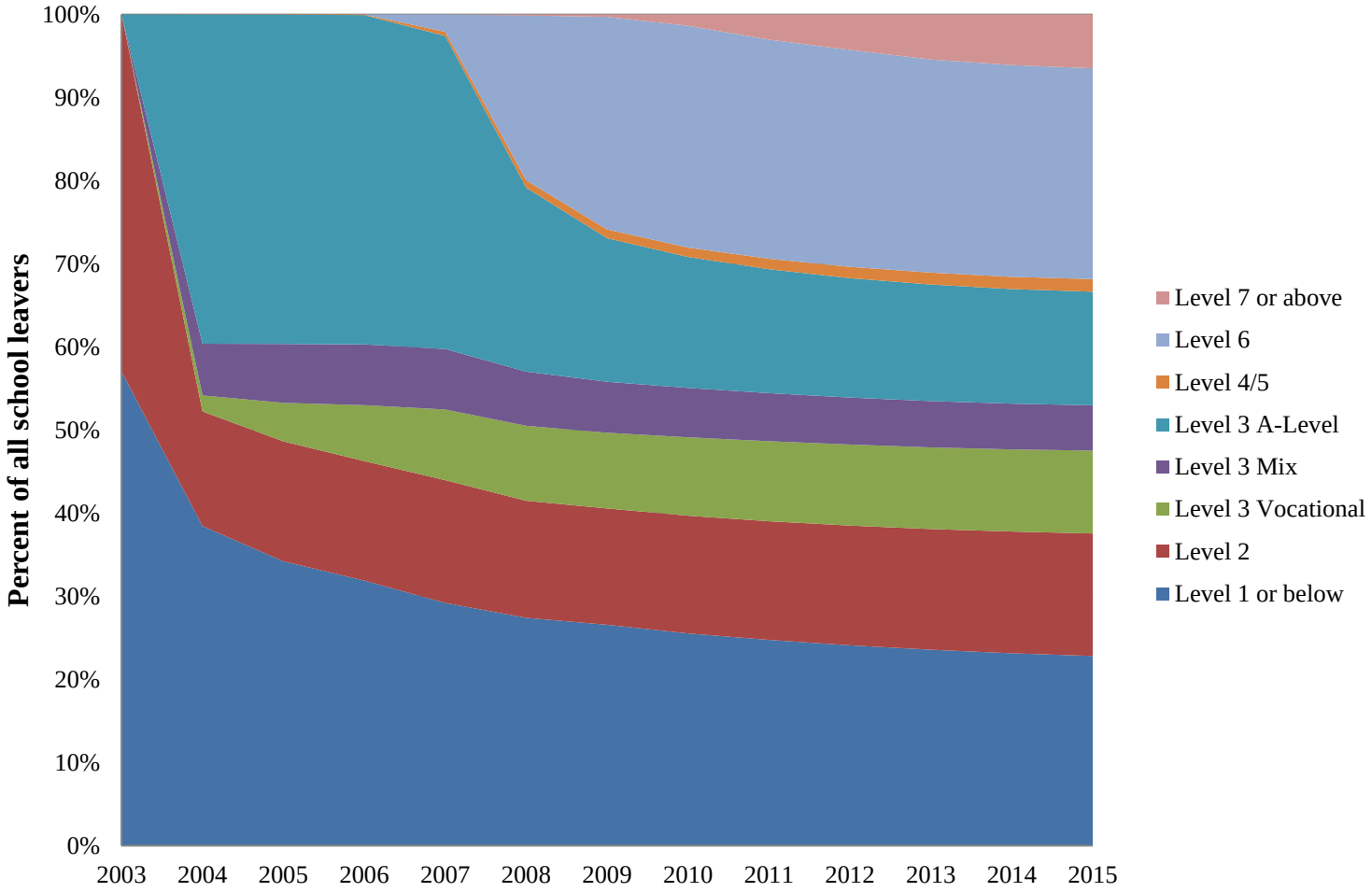
Highest Level of Education[13]	Frequency	Percent
Level 1 or below	141,900	22.82
Level 2	79,100	12.72
Level 2 + Apprenticeship	12,500	2.01
Level 3 Vocational	41,700	6.7
Level 3 Mix	33,900	5.45
Level 3 A-Level	84,900	13.66
Level 3 + Apprenticeship	20,400	3.27
Level 4 Vocational	4,700	0.76
Level 4 Academic	100	0.01
Level 5 Vocational	4,700	0.76
Level 5 Academic	<100	0.01
Level 6 Vocational	<100	0
Level 6 Academic	157,700	25.35
Level 7 or above	40,300	6.48
Total	622,000	100

Source: NPD-ILR-HESA linked data. Totals have been rounded to the nearest hundred

Figure 1 shows the highest level of education held by the cohort in any individual year. Most importantly, we observe education progression over time, but the improvement in education outcomes differs depending on the route chosen. The majority of students with successful academic education at Level 6 achieved this in 2008/09 (that can be considered the "standard" academic route). In contrast, Level 4-5 vocational and technical qualifications increase continuously during this period. Of those who end up with these qualifications, a substantial number achieve them after 2008. Similar early progression patterns have been found studying a more recent cohort (2010/11). We present the main findings of this research in Appendix 2.

[13] Level 1 and below with less than 5 GCSEs A*-C; Level 2 with 5 or more GCSE A*-C (including English and Maths); Level 2 (any) + Apprenticeship; Level 3 A-Level; Level 3 vocational; Level 3 mix (including both A-level and non A-level courses); Level 3 (any) + Apprenticeship; Level 4 vocational (e.g. HNCs, BTEC level 4, and NVQ Level 4); Level 4 academic (1st year Undergraduate and CertHE); Level 4 (any) + Apprenticeship; Level 5 vocational (e.g. HNDs, Foundation Degrees, BTEC Level 4, and NVQ Level 4 that usually last 2 years); Level 5 academic (2nd Year Undergraduate and DipHE); Level 5 (any) + Apprenticeship (empty category according to this hierarchical classification); Level 6 vocational (such as RFQ Level 6 Diploma); Level 6 academic (e.g. BSc, BA, LLB, etc.); and Level 7 or above stands for MSc/MRes and PhD, for example.

Figure 1: Highest level of education acquired per year during period 2003-2015



Source: NPD-ILR-HESA linked data. The Totals have been rounded to the nearest hundred.

Table 2: Achievement age

	Highest level of education			
Age	L4V	L5V	L6A	Total
16-17	2.0	0.3	0.0	0.1
17-18	1.5	0.1	0.0	0.1
18-19	2.4	0.4	0.0	0.1
19-20	16.9	15.4	6.4	6.9
20-21	16.6	23.1	54.5	52.5
21-22	14.1	6.7	19.1	18.6
22-23	12.3	6.3	7.5	7.6
23-24	9.4	10.9	4.5	4.8
24-25	7.1	11.4	3.2	3.5
25-26	7.0	9.4	2.2	2.5
26-27	5.2	8.3	1.7	2.0
27-28	5.6	7.7	0.9	1.3
Total	100	100	100	100
Total	4,700	4,700	157,700	167,100

Source: NPD-ILR-HESA linked data. Total rounded to the nearest hundred.

Column headings: L4V: Level 4 Vocational, L5V: Level 5 Vocational, L6A: Level 6 Academic

Next, we describe the age at which the highest level of education was achieved, where we investigate the differences between Level 4 and 5 vocational/technical and Level 6 academic qualifications. We see that the majority of 18-year-olds, who remain in the education system, pursue university degrees, with 80% achieved by the age of 22[14].

In contrast, very few students pursue Level 4-5 programmes. Students showing such qualifications as their highest attainment look different from "standard" academic trajectories combining A-Levels with swift university entry and achievement of the qualification in the early 20s (Table 2): While most degrees are achieved by age 22, successful completion of Level 5 vocational qualifications is much more spread out between the ages of 19 and 26. For those having Level 4 vocational as their highest qualification, the age at which this is achieved is also more spread out between the age of 19 and 24, with some achieving this between age 25 and 28).

[14] Based on the expected length of the programme.

Table 3: Highest level of education, by previous level of education held

Previous Level	L2	L2+App	L3V	L3Mix	L3ALev	L3+App	L4V	L4A	L5V	L5A	L6V	L6A	L7+	Total
Level 1 or below	100.0	88.2	77.3	52.2	40.0	11.6	13.4	50.6	2.4	25.0	0.0	1.7	1.0	34.7
Level 2		11.8	22.7	20.1	58.1	20.4	17.9	7.6	4.4	12.5	0.0	5.2	1.0	17.6
Level 2 + Apprenticeship			0.0	1.1	0.0	29.0	1.0	0.0	0.6	0.0	5.0	0.1	0.1	1.4
Level 3 Vocational				26.7	0.0	8.1	15.8	6.3	16.7	3.1	10.0	5.4	0.5	4.6
Level 3 Mix					1.9	6.9	17.7	5.1	17.4	9.4	20.0	6.6	0.7	3.4
Level 3 A-Level						24.0	25.3	26.6	47.6	15.6	45.0	77.9	8.1	29.3
Level 3 + Apprenticeship							8.8	0.0	4.9	0.0	5.0	1.1	0.2	0.5
Level 4 Vocational								3.8	6.0	6.3	0.0	0.6	0.1	0.3
Level 4 Academic									0.0	0.0	0.0	0.0	0.0	0.0
Level 4 + Apprenticeship									0.0	0.0	0.0	0.0	0.0	0.0
Level 5 Vocational										28.1	15.0	1.5	0.3	0.5
Level 5 Academic												0.0	0.0	0.0
Level 6 Academic													88.0	7.7
Total (%)	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Total (Frequency)	58,400	12,500	41,700	33,900	84,900	204,000	4,700	<100	4,700	<100	<100	157,700	40,300	459,400

Source: NPD-ILR-HESA linked data. The totals have been rounded to the nearest hundred.

Qualifications (acronyms) in the column headings: L2: Level 2, L2+App: Level 2 combined with Apprenticeship, L3V: Level 3 Vocational, L3 Mix: Level 3 A-Level and vocational combined, L3ALev: Level 3 A-Level, L3+App: Level 3 combined with Apprenticeship, L4V: Level 4 Vocational, L4A: Level 4 Academic, L5V: Level 5 Vocational. L5A: Level 5 Academic, L6V: Level 6 Vocational, L6A: Level 6 Academic, and L7+: Level 7 or above

In Table 3, we show how people's highest observed education during the period 2003-2015 relates to the previous highest level of education held.

The column headings relate to the highest qualification a person achieved and the rows to the qualification they held previously (irrespective of the number of years between achieving the highest and previous qualification). As expected, the majority of young people who achieve a university degree (L6A) previously achieved A-levels (about 77%), with most others achieved Level 3 vocational qualifications such as BTECs (about 5.4%) or some combination (6.6%). The previous attainment of those achieving Level 5 (vocational – L5V) as their highest qualification is again mainly A-levels (47.6%), with a significant minority with Level 3 vocational qualifications (17%) or a mix (17%). About 8% of those achieving Level 5 (vocational – L5V) as their highest qualification have only previously achieved Level 1 or Level 2 qualifications which in most cases were achieved more than 5 years before. Interestingly, there is only modest transition from Level 4 to Level 5 vocational qualifications. In contrast, people achieving Level 4 vocational/technical qualification (L4V) as their highest attainment come from a very diverse background with pre-existing qualifications ranging from Level 1 to Level 3.

The description of the last highest education attainment ahead of obtaining higher level vocational/technical or academic qualifications shows a very diverse educational background, suggesting that unlike Level 6 academic qualifications, higher level vocational may offer more flexible entry routes. These could include crediting work experience or some other access qualification.

Table 4: Demographic characteristics and achievement in compulsory education

		Highest Level of Education Achieved in 2015			
		L4V	L5V	L6 Non-Russell	L6 Russell
Gender	Female	44.6	48.4	54.2	51.8
	Male	55.3	51.5	45.7	48.1
Ethnicity	Other	15.3	19.9	33.7	47.3
	White British	84.6	80.0	66.2	52.6
FSM Eligibility	No	93.0	91.0	91.8	96.2
	Yes	7.0	8.9	8.1	3.7
GCSE A*-C	0	5.2	7.2	2.3	0.5
	1	4.1	6.5	2.0	0.6
	2	4.4	5.8	2.0	0.2
	3	5.1	6.0	2.3	0.2
	5	11.8	14.9	6.9	0.8
	5+	69.2	59.3	84.2	97.5
KS3 English	Above expected	5.8	4.4	13.3	31.0
	Unobserved	6.0	6.5	11.7	31.3
KS3 Math	Above expected	23.5	13.6	27.7	49.9
	Unobserved	6.0	6.2	11.3	30.2
KS2 English	Above expected	17.0	12.4	26.1	47.9
	Unobserved	5.0	7.95	9.9	18.5
KS2 Math	Above expected	22.7	13.8	24.4	50.6
	Unobserved	5.1	8.8	10.1	18.5
Totals		4,700	4,700	117,000	38,000

Source: NPD-ILR-HESA linked data. Totals have been rounded to the nearest hundred.

Table 4 shows background characteristics and achievement in compulsory education for the different groups of students in tertiary education. First, the statistics by gender show that women have a lower probability of acquiring Level 4/5 vocational/technical qualifications, in contrast to academic programmes, where they represent significantly above 50% of the achievers (54.2% in non-Russell Group universities, 51.8% elsewhere). Second, within Level 4-5 qualifications, students classified as White British represent the majority, similar to the composition of the total cohort of KS4 leavers (above 80%), while we observe a much higher proportion of students of other ethnic backgrounds within Level 6 academic qualifications (especially amongst Russell group universities)[15]. Third, those with Levels 4 or 5 as their highest qualification are comparable to those with degrees in terms of the probability of being eligible for free school meals when at school – below 10% in each

[15] Since all students looked at in this analysis had their GCSEs in England in 2003, about four out of five being white British, the under-representation of these students in the best universities results from the attainment gap during secondary education, see e.g. https://www.ethnicity-facts-figures.service.gov.uk/education-skills-and-training/11-to-16-years-old/gcse-results-attainment-8-for-children-aged-14-to-16-key-stage-4/latest

case. Fourth, amongst those with Level 4/5 (vocational) a lower proportion achieved 5 or more “good” GCSEs at Grades A*-C compared to Level 6 academic. Furthermore, those achieving degrees (Russell and non-Russell group) have obtained better results at Key Stage 3 and Key Stage 2, compared to those with Level 4/5 higher technical qualifications. All these results suggest that there are important differences between groups that we need to consider in the empirical section, controlling for key variables in the models and exploring subsequent heterogeneity of effects.

Figure 2: Earnings Trajectories*, by type of qualification during the period 2004-2017

** in £ at 2015 price levels (CPI adjustment), academic qualifications by Russell/Non-Russell, see below*
Source: NPD-ILR-HESA-HMRC linked data

Figure 2 shows the unconditional/descriptive association between average earnings and qualification type. It depicts an early advantage (in monetary terms) of vocational education compared to academic education. The earnings path can be understood essentially as a trade-off between work experience and education investments. In this scenario, it is worth noting that only after a few years, the lines depicting earnings trajectories, intersect. The average earnings growth for those who attended Russell group institutions is particularly striking. For men, average earnings for vocational and academic (non-Russell group) graduates converge by the age of 30 whereas for women, earnings of the former group increase at a faster rate. The gender differences largely reflect different subject choices made by men and women, which we discuss below.[16]

Econometric Analysis

Main findings

This section presents the OLS and IPWRA results. Both approaches are improved by the LASSO method of model selection, which discards redundant variables in a first step[17]. We present the LASSO estimates for comparison in Tables 5 to 7. All these models take the natural log of earnings in year 2017 as the dependent variable, and making use of our rich dataset, we separate the Level 6 academic group into programmes of study at the research-intensive "Russell Group" universities and other universities ("Non-Russell" group). Hence, in the following specifications, 'Education' is represented by two regressors that take value 1 if the individual follows an academic path (e.g. Russell or non-Russell) or 0 if the highest level of education acquired is Level 4/5 vocational (i.e. HNCs, HNDs, and Foundation Degrees are the baseline group in the model).

Results are shown in Tables 5 and 6 for men and women respectively. The coefficients are stable across different specifications and methods, and the magnitudes

[16] Earnings levels observed here likely differ by gender primarily because of different subject choices made by male and female students of Level 4-5 and Level 6 qualifications; see Table A.5 and discussion below. We would not expect much variation because of different labour market attachment of men and women by age 30 as labour market participation is generally high for both genders at this age (82-86% of all Level 6 achievers in employment). For those with Level 4-5 technical/vocational qualifications, employment rates remain consistently high for men from their mid-twenties, but decrease slightly for women with Level 4-5 qualifications from 87% by age 25/26 to 83% when women are around 29/30, likely to reflect reduced employment due to having children (see Table A.3). However, like Belfield et al. (2018), we are limited in providing a full picture of labour market activity as there are no further variables for e.g. part-time work in the data, which might affect specifically women at this point of the career, and because female Level 4-5 achievers might have children earlier, could create some of the gap observed between female groups by the end-twenties.

[17] For instance, variables 'urban' (vs rural) and 'school fixed effects' are removed from the model in the context that we also control for region, school type and Index of Multiple Deprivation at the Lower Level Super Output Area. Also, 'prior earnings' (i.e. earnings obtained before achieving the highest qualification) is removed with an observed coefficient not different from zero. It is worth noting that KS2 results (English and Maths) were not removed by the LASSO, however these were not included in the model to avoid losing around 10,000 observations and to make IPWRA estimation feasible. Removing the variable 'KS2 scores' did not change the results and conclusions in this study. The full set of variables included in these models is available in the appendix.

change only when we control for previous attainment (GCSEs results) in Columns 6 and 7, which suggests underlying heterogeneous effects that we further explore below.

The coefficients in Column 8 ("average treatment effect-on-the-treated") are strongly significant in our preferred method, the IPWRA estimation and show the average gain from treatment (e.g. academic compared to vocational education) for those who were actually treated. That is, we estimate that the average incomes for male degree holders from Non-Russell group universities are not significantly different to earnings they would have achieved if they had acquired a higher level vocational qualification.

In contrast, for women, the earnings differential from attending a Non-Russell group university is always high compared to holding a higher level vocational qualification. For both men and women, the estimated earnings differential from achieving a qualification from a Russell group university is extremely high.

Table 5: Returns to high-level vocational/technical and academic education, males

	Ordinary Least Squares (OLS)						LASSO	IPWRA
(Baseline: Vocational L4/5)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Academic (Non-Russell)	-0.00	-0.00	-0.02*	-0.03**	0.00	-0.05***	-0.04	0.02
	(0.012)	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)	-	(0.021)
Academic (Russell)	0.33***	0.32	0.29***	0.28***	0.29***	0.18***	0.20	0.18***
	(0.014)	(0.014)	(0.014)	(0.015)	(0.014)	(0.015)	-	(0.027)
Work experience	✓	✓	✓	✓	✓	✓	✓	✓
Personal Characteristics		✓	✓	✓	✓	✓	✓	✓
Geographical Characteristics			✓	✓	✓	✓	✓	✓
KS4 School Type				✓	✓	✓	✓	✓
Broad Subject Area					✓	✓	✓	✓
Previous attainment (GCSEs)						✓	✓	✓
N	55,828	55,828	55,828	55,828	55,828	55,828	55,828	55,828
Adj. R^2	0.147	0.149	0.157	0.157	0.182	0.193		

*Source: NPD-ILR-HESA-HMRC linked data. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001*

The Russell Group (Russell) refers to an association of 24 universities in the United Kingdom with a focus on world-leading research and a reputation for academic achievement. The full list of current members is presented in Table A.4.

Table 6: Returns to high-level vocational/technical and academic education, females

	Ordinary Least Squares (OLS)						LASSO	IPWRA
(Baseline: Vocational L4/5)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Academic (Non-Russell)	0.17***	0.17***	0.14***	0.14***	0.19***	0.13***	0.14	0.16***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	-	(0.027)
Academic (Russell)	0.53***	0.52***	0.47***	0.47***	0.51***	0.39***	0.42	0.41***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(0.016)	-	(0.031)
Work experience	✓	✓	✓	✓	✓	✓	✓	✓
Personal Characteristics		✓	✓	✓	✓	✓	✓	✓
Geographical Characteristics			✓	✓	✓	✓	✓	✓
KS4 School Type			✓	✓	✓	✓	✓	✓
Broad Subject Area				✓	✓	✓	✓	✓
Previous attainment (GCSEs)					✓	✓	✓	✓
N	65,647	65,647	65,647	65,647	65,647	65,647	65,647	65,647
Adj. R^2	0.122	0.124	0.132	0.133	0.146	0.158		

*Source: NPD-ILR-HESA-HMRC linked data. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001*

The Russell Group (Russell) refers to an association of 24 universities in the United Kingdom with a focus on world-leading research and a reputation for academic achievement. The full list of current members is presented in Table A.4.

Heterogeneity: Subjects and mature students

Table 7 shows further estimates[18] using the most exhaustive specification to show differences across subjects. For the group of male students, we observe that returns to Level 4/5 STEM qualifications are higher than returns to degrees taken outside the Russell group universities. That is, we estimate that the average income for those who obtained a degree from a non-Russell institution is 11 per cent lower by the age of 30 than it would have been had those same individuals not graduated from the programme (i.e. achieving higher-technical qualifications instead) when controlling for differences in observable characteristics. As many male students study STEM programmes, compared to female students, whose subject choices tend to favour business, education and care (see Table A.5 in the appendix), the subject choice results in the overall insignificant differential between Level 4-5 vocational and Level 6 (non-Russell group) qualifications. Returns to Level 4-5 vocational/technical education in construction are also very similar returns to Level 6 academic programmes, but relatively few individuals achieve these qualifications. In all other subject areas, academic qualifications (whether achieved in Russell group universities or not) are associated with higher returns by age 30.

[18] IPWRA is our preferred method.

Table 7: Returns to high-level vocational/technical and academic education, by subjects

	IPWRA							
	STEM	**Construction**	**ALH**[19]	**S. Sciences**	**Business**	**Education**	**Health**	**Other**
(Baseline: Vocational L4/5)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Full sample								
Academic (Non-Russell)	-0.05	0.05	0.17***	0.24***	0.07***	0.32***	0.29***	-0.02
	(0.037)	(0.035)	(0.040)	(0.062)	(0.021)	(0.041)	(0.048)	(0.043)
Academic (Russell)	0.14***	0.00	0.34***	0.47***	0.36***	0.51***	0.53***	0.22***
	(0.040)	(0.048)	(0.048)	(0.066)	(0.033)	(0.043)	(0.069)	(0.055)
Males								
Academic (Non-Russell)	-0.11***	0.06	0.11**	-0.22**	0.04	0.26***	-	-0.14
	(0.029)	(0.036)	(0.043)	(0.065)	(0.029)	(0.078)		(0.079)
Academic (Russell)	0.04	-0.02	0.24***	0.016	0.31***	0.45***	-	-0.01
	(0.035)	(0.054)	(0.055)	(0.069)	(0.047)	(0.084)		(0.114)
Females								
Academic (Non-Russell)	0.06	-	0.20***	0.36***	0.09***	0.35***	0.30***	0.04
	(0.090)		(0.057)	(0.053)	(0.028)	(0.042)	(0.054)	(0.045)
Academic (Russell)	0.31***	-	0.38***	0.58***	0.42***	0.55***	0.51***	0.24***
	(0.092)		(0.067)	(0.060)	(0.046)	(0.045)	(0.072)	(0.072)
N Total	33,503	3,275	32,305	12,176	20,977	10,638	4,920	3,681
N Males	21,077	2,704	12,387	4,894	9,850	2,493	-	1,686
N Females	12,426	-	19,918	7,282	11,127	8,145	4,183	1,995

*Source: NPD-ILR-HESA-HMRC linked data. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001*

All models in Table 7 are using the most exhaustive specification.

The Russell Group (Russell) refers to an association of 24 universities in the United Kingdom with a focus on world-leading research and a reputation for academic achievement. The full list of current members is presented in Table A.4.

In Column headings: STEM: Science, Technology, Engineering and Mathematics, Construction, ALH: Art, Languages and Humanities, S.Sciences: Social Sciences, Business, Education, Health, and Other.

As shown in the descriptive analysis, students of higher vocational/technical education tend to be older than people going into academic programmes, who mostly come directly from KS5. Therefore it's worth considering a comparison between those students who achieve these qualifications in their mid-twenties. We focus on the subgroup of those who achieved qualifications after 2008. By providing estimates for this subgroup of students, we hope to estimate the earnings differentials for a more comparable group.

[19] Arts, Languages and Humanities.

The vast majority (more than 90%) of those taking academic qualifications later in life study in Non-Russell group universities. Therefore, in the following we do not distinguish between types of university. We present the results in Table 8.

Table 8: Returns to high-level vocational/technical and academic education, mature students

	IPWRA							
	STEM	**Construction**	**ALH**	**S. Sciences**	**Business**	**Education**	**Health**	**Other**
(Baseline: Vocational L4/5)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Full sample								
Academic	-0.14***	0.08	0.19**	0.33***	0.00	0.36***	0.25***	0.14
	(0.039)	(0.044)	(0.065)	(0.078)	(0.036)	(0.041)	(0.053)	(0.052)
Males								
Academic	-0.16***	0.10*	0.11	-	-0.01	-	0.25**	-0.06
	(0.040)	(0.047)	(0.070)		(0.050)		(0.074)	(0.093)
Females								
Academic	-0.03	-	0.25**	0.28**	0.03	0.40***	0.24***	0.13*
	(0.114)		(0.081)	(0.097)	(0.044)	(0.044)	(0.060)	(0.058)
N Total	6,954	1,365	6,597	2,809	4,094	3,174	1,871	1,636
N Males	5,068	1,213	3,076	945	1,897	-	365	790
N Females	1,886	-	3,521	1,864	2,197	2,591	1,506	846

*Source: NPD-ILR-HESA-HMRC linked data. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.*

In Column headings: STEM: Science, Technology, Engineering and Mathematics, Construction, ALH: Art, Languages and Humanities, S.Sciences: Social Sciences, Business, Education, Health, and Other. The regression analysis uses the most exhaustive specification and here also controls for prior earnings.

The IPWRA estimates for mature students confirm that high-level vocational/technical qualifications in STEM subjects have higher returns compared to Level 6 academic amongst males, but not amongst females. Similarly, in the full sample of mature students, we fail to find significant differences in Business, which suggests that higher vocational/technical qualifications might be a valuable alternative to academic qualifications within this group, at least based on earnings observed by the age of 30. These results are robust to the inclusion of prior earnings as an additional control. As before, results have been obtained when controlling for work experience and other observable characteristics, but again sources of unobserved heterogeneity such as employer support and available options for career progression affecting the selection into Level 4-5 vocational programmes cannot be controlled for in our data. Furthermore, Level 6 academic students might have

other motivations to engage in academic study, which are not directly related to career progression. Thus, we need to be cautious about interpreting the results. They only reflect the causal effect of undertaking academic versus vocational education under the assumption that all relevant factors influencing qualification choice as well as earnings are captured by the control variables.

A final point related to the wide subject differences in estimated effects, including those for mature students, is that our data do not offer any information about occupational roles undertaken, differences in the long-term career progression, etc., which may differ between people with tertiary level vocational or academic qualifications. Clearly, further research on labour utilisation will be required to understand the match of graduates to graduate jobs and/or how much people with Level 4-5 vocational qualifications provide equivalent skills and/or how their longer term earnings and careers differ. However, the recent increase in Apprenticeships, where Level 4-5 starters represent three quarters of those beginning tertiary education funded via the Apprenticeship Levy[20], suggests that people holding Level 4-5 vocational qualifications are likely to fulfil a great variety of high skilled job roles in the economy.

5. Conclusions

Previous empirical research in England on the earnings return to higher education has been conducted primarily by estimating effects for degree holders, i.e. people achieving Bachelor of Arts (BA) and Bachelor of Science (BSc) degrees, which represent the vast majority of higher education. However, there are also "higher vocational and technical programmes" of tertiary education, both in universities and further education colleges.

This paper provides estimates for these programmes (relative to university degrees), using the new Longitudinal Education Outcomes (LEO) data, which link earnings from administrative data at census-level to individual records from England's central education register covering education from primary schools all the way up to university. We use LEO for students completing their compulsory education in 2002/3 in order to track their education biography and earnings until the age of 30.

In the descriptive analysis, we look into education progression and describe the highest level of education of individuals over time. This description shows that higher level vocational/technical qualifications (Level 4-5) are the highest education achievement of very few people in the cohort (around 2%). While tertiary education attainment increases over time, Level 4-5 vocational qualifications tend to be acquired relatively late compared to degrees, which are mostly achieved up to the age of 22/23 (i.e. until the year 2009). We also observe that students with Level 4-5 vocational/technical qualifications have a more diverse

[20] http://researchbriefings.files.parliament.uk/documents/SN06113/SN06113.pdf

background in terms of qualifications, ranging from Entry levels to Level 3 A-level, very different to students aiming for degrees. In addition, male and female students make very different subject choices. Looking into their earnings over time, we find comparatively similar earnings trajectories of Level 4-5 students, which look very different from with Level 6 academic qualifications.

In the econometric analysis, we estimate whether earnings of achievers of Level 4-5 vocational/technical qualifications in 2017 differ from those having acquired degrees. We find that earnings for male degree holders are similar to higher vocational/technical education if they studied in non-Russell group universities, and higher for those from Russell group universities. Earnings for female degree holders are found to be higher regardless of the university type compared to those who achieved higher vocational/technical education.

Within these overall findings, there is strong heterogeneity of effects by subject area and also by gender. When looking into results by subject, we find that earnings of males with Level 4-5 Science, Technology, Engineering and Mathematics ("STEM") qualifications are comparable or higher than earnings of STEM degree holders. Results for male students with qualifications in construction are similar, showing high returns for both academic and vocational qualifications. These empirical findings remain valid after controlling for prior attainment, and estimating human capital models with more sophisticated econometric techniques such as the IPWRA combined with LASSO method.

Although we need to be cautious about interpreting the results in this paper as reflecting the causal impact of qualifications, the estimated earnings differentials between university degrees and higher-level vocational education provide useful information for young people and for public policy.

In the next stage of this work, we will use several more recent cohorts of secondary school leavers in order to provide a comprehensive set of estimates of earnings differentials associated with the full range of higher education options (vocational compared to academic) compared to counterfactual Level 3 attainment. The use of multiple cohorts, which – given the small number of individuals studying for Level 4 and 5 qualifications – will be important in boosting the sample and improving the precision of our estimates. This will be particularly useful for a range of sub-group analyses. We will also consider the potential implications of differential drop-out of students pursuing different routes.

6. References

Aghion, P., Besley, T., Browne, J., Caselli, F., Lambert, R., Lomax, R., Pissarides, C., Stern, N. and Van Reenen, J. (2013). Investing for Prosperity: Skills, Infrastructure and Innovation. LSE Growth Commission Report.

Autor, D., Dorn, D., and Hanson, G. (2015). Untangling Trade and Technology: Evidence from Local Labour Markets. The Economic Journal, 2015, 125 (May), 621–646.

Belfield, C., Britton, J., Buscha, F., Dearden, L., Dickson, M., van der Erve, L., Sibieta, L., Vignoles, A., Walker, I. and Zhu, Y. (2018). The impact of undergraduate degrees on early-career earnings. Department for Education (Ref: RR808) and Institute for Fiscal Studies.

Belloni, A., Chernozhukov, V. and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. Review of Economic Studies, 81: 608–650.

Brunello, G., and Rocco, L. (2017). The labour market effects of academic and vocational education over the life cycle: Evidence from two British cohorts. Journal of Human Capital

Card, D. (1999). The Causal Effect of Education on Earnings," in Handbook of Labor Economics, Volume 3A, ed, by Orley Ashenfelter and David Card. Amsterdam and New York: North Holland.

Carneiro, P. and Heckman, J. (2002). The Evidence on Credit Constraints in Post-Secondary Schooling. Economic Journal, 112(482):705–734.

Crawford, C. Macmillan, L. and A. Vignoles (2014), High-attaining children from disadvantaged backgrounds, Social Mobility and Child Poverty Commission, IFS Working Paper

Dearden, L., McIntosh, S., Myck, M. and Vignoles, A. (2002). The Returns to Academic and Vocational Qualifications in Britain, Bulletin of Economic Research 54(3): 249-74.

Golsteyn, B. and Stenberg, A. (2017). Earnings Over the Life Course: General vs Vocational Education. IZA – Institute of Labour Economics. Discussion Paper

Hall, C. (2012). The Effects of Reducing Tracking in Upper Secondary School: Evidence from a Large-Scale Pilot Scheme. Journal of Human Resources 47(1), 237-269.

Hanushek, E., Schwerdt, G., Wößmann, L. and Zhang, L. (2017). General Education, Vocational Education, and Labour market Outcomes over the Life-Cycle. Journal of Human Resources.

Heckman, J., Lochner, L. and Todd, P. (2008). Earnings Functions and Rates of Return, Journal of Human Capital, University of Chicago Press, vol. 2(1), pages 1-31.

Hedges, S. and Speckesser, S. (2017), Peer Effects and Social Influence in Post-16 Educational Choice, forthcoming in the *CVER Discussion Paper*

Hupkau, C., McNally, S., Ruiz-Valenzuela, J. and Ventura, G. (2016). Post-Compulsory Choices in England: Choices and Implications. Centre for Vocational Education Research (CVER), London School of Economics. Discussion Paper 001.

Jackson, S. (2014). Educational System in the UK. Available on www.gov.uk

Krueger, D. and Kumar, K. (2004). Skill-specific rather than general education: A reason for US-Europe Growth Differences? Journal of Economic Growth 9:167-207.

Makepeace, G. and Dolton, P. (2001). What can the National Child Development Study and the 1970 British Cohort Study tell us about the individual's acquisition and use of education and training? CLS Working Paper.

Mincer, J. (1974). Schooling, Experience and Earnings. New York: National Bureau of Economic Research.

Oosterbeek, H. and Webbink, D. (2007). Wage Effects of an Extra Year of Basic Vocational Education. Economics of Education Review 26(3), 408-419.

Pischke, J., and Von Wachter, T. (2008). Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation. The Review of Economics and Statistics 90(3): 592-98.

Ryan, P. (2001). The school-to-work transition: a cross-national perspective. Journal of economic literature, March 2001, Vol. 39, No 1, pp. 34-92.

Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. Journal of the Royal Statistical Society, Vol.58(1), pp. 267-288.

Wolf, A. (2011). Review of Vocational Education: The Wolf Report, mimeo

Wooldridge, J. (2007). Inverse probability weighted estimation for general missing data problems. Journal of Econometrics, Vol. 141(2). pp. 1281-1301.

Appendix 1: Further descriptive statistics

Tables A.1 and A.2: Summary statistics of earnings, years 2004-2017

		Full cohort				
Income per tax year:	Observations	Mean	SD	Min	Median	Max
2004	179,310	2,371	1,690	1	1,746	54,153
2005	291,294	4,352	3,440	1	3,283	765,771
2006	368,840	6,184	15,530	1	4,795	11,173,986
2007	380,192	7,948	9,296	1	6,185	4,381,150
2008	376,325	9,395	8,596	1	7,700	1,969,111
2009	392,741	10,767	10,009	1	9,440	2,846,358
2010	419,503	12,504	12,228	1	11,918	3,931,388
2011	449,140	14,361	13,238	1	13,941	3,004,445
2012	461,261	16,030	18,248	1	15,439	6,524,454
2013	451,683	17,485	17,218	1	16,517	3,610,136
2014	457,353	18,722	18,316	1	17,460	3,659,157
2015	459,359	20,211	20,729	1	18,533	3,354,250
2016	458,123	21,641	23,898	1	19,527	4,744,336
2017	454,641	22,416	27,912	1	19,835	5,382,851

Source: NPD-ILR-HESA-HMRC linked data. Real income (CPI 2015=100) in sterling pounds.

		Sample Level 4 plus				
Income per tax year:	Observations	Mean	SD	Min	Median	Max
2004	46,548	1,795	1,162	1	1,121	39,216
2005	76,224	3,039	1,944	1	2,037	42,307
2006	103,121	4,009	2,830	1	2,440	100,174
2007	105,307	4,746	3,722	1	2,826	158,924
2008	103,571	5,667	6,961	1	3,311	1,969,111
2009	115,559	7,583	5,876	1	5,288	740,647
2010	130,020	11,250	7,897	1	8,718	373,965
2011	139,411	15,104	10,193	1	13,602	1,125,234
2012	144,941	18,110	11,589	1	16,974	990,542
2013	142,880	20,782	13,825	1	19,475	685,147
2014	143,678	22,859	16,434	1	21,455	1,187,817
2015	143,267	25,191	20,371	1	23,204	2,636,981
2016	142,501	27,410	26,964	1	24,909	4,744,336
2017	140,909	28,728	30,149	1	26,250	5,382,851

Source: NPD-ILR-HESA-HMRC linked data. Real income (CPI 2015=100) in sterling pounds.

Table A.3: Employment rates (%) by qualification types and gender

	Males			Females		
Year	Level 4/5	Level 6 Non-Russell	Level 6 Russell	Level 4/5	Level 6 Non-Russell	Level 6 Russell
2004	0.35	0.27	0.21	0.34	0.30	0.24
2005	0.55	0.46	0.33	0.55	0.50	0.39
2006	0.71	0.61	0.52	0.72	0.65	0.57
2007	0.74	0.62	0.50	0.74	0.68	0.57
2008	0.75	0.62	0.47	0.74	0.67	0.54
2009	0.77	0.69	0.57	0.77	0.74	0.63
2010	0.81	0.76	0.70	0.81	0.82	0.75
2011	0.87	0.82	0.78	0.86	0.86	0.81
2012	0.89	0.86	0.84	0.87	0.88	0.87
2013	0.87	0.85	0.83	0.85	0.87	0.86
2014	0.88	0.85	0.83	0.85	0.87	0.86
2015	0.89	0.85	0.83	0.85	0.87	0.85
2016	0.88	0.85	0.82	0.84	0.86	0.84
2017	0.88	0.84	0.82	0.83	0.85	0.83

Source: NPD-ILR-HESA-HMRC linked data

Table A.3 shows employment rates (%) by qualification types and gender. In any given year, employment is defined by observing (or not) positive income in the HMRC data. By year 2017, the employment rate is slightly lower amongst degree holders, and males have higher participation rates than females only amongst those achieving Level 4/5 higher-technical qualifications.

Table A.4: List of Russell Group universities

Russell Group universities	
Cardiff University	University of Edinburgh
Durham University	University of Exeter
Imperial College London	University of Glasgow
King's College London	University of Leeds
London School of Economics	University of Liverpool
Newcastle University	University of Manchester
Queen Mary University of London	University of Nottingham
Queen's University Belfast	University of Oxford
University College London	University of Sheffield
University of Birmingham	University of Southampton
University of Bristol	University of Warwick
University of Cambridge	University of York

Source: https://russellgroup.ac.uk/

In the empirical section we distinguish between Level 6 academic institutions: Russell and non-Russell. The Russell Group, originally formed in 1994, is an association of 24 universities in the United Kingdom with a focus on world-leading research and a reputation for academic achievement. The full list of current members is presented in Table A.4.

Table A.5: Level 4 plus qualifications by broad subject area

	Males			Females			All			
Subject Area	L4/5 Voc.	L6 Non-Russell	L6 Russell	L4/5 Voc.	L6 Non-Russell	L6 Russell	L4/5 Voc.	L6 Non-Russell	L6 Russell	Total
STEM	37.9	34.2	40.2	6.7	18.4	22.6	23.4	25.6	31.0	26.7
Construction	14.2	4.7	1.3	*	0.9	0.8	8.5	2.6	1.0	2.6
ALH	14.8	26.1	19.2	16.0	33.2	30.4	15.3	30.0	25.0	28.0
Social Sciences	*	8.2	16.0	2.6	10.7	14.0	1.9	9.6	15.0	10.4
Business	17.3	19.2	10.4	33.4	16.9	10.5	24.7	17.9	10.5	16.6
Health	3.4	3.2	10.4	17.0	10.2	18.6	9.7	7.0	14.7	8.9
Education	3.7	1.2	0.2	14.5	6.5	0.8	8.7	4.1	0.5	3.5
Other	7.1	2.8	1.8	7.7	2.9	1.9	7.4	2.9	1.9	2.9
Total (Freq.)	5,000	53,800	18,200	4,200	64,300	19,600	9,200	118,100	37,800	165,100
Total (%)	100	100	100	100	100	100	100	100	100	100

Source: NPD-ILR-HESA linked data. In Column 1, STEM: Sciences, Technology, Engineering and Mathematics, ALH: Arts, Languages and Humanities

** fewer than 100 in cell*

Table A.5 shows the spread of subject areas across Level 4/5 and 6 qualifications, and a breakdown by gender. The most striking aspect of this table is that Level 4/5 (vocational) is concentrated between Business and STEM subjects. Males have a clear preference for STEM subject. Females concentrate on Business, but also on Health and Education (very few females take STEM subjects). These graphs suggest that subject area is an important element to consider when analysing L4+ qualifications, which is confirmed in the empirical analysis.

Table A.6: Subject classification codes

Broad Subject Areas	Code classification in the ILR (a_ssa_t1) and HESA (JACS)	
1. STEM Subjects	ILR	(2) Science and Mathematics
		(4) Engineering and Manufacturing Technologies
		(6) Information and Communication Technology
	HESA	(6) Physical Sciences (F0, F1, F2, F3, F4, F5, F6, F7, F8, F9)
		(7) Mathematical Sciences (G1, G2, G3, G9)
		(8) Computer Science (I1, I2, I3, I4, I5, I6, I7, I9)
		(9) Engineering and Technology (H0, H1, H2, H3, H4, H5, H6, H7, H8, H9, J1, J2, J3, J4, J5, J6, J7, J9)
2. Construction	ILR	(5) Construction, Planning and the Built Environment
	HESA	(10) Building, Planning and Architecture (K0, K1, K2, K3, K4, K9)
3. Arts, Languages and Humanities	ILR	(9) Arts, Media and Publishing
		(10) History, Philosophy and Theology
		(12) Languages, Literature and Culture
	HESA	(14) Mass communications and Documentation (P0, P1, P2, P3, P4, P5, P9)
		(15) Languages (Q0, Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8, Q9, R1, R2, R3, R4, R5, R6, R7, R8, R9, T1, T2, T3, T4, T5, T6, T7, T8, T9)
		(16) Historical and Philosophical Studies (V0, V1, V2, V3, V4, V5, V6, V7, V9)
		(17) Creative Arts and Design (W0, W1, W2, W3, W4, W5, W6, W7, W8, W9)
4. Social Sciences	ILR	(11) Social Sciences
	HESA	(11) Social Studies (L0, L1, L2, L3, L4, L5, L6, L7, L8, L9)
5. Business	ILR	(15) Business, Administration and Law
	HESA	(12) Law (M0, M1, M2, M9)
		(13) Business and administrative studies (N0, N1, N2, N3, N4, N5, N6, N7, N8, N9)
6. Health	ILR	(1) Health, Public Services and Care
	HESA	(1) Medicine and Dentistry (A0, A1, A2, A3, A4, A9)
		(2) Subjects allied to Medicine (B0, B1, B2, B3, B4, B5, B6, B7, B8, B9)
		(3) Biological Sciences (C0, C1, C2, C3, C4, C5, C6, C7, C8, C9)
7. Education	ILR	(13) Education and Training
	HESA	(18) Education (X0, X1, X2, X3, X9)
8. Other	ILR	(3) Agriculture, Horticulture and Animal Care
		(7) Retail and Commercial Enterprise
		(8) Leisure, Travel and Tourism
		(14) Preparation for Life and Work
	HESA	(4) Veterinary Science (D1, D2)
		(5) Agriculture and related subjects
		(19) Combined (Y0)

Source: ILR-HESA data

Table A.7: LASSO method of model selection. Dependent variable log earnings

Variables selected		Variables dropped	
Experience 2004	Gender	Experience 2007	Ethnicity
Experience 2005	FSM eligibility	Experience 2008	Urbanisation Index
Experience 2006	Type of KS4 School	Prior earnings	School fixed effects
Experience 2009	Region		
Experience 2010	Index of Multiple Deprivation		
Experience 2011	Number of GCSE A*-C		
Experience 2012	Broad Subject Area		
Experience 2013	KS2 Math result		
Experience 2014	KS2 English result		
Experience 2015			
Experience 2016			

Source: NPD-ILR-HESA linked data

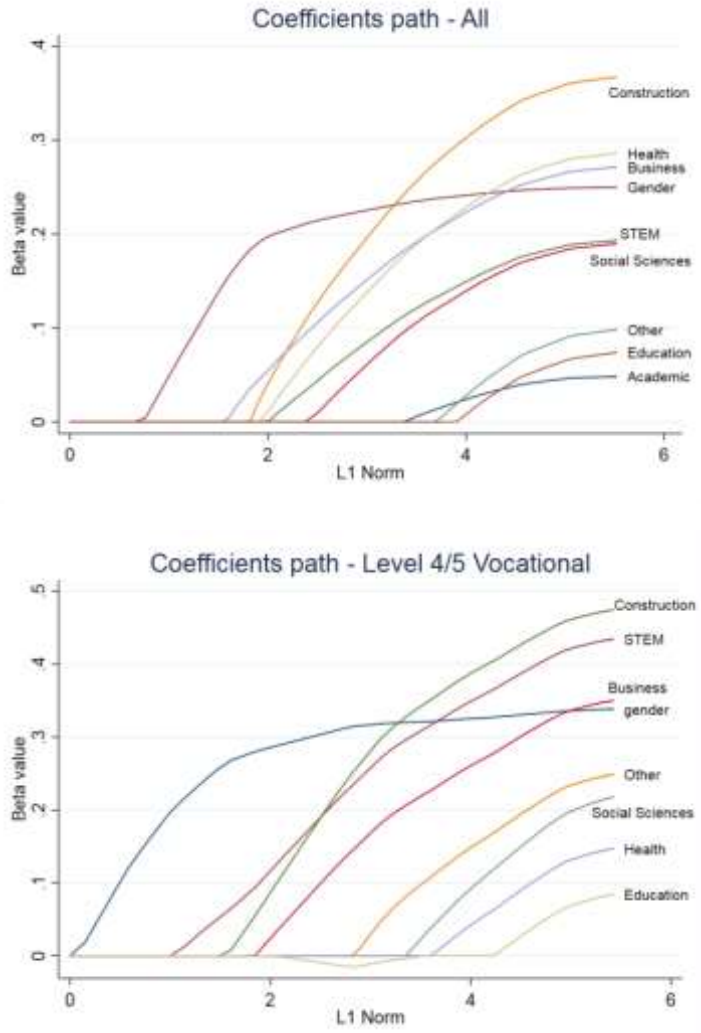
Table A.8: LASSO method of model selection. Dependent variable education (vocational or academic)

Variables selected		Variables dropped	
Experience 2004	Ethnicity	Experience 2010	Urbanisation Index
Experience 2005	Type of KS4 School	Experience 2013	Prior earnings
Experience 2006	Region	Experience 2014	School fixed effects
Experience 2007	Index of Multiple Deprivation	Experience 2015	
Experience 2008	Number of GCSE A*-C	Experience 2016	
Experience 2009	Broad Subject Area	Experience 2017	
Experience 2011	KS2 Math result		
Experience 2012	KS2 English result		
Gender			

Source: NPD-ILR-HESA linked data

We use the LASSO method to remove variables that are redundant and do not add any information to our theory-driven Mincerian model (e.g. keeping the explanatory variables with more predictive power), and thus, reduce the likely overfitting problem. Tables A.7 and A.8 show the set of variables used in the empirical analysis. As can be seen, most of the variables successfully pass the LASSO test and are then included in our OLS and IPWRA estimations.

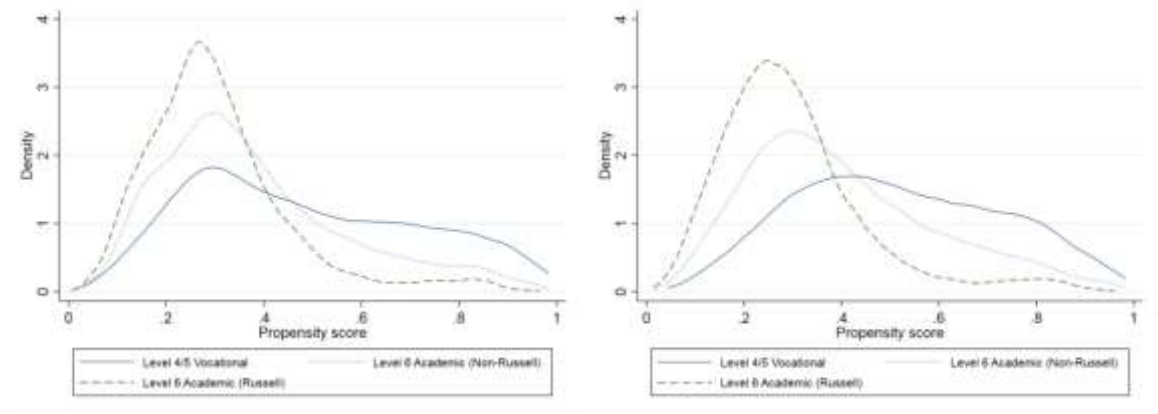
Figure A.1: LASSO Coefficients path



Source: NPD-ILR-HESA linked data

Figure A.1 plots beta coefficients against the L_1 norm. These models are ordered from strongest regularised (left) to least regularised (right). On the left-hand side of the figures (strong regularisers), all the coefficients are exactly 0. When regularisation gets progressively looser, coefficients can get non-zero values one after the other. If we take variable 'gender' as an example, we observe that it enters into the model first (the line is always to the left compared to other lines), which suggests that it is a good predictor (i.e. different from zero). Furthermore, Figure A.1 shows that gender is also correlated with 'subject' because it changes its slope exactly when 'subject' enters into the model. Our key causal variable of interest 'academic' (vs vocational) enters later into the model and with a positive coefficient that is significantly different from zero. The rest of the covariates have been omitted in these graphs.

Figure A.2: Propensity scores (IPWRA first stage)
Females (left) and Males (right)



The Common Support assumption (also known as the overlap assumption) states that for each value of observable covariates X, there is a positive probability of being both treated and untreated. This implies that the probability of receiving a particular treatment for each possible value of the vector X is strictly within the unit interval. This assumption guarantees that there is sufficient overlap in the characteristics of treated and untreated units to find adequate matches (therefore, a comparison is made between similar individuals). In our multivalued case, we observe individuals with similar propensity scores and, therefore, the matching process (via IPWRA) becomes feasible.

Appendix 2: Description of the 2010/11 cohort of school leavers (online only)

In this section we use rich administrative data (NPD-ILR-HESA-HMRC) for a recent cohort (KS4 2010/11) to draw a comprehensive picture of education participation and labour market outcomes of young people in England until the age of 22/23, with a particular focus on high-level technical/vocational versus academic education (Level 4-6). We choose this cohort because it's the most recent cohort that we can feasibly explore early progression routes. This complements our main analysis where we compare the earnings trajectories of these levels of education for an older cohort (KS4 2002/03).

Similar to Brunello and Rocco (2017), we classify individuals to vocational or academic education depending on their highest qualification acquired[21] during the period 2011-2017. In doing this, we implicitly assume that most people of this cohort will have completed full time education by age 22/23. We also include whether the qualification was obtained from an apprenticeship (Table A.9).

Table A.9: Highest level of education achieved by 2017

Highest level of education	Frequency	Percent
Level 1 or below	155,400	24.4
Level 2	64,300	10.1
Level 2 + Apprenticeship	8,500	1.3
Level 3 Vocational	66,400	10.4
Level 3 Mix	22,400	3.5
Level 3 A-Level	113,700	17.8
Level 3 + Apprenticeship	6,400	1.0
Level 4 Vocational	1,800	0.3
Level 4 Academic	200	0.0
Level 4 + Apprenticeship	100	0.0
Level 5 Vocational	1,700	0.3
Level 5 Academic	100	0.0
Level 6 Vocational	<100	0.0
Level 6 Academic or above	196,100	30.8
Total	637,000	100.0

Source: NPD-ILR-HESA linked data. The totals have been rounded to the nearest hundred

[21] Level 1 and below if highest level of education is less than 5 GCSEs A*-C; Level 2 with 5 or more GCSE A*-C (including English and Maths); Level 2 (any) + Apprenticeship; Level 3 A-Level; Level 3 vocational; Level 3 mix (including both A-level and non A-level courses); Level 3 (any) + Apprenticeship; Level 4 vocational (e.g. HNCs, BTEC level 4, and NVQ Level 4 that usually last one year); Level 4 academic (1st year Undergraduate and CertHE); Level 4 (any) + Apprenticeship; Level 5 vocational (e.g. HNDs, Foundation Degrees, BTEC Level 4, and NVQ Level 4 that usually last 2 years); Level 5 academic (2nd Year Undergraduate and DipHE); Level 5 (any) + Apprenticeship (empty category according to this hierarchical classification); Level 6 vocational (such as RFQ Level 6 Diploma); and Level 6 academic (e.g. BSc, BA, LLB, etc.).

First, and most remarkably, Table A.9 shows that almost a quarter of young people only reach education at Level 1 or below by age 22. As summarised in the Regulated Qualifications Framework, Level 1 education provides basic factual knowledge, cognitive and practical skills to complete routine tasks and procedures. Attainment of such qualifications prepares young people to engage in occupational roles of limited complexity. In contrast, and 31% of the cohort hold Level 6 academic qualifications by the age of 22/23 – i.e. highest level conceptual and technological knowledge, including analytical skills to evaluate complex information, to creatively design and produce in their working area.[22] To complete the description, we also observe a sizable group (around 1/3 of the cohort) with achievement, both academic or vocational/technical education, at Level 3 and some successful completion of apprenticeships at Levels 2 and 3. Generally, attainment in low and intermediate level qualifications observed by the age of 22/23 are quite similar to those observed by age 29/30 as show in the main text of this discussion paper. The main difference is that some further acquisition of Level 2 qualifications over time reduces people observed with Level 1 and below. In addition, we find more people with higher education (Levels 6 and 7), which results from both increased university access and the later time point (29/30 as opposed to 22/23 years of age) when looking at the 2002/03 cohort (as shown in Table 1).

Table A.10: Highest level of education achieved per year during period 2011-2017

Highest level of education	2011	2012	2013	2014	2015	2016	2017
Level 1 or below	46.6	31.1	26.8	25.4	24.8	24.6	24.4
Level 2	51.7	11.2	10.0	10.0	10.0	10.1	10.1
Level 2 + Apprenticeship	1.7	1.4	1.3	1.3	1.3	1.3	1.3
Level 3 Vocational		14.6	15.4	16.4	16.7	11.2	10.4
Level 3 Mix		2.0	3.8	3.8	3.7	4.2	3.5
Level 3 A-Level		39.4	42.3	42.2	42.2	25.3	17.8
Level 3 + Apprenticeship		0.3	0.5	0.7	0.9	1.0	1.0
Level 4 Vocational		0.0	0.0	0.0	0.1	0.3	0.3
Level 4 Academic		0.0	0.0	0.0	0.0	0.0	0.0
Level 4 + Apprenticeship			0.0	0.0	0.0	0.0	0.0
Level 5 Vocational				0.2	0.2	0.3	0.3
Level 5 Academic				0.0	0.0	0.0	0.0
Level 6 Vocational				0.0	0.0	0.0	0.0
Level 6 Academic				0.0	0.1	21.8	30.8
Total (%)	100	100	100	100	100	100	100
Total (freq.)	637,000	637,000	637,000	637,000	637,000	637,000	637,000

Source: NPD-ILR-HESA linked data. The totals have been rounded to the nearest hundred.

[22] This figure is based on the expected length of the Level 6 Academic programme. Females and students non-eligible for Free School Meal (KS4) are more likely to achieve Level 6 Academic qualification, compared with males and those FSM eligible, respectively (not shown here).

Table A.10 shows how education outcomes evolve over time for those finishing compulsory education (i.e. age 16) in 2010/11. Here, we provide a description of the highest academic or technical/vocational education achieved by the end of each academic year. Obviously, people gain qualifications continuously, but the change over time differs depending on the route chosen[23]. The majority of 18 year olds who remain in the education system a year or two later pursue university degrees (i.e. by considering the Level 3 A-Level row and the Level 6 academic row). In contrast, very few students enrol for courses at Level 4 or 5. Most tertiary education attainment is Level 6 academic education achieved in the 2016 and 2017 academic years (i.e. by age 21-23). In contrast, Level 4 and Level 5 vocational/technical qualifications increase more slowly. This suggests that the small number of high-level technical qualifications in England tend to be acquired relatively late compared to degrees[24].

Table A.11 shows the highest level of qualification acquired between 2011 and 2017 by the qualification an individual had just before that.[25] As expected, the majority of young people achieving a degree previously had A-levels (about 73%), with most others achieving Level 3 vocational qualifications (about 12%) or some combination of both (13%). Many people achieving Level 5 vocational/technical qualifications also had A-levels as their highest qualification previously (40%), but a higher proportion of young people held Level 3 vocational qualifications previously (46%).

About 8% of those achieving Level 5 (vocational) as their highest qualification have only previously achieved Level 1 or Level 2 qualifications, which in most cases were achieved more than 5 years previously. Also, we see almost no transition from Level 4 to Level 5 vocational qualifications. We find more diversity among students achieving Level 4 vocational qualification, who come from a mixed background of Level 1 to Level 3 qualifications.

[23] It is worth noting that the Table underreports some education outcomes over time as e.g. A-Level graduates, who are subsequently gaining further technical/vocational qualifications either at Level 3 or below, are not reported as the A-Level qualification is likely to be the most significant for progression to HE, and therefore, subsequent vocational/technical education is not shown in the Table.

[24] This is confirmed by the main findings of this discussion paper.

[25] The column headings relate to the highest qualification a person achieved and the rows to the qualification they held previously (irrespective of the number of years between achieving the highest and previous qualification)

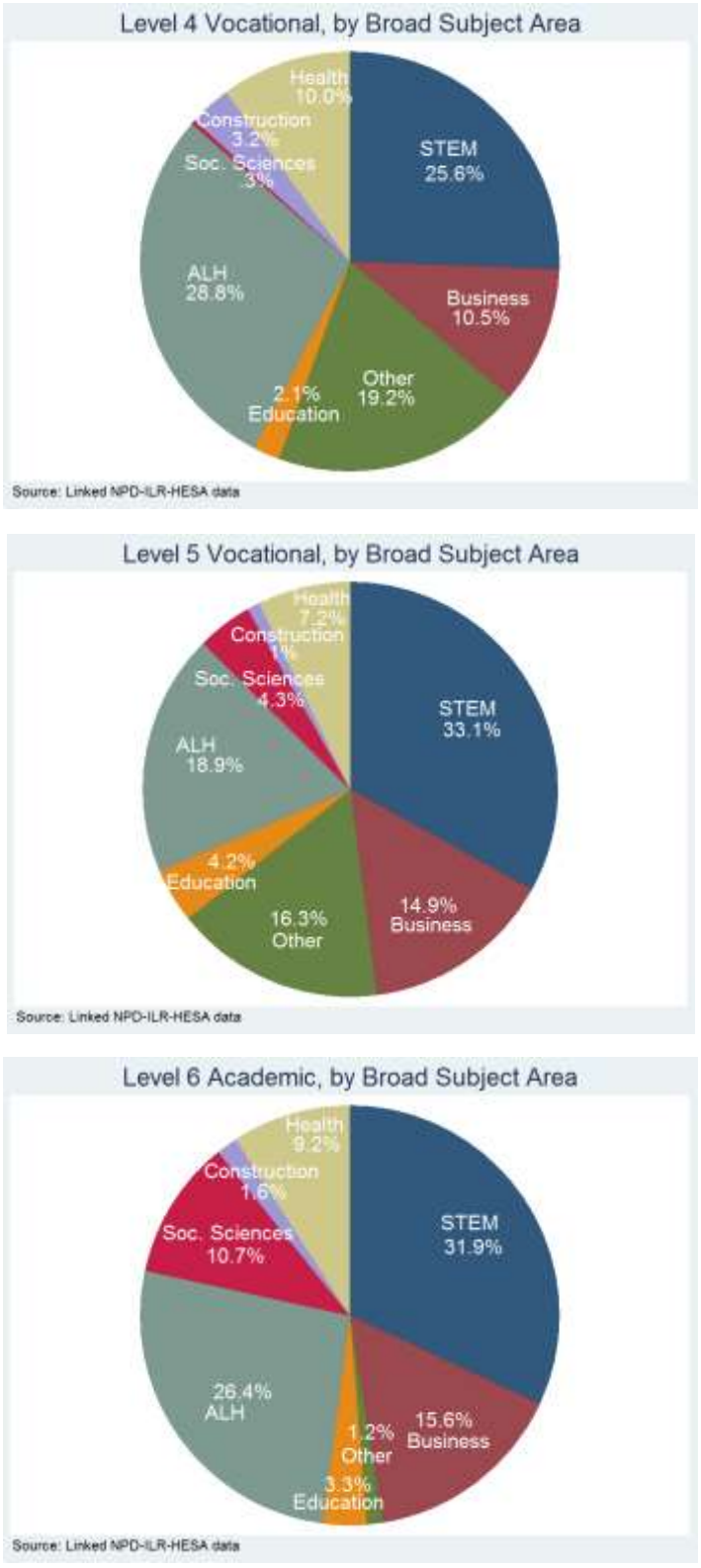
Table A.11: Highest level of education in 2017 by previous highest level held

	Highest Level of Education achieved up to 2017												
Previous level of education	Level 2	Level 3 Voc.	Level 3 Mix	Level 3 A-Level	Level 3+App.*	Level 4 Voc.*	Level 4 Acad.*	Level 4+ App.*	Level 5 Voc.	Level 5 Acad.	Level 6 Voc.	Level 6 Acad.	Total
Level 1 or below	100.0	66.0	29.6	25.1	55.1	14.9	3.2	7.4	4.3	0.0	8.0	0.1	25.4
Level 2		33.3	33.1	68.0	44.9	12.5	0.0	25.0	3.4	0.0	16.0	0.2	25.3
Level 2 + Apprenticeship		0.7	0.9	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.5
Level 3 Vocational			36.4	5.7	0.0	33.3	24.6	19.4	45.5	21.1	32.0	12.7	9.4
Level 3 Mix				0.2	0.0	8.0	5.0	6.5	5.6	11.4	16.0	13.4	6.1
Level 3 A-Level					0.0	30.7	67.3	41.7	40.0	66.7	28.0	73.3	33.3
Level 3 + Apprenticeship						0.6	0.0	0.0	1.2	0.9	0.0	0.1	0.0
Level 4 Vocational							0.0	0.0	0.1	0.0	0.0	0.2	0.1
Level 5 Vocational										0.0	0.0	0.0	0.0
Level 5 Academic											0.0	0.0	0.0
Level 6 Vocational												0.0	0.0
Total	100	100	100	100	100	100	100	100	100	100	100	100	100
Total (freq.)	27,700	66,400	22,400	113,700	6,400	1,800	200	100	1,700	100	<100	196,100	436,500

** App: qualification related to apprenticeship at related level: voc. denotes vocational and acad. denotes academic qualifications*

Source: NPD-ILR-HESA linked data. The totals have been rounded, and those achieving their highest level of education in year 2011 have been omitted (i.e. no transition).

Figure A.3: Level 4-6 qualifications, by broad subject area



Source: NPD-ILR-HESA-HMRC linked data

Figure A.3 shows the subject areas of Level 4, 5 and 6 qualification taken by the 2010/11 KS4 leavers. Comparing Level 4 to Level 6, the broad pattern of subject specialism is not dissimilar, with STEM, Art, Languages and Humanities (ALH) and Business being the most popular subjects.

Table A.12: Level 4-6 qualifications in subject groups by gender

	Level 4V		Level 5V		Level 6A	
Broad Subject Area	Female	Male	Female	Male	Female	Male
STEM	12.0	88.0	19.8	80.2	40.0	60.0
Construction	7.1	92.9	14.3	85.7	35.5	64.5
ALH	58.2	41.8	46.0	54.0	64.8	35.2
Soc. Sciences	83.3	16.7	65.1	34.9	55.8	44.3
Business	53.7	46.3	49.5	50.5	52.0	48.0
Health	81.5	18.5	71.4	28.6	73.3	26.7
Education	77.8	22.2	95.1	4.9	88.2	11.8
Other	67.7	32.4	59.1	40.9	64.3	35.7
Total (%)	49.0	51.0	44.4	55.6	55.0	45.0
Total (Freq.)	900	900	600	800	107,400	87,900

Source: NPD-ILR-HESA linked data. Totals rounded, also note N<100 for Education and Construction L4/5V

While the overall differences in subject areas between Levels 4, 5 and 6 do not look too dissimilar, Table A.12 shows that subject choice is to a large extent driven by different choices of male and female students. It is worth noting that across Level 4 plus qualifications, STEM and Construction are heavily dominated by males, with Health, Social Sciences and Education dominated by females. As a consequence, the differences in earnings trajectories between male and female students are to a large extent caused by the different subjects.

Table A.13: Earnings Trajectories 2011-2017, by type of qualification, gender and Free School Meal Eligibility, based on positive income in 2015 £

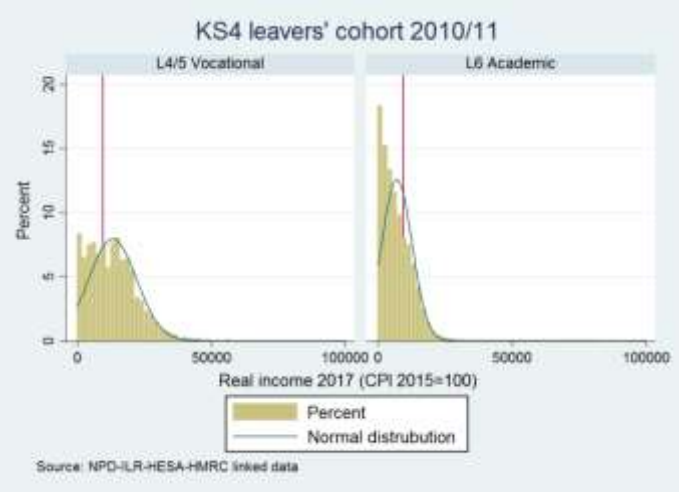
Income per tax year	All		Subgroups							
			Males		Females		FSM: No		FSM: Yes	
	L4/5 Voc.	L6 Acad.	L4/5 Voc.	L6 Acad.	L4/5 Voc.	L6 Acad.	L4/5 Voc.	L6 Acad.	L4/5 Voc.	L6 Acad.
2012	1,997	1,417	2,492	1,486	1,543	1,375	2,032	1,409	1,651	1,335
2013	3,636	2,231	4,444	2,241	2,884	2,225	3,735	2,271	3,207	2,101
2014	5,699	2,931	6,518	2,980	4,859	2,896	5,941	2,969	4,444	2,846
2015	8,556	3,189	9,517	3,218	7,564	3,170	8,957	3,237	6,510	3,379
2016	10,923	4,328	12,331	4,641	9,444	4,101	11,398	4,391	8,609	4,441
2017	13,137	7,001	14,898	7,025	11,281	6,983	13,742	7,109	10,669	6,920
N (in year 2017)	1,886	159,609	968	68,023	918	91,586	1,519	121,068	277	17,859

Source: NPD-ILR-HESA-HMRC linked data, ONS Consumer Price Index

Table A.13 shows earnings trajectories by type of qualification and in addition, subgroups by gender and Free School Meal (FSM) eligibility. The description of the earnings profiles shows that people with vocational qualifications have higher earnings in early adulthood compared to those with academic education.

However, to a large extent this pattern is likely to result from the fact that by age 22/23, many young people – even when showing some earnings in the data – remain involved in education. Those engaged in degree programmes in particular are likely to work fewer hours, something administrative data on earnings cannot capture, but which can be shown to some degree when describing the earnings distributions in histograms (Figure A.4). Here, we show all people reporting non-zero earnings in the most recent financial year available, which is when people are around 22 or 23 years old. The skewed distribution at the lower end for people following the academic route shows that many people here – compared to those in the vocational route – are likely to work few hours. In contrast, those with Level 4 and 5 vocational education have a less skewed distribution and a higher median (red line in the diagram). The higher incomes of those with vocational/technical routes likely result from more hours spent working compared to those in the academic route and more work experience of those obtaining Level 4 and 5 vocational qualifications in early adulthood compared to those studying for degrees.

Figure A.4: Earnings distribution in 2017 by educational route



Source: NPD-ILR-HESA-HMRC linked data

As a consequence, a systematic comparison of earnings profiles until age 22 is too early to be informative to understand earnings differentials from making the decision to invest into either vocational or general tertiary education. Labour market participation itself is still affected by education participation at this age. Furthermore, those with Level 4/5 vocational education have more work experience by this time (programmes take less time and there is more time to engage in paid work).

To conclude, the main finding of this complementary analysis is that progression to high-level technical/vocational education (Level 4-5) is only taken by a tiny fraction of the full cohort, and tends to be acquired late relative to academic (Level 6) education. Students aiming for such qualifications have a mixed educational background ranging from Basic Skills to 3 A-levels and very different subjects are chosen by male and female students, which are the key drivers of subsequent earnings trajectories[26]. Although here we describe important early education progression patterns and earnings up to age 23, we refer the reader to the main text of the discussion paper to properly understand how earnings may be related to different educational pathways. Earnings at age 23 are still very much affected by ongoing education involvement amongst the groups of interest in this analysis.

[26] While the descriptive work here is limited to 2010/11 KS4 leavers, the discussion paper focusing on earnings differentials between groups based on the 2002/03 cohort shows very similar early patterns of education progression and individual characteristics associated with particular routes.