EDGE FOUNDATION

In collaboration with

NFER
National Foundation for Educational Research

Investigating the potential use of long-term school and college destination measures

Technical report

February 2022

Making Education Relevant

www.edge.co.uk

In collaboration with

Investigating the potential use of long-term school and college destination measures

Technical report

Jenna Julius
Jude Hillary
Henry Faulkner-Ellis

February 2022

Published in February 2022 by the National Foundation for Educational Research, The Mere, Upton Park, Slough, Berkshire SL1 2DQ
www.nfer.ac.uk

ISBN: 978-1-912596-53-9

How to cite this publication: Julius, J., Hillary, J. and Faulkner-Ellis, H. (2022). *Investigating the Potential use of long-term School and College Destination Measures. Technical Report.* Slough: NFER.

Acknowledgements:

We wish to thank officials at the Department for Education (DfE) for providing us with access to the Longitudinal Educational Outcomes (LEO) dataset through the LEO pilot and making this research possible. We would particularly like to thank David Burnett and Philippa Norgrove for their invaluable support.

We would also like to thank Anna Vignoles CBE, Director of the Leverhulme Trust who provided detailed feedback on an earlier draft of this report.

The National Foundation for Educational Research (NFER) is an independent, not-for-profit organisation dedicated to producing high quality, independent research and insights to inform key decision-makers about issues across the education system. Our mission is to generate evidence and insights that can be used to improve outcomes for future generations everywhere, and to support positive change across education systems. Visit https://www.nfer.ac.uk.

This research was funded by the Edge Foundation, an independent foundation working to inspire the education system to give all young people across the UK the knowledge, skills and behaviours they need to flourish in their future life and work. Our mission is to make education relevant. Visit www.edge.co.uk.

Disclaimer:

Contents

1 Our research

One of the key purposes of education is to enable young people to gain knowledge and develop the skills and behaviours to support their progression to further study, training, and employment. This enriches the young person's experience by making education meaningful and demonstrating its relevance to their current and future life experience (Edge Foundation, 2018; OECD; 2018). However, the main focus of schools and colleges has historically been on attainment outcomes and less on specific measures of future labour market success or the wider benefits to society and individuals.

Over the last decade, there has been an increasing focus on looking at what young people move on to after they leave the education system and the vital role schools and colleges play in supporting a young person's journey towards the labour market. New destination measures have been developed and published by the Department for Education (DfE), which provide information about what young people were doing in the 12 months after leaving their post-16 school or college.

While these destination measures have been an important development, they only provide a view of what is happening in the short term, when many young people are still in a transitionary state. Information about longer-term destination outcomes may provide greater insights for schools and colleges, who could potentially use them to help improve young peoples' labour market outcomes.

The purpose of this exploratory study is to investigate the potential for using DfE's Longitudinal Educational Outcomes (LEO) dataset to historically track the labour market outcomes of young people (up to age 30) who completed their education in the English education system. It also seeks to provide information about the value that schools and colleges can add to young people's destinations.

While there are many factors which impact a young person's labour market outcomes both up to post-16 (e.g. their family circumstances, etc.), and after post-16 (e.g. the institution where they complete any subsequent qualifications, etc.), we focus on where young people study their post-16 qualifications. Specifically, our research considers to what extent destination measures can provide information about the 'value-added' by schools and colleges in supporting young people to progress onto high quality destinations. In other words, by 'value-added', we focus on isolating the impact of the institution where they studied their main 16-18 qualification, over and above their prior attainment and wider contextual factors.

This technical report provides an overview of the datasets and methodology used in our main report. As this project was of several chosen to pilot opening up the LEO data to third party researchers, it also outlines a number of additional recommendations for improving and enhancing the LEO data.

Recommendations for improving the LEO data

NFER was provided access to the LEO data through the Department for Education's LEO pilot. A key goal of the pilot was to extend the use of the LEO data to the wider research community, and support continued improvements in the access and use of the LEO data. To support this goal, we have outlined a number of recommendations relating to the future development of the LEO data below.

The current LEO dataset presents a wealth of new opportunities for researchers to explore the progression routes of young people. However, there are a number of improvements that would expand the potential of LEO, and improve the quality of research using LEO.

Recommendation 1: Improve the integration of post-16 apprenticeship data into the National Pupil Database to support the analysis and comparison of all post-16 routes and trajectories

It is currently difficult to compare data for young people who progress to apprenticeships, compared to other routes. As part of the Government's plans to streamline post-16 provision, the Government should also integrate apprenticeship data into post-16 National Pupil Database data to facilitate comparisons between young people who pursue apprenticeships and those undertaking other routes.

Recommendation 2: The Government needs to invest in the collection of new administrative data items which could be added to LEO to help generate much stronger insights and support evidence based policy decisions

Our research highlights a number of issues with the LEO data such as lack of information on hours worked, occupation and information on wider background characteristics. These key pieces of information are essential for better understanding labour market dynamics, and for improving policy development to support mobility and opportunities for all groups within the labour market, including those who are socio-economically disadvantaged.

However, these key pieces of information are not currently consistently collected as part of the administrative data that Government departments collect in discharging their functions. The Government need to develop a strategic cross-Government plan to collect and integrate this data into LEO and other data sources.

Recommendation 3: The Government should continue to dedicate resources towards improving the quality and completeness of data within LEO

While the LEO dataset offers novel opportunities for research, and there have been significant improvements to the quality of LEO data in recent years, there is a need for the Government to continue to invest in improving the quality and reliability of the LEO data. In particular, the Government should focus on improving the employment and earnings information within LEO. For example, there are recorded employment spells in LEO with missing earnings information.

2 Data

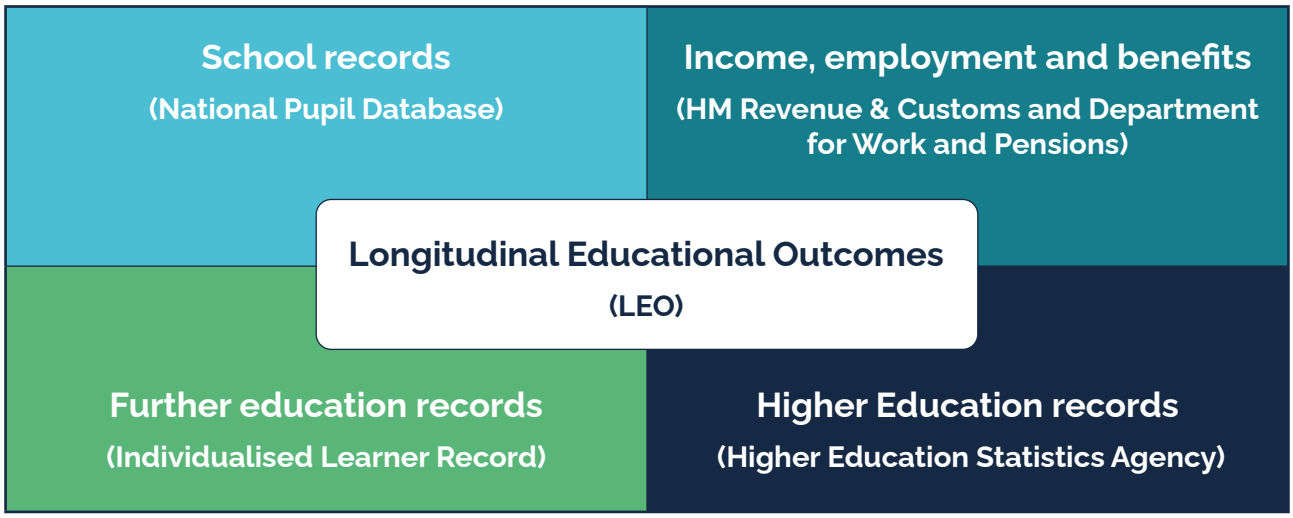
This section provides an overview of the LEO data, which was used to do all of the analysis in this research report, and outlines the approach used to construct our analysis dataset (see section 1.2 of the main report for a summary). It also highlights the key caveats and limitations associated with the LEO data.

2.1 Overview

Our analysis uses the LEO dataset, which combines school, further and higher education information with earnings and benefits data, as shown by Figure 1 . As the LEO dataset is based on administrative data, it can be used to track the labour market outcomes of whole cohorts of young people who completed their education in the English education system. This enables us to identify the association between the post-16 schools and colleges within which they studied and their labour market outcomes. Further, the LEO data contains a rich set of contextual information that can be used to investigate a range of other factors that affect young people's labour market outcomes alongside their qualification level and the institution they attended.

More specifically, we analyse the labour market trajectories until 2016/17[1] of all the cohorts of young people who started their post-16 qualifications in mainstream schools and colleges in England between 2003/04 and 2012/13. However, we mainly focus on describing the trends within the 2003/04 cohort, as we are able to track their labour market histories for the longest.

Figure 1: The components of the LEO data



[1] This is the latest data available to researchers at the time which data access was granted.

2.2 Key caveats

While the LEO data enables analysis that would not have been possible using previously available data, there are a number of important caveats associated with the LEO data, which include:

- **Coverage:** The school and further education data within LEO covers England, whilst the labour market information covers Great Britain. In turn, the data only allows us to analyse the long-term labour market outcomes of young people who complete their schooling within England, and remain in Great Britain during their employment. In other words, it is not possible to track the educational trajectories of people who were educated overseas who move to Great Britain for work or the labour market outcomes of young people educated in England who move overseas.
- **Hours worked:** The data does not include information on hours worked, which means we are not able to adjust earnings for the intensity of an individual's role. As a result, our analysis under-estimates the earnings of part-time workers compared to full-time workers. As such, our analysis implicitly assumes that individuals who attend a specific school or college are no more likely to subsequently work part-time than those at other institutions. If this assumption is incorrect, then our analysis will tend to under-estimate the earnings of institutions where individuals are more likely to work part-time. This is a particular concern for women, who are more likely than men to work part-time as they progress in their careers. This will underestimate the value-added of institutions where women are systematically more likely to choose to work part-time once they enter the labour market, even once background characteristics are taken into consideration.
- **Self-employment data:** We are only able to identify self-employed workers in our data from the 2014/15 financial year onwards. In earlier years, these workers are not identified as being in employment or having earnings. In turn, where our analysis relates to data from earlier years, we may be penalising young people (and the schools and colleges which they attend) who are more likely to be working in sectors with high self-employment such as skilled trades or the creative industries.
- **Data quality:** The LEO data combines information from a number of administrative sources that require matching across individuals, and vary in data quality. For example, only 93 per cent of records could be matched for the KS4 cohort starting their post-16 qualifications. This could affect the magnitude of our estimates if different types of individuals from some schools or colleges are more or less likely to be observed in the data, albeit the impact is likely to be small. For further detail, see Department for Education LEO guidance documentation, and previous reports (Anderson and Nelson, 2021).
- **Availability:** By their very nature, one of the limitations of looking at destinations is the length of time needed to link a young person aged 16 to 18 to their labour market outcomes. Currently, this is compounded by the fact that there is a long delay associated with earnings and employment data becoming available in LEO. Indeed, the latest data available to researchers at the time which data access was granted for our research was 2016/17. While it is currently possible to apply to access LEO data up to 2018/19, this still represents a long lag.
- **Background information:** The LEO data includes a less extensive, rich set of contextual and background characteristics compared to other large longitudinal surveys, such as the Millennium Cohort Study and Next Steps. For example, a young person's social class can be observed in these datasets, but not in LEO. In future, it may be possible to link the LEO data with these surveys to improve the robustness of analysis conducted using LEO.

2.3 Dataset construction

This section provides an overview of the main steps involved in the construction of our analysis dataset.

Earnings information

Our central estimates are based on measures of annual employment and self-employment earnings that have been constructed by the Department for Education (DfE). Earnings were cleaned to remove extreme and inconsistent values. For example, earnings which were unreasonably low given the number of days worked were excluded. The robustness of our results was checked using alternative measures of earnings (annualised and daily earnings).

Earnings information was only used for individuals who were recorded as being in either sustained employment or self-employment, but not in sustained education. Given that, as discussed in section 2.2, it is not possible to identify which individuals are working part-time in LEO, this was intended to minimise the impact that including individuals who might be working part-time while studying could have on our estimates.

Sustained destination

Our definitions of sustained education and employment destinations are based on those developed by the DfE (see Anderson and Nelson, 2021). In turn, at each age, every young person in the cohort is identified as being in one of the following categories:

- **Sustained education:** In education at least one day in each of the 12 months of that financial year.
- **Sustained employment:** In employment at least one day in each of the 12 months of that financial year, and not identified as being in sustained education.
- **Sustained self-employment:** Self-employed at least one day in each of the 12 months of that financial year, and not identified as being in sustained education or employment. This information is only available for the last three years of data.
- **Claiming benefits:** On benefits at least one day in each of the 12 months of that financial year[2].
- **Without a sustained education or employment destination:** In education, employment or on benefits which was not sustained in the previous 12 consecutive months. This category would, for example, include young people who had just transitioned from education to employment.
- **No destination identified:** Not identified as being in any of the categories above. This group includes those who are not in education or employment or claiming benefits, who have moved overseas and those who are deceased.

National Pupil Database (NPD) data

Key stage 2 and 4 data were cleaned to produce a summary of each young person's educational history. Where a young person was identified as having multiple records, we used the number of points scored in their qualifications to identify a main record wherever possible.

[2] Without being identified as being in sustained education or employment.

Our analysis draws on KS5 student and exam files to identify the main qualifications undertaken within schools at KS5[3]. It identifies all qualifications over which an individual has been enrolled in an exam (including those qualifications where an individual was not awarded a grade). Only qualifications started at ages 16 and 17 were kept in the analysis. By construction, this data excludes all young people who have withdrawn from their qualification before enrolling in an exam. As with the KS2 and KS4 data, records were cleaned to identify a unique institution for each individual. For example, where the same young person had records linking them to multiple institutions, their main institution was identified using the KS5 point scores associated with their qualifications.

Ethnic group information for pupils was drawn from the NPD, and pupils were classified across the following categories: Asian (excluding Indian and Chinese), Black (excluding African), Chinese, mixed race, Indian, African, and other non-white ethnicity. Indian, Chinese and African pupils were considered separately to other Asian and African ethnic groups, as there are sufficiently large numbers of pupils in these groups for their outcomes to be analysed separately.

Individualised learner record (ILR) data

ILR data was cleaned to identify a single observation per qualification, and a single learning aim per apprenticeship and traineeship qualification. Only qualifications started at ages 16 and 17 by the LEO cohorts were kept in the analysis. All AS and A-Level qualifications were excluded as these are included in the KS5 NPD data above.

Only completed (regardless of whether the qualification was passed or failed) qualifications were included in the analysis. This was in order to maximise comparability of the analysis across vocational and academic qualifications, and to simplify the identification of the main qualification undertaken for different individuals. For example, attrition rates for apprenticeship qualifications are very high, relative to most other qualifications.

Institutional information

Historical information for schools, colleges and other providers is patchy. In order to identify unique schools and colleges in the data, we developed a consistent mapping between the Unique Reference Number (URN), Unique Provider Number (UPIN) and UK Provider Reference Number (UKPRN) drawing on individual-level NPD and ILR records, combined with the Get Information About Schools (GIAS) dataset, performance tables, Ofsted records and other DfE publications. School- and college-level information was also drawn from the above sources. It should be noted that Ofsted information from 2005 and performance tables from 2006 onwards were used to impute institutional characteristics in earlier years.

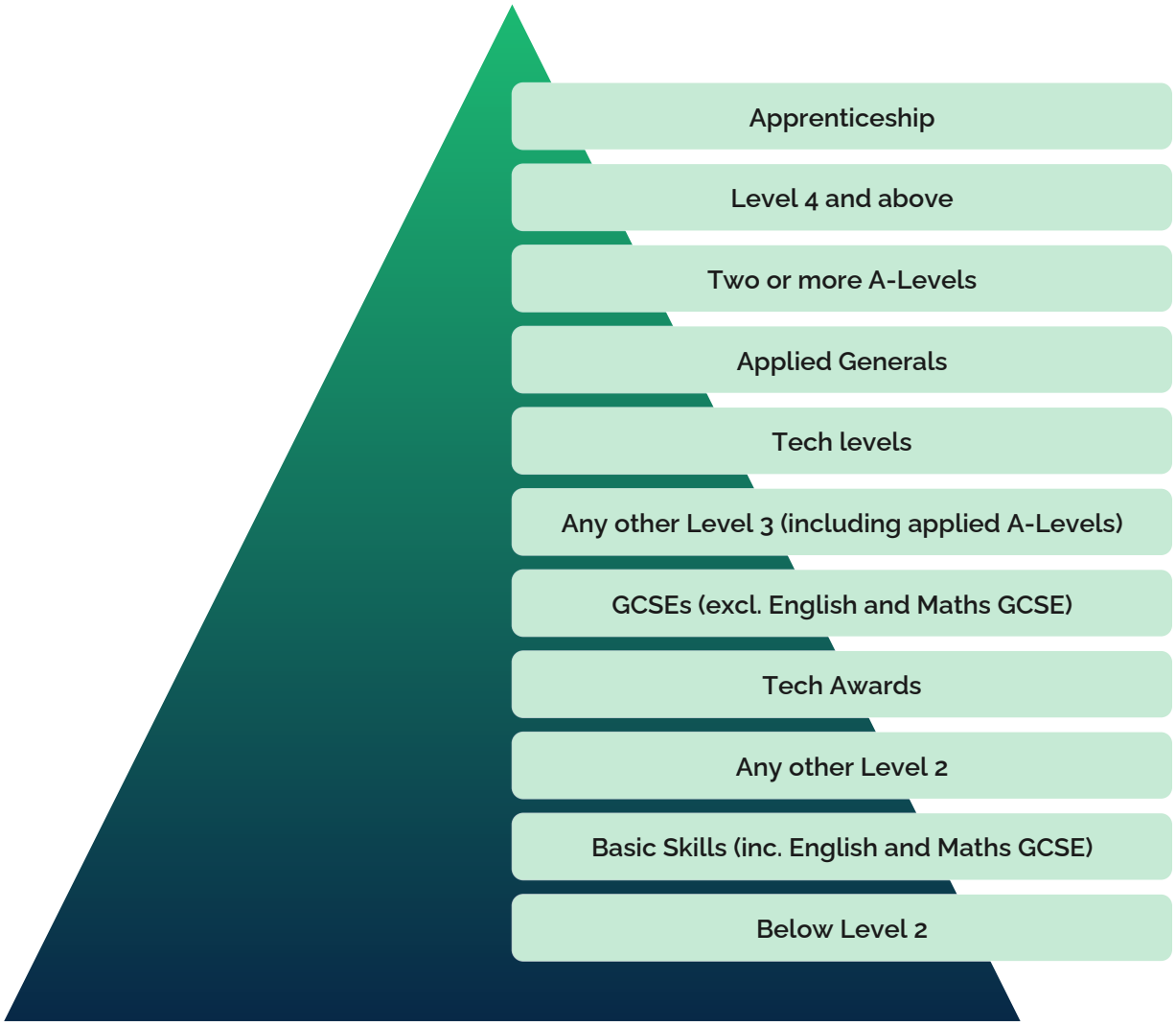
Identifying the main qualification studied and institution between age 16 and 18

In order to identify the main qualification and institution studied between the ages of 16 and 18, our analysis combined the data processed from the KS5 data in the NPD and the ILR, as outlined above. We use the ordering outlined in Figure 2 to identify the main qualification undertaken by the individuals across both the ILR and the NPD.

For each qualification studied, we then classify the subjects studied in each of the following categories: Science and Mathematics (incl. Engineering), Music and Drama, Art and Design, Humanities, English literature and language and other.

[3] Post-16 learning aims (PLAMS) data was not used as it is only available from 2007/08 onwards.

Figure 2: Main qualification undertaken by an individual



Linking across datasets

While the LEO data can be used to track the labour market outcomes of whole cohorts of young people who completed their education in the English education system, we restrict the sample of individuals included in our analysis to reflect missing data, and to maximise the comparability of the young people included in our analysis. We also restrict our sample to individuals for whom we are able to identify the institution where they studied their main post-16 qualification. Table 1 presents the criteria used for restricting records in the KS4 and post-16 analysis.

Table 1: Criteria for excluding records from KS4 and post-16 analysis

	Criteria	Rationale
Excluding records from both the KS4 and post-16 analysis	• Recorded as studying in anything other than mainstream state-funded school at KS4 • Excluded from the national key performance table estimates at KS4 (for all years where this applies) • No recorded KS4 attainment information	• Detailed demographic information is only available for pupils studying in a mainstream state-funded school at KS4 • KS4 information is required to account for prior attainment • School or college information is required to identify institutional value-added • We focus on pupils in state-funded mainstream schools and colleges to ensure comparability across pupils given that there is a wide range of experience and capability across pupils outside of mainstream provision • We focus on pupil's studying at state-funded schools given that detailed demographic information is not available for pupils in independent schools • We focus on pupil's studying Level 2 and above qualifications and on pupils who have completed or enrolled in an exam for their qualification to maximise comparability across pupils studying at different institutions and across different routes
Excluding records from the KS4 analysis only	• Identified as attending multiple institutions during final year of KS4 • Included in the school key performance table estimates at KS4 (for all years where this applies)	
Excluding records from the post-16 analysis only	• No school or college identified with a qualification started at ages 16 and 17 • Recorded as studying in a non-mainstream or independent school for their post-16 qualifications • Studying a below Level 2 qualification between ages 16 and 17 • Pupils who have not completed or enrolled in an exam for their qualification	

The impact of these criteria on the sample included in our analysis is shown in Table 2 which illustrates what proportion of the cohorts' records were identified to be included in each of the analyses. It shows that around ninety per cent of the cohort records were in scope for the KS4 analysis, and around two-thirds of the cohort records were identified for the post-16 analysis.

Table 2: Records for inclusion in the KS4 and post-16 analysis

Year of KS4	Number of learners in LEO	% of learners with an institution identified in scope for KS4 analysis	% of learners identified for post-16 analysis	% of learners with a qualification and institution identified for post-16 analysis
2002	580,070	92%	87%	65%
2003	602,245	92%	87%	67%
2004	609,965	90%	85%	66%
2005	627,285	90%	85%	67%
2006	636,995	90%	85%	67%
2007	639,550	90%	86%	70%
2008	626,265	89%	86%	71%
2009	632,465	89%	85%	71%
2010	621,325	89%	86%	74%
2011	616,175	89%	85%	74%

Source: NFER analysis of LEO data

3 Econometric analysis

This section outlines the econometric approach used to estimate the returns associated with attending different types of post-16 institutions, and the value-added associated with attending a given institution.

3.1 Analysis of post-16 institution type

The institutions and qualifications young people select into at KS4 and ages 16-19 depend on factors such as their prior attainment, family background, regional labour markets, subject interests and individual characteristics. This makes it difficult to disentangle, using descriptive statistics, the extent to which differences between institutions are driven by differences in the characteristics and background factors of pupils who attend different types of institutions.

We use an econometric model to account for these factors. This enables us to identify the impact of studying at a school or college on later life earnings or future employment status, over and above the potential outcome from attending an alternative institution – which may be realistic for a young person with such characteristics to attend.

To do so, we run a multivariate linear regression for earnings outcomes and a multivariate logit estimation for whether an individual is in a sustained education, employment or self-employment destination, controlling for individual-level and qualification-level characteristics with errors clustered at the institution level.[4]

In order to estimate the impact of post-16 institution type of earnings, we use the log of average earnings at ages 28 and 29 (including self-employment earnings) as our main outcome variable. We use averaged earnings over two years to smooth fluctuations in income over time, and to maximise the number of young people who can be included in our analysis. Further, the log of earnings is used to minimise the impact of our extreme observations on our results. As flagged in section 2.3, we only include earnings for individuals who are in sustained employment and self-employment, but who are not also identified as being in sustained education.

In order to estimate the impact of post-16 institution type on employment, we use a flag for whether each individual has been identified as being in sustained education, employment or self-employment as our main outcome variable. As outlined in section 2.3, an individual is considered to be in sustained education, employment or self-employment if they have been identified as doing that activity at least one day per month throughout that year.

The individual-level characteristics that we control for include region at KS4, whether a pupil is in an urban or rural area at KS4, GCSE attainment, disadvantage at KS4 (location-based, through IDACI, and pupil-level, through FSM eligibility), special educational needs at KS4, ethnicity and first language. We also include controls for the characteristics of a pupil's peers in the institution where a young person studied at KS4. This is to account for the fact that peer characteristics (e.g. average attainment of peers) at KS4 have been shown to be associated with the institution where a young person chooses to study their post-16 qualification (Crawford *et al.*, 2011).

[4] These are whether English is an additional language, average KS4 attainment, whether achieved a C or above in GCSE English and maths, whether eligible for free school meals at age 16, whether eligible for special educational needs at age 16, IDACI score of a young person's home postcode at age 16, government office region at age 16, ethnicity and whether institution is located in an urban or rural location.

We also account for the broad qualification type and subjects undertaken by each young person. This means that our analysis is identifying the impact of attending a certain type of institution over and above the fact that pupils may be more or less likely to study a given qualification at one institution, compared to another. Given that there is already a relatively large literature looking at the returns associated with studying different qualifications (Patrignani *et al.*, 2017; Battiston *et al.*, 2019), this enables us to focus on the extent to which outcomes vary across institutions, once differences in qualifications are accounted for. We are only able to control for qualification characteristics at a broad level because there are substantial differences in the qualifications that learners study at different institutions. To account for this, we also run robustness checks by estimating institutional value-added on sub-samples of learners which are more comparable.

Regressions are run separately by gender, and we focus on reporting estimates for men. This is due to us being unable to observe part-time work in our analysis and the fact that women are more likely to work part-time (UK Parliament, 2021), which means our estimates for women are less robust. More specifically, our estimates for women will tend to underestimate the value-added associated with attending post-16 schools or colleges that are more likely to attract pupils who will choose to work part-time after they finish their education. For example, women in some communities may choose to have children earlier, and work part-time.

We check for the robustness of our findings across a number of specifications (including varying the controls included in our analysis, the sub-samples considered, and accounting for local authority fixed-effects). We also test for multicollinearity to account for the large number of controls included in our analysis.

A high-level summary of our regression findings from our earnings estimations is provided in Table 3 below. These are discussed in further detail in Section 4.2 of the main report. The coefficients presented in the table have been transformed using an exponential approximation. This means that each of the coefficients in the table can be interpreted as the percentage change in earnings associated with attending a given type of institution compared to a school. Table 3 shows that, while the magnitudes of our point estimates vary substantially across specifications, pupils who attend general further education (FE) colleges or sixth form colleges tend to have lower earnings than their counterparts attending schools do.

We find that the relative returns associated with attending a general FE college compared to a sixth form college or a school are lower for females compared to males. While this is a finding that merits further investigation, one plausible explanation could be that women who study at general FE colleges are more likely to subsequently work part-time than women who study in schools and sixth form colleges. However, as highlighted in section 2.2, we are unable to test this theory, as we are unable to observe part-time work in the LEO data. This again highlights the need to improve the earnings information that is currently available in LEO.

Table 3: Earnings differentials associated with attending different types of institutions

Outcome	Including self-employment earnings	General FE college	Sixth form college	Number of observations	Specification
Male					
Log of average annual earnings 28-29	Yes	-3.2***	-3.3***	294,925	Main specification
Log of average annual earnings 28-29	Yes	-2.9***	-2.5***	294,925	Including only basic pupil-level and prior attainment controls
Log of average annual earnings 28-29	Yes	-6.1***	-2.9***	140,122	Restricting sample to A-Level pupils only
Log of average annual earnings 28-29	Yes	-2.8***	-2.9***	294,912	With local authority fixed-effects
Log of average annual earnings 26-27	No	-1.9***	-3.1***	284,151	Using 2005/06 and 2006/07 cohorts
Female					
Log of average annual earnings 28-29	Yes	-6.1***	-1.8***	293,777	Main specification
Log of average annual earnings 28-29	Yes	-12.6***	-1.9	293,777	Including only basic pupil-level and pri-or attainment con-trols
Log of average annual earnings 28-29	Yes	-4.5***	-1.4***	163,368	Restricted to A-Level pupils
Log of average annual earnings 28-29	Yes	-5.7***	-1.2**	293,777	With local authority fixed-effects
Log of average annual earnings 26-27	No	-4.5***	-1.7***	301,767	Using 2005/06 and 2006/07 cohorts

Source: NFER analysis of LEO data
Note: Analysis includes all individuals in a mainstream state-funded school in England at KS4 and in a mainstream institution for their post-16 qualifications with non-missing KS4 attainment records and at least one completed Level 2 and above post-16 qualification. The coefficients presented in the table have been transformed using an exponential approximation. *** indicates that coefficient is statistically significant at the one per cent level, ** indicates that coefficient is statistically significant at the five percent level, * indicates that coefficient is statistically significant at the one per cent level.

A summary of our regression findings from our employment estimations is also provided in Table 4. Coefficients are provided in terms of odds ratios. In other words, they provide the relative odds of a young person being in sustained employment, self-employment or education at age 29 if they attended a general FE college or a sixth form college compared to a school. An odds ratio which is less (greater) than one indicates that a young person attending a given type of institution is less (more) likely to be in a sustained destination compared to a young person attending a school.

The table shows that pupils who attend general FE colleges tend to have a lower likelihood of being in sustained education, employment or self-employment than their counterparts attending schools. However, as shown by the table, this result is not statistically significant across all specifications so results should be interpreted with caution (as outlined in section 3.2 of the main report). Similarly, we do not find evidence that pupils who attend sixth form colleges are less likely to be in sustained employment, self-employment or education compared to pupils attending schools.

Table 4: Odds ratios of being in sustained employment, self-employment or education associated with attending different types of institutions

Outcome	Including self-employment earnings	General FE college	Sixth form college	Number of observations	Specification
Male					
Probability of being in sustained employment, self-employment or education at age 29	Yes	0.96**	0.97*	373,063	Main specification
	Yes	1.00	1.00	373,063	Including only basic pupil-level and prior attainment controls
	Yes	0.92***	0.99	177,147	Restricting sample to A-Level pupils only
Female					
Probability of being in sustained employment, self-employment or education at age 29	Yes	0.96***	0.98	391,413	Main specification
	Yes	0.94***	1.00	391,413	Including only basic pupil-level and prior attainment controls
	Yes	0.94	1.00	205,691	Restricting sample to A-Level pupils only

Source: NFER analysis of LEO data
Note: Analysis includes all individuals in a mainstream state-funded school in England at KS4 and in a mainstream institution for their post-16 qualifications with non-missing KS4 attainment records and at least one completed Level 2 and above post-16 qualification. The coefficients presented in this table are odds ratios. *** indicates that coefficient is statistically significant at the one per cent level, ** indicates that coefficient is statistically significant at the five percent level, * indicates that coefficient is statistically significant at the one per cent level.

3.2 Value-added estimates

Recognising that there are many factors which impact a young person's labour market outcomes both up to 16 (e.g. their family circumstances, where they grow up, etc.), and post-16 (e.g. the institution where they complete any subsequent qualifications, training offered by their employer, etc.), our analysis considers to what extent destination measures can provide information about the 'value-added' by schools and colleges in supporting young people to progress onto high-quality destinations. In other words, by 'value-added', we focus on isolating the impact of the institution where they studied their main 16-18 qualification, over and above their prior attainment and wider contextual factors. This section expands on the analysis and discussion in section 3.3 of the main report.

We estimate institutional value-added using a two-stage fixed-effects model, following Aucejo *et al*. (2020). We use a fixed-effects model rather than a multi-level model, as multi-level modelling relies on the critical assumption that all of our covariates are uncorrelated to our institutional-level effects. As we find that this assumption is violated using a Hausman test, a fixed-effects model is used.

In the first stage, we estimate a linear multivariate model with either earnings or sustained education or employment destination as the outcome variable. As highlighted in section 3.1, the institutions and qualifications young people select into between ages 16-18 depend on factors such as their prior attainment, family background and individual characteristics. In turn, we also include a detailed set of pupil- and qualification-level controls to account for differences in pupil characteristics and qualifications undertaken across institutions. This includes prior attainment of learners, KS4 region of learners and disadvantaged status, alongside the other variables outlined in the previous section. As outlined in section 3.1, we also include controls for the characteristics of a pupil's peers in the institution where a young person studied at KS4, and regressions are run separately by gender. For references, results from our first stage estimation from our main earnings regression are presented in Table 10.

Using the derived residuals, we then estimate a fixed-effects model to construct a value-added measure for each institution. Finally, our fixed-effects are updated using a Bayesian prior following Koedel *et al*. (2015).

All institutions with fewer than 30 learners are grouped by institution type, as there is not sufficient statistical power to analyse the outcomes of these institutions separately. Pupils attending institutions other than a sixth-form college, school or general FE colleges are also grouped by type. We do not include these grouped estimates in our reported findings.

Robustness checks

While we control for individual pupil- and qualification-level characteristics, there is still a concern that our estimates may be affected by the fact that certain types of individuals are more likely to attend a given institution. Following Belfield *et al*. (2018), we use an Inverse Probability Weighted Regression Adjustment (IPWRA) to check the robustness of our findings to weighting by the likelihood of an individual selecting to attend a given type of institution. This method weights young people in each treatment group (provider type) in relation to how similar they are to other young people. We find that the relative rankings of our value-added estimates are robust to the inverse probability weighting, suggesting that the fact that different types of learners select into different institutions (at least based on characteristics that can be observed within our analysis) is not a strong driver of our observed impacts.

We also test the robustness of our findings to using a number of different measures of earnings and employment measures, such as daily and annualised earnings, and to a number of different specifications to ensure that our findings are not being driven by our exact specification.

Further, although we include qualification and subject-level controls in our specification, it is important to recognise that we are only able to control for these at a grouped-level because there are substantial differences in the qualifications that learners study at different institutions. To account for this, we also run robustness checks by estimating institutional value-added on sub-samples of learners which are more comparable.

A summary of our regression findings is provided in Table 5 below. The table shows that our findings are largely stable across specifications, and similar across both female and male learners. These are discussed further in section 3.3.1 of the main report.

It is important to recognise that there are, nonetheless, a number of factors that we are unable to account for in our analysis related to both an individual's outcomes and the institution that they attend. For example, if more motivated pupils are more likely to select into a given institution and achieve better destinations, than that institution's value-added will appear to be larger than it actually is. These factors could be biasing our estimates. While this bias could be in any direction, we would generally expect that omitting these factors would tend to mean that we are over-estimating the amount variation across institutions.

Table 5: Value-added estimates by specification

			Men				Women			
Type of model	**Outcome variable**	**Difference to main model**	**S.D**	**P(25)**	**P(75)**	**N**	**S.D**	**P(25)**	**P(75)**	**N**
Main earnings model	Log of total earnings at age 28 and 29	-	4.4	- 2.9	2.5	291,819	4.4	- 2.7	2.7	290,761
Main employment model	Dummy for whether in sustained employment at age 29	-	1.8	- 1.2	1.2	375,338	1.7	- 1.1	1.2	393,519
Sensitivity checks										
Earnings model	Log of total earnings at age 28 and 29	Excluding controls for qualifications	5.3	- 3.4	3.2	291,819	5.4	- 3.5	3.6	290,761
Earnings model	Log of total earnings at age 28 and 29	Including controls for higher and further education participation	4.2	- 2.7	2.4	291,820	4.3	- 2.7	2.7	290,760
Earnings model	Log of earnings outcomes at age 22 and 23	Outcome variable; self-employment earnings are excluded	4.5	- 3.1	3.0	157,670	4.3	- 3.0	3.0	186,280
Earnings model	Log of earnings outcomes at age 24 and 25	Outcome variable; self-employment earnings are excluded	4.2	- 2.8	2.6	233,554	3.9	- 2.4	2.4	255,219

			Men				Women			
Type of model	**Outcome variable**	**Difference to main model**	**S.D**	**P(25)**	**P(75)**	**N**	**S.D**	**P(25)**	**P(75)**	**N**
Earnings model	Log of earnings outcomes at age 26 and 27	Outcome variable; self-employment earnings are excluded	4.3	- 2.9	2.5	257,578	4.3	- 2.8	2.7	273,833
Earnings model	Log of earnings at age 28 and 29 excluding self-employment earnings	Self-employment earnings are excluded	4.4	- 2.8	2.5	272,426	4.4	- 2.8	2.8	283,503
Earnings model	Daily earnings	Different outcome variable	4.1	- 2.7	2.4	269,881	4.3	- 2.8	2.7	281,495
Earnings model	Annualised earnings	Different outcome variable	4.3	- 2.8	2.4	268,919	4.4	- 2.8	2.8	280,732
Earnings model	Log of total earnings at age 28 and 29	Restricted the sample to A-Level learners only	4.3	- 2.6	2.5	207,333	5.5	- 3.5	3.6	262,697
Employment model	Dummy for whether in employment at least 90 days of the year	Different outcome variable	1.5	- 1.0	1.1	350,498	1.5	- 0.9	1.0	357,086
Higher education model	Dummy for whether identified in sustained higher education at any point between the ages of 18 and 21	Different outcome variable	6.0	- 3.9	3.9	272,425	4.8	- 3.2	3.3	283,505

Source: NFER analysis of LEO data
Note: Analysis includes all individuals in a mainstream state-funded school in England at KS4 and in a mainstream institution for their post-16 qualifications with non-missing KS4 attainment records and at least one completed Level 2 and above post-16 qualification. Sample sizes for specifications including HESA variables have been rounded to the nearest five in order to adhere to statistical disclosure control rules.

Mechanisms

There are a range of mechanisms through which post-16 institutions can help young people to achieve better destinations, as outlined in sections 3.3.2 and 3.3.3 of the main report. These include helping them to attain better grades in their qualifications, assisting them to progress onto higher or further education, and supporting them to progress onto a better quality employer who provides additional support or training.

Our main value-added estimates capture the value associated with studying at a given institution associated with each of these mechanisms. As such, an institution could have higher value-added for a large number of reasons. Indeed, it may be that the institution is better at supporting young people progress onto higher education (which may then improve their labour market prospects) or it could be the institution is supporting young people to progress onto higher quality destinations by providing high-quality careers guidance and support.

Given that many of these mechanisms are of interest in their own right, our analysis also investigates to what extent different mechanisms explain the variation in our value-added estimates across institutions using correlations and multivariate regression techniques to explore the extent to which different mechanisms can explain the variation in value-added across institutions. We focus on earnings rather than employment value-added, given the greater differentiation across institutions in earnings outcomes compared to employment outcomes.

More specifically, we explore the mechanisms associated with institutions having higher value-added in two ways. First, we investigate the correlations between different institution-level variables and our estimates for institutional value-added[5]. A summary of our findings is provided in Table 6 below (see section 3.3.3 of the main report for further discussion).

Table 6: Factors relating to institutions with higher value-added

Variable	School	General FE college
	Significant at 5% level	Significant at 5% level
Attending a grammar school	Positively associated	N/A
Attending a religious school	Positively associated	N/A
Higher average IDACI score of post-16 pupils	Negatively associated	No significant association at the 5% level
Higher KS5 point score	Positively associated	No significant association at the 5% level
Higher share of pupils progressing onto a undergraduate degree	Positively associated	No significant association at the 5% level
Higher share of pupils progressing onto a postgraduate degree	Positively associated	No significant association at the 5% level
Higher share of pupils progressing into Oxbridge universities	Positively associated	No significant association at the 5% level
Higher share of pupils progressing into Russell Group universities	Positively associated	No significant association at the 5% level

[5] For continuous variables, we test correlations using a t-test. For discrete variables, we test correlations using the F-test of a multivariate regression containing only that variable and a constant term.

Variable	School	General FE college
	Significant at 5% level	Significant at 5% level
Higher share of pupils progressing into Million Plus universities	Negatively associated	No significant association at the 5% level
Higher share of pupils progressing into Northern Association universities	Positively associated	No significant association at the 5% level
Higher share of pupils progressing into different university subjects	Some significant associations, but findings are variable	Some significant associations, but findings are variable
Higher share of pupils getting different degree classifications	Positive association between more pupils achieving a 2.1 and higher institutional value-added	No significant association at the 5% level
Higher share of free school meal eligibility	Negatively associated	Negatively associated
Higher average KS4 score of their intake	Positively associated	Negatively associated
Higher share of pupils with English as an additional language	Higher disadvantage is associated with lower value-added	No significant association at the 5% level
Ofsted rating	Positively associated with achieving a better Ofsted rating	No significant association at the 5% level

Source: NFER analysis of LEO data
Note: Analysis includes all individuals in a mainstream state-funded school in England at KS4 and in a mainstream institution for their post-16 qualifications with non-missing KS4 attainment records and at least one completed Level 2 and above post-16 qualification. All estimates based on regression for male learners, albeit results for females are largely similar. Second, we run multivariate regression models to explore how far different institutional level characteristics are able to account for differences in institutional value-added across schools and colleges.

Second, we examine the extent to which different factors can explain variation in value-added across institutions by running multivariate regression models to explore how far different institutional level characteristics are able to account for differences in institutional value-added across schools and colleges. Our findings are presented in Table 7 (see section 3.3.2 of the main report for further discussion).

Table 7: Factors explaining the variation in value-added outcomes

	Male learners		Female learners	
Controls	**Share of variation explained**	**Number of institutions**	**Share of variation explained**	**Number of institutions**
Higher education controls (including subject and degree class)[6]	22%	725	12%	910
Higher education controls (excluding subject and degree class)	9%	1,665	8%	1,720
Progression onto higher level further education qualifications	1%	1,665	1%	1,720
Key Stage 5 attainment score	5%	1,520	6%	1,605

Source: NFER analysis of LEO data
Note: Analysis includes all individuals in a mainstream state-funded school in England at KS4 and in a mainstream institution for their post-16 qualifications with non-missing KS4 attainment records and at least one completed Level 2 and above post-16 qualification. Sample sizes for specifications including HESA variables have been rounded to the nearest five in order to adhere to statistical disclosure control rules. The share of variation explained is based on the adjusted R-squared from a regression which includes the listed controls.

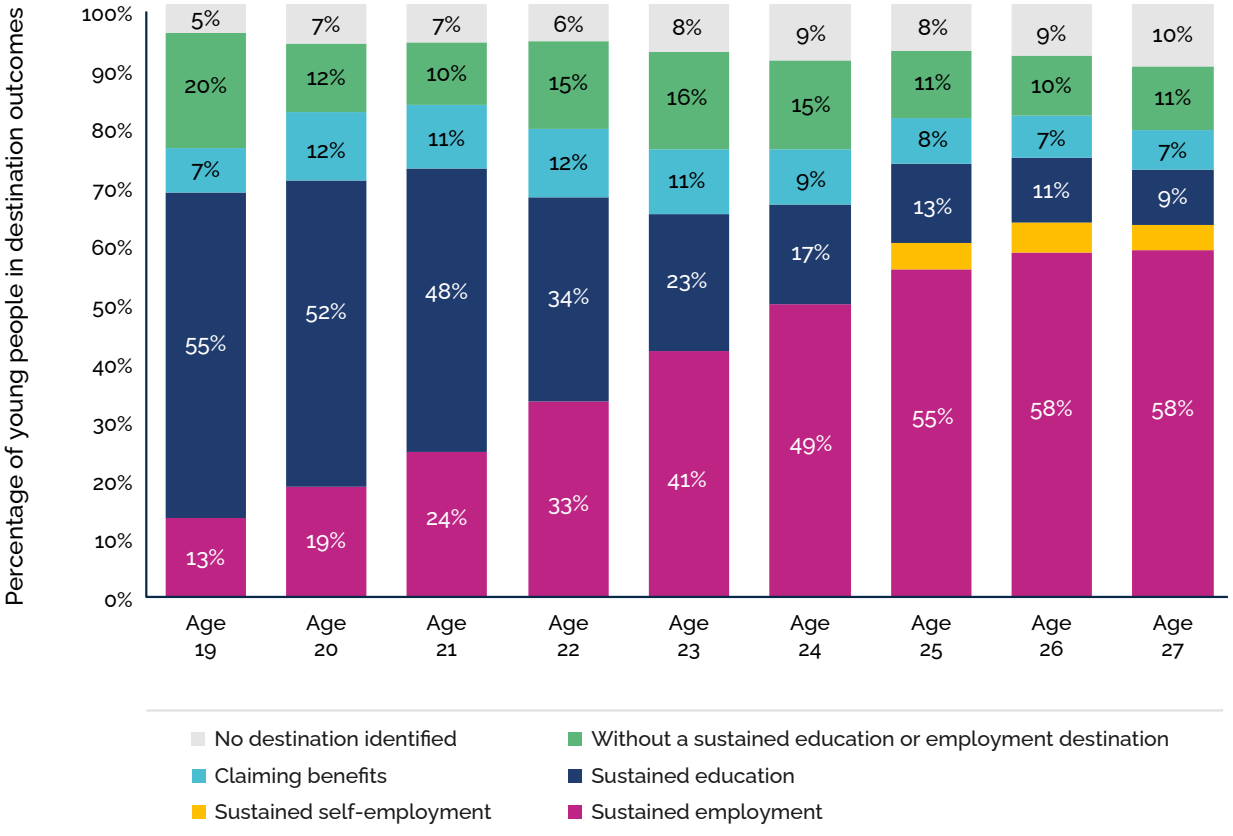
[6] Note that this only includes a sub-sample of institutions where at least 30 learners progressed onto higher education. In turn, it is not a representative sample of institutions.

4 Additional tables

4.1 Comparing destination measures across cohorts

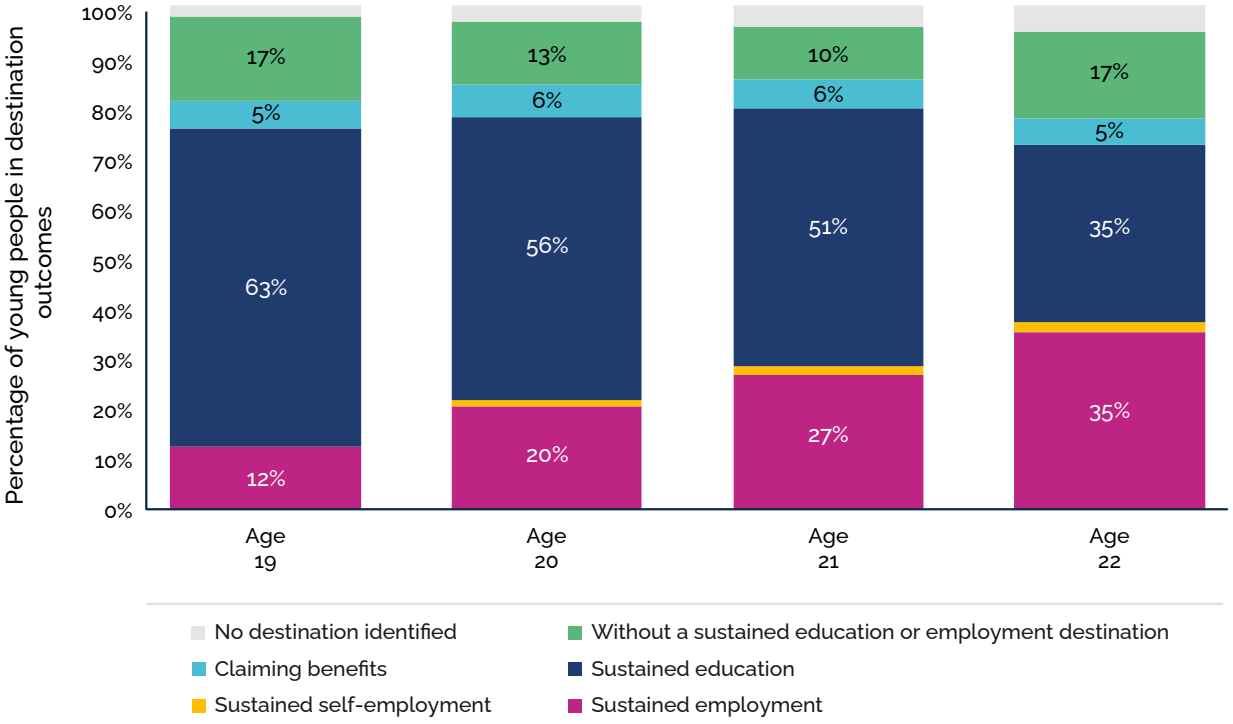
Figure 3 and Figure 4 present the destination outcomes for the 2006/07 and 2011/22 post-16 cohorts respectively. They show that, compared to the 2003/04 cohort (as presented in Figure 2 of the main report), patterns in progression from compulsory education to sustained employment, and sustained further or higher education have remained largely comparable over time. There are, however, also differences across cohorts. The figures show that, since 2003/04, the proportion of pupils progressing onto sustained education destinations has increased.

Figure 3: Destination outcomes up to age 27, 2006/07 post-16 cohort



Source: NFER analysis of LEO data
Note: Analysis includes all individuals in a mainstream state-funded school in England at KS4 with non-missing KS4 attainment records

Figure 4: Destination outcomes up to age 22, 2011/12 post-16 cohort



Source: NFER analysis of LEO data
Note: Analysis includes all individuals in a mainstream state-funded school in England at KS4 with non-missing KS4 attainment records

4.2 By institution type

Table 8 and Table 9 present the full econometric models for our baseline models of the returns associated with attending different types of institutions at post-16.

Table 8: Baseline econometric model of earnings returns by institution type

Dependant variable: Log of average total earning between ages 28 and 29	Male	Female
Variable	**Coefficient (Standard error)**	**Coefficient (Standard error)**
Type of institution (Reference category: School)		
General FE college	-0.0330*** (0.00569)	-0.0626*** (0.00560)
Private Sector Public Funded	0.00535 (0.0113)	-0.0526*** (0.0107)
Sixth form college	-0.0340*** (0.00638)	-0.0185*** (0.00572)

Dependant variable: Log of average total earning between ages 28 and 29	Male	Female
Variable	**Coefficient (Standard error)**	**Coefficient (Standard error)**
Year fixed-effects (Reference category: 2003/04)		
Dummy for whether started post-16 qualification in 2004/05	0.0287*** (0.00220)	0.0145*** (0.00216)
Main qualification studied at post-16 (Reference category: Level 3 - A-Level)		
Level 2 - Apprenticeship	0.0277*** (0.00677)	-0.164*** (0.00741)
Level 2 - Other	-0.00265 (0.00765)	-0.135*** (0.00637)
Level 3 - Apprenticeship	0.147*** (0.0155)	-0.0807*** (0.0132)
Level 3 - Other qualification	-4.71e-05 (0.00521)	-0.0541*** (0.00462)
Region where pupil completed their KS4 qualification (Reference category: London)		
East Midlands	-0.157*** (0.00796)	-0.197*** (0.00826)
East of England	-0.0771*** (0.00852)	-0.120*** (0.00932)
North East	-0.178*** (0.00963)	-0.206*** (0.00925)
North West	-0.195*** (0.00821)	-0.182*** (0.00787)
South East	-0.0883*** (0.00794)	-0.126*** (0.00793)
South West	-0.178*** (0.00869)	-0.212*** (0.00957)
West Midlands	-0.159*** (0.00761)	-0.179*** (0.00759)
Yorkshire and the Humber	-0.180*** (0.00871)	-0.207*** (0.00788)
Ethnicity (Reference category: White)		
Asian (Excluding Indian and Chinese)	-0.0731*** (0.0102)	-0.0457*** (0.00899)
Black (Excluding African)	-0.106*** (0.0100)	-0.0260*** (0.00890)
Chinese	-0.0153 (0.0177)	0.0766*** (0.0158)
Mixed	-0.0416*** (0.00967)	0.00727 (0.00843)

Dependant variable: Log of average total earning between ages 28 and 29	Male	Female
Variable	**Coefficient (Standard error)**	**Coefficient (Standard error)**
Ethnicity (Reference category: White) - *Continued*		
Other non-white ethnicity	-0.0125** (0.00506)	-0.00368 (0.00568)
Indian	0.0481*** (0.00957)	0.0893*** (0.00824)
African	-0.0966*** (0.0127)	-0.00199 (0.0107)
Pupil characteristics		
Dummy for whether pupil was eligible for free school meals at age 16	-0.0399*** (0.00407)	-0.0437*** (0.00414)
Dummy for whether pupil had special educational needs at age 16	-0.0554*** (0.00367)	-0.0212*** (0.00516)
Income Deprivation Affecting Children Index (IDACI) associated with home postcode at KS4	-0.301*** (0.0225)	-0.429*** (0.0214)
Dummy for whether pupil had English as an additional language	-0.0357*** (0.00698)	0.00608 (0.00670)
Dummy for square of the IDACI score	0.337*** (0.0349)	0.485*** (0.0329)
Standardised mean GCSE point score	0.142*** (0.00350)	0.174*** (0.00330)
Dummy for whether pupil achieved at least a C in GCSE English	-0.00509 (0.00342)	0.000136 (0.00425)
Dummy for whether pupil achieved at least a C in GCSE Maths	0.0257*** (0.00364)	0.0517*** (0.00319)
Dummy for whether pupil was in a rural area at KS4	-0.0230*** (0.00422)	-0.0276*** (0.00408)
Cohort characteristics at KS4 institution		
KS4 cohort - Share of pupils in KS4 cohort who were eligible for FSM	-0.0558** (0.0269)	-0.0366 (0.0278)
KS4 cohort - Average IDACI score in pupil's KS4 cohort	-0.0856*** (0.0290)	-0.171*** (0.0307)
KS4 cohort - Share of pupils in KS4 cohort with English as an additional language	-0.0347** (0.0169)	-0.107*** (0.0183)
KS4 cohort - Share of pupils in KS4 cohort attaining five A* to C at GCSE	-0.0452*** (0.0109)	-0.0243** (0.0104)
KS4 cohort - Share of pupils with a SEN statement	-0.00660*** (0.00104)	-0.00953*** (0.00105)

Dependant variable: Log of average total earning between ages 28 and 29	Male	Female
Variable	**Coefficient (Standard error)**	**Coefficient (Standard error)**
Cohort characteristics at KS4 institution - *Continued*		
KS4 cohort - Share of pupils who are white British	9.37e-05 (0.000141)	-0.00121*** (0.000161)
Dummy for subject studied at post-16		
Dummy for whether pupil completed at least one maths or science qualification	0.0652*** (0.00330)	0.0673*** (0.00348)
Dummy for whether pupil completed at least ICT qualification	0.0184*** (0.00342)	0.0343*** (0.00366)
Dummy for whether pupil completed at least one humanities or social science qualification	0.00454 (0.00297)	0.0345*** (0.00292)
Dummy for whether pupil completed a qualification in accounting, finance or business	0.0500*** (0.00373)	0.0869*** (0.00341)
Dummy for whether pupil completed a qualification in an arts subjects	-0.134*** (0.00354)	-0.0327*** (0.00297)
Constant	10.21*** (0.0175)	10.13*** (0.0186)
Cohort characteristics at KS4 institution		
Number of observations	294,925	293,777
R-squared	0.129	0.237

Source: NFER analysis of LEO data
Note: Analysis includes all individuals in a mainstream state-funded school in England at KS4 and in a mainstream institution for their post-16 qualifications with non-missing KS4 attainment records and at least one completed Level 2 and above post-16 qualification. *** indicates that coefficient is statistically significant at the one per cent level, ** indicates that coefficient is statistically significant at the five percent level, * indicates that coefficient is statistically significant at the one per cent level.

Table 9: Baseline econometric model of being in sustained education, employment or self-employment destination by institution type

Dependant variable: Dummy for whether in sustained employment, education or self-employment destination	Male	Female
Variable	**Coefficient (Standard error)**	**Coefficient (Standard error)**
Type of institution (Reference category: School)		
General FE college	0.960** (0.016)	0.960*** (0.015)
Private Sector Public Funded	0.932** (0.031)	0.883*** (0.021)
Sixth form college	0.970* (0.016)	0.981 (0.015)
Year fixed-effects (Reference category: 2003/04)		
Dummy for whether started post-16 qualification in 2004/05	0.827*** (0.010)	0.896*** (0.007)
Main qualification studied at post-16 (Reference category: Level 3 - A-Level)		
Level 2 - Apprenticeship	1.022 (0.027)	0.879*** (0.018)
Level 2 - Other	0.955** (0.021)	0.833*** (0.019)
Level 3 - Apprenticeship	1.243*** (0.049)	1.185*** (0.046)
Level 3 - Other qualification	1.162*** (0.021)	1.139*** (0.018)
Region where pupil completed their KS4 qualification (Reference category: London)		
East Midlands	1.112*** (0.026)	1.039* (0.024)
East of England	1.043* (0.025)	0.963* (0.021)
North East	0.928*** (0.027)	0.981 (0.029)
North West	0.950** (0.022)	1.034 (0.024)
South East	1.019 (0.023)	0.940*** (0.019)
South West	1.033 (0.027)	0.990 (0.024)
West Midlands	1.111*** (0.024)	1.042* (0.023)

Dependant variable: Dummy for whether in sustained employment, education or self-employment destination	Male	Female
Variable	**Coefficient (Standard error)**	**Coefficient (Standard error)**
Region where pupil completed their KS4 qualification (Reference category: London) - *Continued*		
Yorkshire and the Humber	1.001 (0.024)	1.011 (0.023)
Ethnicity (Reference category: White)		
Asian (Excluding Indian and Chinese)	0.891*** (0.027)	0.740*** (0.024)
Black (Excluding African)	0.748*** (0.024)	1.057* (0.032)
Chinese	0.557*** (0.029)	0.657*** (0.043)
Mixed	0.766*** (0.024)	0.922*** (0.028)
Other non-white ethnicity	0.877*** (0.017)	0.950*** (0.019)
Indian	1.053 (0.033)	1.097*** (0.034)
African	0.667*** (0.027)	0.890*** (0.030)
Pupil characteristics		
Dummy for whether pupil was eligible for free school meals at age 16	0.815*** (0.012)	0.766*** (0.011)
Dummy for whether pupil had special educational needs at age 16	0.847*** (0.013)	0.836*** (0.012)
Income Deprivation Affecting Children Index (IDACI) associated with home postcode at KS4	0.995 (0.075)	0.608*** (0.046)
Dummy for whether pupil had English as an additional language	0.792*** (0.018)	0.812*** (0.019)
Dummy for square of the IDACI score	0.959 (0.109)	1.331** (0.152)
Standardised mean GCSE point score	1.199*** (0.015)	1.253*** (0.012)
Dummy for whether pupil achieved at least a C in GCSE English	0.964*** (0.013)	1.115*** (0.014)
Dummy for whether pupil achieved at least a C in GCSE Maths	0.995 (0.013)	1.048*** (0.012)
Dummy for whether pupil was in an urban area at KS4	0.962*** (0.013)	0.965** (0.013)

Dependant variable: Dummy for whether in sustained employment, education or self-employment destination	Male	Female
Variable	**Coefficient (Standard error)**	**Coefficient (Standard error)**
Cohort characteristics at KS4 institution		
KS4 cohort - Share of pupils in KS4 cohort who were eligible for FSM	0.815** (0.076)	0.784*** (0.069)
KS4 cohort - Average IDACI score in pupil's KS4 cohort	0.692*** (0.065)	0.929 (0.091)
KS4 cohort - Share of pupils in KS4 cohort with English as an additional language	1.355*** (0.082)	1.224*** (0.071)
KS4 cohort - Share of pupils in KS4 cohort attaining five A* to C at GCSE	0.576*** (0.021)	0.697*** (0.026)
KS4 cohort - Share of pupils with a SEN statement	1.001 (0.003)	0.993** (0.003)
KS4 cohort - Share of pupils who are white British	1.002*** (0.001)	1.002*** (0.001)
Dummy for subject studied at post-16		
Dummy for whether pupil completed at least one maths or science qualification	1.026** (0.010)	1.044*** (0.011)
Dummy for whether pupil completed at least ICT qualification	1.059*** (0.013)	1.066*** (0.014)
Dummy for whether pupil completed at least one humanities or social science qualification	0.982 (0.011)	1.076*** (0.012)
Dummy for whether pupil completed a qualification in accounting, finance or business	1.081*** (0.012)	1.114*** (0.011)
Dummy for whether pupil completed a qualification in an arts subjects	0.919*** (0.011)	0.959*** (0.009)
Constant	5.281*** (0.348)	4.236*** (0.267)
Number of observations	373,063	391,413

Source: NFER analysis of LEO data
Note: Analysis includes all individuals in a mainstream state-funded school in England at KS4 and in a mainstream institution for their post-16 qualifications with non-missing KS4 attainment records and at least one completed Level 2 and above post-16 qualification. Coefficients are presented as odds ratios. *** indicates that coefficient is statistically significant at the one per cent level, ** indicates that coefficient is statistically significant at the five percent level, * indicates that coefficient is statistically significant at the one per cent level.

4.3 Value-added estimation

The table below presents the first stage regression model for our main earnings specification that is described in section 3.2. The values in brackets refer to standard errors. The residuals from this specification were used to estimate value-added measures at the institution-level.

Table 10: First-stage regression model for main value-added earnings modele

Dependant variable: Log of average total earning between ages 28 and 29	Male	Female
Variable	**Coefficient (Standard error)**	**Coefficient (Standard error)**
Year fixed-effects (Reference category: 2003/04)		
Dummy for whether started post-16 qualification in 2004/05	0.0297*** (0.00242)	0.0137*** (0.00233)
Region where pupil completed their KS4 qualification (Reference category: London)		
East Midlands	-0.156*** (0.00830)	-0.196*** (0.00850)
East of England	-0.0780*** (0.00891)	-0.117*** (0.00930)
North East	-0.175*** (0.00960)	-0.206*** (0.00925)
North West	-0.198*** (0.00880)	-0.182*** (0.00801)
South East	-0.0896*** (0.00806)	-0.125*** (0.00794)
South West	-0.179*** (0.00878)	-0.214*** (0.00931)
West Midlands	-0.160*** (0.00761)	-0.180*** (0.00773)
Yorkshire and the Humber	-0.180*** (0.00970)	-0.206*** (0.00768)
Ethnicity (Reference category: White)		
Asian (Excluding Indian and Chinese)	-0.0752*** (0.0104)	-0.0440*** (0.00902)
Black (Excluding African)	-0.107*** (0.00974)	-0.0306*** (0.00857)
Chinese	-0.0176 (0.0178)	0.0742*** (0.0161)
Mixed	-0.0413*** (0.00889)	0.00596 (0.00871)
Other non-white ethnicity	-0.0127** (0.00508)	-0.00406 (0.00553)

Dependant variable: Log of average total earning between ages 28 and 29	Male	Female
Variable	**Coefficient (Standard error)**	**Coefficient (Standard error)**
Ethnicity (Reference category: White) - *Continued*		
Indian	0.0474*** (0.00937)	0.0933*** (0.00812)
African	-0.0985*** (0.0129)	-0.00463 (0.0112)
Pupil characteristics		
Dummy for whether pupil was eligible for free school meals at age 16	-0.0398*** (0.00506)	-0.0436*** (0.00418)
Dummy for whether pupil had special educational needs at age 16	-0.0552*** (0.00469)	-0.0217*** (0.00549)
Income Deprivation Affecting Children Index (IDACI) associated with home postcode at KS4	-0.303*** (0.0290)	-0.440*** (0.0217)
Dummy for whether pupil had English as an additional language	-0.0349*** (0.00716)	0.00574 (0.00702)
Dummy for square of the IDACI score	0.343*** (0.0433)	0.499*** (0.0335)
Standardised mean GCSE point score	0.0963*** (0.00483)	0.117*** (0.00489)
Dummy for whether pupil achieved at least a C in GCSE English	-0.00747* (0.00385)	0.00103 (0.00508)
Dummy for whether pupil achieved at least a C in GCSE Maths	0.0248*** (0.00293)	0.0521*** (0.00509)
Dummy for whether pupil was in a rural area at KS4	-0.0241*** (0.00401)	-0.0277*** (0.00415)
Cohort characteristics at KS4 institution		
KS4 cohort - Share of pupils in KS4 cohort who were eligible for FSM	-0.0635** (0.0269)	-0.0396 (0.0275)
KS4 cohort - Average IDACI score in pupil's KS4 cohort	-0.0773** (0.0342)	-0.163*** (0.0302)
KS4 cohort - Share of pupils in KS4 cohort with English as an additional language	-0.0407** (0.0166)	-0.114*** (0.0180)
KS4 cohort - Share of pupils in KS4 cohort attaining five A* to C at GCSE	-0.0361*** (0.0120)	-0.0113 (0.0106)
KS4 cohort - Share of pupils with a SEN statement	-0.00694*** (0.00101)	-0.00970*** (0.00114)
KS4 cohort - Share of pupils who are white British	1.21e-05 (0.000144)	-0.00129*** (0.000159)

Dependant variable: Log of average total earning between ages 28 and 29	Male	Female
Variable	**Coefficient (Standard error)**	**Coefficient (Standard error)**
Main qualification studied at post-16 (Reference category: Level 3 - A-Level)		
Level 2 - Apprenticeship	0.0163*** (0.00506)	-0.190*** (0.00569)
Level 2 - Other	-0.0357*** (0.0120)	-0.166*** (0.00652)
Level 3 - Apprenticeship	0.145*** (0.00511)	-0.106*** (0.00854)
Level 3 - Other qualification	-0.0257*** (0.00735)	-0.0759*** (0.00463)
Dummy for subject studied at post-16		
Dummy for whether pupil completed at least one maths or science qualification	0.0603*** (0.00307)	0.0714*** (0.00560)
Dummy for whether pupil completed at least ICT qualification	0.0151*** (0.00368)	0.0366*** (0.00372)
Dummy for whether pupil completed at least one humanities or social science qualification	0.00570* (0.00306)	0.0348*** (0.00291)
Dummy for whether pupil completed at least one language or literature qualification	-0.0374*** (0.00303)	0.0269*** (0.00317)
Dummy for whether pupil completed a qualification in accounting, finance or business	0.0485*** (0.00666)	0.0865*** (0.00657)
Dummy for whether pupil completed a qualification in an arts subjects	-0.134*** (0.00375)	-0.0328*** (0.00320)
Constant	9.782*** (0.0228)	9.523*** (0.0224)
Number of observations	291,819	290,761
R-squared	0.130	0.237

Source: NFER analysis of LEO data
Note: Analysis includes all individuals in a mainstream state-funded school in England at KS4 and in a mainstream institution for their post-16 qualifications with non-missing KS4 attainment records and at least one completed Level 2 and above post-16 qualification. *** indicates that coefficient is statistically significant at the one per cent level, ** indicates that coefficient is statistically significant at the five percent level, * indicates that coefficient is statistically significant at the one per cent level.

References

Anderson, O. and Nelson, M. (2021). *Education and Labour Market Pathways of Individuals (LEO) Technical report:* [online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/993969/Technical_Report_for_Education_and_Labour_Market_Pathways_of_Individuals__LEO_.pdf [24 January 2022].

Aucejo, E., Hupkau, C. and Ruiz-Valenzuela, J. (2020). *Where Versus What: College Value-Added and Returns to Field of Study in Further Education. Research Discussion Paper 030* [online]. Available: https://cver.lse.ac.uk/textonly/cver/pubs/cverdp030.pdf [24 January 2022].

Battiston, A., Patrignani, P., Hedges, S. and Conlon, G. (2019). *Labour Market Outcomes Disaggregated by Subject Area Using the Longitudinal Education Outcomes (LEO) data. CVER Discussion Paper 021* [online]. Available: https://cver.lse.ac.uk/textonly/cver/pubs/cverdp021.pdf [24 January 2022].

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 Relative Labour Market Returns to Different Degrees: Research Report: June 2018* [online]. Available: https://ifs.org.uk/publications/13036 [24 January 2022].

Crawford, C., Meschi, E. and Vignoles, A. (2011). *Post-16 Educational Choices and Institutional Value Added at Key Stage 5. CEE DP 124* [online]. Available: https://files.eric.ed.gov/fulltext/ED529851.pdf [24 January 2022].

Edge Foundation (2018). *Towards a Twenty-First Century Education System: Edge Future Learning* [online]. Available: https://www.edge.co.uk/documents/12/edge_future_learning_report_final.pdf [24 January 2022].

Great Britain. Parliament. House of Commons. Briefing Paper (2018). *Women and the Economy*. (CBP06838) [online]. Available: https://researchbriefings.files.parliament.uk/documents/SN06838/SN06838.pdf [24 January 2022].

Koedel, C., Mihaly, K. and Rockoff, J. E. (2015). 'Value-added modelling: A review', *Economics of Education Review*, 47, 180-195. [online]. DOI 10.1016/j.econedurev.2015.01.006.

Organisation for Economic Cooperation and Development (OECD). (2018). *The Future of Education and Skills: Education 2030* [online]. Available: https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf [24 January 2022].

Patrignani, P., Conlon, G. and Hedges, S. (2017). *The Earnings Differentials Associated with Vocational Education and Training using the Longitudinal Education Outcomes Data. Research Discussion Paper 007* [online]. Available: https://cver.lse.ac.uk/textonly/cver/pubs/cverdp007.pdf [24 January 2022].

The Edge Foundation
44 Whitfield Street
London, W1T 2RH

T +44 (0)20 7960 1540
E enquiry@edge.co.uk

www.edge.co.uk