IFS

Institute for Fiscal Studies

Jack Britton
Laura van der Erve
Chris Belfield
Anna Vignoles
Matt Dickson
Yu Zhu
Ian Walker
Lorraine Dearden
Luke Sibieta
Franz Buscha

Working paper 21/24

How much does degree choice matter?

Economic and Social Research Council

How much does degree choice matter?[*]

Jack Britton *(Institute for Fiscal Studies)*[†]
Laura van der Erve *(IFS & University College London)*,
Chris Belfield *(IFS)*,
Anna Vignoles *(University of Cambridge)*,
Matt Dickson *(University of Bath)*,
Yu Zhu, *(University of Dundee)*,
Ian Walker *(Lancaster University)*,
Lorraine Dearden *(University College London)*,
Luke Sibieta *(IFS)*,
Franz Buscha *(University of Westminster)*

Abstract

Using a large and novel administrative dataset, this paper investigates variation in returns to different higher education 'degrees' (subject-institution combinations) in the United Kingdom. Conditioning on a rich set background characteristics, it finds substantial variation in returns, even within subject, across universities with very similar selectivity levels, suggesting degree choices matter a lot for later-life earnings. Selectivity is weakly related to returns through most of the distribution but strongly positively correlated at the top end. Other than selectivity, returns are poorly correlated with observable degree characteristics, which has implications for student choices and the incentives of universities.

Keywords: Returns to education, Degree choice
JEL: I23, I24, I26

[*]The authors are grateful for the Department for Education for data access and a huge amount of help and support in developing this paper. We thank Elaine Drayton and Ben Waltmann for research support and Richard Blundell, Pedro Carneiro, Damon Clark and Eric French for detailed comments, as well as seminar participants at the IZA Economics of Education Workshop, Royal Holloway, Institute of Education, UC Irvine and Lancaster University. We also thank the Economic and Social Research Council for funding through grant ES/S010718/1 and the Centre for Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies (RES-544-28-50001) and NORFACE for funding through grant HuCIAW 462-16-090. All errors are our own.

[†]Corresponding author, Email: `jack.britton@ifs.org.uk`.Institute for Fiscal Studies, 7 Ridgmount St., London, UK

1 Introduction

As in many countries around the world, prospective higher education students in the United Kingdom (UK) choose between a vast number of different degree options when entering university. This paper exploits a pioneering new administrative dataset to look at labour market outcomes at the degree level - that is, the interaction of subject field and institution. We explore the variation in earnings returns and investigate the predictability of those returns based on other observable characteristics of the degree. To our knowledge, our paper is the first to estimate returns for individual degrees across an entire higher education market.

We find substantial variation in returns, even within relatively tight selectivity bands and within subject. This implies degree choice matters much more than some of the previous evidence has suggested. We find only a weak relationship between selectivity and returns through much of the distribution, but a strong overall positive relationship at the top end of the selectivity distribution, suggesting that there is a large payoff to high ability students attending elite universities. However, this is not true for all subject areas - for some, such as creative arts, there is only a very weak relationship between selectivity and returns throughout the distribution. Finally, aside from selectivity, we find that existing measures of degree quality are not well related to returns. This matters because these measures influence both student choices and university behaviour.

We exploit a new administrative data linkage that was developed in partnership with the UK Department for Education. The dataset links together administrative school, university and tax records for the more than three million individuals who completed secondary school in England between 2002 and 2007. The tax records include annual earnings from 2005/06 to 2016/17, meaning we observe the oldest cohort in the data up until age 30. The school records allow us to condition on an extremely detailed set of prior attainment controls that include exam grades in specific subjects at ages 11, 16 and 18, as well as rich information on student background and secondary (high) school fixed effects. Unlike some of the recent papers in this literature, the dataset tracks all of students through all of the available higher education institutions in the country, and captures anyone who is filing for taxes anywhere in the country.

Our data contains more detailed background information on students than many previous papers have been able to use. We exploit this to test the likely role of unobservable factors in driving our results. We show that our headline findings are robust to the exclusion of subsets of our control variables, suggesting that unobservable factors are not likely to affect our main conclusions. We also show that the main findings are not sensitive to reasonable changes in the sample selected or the regression specification we choose.

We start by estimating overall returns to higher education, before looking at how returns vary by institution and subject, within the set of people who go. We find fairly low returns in overall returns for men, but much higher returns for women. However, when we look within the set of those who go to higher education, gender differences are less important. Across institutions, we find a weak association between selectivity and returns through much of the selectivity distribution, but a much stronger relationship at the top end of the distribution, suggesting large payoffs to attending the most elite universities in the UK, in particular the Universities of Oxford and Cambridge, the London School of Economics, and Imperial College London. We estimate big differences by subject, with medicine, economics and law doing particularly well and social care and creative arts courses doing poorly. In general, these findings are consistent with the previous literature, which is reassuring for our degree-level estimates.

We then turn to the most novel contribution of the paper by estimating returns at the 'degree' level, which is the interaction of institution and subject. We are able to estimate returns for almost 2000 subject-university combinations (for example, mathematics at the University of Warwick). This is a natural level of granularity to focus on for the UK, where people choose specific subject-university combinations for their degrees prior to starting, and is only viable because of the unique dataset at our disposal. There is substantial variation in raw earnings outcomes across different degrees: the standard deviation of the degree-level fixed effects, without any controls, is 32 percentage points (ppts) and the 90:10 range is 75 ppts. These figures drop to 22 ppts and 52 ppts respectively once we include the full set of controls for prior attainment, student characteristics, and secondary school fixed effects.

There is still substantial variation in returns, even when looking within relatively tight selectivity bands. Amongst the least selective degrees, the standard deviation in returns is still more than 15 ppts, increasing to 29 ppts amongst the most selective set of degrees. It is also the case that a large share of the variation in returns is *within* subject, even within our selectivity bands. Roughly 50% of the variation in degree returns for the least selective band of degrees is within subject, rising to more than 70% of the variation for the most selective degrees. Combined, these results strongly suggest that degree choice is crucial for subsequent earnings outcomes, right across the selectivity distribution, even holding subject choice fixed. For example, it is not at all uncommon to see differences in returns of 40 ppts between degrees in the same subject at similarly selective universities.

Given the importance of degree choice in determining earnings outcomes, in the final part of the paper we consider the predictability of returns across different institutions, within subject. We find that existing measures of degree quality are not well correlated with returns. As with the institution estimates, on average there is only a weak relationship between degree selectivity and returns through much of the distribution but a much stronger relationship at the top end. However, this varies a lot by subject area: for economics, law and business, returns increase rapidly with university selectivity, while for others, such as sociology and the creative arts, they do not.[1] We then see that other measures of degree quality including publicly available subject-specific university rankings, completion rates and degree performance are all correlated with returns, but this almost completely disappears once we control for selectivity. This suggests that observable measures of degree performance contain little information over and above a simple measure of selectivity. Student satisfaction ratings and early career earnings are not well correlated with returns, even unconditionally.

These observable degree characteristics matter. For example, Gibbons et al. (2015) shows that public league table rankings are a key driver of student choices, while many of the other measures we look at (such as very early career earnings and student survey scores) are used as inputs for centralised evaluation of teaching quality in the UK, through the 'Teaching Excellence Framework'. The result that public information on degrees is not well correlated with the earnings outcomes of students has several important implications. First, it will matter for productivity if students select degrees that are not highly valued in the labour market. Second, it will affect in-

[1] We also see no relationship between returns and selectivity for medicine and education, which is not surprising, as so many graduates from these subjects go into careers with centrally regulated wages.

equality, as students from more disadvantaged backgrounds are more likely to have to rely on public information when making their higher education choices. Indeed, Campbell et al. (2019) highlight that poorer students are more likely to choose degrees associated with lower earnings outcomes, conditional on prior attainment. Third, it is likely to incentivise universities to focus on metrics that may not be beneficial to the long-term outcomes of students, as doing well on those metrics helps them to achieve good scores in teaching evaluations and to attract students.

This rest of this paper is set out as follows. Section 2 reviews the related literature and discusses how our paper fits into it. Section 3 then describes the dataset we use and gives more detail on the institutional background in the UK and Section 4 outlines our methodology. Our results are then presented in Sections 5 and 6. Section 5 provides estimates of the overall earnings returns of attending university versus not attending, and looks at heterogeneity in returns across institutions and subjects. Section 6 then focusses on the degree level returns estimates and shows the relationship between degree level returns and selectivity, as well as with other observable characteristics of the degree. Finally, Section 7 concludes.

2 Literature

Our work draws upon and contributes to a substantial academic literature which investigates returns to higher education. This literature can be divided into three main branches. The first investigates the overall returns to higher education using selection-on-observables, finding that returns are high on average (Webber, 2016; Walker and Zhu, 2011; Blundell et al., 2000). Our findings also suggest good overall returns to university and the magnitudes of the estimates are consistent with the previous UK literature based on estimates before the rapid expansion of higher education that occurred in the UK between during the 1980s and 1990s, suggesting the returns held up well through the expansion. This finding is also consistent with descriptive evidence from Blundell et al. (2016).

The second strand investigates heterogeneity in returns by university attended. Most of this literature looks at heterogeneity across broad groups of institutions (Chevalier and Conlon, 2003; Andrews et al., 2017; Walker and Zhu, 2018) or at the relationship between returns and a continuous measure of university quality or selectivity (Hussain et al., 2009; Broecke, 2012; Black and

Smith, 2006; Dale and Krueger, 2002, 2014; Dillon and Smith, 2020). However, more recent papers have started to investigate heterogeneity in returns across individual institutions (Cunha and Miller, 2014; Mountjoy and Hickman, 2020; Chetty et al., 2020). Many of these papers, like ours, identify returns based on OLS estimation with rich background characteristics. Some of them attempt to address selection issues by controlling for the set of colleges students applied to or were accepted at. While we do not observe application sets, our data contains much more detailed background information on students than previous work has been able to use. In particular, we observe full and detailed academic histories of each student, including specific grades in specific subjects based on national tests taken at ages 11, 16 and 18, alongside rich background characteristics allowing us to control for the local area in which people grow up as well as for the school they attended. Hastings et al. (2013) and Hastings et al. (2018) instead exploit discontinuities in university entry cutoffs to identify returns to different institutions find their results to be consistent with those obtained using OLS conditioning on rich observables, without controlling for application sets. Drawing on evidence from their own experimental work as well as that of Wiswall and Zafar (2014) they argue that students do not know much about earnings outcomes and select their university largely based on factors that are unlikely to be correlated with later outcomes. Dillon and Smith (2020) make a similar argument in a recent paper that focuses on match effects in higher education. These papers further strengthen our confidence in our identification strategy.

The evidence from this literature on the relationship between returns and selectivity is mixed. The UK evidence consistently finds a strong relationship between university selectivity and returns (for example, Walker and Zhu, 2018), as do many of the previous papers from the United States, but Dale and Krueger (2002, 2014) and Mountjoy and Hickman (2020), which all control for the application sets of students, all suggest a very weak relationship. However, the work by Cunha and Miller (2014), which exploits very similar data and uses a similar approach to Mountjoy and Hickman (2020) find a strong relationship for universities in Texas, while several papers which have exploited discontinuities in university entry cutoffs to identify returns to specific universities have also found large effects (Anelli, 2020; Hoekstra, 2009; Hastings et al., 2013; Saavedra, 2008; Zimmerman, 2019).[2] We find a weak association between selectivity and returns through-

[2]Our paper also relates more tenuously to papers that have estimated effects for students at the margin of going and

out much of the selectivity distribution, but this becomes much stronger at the top end of the distribution. This suggests very large payoffs to attending the most elite universities in the UK.[3]

The third strand of related literature investigates heterogeneity in returns by subject studied. Altonji et al. (2012) reviews the evidence to that date, highlighting that the majority of papers estimating returns assume selection on observables (Walker and Zhu, 2011, 2018; Blundell et al., 2000; Chevalier, 2011). However, again there are some papers which have exploited discontinuities in entry cutoffs to identify returns to different subject choices (Kirkeboen et al., 2016; Hastings et al., 2013). Kirkeboen et al. (2016) presents a compelling case that there is a large amount of selection into different subjects based on comparative advantage, suggesting returns based on OLS regressions would overstate the causal effects. This finding leads us to be cautious about our cross-subject returns estimates, although we note that our data on subject-specific prior attainment is a considerable improvement on much of the literature. With this caution in mind, we find that economics and medicine are the highest returning subjects, with conditional returns of more than 30% relative to history (our base case). Computing, business and architecture also do well, with returns of 15-20% above the base. At the other end of the scale, social care, creative arts, agriculture and veterinary sciences are the lowest-returning subjects with returns of 10-15% below history (social care is the lowest). Psychology, English, languages and biological sciences also perform poorly (notably, many of these subjects are much more likely to be chosen by women). Our results are more mixed for STEM degrees, as we find a lot of variation in returns within this broad subject area, which is an important result given the context of large pro-STEM agendas in several countries.

As described above, we believe our main contribution to the literature is to investigate returns at the degree level. The only previous paper that has had sufficiently high quality data do this

not going to university. For example, Zimmerman (2014) finds very large earnings returns for academically marginal students. While our estimates are in relative terms and are therefore not compared to the outside option of not going, we still see that the average returns for the least selective institutions are no lower than returns for middling universities, suggesting that our results are consistent with the idea that returns are reasonably good for universities that accept students with low prior attainment who are therefore likely to be close to the margin of going and not going.

[3]The discrepancy between this finding and that of Mountjoy and Hickman (2020) could be explained by the fact that they do not observe anything like the range in university quality that we do. They look at 27 four-year colleges in Texas, where the top institution is University of Texas, Austin. This is a considerably less selective, and less elite institution than the top UK universities. We also note that they suffer from out-of-state selection problems (both for university and for work), while this is dramatically less important in the UK, where very few students study abroad or work abroad after graduation.

is Hastings et al. (2013), which is able to exploit discontinuities in entry cutoffs to around 1,100 different degree programmes in Chile in order to identify returns. However, the focus of that paper is not on individual degree returns, but rather the relationship between returns and university selectivity, the returns by subject, and the returns by subject interacted with a binary indicator for high selectivity. Like our paper, selectivity is found to be strongly related to subject returns for some subjects but not for others. To our knowledge, our paper is the first to estimate individual returns for individual university degrees across the whole of a higher education system. This exercise is extremely revealing about the extent to which degree choices can potentially impact later-life outcomes.[4]

Looking at individual degree level returns also enables us to look more carefully at the relationship between observable degree-level characteristics and returns than any previous paper.[5] This enables us to consider the relevance of the *ex-ante* information on degree quality available for students and regulators. Our finding that the measures of degree quality that we consider are unrelated to returns is highly pertinent as this information influences student's degree choices, regulator's ratings of teaching quality, and also the priorities of universities.

3 Data

We use the Longitudinal Educational Outcomes (LEO) dataset, which was developed in collaboration with the UK Department for Education for the purposes of this paper. In this section, we define our analysis sample, give more information about each of the composite datasets of LEO, and show summary statistics of our analysis sample.

[4]Our institutional setting is also quite different to that of Hastings et al. (2013). The UK has a much larger higher education sector than Chile (OECD, 2014), with a much broader range of institutions, including many that cater to students with relatively low prior attainment as well as several internationally renowned institutions that regularly feature in the top ten of world university rankings. Our findings are therefore likely to be more relevant to higher education systems of countries with more advanced economies such as the US, Australia and several European countries.

[5]Mountjoy and Hickman (2020) and Chetty et al. (2017) look at this, for example. But their comparisons are quite limited as they are only able to look at overall university characteristics of the university, rather than characteristics at the subject-institution level.

3.1 Sample

Our base sample of students consists of all individuals who: (1) attended school in England; (2) passed their age 16 exams between 2002 and 2007,[6]; (3) are linked to UK tax records for any of the tax years 2013-14 to 2016-17; and (4) started an undergraduate degree in the UK between the ages of 17 and 21 as a full-time student. This gives us between 161,000 and 204,000 individuals in each cohort (as defined by the year they took their age 16 exams), giving a total of over one million individuals.

When estimating the overall returns to a degree, we will compare these individuals to a control group of individuals who satisfy criteria (1) to (3) above, but did not attend university at any point in our dataset (we drop part time and mature students from the analysis completely). We identify individuals in each group by taking all individuals who appear in administrative records of age 16 exams, and linking them to administrative tax and university records. More information on match rates and sample selection is provided in Appendix A1.

3.2 Demographics and school attainment

We obtain information on background characteristics and school attainment of individuals from the National Pupil Database (NPD), which contains exams files as well as a census of English schools.

In England, students take national, externally marked examinations at age 11, 16 and 18, and we have all three records in our data. The age 11 tests, taken at the end of primary school, are the Year 6 Standard Assessment Tests (SATs). They are taken in three subjects - English, mathematics and science - and we have detailed scores from each. The age 16 tests are based on 'General Certificate of Secondary Education' (GCSE) exams, the majority of which are taken in the summer of the school year people turn 16 (Year 11).[7] GCSEs during this period were taken in English (literature and language), mathematics and science plus typically five to seven additional subjects and were graded from A*-G. A grade C was generally considered to be a pass - indeed, a key metric for progression onto further education or training was often whether an individual had at

[6] We define "passing" age 16 exams as obtaining at least 5 A*-C grades in GCSE exams - see below for more detail on these exams.

[7] The school year in England runs from September 1 to August 31.

least five GCSEs graded between A* and C. We observe all of the subjects taken and the grades achieved in each. For presentational purposes, the GCSE exam grades are converted into a single points index (where an A* is worth 58 points, an A is worth 52 points and so on down to the lowest scored grade of G, which is worth 16 points). The age 18 assessment data are based primarily on scores in A-level exams, which are usually taken two years after GCSEs (Year 13). For A-levels, students take exams in the typically three or four subjects they chose to study after GCSEs. A-levels were graded from A-F during this period, with a D grade often considered to be the minimum pass. Again, we observe the subjects taken and the grades achieved. Students during this period can take equivalent vocational qualifications instead of (or as well as) A-levels, such as courses in retail or hospitality, which we also observe.

The school census contains school identifiers and student level demographics, including gender, age, ethnicity, special educational needs and an indicator for English not being the student's first language. We further observe whether a student is eligible for Free School Meals (FSM) and have access to detailed measures of deprivation in the small local area (approximately 130 households) where the child lives at age 16.[8] Following several previous papers (e.g. Chowdry et al., 2013), we combine these multiple measures into one continuous index of socio economic status (SES) at age 16 using principal components analysis. The approximately 7% of pupils who attend private secondary schools are missing the school census data (but we do observe their exam records).[9] We keep this group in the analysis and include missing dummies for any missing school census information.

The earliest cohort for whom we have individual level school records are the students who took their age 16 exams in 2002. The vast majority of these individuals were born between 1 September 1985 and 31st August 1986.

[8]In order for a pupil to be eligible for free school meals their family has to be on means-tested benefits, FSM eligible pupils therefore approximately represent pupils from the poorest 15% of families. Local area level deprivation measures include the proportion of individuals in the pupil's local area of residence with a degree, with no qualifications, in managerial and professional jobs, in routine occupations, long-term unemployed, homeowners, in social housing as well as the proportion of children living in income deprived households (IDACI). All these measures are included at the Output Area level (containing 130 households on average), except IDACI and the proportion of individuals living in social housing, which are both measured at the Lower Super Output Area (around 670 households on average).

[9]Some students who also attended a private primary school have no age 11 exam records, but these students do all have age 16 and age 18 exam records.

3.3 University attendance

We obtain information on higher education attendance from the Higher Education Statistics Authority (HESA) data. For each year an individual attends a university in the UK this administrative dataset records the type of degree, subject studied, university attended, course intensity (part-time vs full-time) and degree performance. We link individuals over time to determine whether they graduate from their degree.

Students who apply to university typically do so in the the academic year they take their A-level (or equivalent) exams. About half of students who go to university do so within a few months after their A-level (or equivalent) exams, while another 30-40% go within the next two years. We focus on university entrants within this three-year window, meaning that the majority of the HESA records we use are from the 2004/05 - 2009/10 academic years. People who we observe going to university after this window are dropped from the analysis. We observe HESA data up until 2015/16, which allows us to remove mature students starting university up until the year they turn 29.

The most common route through university is to attend one institution for an undergraduate degree and to study one subject (although several students study joint degrees with more than one subject). Full-time degrees are usually three years, though some degrees such as languages or sciences are four year degrees. In the HESA records we observe subject, university and course intensity (part-time vs full time), and we are able to link people over time to determine whether they graduate.[10]

Degree subjects are recorded in meticulous detail, with more than 1,500 different subject categories provided. We aggregate these up to around 30 broad subject areas (for example mechanical engineering and civil engineering are aggregated to engineering) based on the official 'Common Aggregation Hierarchy'.[11] To summarise our findings, we sometimes further group these subjects in three groups: LEM (Law, Economics and Management[12]), STEM (Science, Technology, Engineering and Mathematics), and Other, which consists of other social sciences, arts and humanities

[10]For people who did not graduate from their first degree and switched to a second degree, we take their second degree as their undergraduate qualification, so long as it was taken as a full-time, non-mature student.

[11]For a complete list of these, see https://www.hesa.ac.uk/innovation/hecos.

[12]This is common terminology - in practice for our subject classifications this is law, economics and business.

subjects. A complete list of the subjects in each group is provided in the Appendix.

Individuals attend one of more than 100 UK universities which provide undergraduate degrees. For some analysis we classify universities into five broader groups based on prestige and selectivity. The four most selective and prestigious universities in the UK (the University of Oxford, the University of Cambridge, Imperial College London and the London School of Economics) are put together into the 'Elite Russell Group'. These four universities have notably higher prior attainment than any other universities in the country.[13] The next most selective group of universities are the 'Russell Group' which is a well-known self-defined collective of 24 (including the aforementioned four) high-status institutions. This group is followed by the 'Old universities' which includes the remaining 31 institutions which predate the large expansion of universities that occurred in England in 1992. The remaining universities are non-traditional universities, such as art colleges, or are former technical colleges which converted to university status in 1992. This group of around 80 typically less selective institutions is divided into two equal groups ('more selective' and 'less selective') based on the average GCSE points scores of their students. A complete list of the universities in each group is provided in an Online Appendix.

3.4 Earnings

Individuals' earnings are obtained from Her Majesty's Revenue and Customs (HMRC) tax records. Earnings from conventional employment are recorded in Pay As You Earn (PAYE) records, which we have for the 2005/06 - 2016/17 tax years.[14] Earnings from self employment and profits from partnerships are recorded separately in Self Assessment (SA) records. We only have these latter records from 2013/14 - 2016/17. To avoid missing a substantial fraction of total earnings,[15] we only make use of the data from 2013/14 onwards. This has the additional advantage of avoiding the immediate labour market fallout from the 2008 recession. The tax data only includes information only on total annual earnings, and we observe no measures of hours worked.

Tax records have been matched to university and school records by the UK Department for Work and Pensions. They employed fuzzy matching using National Insurance Number,[16] first

[13] This can be seen in Figure A4 in the Appendix which show the average GCSE score of universities' student intake.
[14] In the UK tax years run from April 6th to April 5th of the following year.
[15] By age 30, around 10% of individuals in our sample have self-employment income.
[16] Equivalent to the US Social Security Number.

name, surname, date of birth, postcode and gender. The first cohort for whom this link exists are those who took their age 16 exams in 2002, who were born between 1st September 1985 and 31st August 1986.[17] These individuals will be approximately aged 30 in the last tax year for which we have earnings records (2016/17).

Due to concerns about early career earnings not being representative of later life earnings, we only include earnings from individuals aged 25 or older. As our complete earnings records run from 2013/14 to 2016/17, the age restrictions mean our analysis will include individuals born between 1st September 1985 and 31st August 1991.

3.5 Data descriptives

In Table 1 we show some background characteristics, demographics and prior attainment of individuals who passed their GCSEs split by whether or not they studied for an undergraduate degree. Undergraduates are more likely to have attended a private secondary school, and are more likely to come from higher socio-economic backgrounds. They are also more likely to be non-white, reflecting the higher participation rates of Asian students in particular. As expected, they also have higher prior attainment, being much more likely to have achieved high grades both in age 11 and age 16 exams.

Table 2 summarises our undergraduate sample by the different university and subject groups. More women than men attend university, but slightly more men than women attend the most selective universities. We see around 20,000 men attending one of the four Elite Russell group universities compared to around 17,000 women.[18] Between 100,000 and 170,000 individuals of each gender attend each of the four other university groups, with women outnumbering men in each of these groups. There are also large gender differences in the broad subject areas studied, with more than half of women studying arts, humanities and other social science degrees (labelled 'Other') compared to just 40% of men. Men are more likely to do both LEM (21% vs 16%) and

[17] In practice, individuals who sat their age 16 exams in 2002 but skipped a year in school or were held back a year might be born before or after this. Skipping a year, or being held back a year is however a very rare occurrence in England.

[18] One reason for this is that two the four Elite Russell Group universities (Imperial College and LSE) specialise in only a subset of subject areas that are more commonly chosen by men. However, it is still the case that women attend Russell Group or Elite Russell Group universities at a lower rate than men do.

STEM degrees (39% vs 30%).[19] A comparison of the earnings distribution of graduates and non-graduates, and earnings across subjects and institution groups is shown in Appendix A3.

Table 1: Background characteristics by attainment group

	Women		Men	
	No UG	UG	No UG	UG
Background				
State school	0.95	0.85	0.94	0.83
of which:				
SES Q1 (richest)	0.23	0.33	0.25	0.35
SES Q2	0.25	0.25	0.26	0.25
SES Q3	0.22	0.20	0.22	0.19
SES Q4	0.18	0.13	0.16	0.12
SES Q5 (poorest)	0.13	0.09	0.11	0.08
FSM	0.07	0.05	0.06	0.04
EAL	0.04	0.10	0.04	0.10
SEN	0.03	0.02	0.05	0.03
Ethnicity				
White	0.91	0.80	0.91	0.80
Black	0.01	0.04	0.01	0.03
Asian	0.03	0.09	0.02	0.09
Other	0.06	0.08	0.06	0.08
Attainment				
Age 11 Maths level 5+	0.23	0.40	0.34	0.50
Age 11 English level 5+	0.33	0.52	0.24	0.41
Age 16 Maths A/A*	0.07	0.34	0.09	0.40
Age 16 English A/A*	0.14	0.49	0.08	0.38
N	329,079	602,169	320,506	500,086

Note: UG indicates the individual is treated as an undergraduate in our sample. The No UG group excludes people who did not get five A*-C grades in their GCSE exams. We pool here pooled across the six GCSE cohorts. EAL = English as an additional language, FSM = free school meals, SEN = non-statemented special educational needs. Most of the shares here are based on the state school sample only, except the state-educated share, the age 16 (GCSE) results and some of the age 11 (SAT) exam results (as described in the data section above). The attainment section shows the share of individuals who obtained at least level 5 in their age 11 exams, and the share who obtained an A or A* (the two highest grades) in their maths or English age 16 exams. N is based on the full sample including the independently educated.

[19] For the full list of subjects in each of these categories, see Table A3 in the Appendix.

Table 2: Number of students by university and subject groups

	Women		Men	
	N	Share	N	Share
University group				
Elite Russell	16,965	0.03	20,362	0.04
Russell Group	158,549	0.26	138,453	0.28
Old universities	106,063	0.18	100,149	0.20
Other (more selective)	162,466	0.27	127,865	0.26
Other (less selective)	156,376	0.26	112,363	0.23
Total	600,419	1.00	499,192	1.00
Subject group				
LEM	96,526	0.16	103,372	0.21
STEM	182,378	0.30	194,828	0.39
Other	323,265	0.54	201,886	0.40
Total	602,169	1.00	500,086	1.00

Note: includes individuals in the undergraduate group, pooling across six GCSE cohorts. A very small number of graduates are missing information on the university attended, hence the slightly lower sample size in the top panel. LEM indicates 'Law, Economics and Management' and STEM indicates 'Science, Technology, Engineering and Mathematics'.

4 Earnings model

Our identification strategy relies on selection on observable characteristics. The basic premise follows much of the returns to education literature (for example, Blundell et al., 2000) by estimating a regression model:

$$ln(y_{it}) = \alpha + X_i'\gamma + \sum_j \beta_j D_{ji} + \epsilon_{it} \tag{1}$$

where D_{ji} is an indicator for the type of degree (j) the individual (i) has graduated from and X_i' is a vector of observable characteristics. The outcome measure of interest, $ln(y_{it})$, is the log of annual earnings at time t.[20]

The key assumption here is that there are no variables omitted from this equation that are related to both the higher education choice and subsequent earnings outcomes. Put differently, the assumption is that:

$$cov(D_{ji}, \epsilon_{it}|X_{it}) = 0 \quad \forall \ j \tag{2}$$

which says that conditional on the control variables X there is no correlation between the

[20]We do not adjust for where in the country people are living when they are working as we consider this to be part of the causal pathway from going to university.

earnings residual and the decision to enter higher education. The challenge in estimating the earnings return to university is therefore to account for all the differences between individuals that might affect both their decision to enrol and their earnings prospects. In what follows we set out our approach to dealing with this challenge.

4.1 Pooled earnings model

We start by documenting the regression specification that we use, which extends the model given by equation 1. The oldest cohort in our sample, the 2002 GCSE cohort, has a median age of 30 in 2016/17, our last year of data. For our headline estimates we use age 30 in order to allow for growth in returns with age as much as possible while keeping our estimates within sample. However, to avoid relying solely on observations from one cohort of students, we include several cohorts of students and also multiple earnings observations per individual in a pooled cross-sectional model. This is important because when we look at the degree (subject interacted with institution) level, sample sizes can be small. The pooled model allows us to estimate returns at age 30 while smoothing across several cohorts, reducing the chances of us over-fitting the model.

Specifically, for individual i from GCSE cohort $c \in \{2002, ..., 2007\}$ at time $t \in \{-5, ..., 0\}$, where t is the number of years since the individual took their GCSEs (normalised to zero for the tax year 14 years after GCSEs, or approximately age 30), we model log real earnings as follows:

$$\begin{aligned} ln(y_{ict}) = \quad & X_i'\gamma + I(age_{start} > 18) + \omega_1 t + \omega_2 t^2 + \sum_{c=2003}^{2007} c \\ & + \sum_j \beta_j D_{ji} + \sum_j \beta_{1j}(D_{ji}t) + \sum_j \beta_{2j}(D_{ji}t^2) + \epsilon_{ict} \end{aligned} \quad (3)$$

That is, we model log earnings as a function of observable characteristics X_i' (see more on this below), a dummy for the individual not starting their degree at age 18 (that is, straight after leaving school), a quadratic in t, a set of cohort dummies based on GCSE year (with 2002 the omitted category), the treatment of interest (D_i), a treatment-specific quadratic trend in age ($D_i f(t)$) and a random component (ϵ_{ict}).

We exclude individuals still in education or with earnings below £1,000.[21] We further wind-

[21]We check robustness of our findings to this restriction and find that our results do not quantitatively change if we instead restrict on earnings above £0, nor if we restrict on earnings above £5,000.

sorise earnings at the 99th percentile. The latter restriction is to reduce sensitivity to large outliers, while the former is because we are concerned that people with very low earnings in a given tax year are likely to only be working part of the tax year, or a very low number of hours.[22] All earnings data are put into 2018/19 tax year prices to adjust for inflation.

Our main results focus on earnings at age 30, or $t = 0$. We therefore extract our estimates for the different treatments of interest by plugging $t = 0$ into equation 3.[23] These estimates are point-in-time gross earnings returns meaning they do not adjust for taxes or student loan payments, nor foregone work experience and other costs incurred during study.

We estimate two main sets of models for equation 3. To estimate the overall returns to attending higher education, the treatment of interest D_i is simply a dummy for whether the individual attended higher education. In this case, the control group is individuals who did not attend university.[24] When estimating the returns to different subjects, institutions and degrees, however, we only include individuals who attended higher education and estimate returns relative to a base case.[25] For the subject and institution estimates, D_{ij} in equation 3 is a set of dummies for each of the different subjects and institutions, all included in the same regression additively. For the degree estimates, D_{ij} is a set of dummies for all of the interactions between subjects and institutions. The specification outlined in equation 3 means that we allow all of these treatments to have their own independent time effects.

4.2 Control variables

We believe that the full set of control variables included in the vector X_i in equation 3 is plausibly giving us causal estimates for attending different universities because of the uncommonly rich information we have on each individual in our administrative data.

Specifically, the vector X_i includes three sets of characteristics, all of which are obtained from the NPD data. First, for all children who attended a state secondary school (about 93% of each

[22] We solely observe annual earnings in the tax records, and do not observe hours worked.

[23] This means for our headline estimates, the coefficient of interest is β_j. We also investigate returns at other ages by plugging in different values of t. However, we do not look later than age 30 because this would involve predicting returns out of sample.

[24] As outlined above, for we always restrict to individuals who passed their age 16 high school exams.

[25] These are history, Sheffield Hallam, and history at Sheffield Hallam for subject, institution and degree returns respectively. These were chosen as they have relatively large numbers of students and have earnings close to the middle of the distribution. In practice our estimates are not sensitive to the base case.

cohort) we have a comprehensive set of background controls which includes individual and area based measures of socio-economic background, ethnicity, an indicator for English as an additional language and special educational needs eligibility (see Section 3 for more detail). Second, X_i includes individual secondary school identifiers, which we include as fixed effects. Third, and most importantly, it incorporates extremely detailed information on the prior attainment of each student, specifically the student's grades in specific subjects in national examinations taken at age 11, 16 (GCSE) and 18 (A-level) as described earlier, as well as number of subjects taken and subject mix.[26] Finally, we interact A-level attainment and subject choice variables with quadratic time trends to allow, for example, maths A-levels to have an impact on earnings which grows over time.[27] We do not condition on degree outcomes or on people progressing onto postgraduate study. The estimates therefore include the option value of a good degree and of progressing to postgraduate study (which is not necessarily positive by age 30).

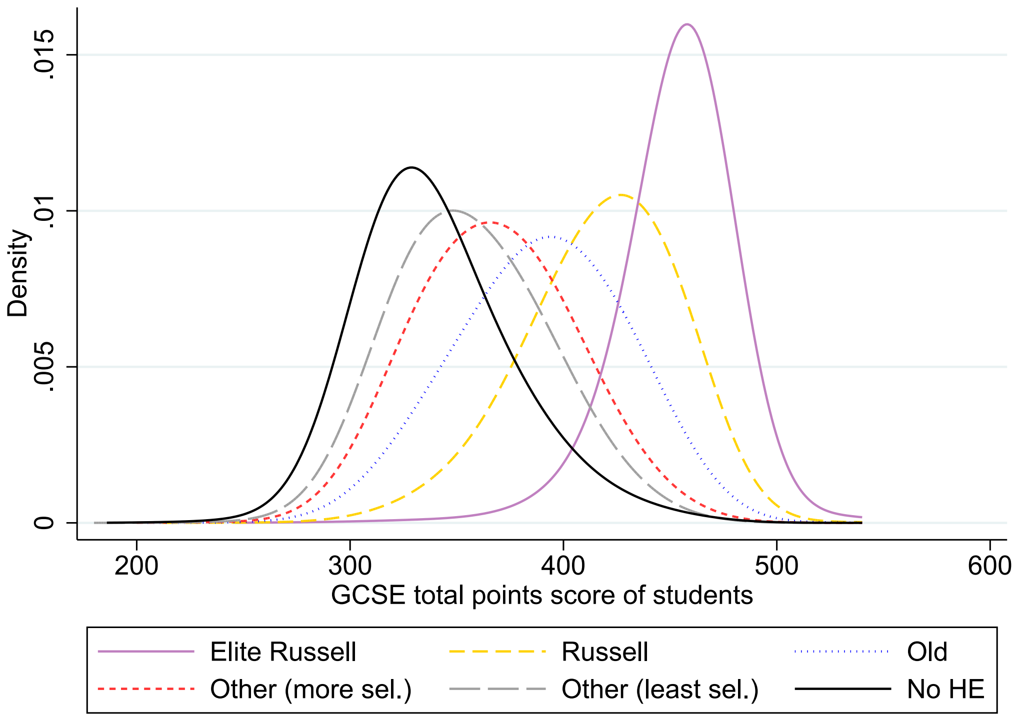
A key issue that we face here is that there is considerable sorting on ability across universities, as we can see in Figure A4 in the Appendix. This raises the question of how we identify returns for the elite institutions for whom there are not many people with similar characteristics who attended the least selective institutions. Figure 1 gives an intuitive idea of how we identify the effects by showing the density of GCSE (age 16) point scores for the different university groups. While there is not a great deal of overlap between the Elite Russell Group and the least selective institutions, there *is* considerable overlap between the Elite Russell Group and the rest of the Russell Group, the rest of the Russell Group and the Old Universities, the Old University and Other (more selective) institutions, and the Other (more selective) institutions and the Other (least selective) institutions. Of course this is at the university group level - in practice there is much more overlap between institutions within these broader groups. This means that we essentially

[26]At age 11 we control separately for scores in all subjects taken (maths, English and science). At age 16 we control for a cubic in total score; scores in maths and English, and scores in science, history, geography, modern languages and vocational courses for those who took these courses; total number of exam entries, as well as total number of full GCSE entries; number of GCSEs with each grade A* to G. At age 18 we control for having any KS5 qualifications; a cubic in total KS5 score; score in vocational courses; number of subjects taken for AS and A-level; number of AS-levels, academic and vocational A-levels with grade A; dummies for taking A levels in maths, sciences, social sciences, arts, humanities, languages and vocational subjects.

[27]We are unable to control for non-cognitive skills. However Buchmueller and Walker (2020) estimate returns to higher education model that is similar to ours using the Longitudinal Survey of Young People in England (LSYPE) and show that the inclusion of rich non-cognitive variables has no effect on the returns estimates conditional on including prior attainment measures.

build sequential common support, and depend on functional form assumptions for identification of returns to elite institutions compared to attending the least selective universities.[28]

Figure 1: Distribution of GCSE points score by university group



Notes: Uses the 2004 GCSE cohort only. No HE consists of individuals who did not take an undergraduate degree, but passed their age 16 exams (obtaining at least 5 A*-C GCSEs).

Several papers in this literature - most recently Mountjoy and Hickman (2020) - have argued that it is necessary to control for the set of higher education institutions individuals apply to, following Dale and Krueger (2002). It is argued that such controls capture both the ability and the preferences of the students to help extract causal estimates. While we are unfortunately not able to observe these choice sets,[29] we have much richer controls than previous papers in this literature. Indeed, it is very common to have just a single measure of attainment prior to college, such as the SAT examination in the United States,[30] whereas we have scores from high-stakes standardised

[28] A very similar argument to this is made in Hoxby (2018). As a robustness check, we narrow the set of institutions that we include in individual regressions and we get extremely similar estimates of relative returns to when we include the full set.

[29] The data exist as all applications to university are through the centralised University and Colleges Admissions Service (UCAS). Although this dataset could in principle be merged into the LEO dataset, it has unfortunately not been possible to obtain it.

[30] Mountjoy and Hickman (2020) also have scores from an additional 10th grade test taken in Texas. However, this is a low-stakes exam.

exams taken in multiple subjects at age 11, 16 and 18. We also have information on subjects taken, which universities factor into their entry requirements. Particularly in the UK, where individuals typically only choose three subjects to study up to age 18, the subject choices will also capture subject-specific skills and preferences. The rich information on the local-level deprivation of the student combined with individual school fixed effects means we are able to effectively compare students from similar backgrounds who attended the same school and chose the same subjects to study in school, and obtained the same results in their exams. We argue that once all of these factors are controlled for, the drivers of differences in choices between different universities are driven by idiosyncratic preferences which are unrelated to subsequent earnings outcomes.

A concern with this argument is that the decision of whether or not to go is distinct from the decision of *where* to go, with the former decision more likely to be driven by factors related to latent earnings potential. To alleviate this concern, our main results are based on 'relative returns', where we show returns relative to a base level degree.

5 Overall returns and variation by subject and institution

In this section we present our findings on the overall returns to higher education in the UK, before turning to how how these returns vary across different universities and subjects (we look at the interaction of subject and institution in the following section).

5.1 Overall returns

We start by presenting our OLS estimates of the overall returns to higher education in Table 3. This shows the estimated impact of university on gross earnings at age 30, in log points by gender. Column 1 displays the unconditional differences in earnings.[31] Unsurprisingly, graduates earn considerably more on average at age 30 than non-graduates. The coefficient estimate for men is 26 log points, or an earnings premium of around 30%, while the equivalent figure for women is 47 log points (60%).

[31] We include in this specification an age-adjustment for those who started university at age 19 or 20, as well as cohort dummies.

Table 3: Overall returns to university at age 30

	(1) Unconditional	(2) + age 16 attain	(3) + full attain	(4) + background	(5) + school FE
Estimate	0.261	0.089	0.057	0.068	0.065
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
R^2	0.0418	0.0799	0.0965	0.1068	0.1189
Adj. R^2	0.0418	0.0799	0.0964	0.1068	0.1174
Individuals	718,339	718,339	718,339	718,339	718,339
Observations	2,206,994	2,206,994	2,206,994	2,206,994	2,206,994
Women					
	(1) Unconditional	(2) + age 16 attain	(3) + full attain	(4) + background	(5) + school FE
Estimate	0.473	0.286	0.222	0.221	0.216
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
R^2	0.0802	0.1254	0.1386	0.1521	0.1646
Adj. R^2	0.0802	0.1254	0.1386	0.1520	0.1633
Individuals	807,131	807,131	807,131	807,131	807,131
Observations	2,501,733	2,501,733	2,501,733	2,501,733	2,501,733

Note: Table reports derived estimates of the overall impact of HE on annual earnings at age 30 based on the 2002-2007 GCSE cohorts, conditional on at least five A*-C GCSEs. Estimates are in log points (/100), standard errors, clustered at the individual level, are in parentheses. All estimates are statistically significant at the 1% level.

We start by controlling for prior attainment, which has a dramatic impact on our returns estimates. In Column 2 we add controls for maths, English and overall GCSE test scores, which roughly halves the returns for women while cutting them by around two-thirds for men. In Column 3 we include the full set of controls we have for prior attainment.[32] One of the key advantages of our data is that we are able to see very rich information on the test scores of students, taken at three different ages (11, 16 and 18) and in different subjects. This is a significant advantage over much of the literature which often relies on a single measure of prior attainment. We see that the inclusion of these additional attainment controls meaningfully affects our estimates for both genders.

However, the same is not true for the remaining columns. In Column 4 we add background characteristics, including socio-economic status, ethnicity and region, while in Column 5 we add school fixed effects.[33] In each case we see that the estimates do not change very much, despite the overall fit of the model improving. The stability of the estimates to the inclusion of rich, relevant

[32] See Section 4.2 for the complete list of controls.
[33] The individuals in our sample attend more than 4,000 different English secondary schools.

conditioning variables adds weight to our assumption that selection into higher education on unobservable factors is not an important driver of the results.

This final estimate for men suggests a return to university of 6.5 log points, or around 7%, at age 30. This estimate on the face of it seems quite low relative to the previous evidence for the UK, most notably Blundell et al. (2000), who estimate a return of around 12 log points for men using data from the British National Child Development Survey, a panel survey of individuals born in a specific week in 1958. However, these estimates actually align fairly closely. Blundell et al. (2000) estimate returns at slightly later age (33), and focus on *graduates*, while we estimate returns for higher education entrants without conditioning on graduation. In Table 4 we show that the estimates are clearly growing quickly with age, and that the returns estimates are considerably higher when we condition on graduates rather than entrants, at 10 log points compared to 7.[34] Given this, estimates for men appear to align quite closely with those from Blundell et al. (2000). The estimates from Blundell et al. (2000) are based on a cohort born 30 years earlier, who went through higher education at a time when higher education attendance was much lower. It is therefore notable that returns have kept up for more recent cohorts, despite a considerable expansion in higher education attendance. This pattern aligns with some recent work using different data which suggests that the graduate earnings premium in the UK has held up through the rapid expansion of the 1990s (Blundell et al., 2016).

The final estimates for women suggest a much larger return of 22 log points (24%). These estimates are quite a lot smaller than those in Blundell et al. (2000), who estimate a returns of 34 log points. An important difference between their approach and ours is that we are estimating returns in terms of annual earnings and are therefore unable to adjust for differences in labour supply. We think this issue is likely to be particularly important for women, as women who do not go to higher education typically have children earlier and are therefore much more likely to be working reduced hours. This is especially true when considering their earnings across a whole year.[35] Based on this, we interpret our overall returns estimates for women with extreme caution,

[34] This result suggests that outcomes for university dropouts are particularly bad, as only 10-15% of students typically do not complete their degrees in the UK. This finding aligns with Ost et al. (2018), which find large causal effects to dropping out based on a regression discontinuity design from a set of 13 public universities in Ohio.

[35] Figure A1 in the Appendix shows that women who did not go to higher education are much more likely to have very low earnings (say, below £8,000 a year) than women who did go, while the same is not true for men.

and for the rest of the paper focus only on *relative* returns to different higher education options amongst the set of people who go. Differential labour supply is likely to be a much less important issue when making comparisons across different degrees than when comparing people who go to higher education with people who do not.

Table 4: Overall returns to university by age and with dropouts

	(1) Age 25	(2) Age 26	(3) Age 27	(4) Age 28	(5) Age 29	(6) Age 30	(7) Excl. Dropouts
Men							
Estimate	0.008	0.020	0.031	0.042	0.054	0.065	0.103
	(0.014)	(0.011)	(0.007)	(0.005)	(0.003)	(0.005)	(0.005)
Women							
Estimate	0.126	0.144	0.162	0.180	0.198	0.216	0.248
	(0.014)	(0.010)	(0.007)	(0.004)	(0.003)	(0.005)	(0.005)

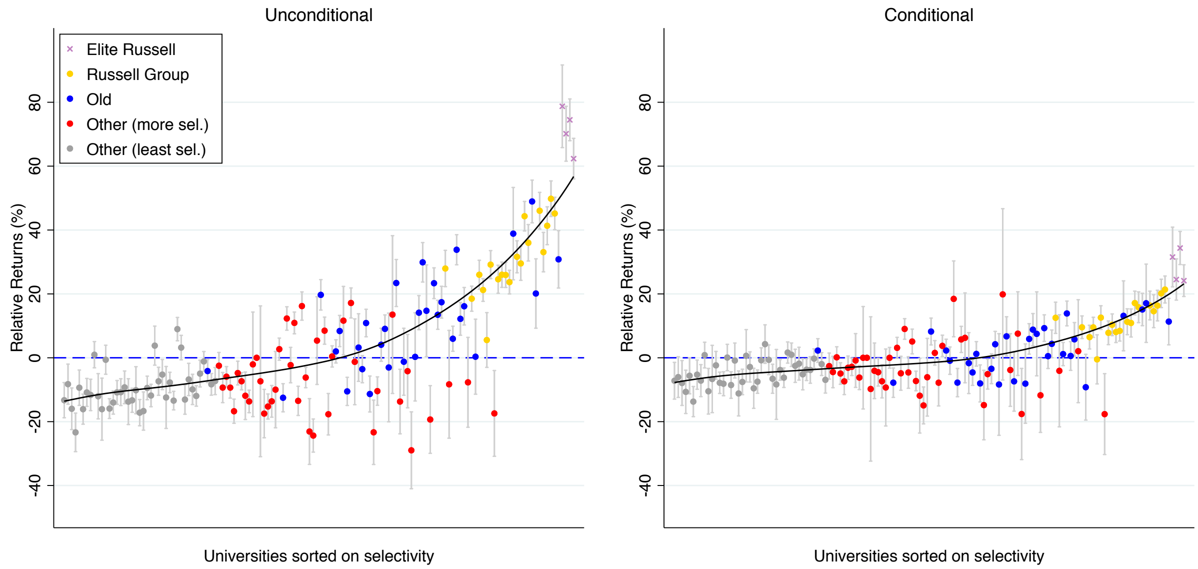
Note: Table reports estimates from the same model as column 5 of Table 3 by age (sample size and fit is the same). The final column shows the age 30 estimates but with university dropouts excluded from the analysis sample. Estimates are in log points (/100), standard errors, clustered at the individual level, are in parentheses. All estimates, other than those in columns 1 and 2 for men, are statistically significant at the 1% level.

5.2 Relative returns by university

We now turn to the estimates of the relative returns to different higher education institutions. Figure 2 displays the institution fixed effects estimates, which are all shown in *relative* terms, with Sheffield Hallam University - a large, mid-ranking institution in the 'Other (more selective)' group - the omitted category. For these estimates, and for all remaining estimates in the paper, we include men and women in the regressions, controlling for gender (separate results by gender are provided in the Appendix). The estimates have been converted into percentage terms from log points, and institutions are sorted on their selectivity rank, as measured by the average GCSE point score of their students.[36] Results are shown both 'unconditionally' in the left hand panel and 'conditionally' on the right. The unconditional estimates include only the university fixed effects in the model, while the conditional estimates include all of the background controls included in column 5 of Table 3, but with controls for subject studied also included. All of the point estimates for this and for subsequent results are provided in an Online Appendix.

[36] While school grades are not the only factors based on which universities select their applications, for most universities it is the most important one. For specialist institutions which tend to rely more on interviews or portfolios, such as music colleges, our measure will reflect their selectivity less well.

Figure 2: Estimated returns at age 30 by institution



Note: Figure reports estimates of the impact of studying at different institutions on annual earnings at age 30 relative to Sheffield Hallam University. Conditional estimates control for year, background, prior attainment and subject. Results have been converted to percentage differences using a log point conversion. Universities are ranked on the average GCSE results of their intake. The black line shows the relationship between returns and selectivity from a locally weighted polynomial regression. 95% confidence intervals are shown by the whiskers and standard errors are clustered at the individual level.

The inclusion of controls substantially flattens the relationship between earnings outcomes and university selectivity. In fact, we see that in the conditional model, the relationship between returns and selectivity is quite weak for universities in the bottom two-thirds of the selectivity distribution. However, the relationship is much steeper amongst the top institutions - returns for the four so-called 'Elite Russell Group' institutions in the conditional model are between 25 and 35% higher than the baseline, while returns for the other Russell Group universities are mostly between 5 and 20% higher than the baseline. This suggests that accessing the very elite institutions can boost outcomes considerably over the next tier of institutions.

At the lower end of the scale, returns amongst the least selective institutions are, on average, around -5% relative to the baseline, which is very similar to the average of the more selective other institutions and a few percentage points below the returns for the 'Old' (more established) institutions. Interestingly, only four of the bottom ten institutions for returns are from the set of least selective institutions - while six of the bottom ten institutions are specialist arts colleges.

Our findings on the relationship between selectivity and returns aligns with the previous UK evidence. Chevalier and Conlon (2003); Hussain et al. (2009); Broecke (2012) and Walker and Zhu (2018) all report similar results. There are some inconsistencies in prior findings in that Broecke (2012) suggests that there is a linear relationship between returns and selectivity, while the other papers align more closely with our findings of a stronger relationship amongst the more elite universities.[37] Broecke (2012) suggest one possible explanation for this is the comparison of a broad range of institutions all within one model. We assess the robustness of our result by re-running our specification using only subsets of universities and find that the estimates are extremely highly correlated across the alternative samples (we show this, plus the fact that the results are robust to some other alternative specifications in Appendix Table A5).

The only previous paper to estimate returns for individual universities in the UK is Walker and Zhu (2018), which uses the Labour Force Survey (LFS), which relies on self-reported earnings and only allows for the most basic control variables. It is notable that our raw differences in earnings are much greater than their estimates, and unsurprisingly, our control variables make a much larger difference to the university fixed effects. Nevertheless, we end up with a similar range of final estimates.

Finally, we note that in our final specification we condition on the subject studied, unlike most of the previous evidence on institution returns from the US (for example, Mountjoy and Hickman, 2020). This does not dramatically change the final set of results, with the relationship between selectivity and returns, as well as the standard deviation of estimates changing very little. However, there are some institutions that experience very large changes to their estimates. Many of these are specialist arts institutions at the bottom end of the returns distribution which perform considerably better when subject controls are included, reflecting the low returns for creative arts degrees.

[37]When we plot returns on the average GCSE scores of the intake rather than selectivity rank, this non-linearity is less clear. This is because the most selective universities are effectively shifted over to the right as they are much more selective than the rest. However, we still observe a steeper relationship at the top end of the selectivity distribution than at the bottom.

5.3 Relative returns by subject

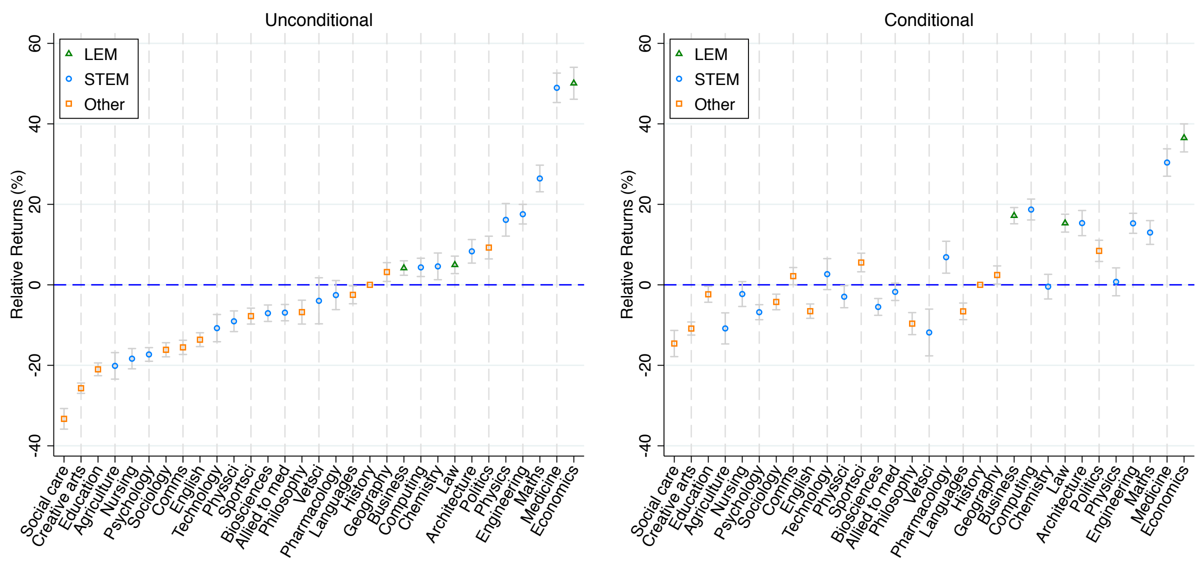
We explore the returns for specific subjects in Figure 3. The estimates are again converted into percentage terms and are now reported relative to the returns for history, which is the omitted category. We again show the unconditional estimates and the fully conditional estimates in the Figure to show the full effect of the control variables. We see that they again make a substantial difference to the distribution of subject fixed effects, although not to the same extent as for the institution fixed effects. This is not entirely unexpected as sorting on ability is less strong across subjects than across institutions. Nevertheless, we observe some big changes. For example, at the top end, relative returns for medicine and economics drop from close to 50% to 30% and 36% respectively once background controls are included (the patterns when we look at men and women separately are extremely similar - see Figure A7 in the Appendix).

We also see fairly large upward shifts in relative returns estimates for some of the lowest earning subjects, most notably social care, creative arts, communications and education. We also see business and computing returns increase considerably, highlighting the fact that these degrees often admit students with relatively low prior attainment.

For the conditional estimates in the right hand panel, we still see significant variation in relative returns across subjects. Outside of economics and medicine, we see very good returns of around 15% for computing, business, architecture and law. At the bottom end, social care, veterinary sciences, creative arts and agriculture all have estimated returns of -10% lower than history or worse. Philosophy, psychology, English, languages and biological sciences all also perform poorly. The correlation of the returns for men and women is very high (0.91).

We also observe an interesting pattern in the returns across our different broad subject groups. In general, it is the case that the three LEM subjects do very well, the 'Other' subjects (which mostly consist of arts, languages and humanities) tend to do quite poorly, while the returns for STEM subjects are very mixed. Medicine, computing, engineering and maths all do well, while veterinary sciences, agriculture, psychology and biological sciences do not. This is particularly worth noting as some of these subjects - especially psychology and biological sciences - are very popular amongst women. This suggests policies encouraging women to study STEM subjects might not actually always result in positive earnings impacts.

Figure 3: Estimated returns at age 30 by subject



Note: Figure reports derived estimates of the impact of studying different subjects on annual earnings at age 30 relative to studying History. Conditional estimates control for age, background, prior attainment and institution. Results have been converted to percentage differences using a log point conversion. Subjects are ranked based on raw earnings differences. 95% confidence intervals are shown by the whiskers and standard errors are clustered at the individual level.

Although (as discussed above) we are very cautious about treating these subject estimates as causal, we still consider it useful to report them. The identification strategy is similar to the most comparable previous evidence from the UK on subject returns (Chevalier, 2011), which estimates the annual earnings effects of different subjects based on data from 2006, three and a half years after graduation.[38] Broadly speaking, he finds similar estimates, with medicine doing very well and creative arts doing poorly. One major difference is economics, which we have as the highest-returning subject while he has it much further down the distribution. This could be because he is only observing earnings quite soon after graduation, although it mostly likely to be because he is working with much smaller samples - for example, he only observes 110 economics graduates.

[38]The most directly comparable results to ours are reported in column 7 of Table 2 in his paper, noting that his base case is physical sciences, whereas ours is history.

6 Returns to different degrees

We now turn to the main contribution of the paper and focus on estimates of returns at the degree level. We explore the overall distribution of earnings outcomes and returns before looking at the variation within given selectivity bands and the relationship between returns and selectivity. Finally, we consider how well other indicators of university quality are correlated with returns.

6.1 Returns by degree

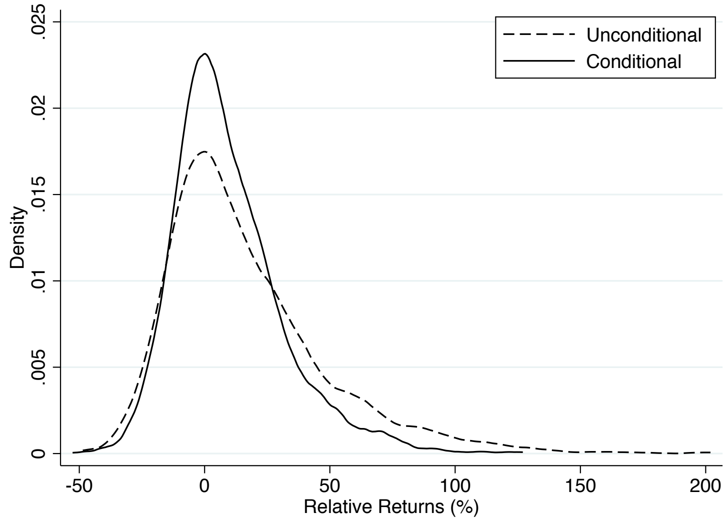
Figure 4 shows the overall distribution of the more than 1,900 degree fixed effect (these are estimated relative a base case of history at Sheffield Hallam).[39] We show the distribution of degree-level fixed effects unconditionally[40] and with the full set of controls. The figure shows substantial variation in the raw estimates, which range from 50% below the base case to 200% above it. The inclusion of the controls considerably reduces this variation. This is summarised in Table 5, which shows the standard deviation and range of the returns estimates for the conditional and unconditional degree level estimates. The standard deviation of our degree returns estimates drops from 32 percentage points to a still very large 22 percentage points with the inclusion of controls. Similarly, the 90:10 range drops from 75 percentage points to 52 percentage points.

In columns 3 and 4 of Table 5 we compare the degree level returns estimates with equivalent estimates from regression models where *only* subject (column 3) or institution (column 4) fixed effects are included, to provide a comparison with estimates when only subjects or institutions are observed. The degree level returns are much more variable, with a standard deviation and 90:10 range around twice as large as the institution and subject estimates. This shows that the variation in institution or subject level returns dramatically understates the variation in returns to higher education degrees. The table also highlights that more of the variation in earnings is explained in the degree-level regressions, with the (adjusted) R^2 increasing from around 0.15 for the subject and institution fixed effects regressions to 0.18 for the degree fixed effects regression with controls.

[39] All individual returns estimates can be found in the Online Appendix. For sample size reasons, not all degrees offered are included in this analysis. Specifically, for inclusion we require the degree to have at least 10 individuals with earnings observations at age 30, and 50 unique individuals with earnings observations at any of the ages 25 to 30. This is to ensure data disclosure requirements are met and that we are not predicting earnings returns 'out-of-sample', which would significantly increase the uncertainty and importance of the underlying assumptions in our estimates.

[40] We do include a dummy for not starting university straight after school, as well as cohort dummies, in this specification.

Figure 4: Estimated returns at age 30 by degree



Note: Figure reports derived estimates of the impact of studying different degrees (subject-institution combinations) on annual earnings at age 30 based on the 2002-2007 GCSE cohorts controlling for age, background and prior attainment. Results have been converted to percentage differences using a log-point conversion.

Table 5: Summary of degree, subject and institution estimates

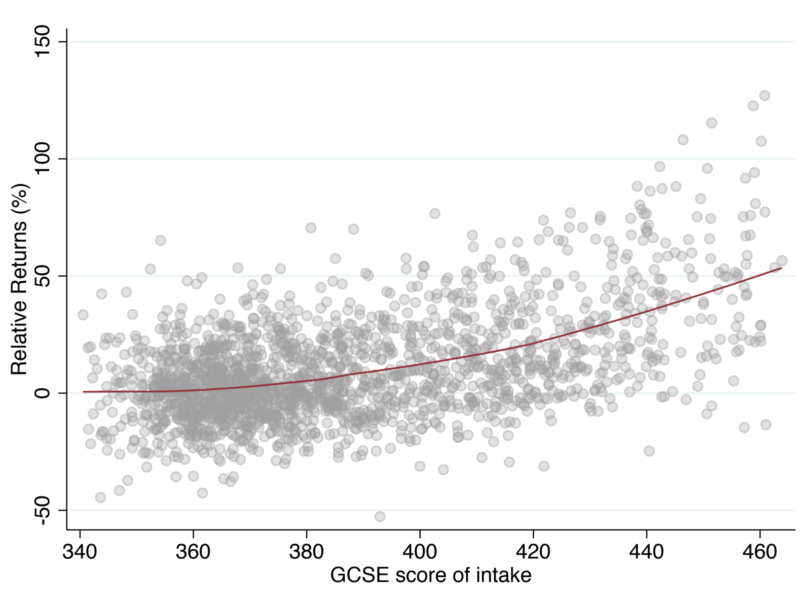
	(1) Degree Unconditional	(2) Degree Conditional	(3) Uni	(4) Subject
σ	32.03	21.89	10.50	12.33
90:10 Range	75.35	51.95	26.26	27.42
Adj. R^2	0.15	0.18	0.15	0.16
Controls	No	Yes	Yes	Yes

Note: σ is the standard deviation of degree returns, the range is the 90th percentile return minus the 10th percentile (all in percentage terms) and the adjusted R^2 is from the underlying earnings regression with degree/HEI/subject fixed effects. The conditional university results *exclude* subject controls and similarly the conditional subject results *exclude* university controls.

To get a sense of the types of degrees that give particularly high or low returns, Table A4 in the Appendix lists the best and worst performing degrees. We find that the top degrees are heavily dominated by law and economics. The top end is also heavily dominated by the high-status Russell Group universities. The worst performing degrees include a wider range of subjects, with social care, philosophy, politics and subjects allied to medicine all appearing in the bottom ten. Most of the lowest performers are from the least selective 'Other' group of universities, although humanities degrees from higher-status institutions do appear. This broad pattern holds throughout the distribution. LEM degrees, and degrees at the most elite institutions perform best, while arts and humanities degrees, and those at low ranked universities perform worst on average.

We show the full set of returns estimates, plotted against their selectivity (as measured by the average age 16 test scores of students), in Figure 5. The first point to note is that average returns increase considerably as we move from the least selective to the most selective degrees, with a difference of more than 50 percentage points in average returns. Again, the relationship between returns and selectivity gets stronger as we move up the selectivity distribution. This is documented more explicitly in Table 6 which reports the slope coefficient from a regression of returns on selectivity within selectivity bands. We see that this increases from -0.02 (meaning a 100 point increase in GCSE points is associated with a 2% decrease in returns) in the least selective band of degrees, to around 1 (meaning a 100 point increase in GCSE points is associated with a 100% increase in returns).

Figure 5: Course returns against selectivity



Notes: Degree level estimates plotted against average GCSE scores of intake. Red line plots the relationship with a locally weighted polynomial.

The second point is that despite this relationship, selectivity by no means explains all the variation in returns. Table 6, highlights the considerable variation in returns across different bands of similarly selective degrees. The standard deviation of returns amongst the least selective band is around 15 percentage points, and this doubles to around 30 percentage points for the most selective band of degrees. For reference, this compares to an estimated overall return to higher education of around 7% for men at the same age.

Table 6: Summary of degree, subject and institution estimates

	GCSE score of intake					
	340-359	360-379	380-399	400-419	420-439	440-459
Main results						
Standard deviation	16.02	15.31	17.79	20.11	23.84	29.29
Share within subject	0.47	0.52	0.46	0.64	0.71	0.77
Selectivity slope	-0.02	0.19	0.28	0.24	1.34	0.72
Excl. school FE						
Standard deviation	18.09	17.21	19.77	21.85	25.79	31.75
Share within subject	0.48	0.51	0.48	0.61	0.70	0.77
Selectivity slope	-0.10	0.27	0.32	0.21	1.48	0.97
Excl. school FE and background						
Standard deviation	16.95	16.23	18.84	21.09	24.68	30.55
Share within subject	0.51	0.57	0.50	0.64	0.71	0.78
Selectivity slope	-0.06	0.18	0.31	0.24	1.38	0.83
Excl. dropouts						
Standard deviation	16.42	16.21	18.21	20.50	23.92	29.15
Share within subject	0.50	0.57	0.48	0.62	0.72	0.77
Selectivity slope	-0.14	0.20	0.28	0.30	1.34	0.69
Cross-sectional						
Standard deviation	14.84	14.61	16.87	19.59	23.51	27.62
Share within subject	0.40	0.42	0.34	0.60	0.67	0.76
Selectivity slope	0.11	0.22	0.24	0.23	1.25	0.69
Shrinkage						
Standard deviation	11.40	11.57	13.27	15.80	18.07	21.94
Share within subject	0.44	0.53	0.50	0.63	0.72	0.74
Selectivity slope	-0.06	0.15	0.20	0.22	1.01	0.58
Number of courses	303	687	362	280	182	106

Notes: Table shows the standard deviation of returns, and the slope of a regression of returns on selectivity (average GCSE score of student intake) within 20 point selectivity bands. The very few degrees with average GCSE scores of the student intake of 460 or more are not shown. These statistics are shown both for the main degree returns, as well as for a further three specifications. 'Excl. dropouts' estimates returns when individuals who do not finish their degree of study are excluded. 'Cross-sectional' estimates returns on individuals at age 30 only, rather than using the panel model used in the main specification. 'Shrinkage' applies shrinkage to the main returns estimates, where estimates are shrunk towards the average degree returns.

As confirmed by Figure 5, the very highest return degrees are dominated by the most selective degrees, yet we also find a number of extremely selective degrees at elite institutions with very low relative returns. Table 6 also shows that these two conclusions are robust to removing subsets of the control variables from our regression models - when school fixed effects and then additional background controls are excluded from the models, the qualitative patterns of the estimates are almost identical. We interpret this as promising evidence that relevant variables that we are excluding from the model would not dramatically change our headline findings. The table further shows that these two findings are robust to the exclusion of dropouts, to using a cross sectional rather than a panel estimation for the regression, and to the application of a shrinkage estimator.

6.2 Within subject returns

As discussed above, we believe that one should be cautious about interpreting the variation in degree level returns across different subject areas, as previous evidence has highlighted the importance of selection on comparative advantage into different fields. However, Table 6 also presents the share of the variation in returns within each selectivity band that occurs within subject. This shows that at least half of the variation is *within* subject, across institutions for all selectivity bands. This increases to around three-quarters of the variation for the most selective degrees.

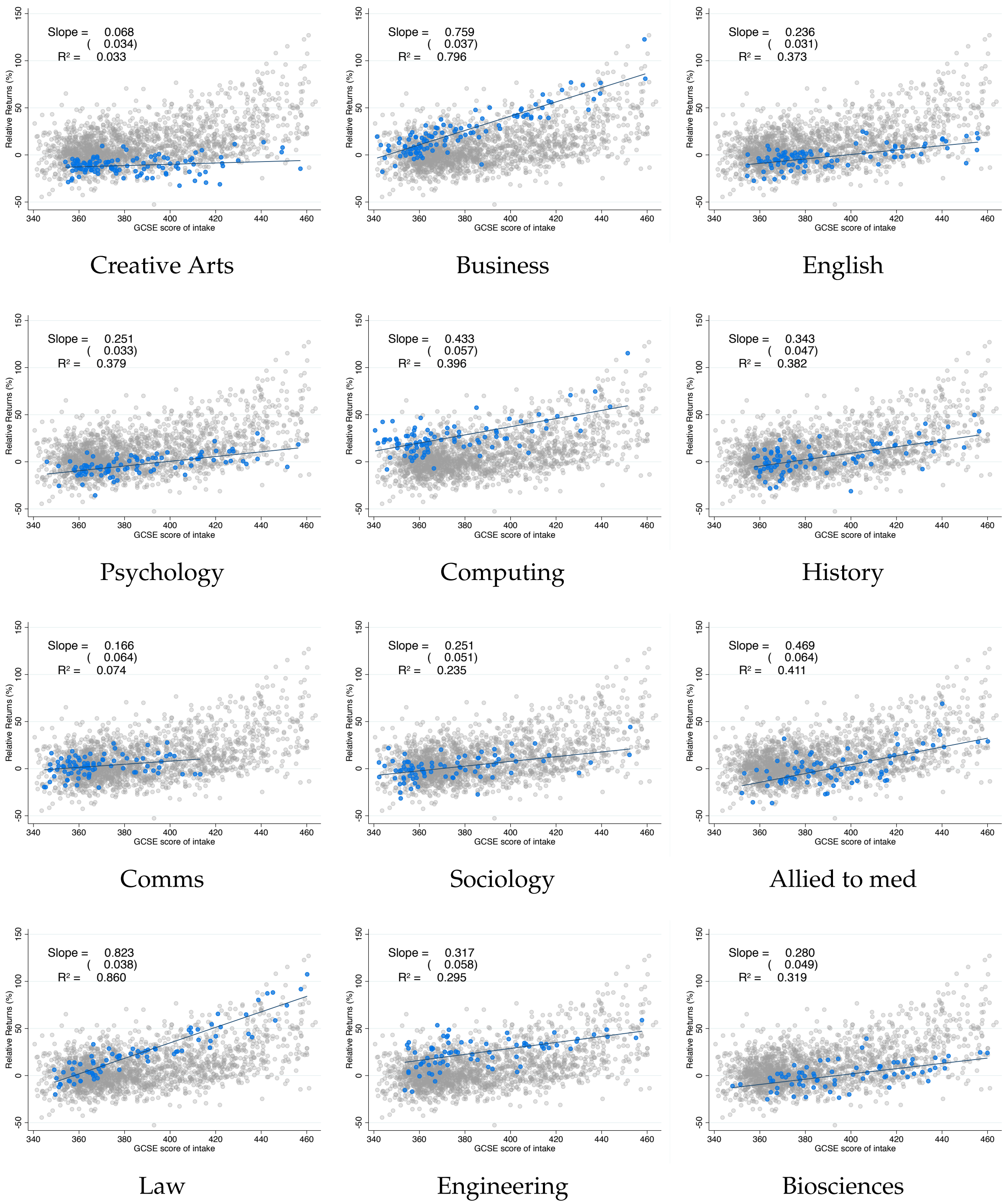
Figure 6 then highlights individual institution estimates for each of the 12 largest subjects.[41] The figure supports the point that there is substantial variation in returns even within given subject areas, even across institutions that are similarly selective. Holding subject choice fixed, attending one university over another, similarly selective university can often lead to 40 percentage point difference in returns. This holds true right through the selectivity distribution.

The figure also documents the relationship between returns and selectivity, as well as the share of the variation in returns that can be explained by selectivity within subject (for a summary of the relationship between selectivity and returns for all subjects, see Figures A8 and A9 in the Appendix). For business and law, the relationship between returns and selectivity is strong, and selectivity can explain more than 75% of the variation. For the other LEM subject, economics, the relationship is also strong and the share of the variation explained is just over 60%. For most STEM and arts and humanities subjects, less than 50% of the variation can be explained by selectivity, however, and moving into more selective universities to study these subjects has a much lower

[41] Correlations with selectivity for all subjects can be found in Table A6 in the Appendix. This Table also shows that the within-subject correlations with selectivity are robust to the precise set of control variables we include in the regression models, again adding weight to the argument that unobserved factors are unlikely to change our qualitative findings here (specifically, we see that the within-subject correlations between returns and selectivity are almost identical when we exclude school fixed effects and other student background characteristics).

pay off. Finally, for subjects which lead to professions with regulated earnings, such as education and medicine, less than 10% of the variation in returns is explained by selectivity.

Figure 6: Relationship of returns and selectivity at the subject level



Notes: Degree level estimates plotted against average GCSE scores of intake.

6.3 Correlates with degree returns

So far we have seen evidence suggesting that degree choices can make a substantial difference to earnings outcomes at age 30 and that selectivity can only explain some of this variation. For many subjects, selectivity only explains a very small share of the variation in returns. In this section we therefore consider whether other characteristics of degrees are predictive of returns outcomes. Continuing the theme of the previous sub-section, we do this *within* subject area.

In addition to showing the correlation between selectivity and returns for each subject, Table 7 shows the relationship between returns and the following set of degree characteristics:

- **League table ranking:** this is the subject-specific league table ranking of universities. We take these from the Complete University Guide (CUG) from 2010, which was the most relevant year we could collect data for. These rankings combine several characteristics of the degree, including student-staff ratios and research intensity and should therefore capture aspects of degree quality reflected in returns which are unrelated to selectivity. Gibbons et al. (2015) highlights the importance of such rankings in driving institution choices of prospective students.

- **Student satisfaction:** this is taken from the National Survey of Students, focussing on overall satisfaction, again based on data from 2010. This measure has recently been included as an input into the governments' Teaching Excellence Framework, a measure of teaching quality.

- **Age 22 returns:** this is the very early-career earnings of students, usually in the first year after graduating from university. We report this as early-career outcomes are often used as a measure of degree quality. For example, labour market outcomes six months after graduation (taken from a graduate survey) have frequently been included as inputs into league table rankings.

- **Completion rate:** this is the share of students starting a degree who complete it. People who start a degree and switch to another full-time degree before the age of 21 are neither treated as completers nor dropouts from the degree they started first (as we just take the degree they switched to as their main degree in those cases).

- **First class degree rate:** this is the share of students who achieved 'First class honours', the highest degree classification. This is not regulated and so varies across different subjects and universities.

Table 7: Correlates with age 30 degree returns

	(1) Selectivity	(2) League table	(3) Student satisfaction	(4) Age 22 returns	(5) Completion rate	(6) First class degree rate
LEM						
Business	0.888	0.806	0.286	0.602	0.772	0.456
Economics	0.792	0.765	-0.087	0.760	0.658	0.550
Law	0.928	0.836	0.073	-0.342	0.789	0.601
STEM						
Allied to med	0.625	0.565	0.323	0.069	0.434	0.555
Architecture	-0.056	0.228	0.357	0.273	0.214	0.180
Biosciences	0.575	0.571	0.257	0.365	0.500	0.449
Engineering	0.521	0.515	0.365	0.022	0.499	0.269
Maths	0.696	0.518	-0.066	0.581	0.590	0.430
Medicine	0.088	0.254	0.182	.	0.455	-0.123
Physsci	0.541	0.306	0.038	0.235	0.369	0.323
Other						
Comms	0.264	0.425	0.272	0.157	0.264	0.147
Creative arts	0.184	0.197	0.090	0.266	0.145	0.081
Education	-0.081	0.037	0.057	0.101	0.104	0.020
History	0.610	0.380	0.018	0.194	0.557	0.517
Languages	0.541	0.490	0.233	-0.049	0.453	0.334
Sociology	0.489	0.407	-0.096	-0.197	0.501	0.115

Note: Descriptions of each of the variables are given in the text. Numbers report the raw correlations. Only subjects for which we could obtain league table rankings and student satisfaction scores are shown.

Column (1) of Table 7 repeats the result from above that returns at age 30 are very highly correlated with selectivity. We then see that league table rankings, completion rates and first class degree rates are all well correlated with returns. This is especially true for the LEM subjects and much less true for the 'Other' subjects, with STEM subjects generally in between. The correlations of returns and student satisfaction ratings with early career (age 22) returns are much noisier and generally weaker. In fact we even see negative correlations between student satisfaction and returns for economic, maths and sociology, suggesting students studying towards these degrees do not value or appreciate things that are well correlated with their subsequent labour market success. For age 22 returns, we see that this is a very unreliable measure of subsequent success in many cases - for example, there is virtually no correlation at all between returns at 22 and returns at 30 for education, and even a negative correlation for law. This suggests that there are large cross-subject differences in the time it takes for career paths to become established.

Table 8: Correlates with age 30 degree returns, controlling for selectivity

	(1) League table	(2) Student satisfaction	(3) Age 22 returns	(4) Completion rate	(5) First class degree rate
LEM					
Business	0.010	0.145	0.158	0.044	0.109
Economics	0.195	0.001	0.235	-0.096	0.055
Law	0.029	-0.054	-0.045	-0.025	0.095
STEM					
Allied to med	0.001	0.019	-0.013	-0.115	0.126
Architecture	0.272	0.377	0.257	0.238	0.204
Biosciences	0.054	-0.064	0.248	0.043	0.098
Engineering	0.076	0.113	-0.016	0.076	-0.057
Maths	-0.234	-0.409	0.242	0.038	0.073
Medicine	0.227	0.209	.	0.438	-0.158
Physsci	-0.320	-0.373	0.159	0.053	-0.061
Other					
Comms	0.209	0.229	0.187	0.065	0.007
Creative arts	0.043	0.023	0.293	0.001	-0.040
Education	0.108	0.079	0.111	0.165	0.069
History	-0.194	0.007	0.213	0.034	-0.047
Languages	0.026	0.124	0.007	-0.010	-0.013
Sociology	0.001	-0.259	-0.133	0.111	-0.258

Note: Descriptions of each of the variables are given in the text. Numbers report the partial correlations, after taking out selectivity. Only subjects for which we could obtain league table rankings and student satisfaction scores are shown.

In Table 8 we then look at how much of the correlations in Table 7 are driven by the correlation of these variables with selectivity. To do this, we regress returns on selectivity and correlate the residual with the variables of interest. We see that conditioning on selectivity removes almost all of the correlations between the university characteristics and returns. This suggests that there is no additional meaningful information in these measures over and above what you get from a simple measure of the selectivity of the degree. This is a disappointing result from the point of view of policy, as it suggests that the information available to students making their choices about where to study is not very well related to their likely outcomes. This could be particularly damaging as our evidence suggest that these choices matter a lot for earnings. It also has concerning implications for the incentives of universities who are competing for students and for regulators trying to incentivise universities to boost the labour market prospects of their students.[42]

7 Conclusion

This paper uses a novel administrative data linkage from the UK to investigate the returns to higher education and how they vary across different degrees. Our key finding is that there is substantial variation in returns at the degree level even within relatively tight selectivity bands. We find that a large share of the variation within selectivity bands is *within* subject, mitigating any concerns that the variation in returns across degrees might be overstated by selection into different fields based on comparative advantage (Kirkeboen et al., 2016). Our results therefore suggest that degree choice matters a lot for earnings outcomes at age 30. We provide suggestive evidence that this finding is robust to the empirical specification used, the exact sample of students included, and to unobserved selection. Since age 30 is still a relatively early point to assess returns to higher education, considerable variation at this age is likely to be indicative of even greater variation later on.

While degree choice appears to matter a lot, we find that once we control for a simple measure of the selectivity of a degree (specifically, the average GCSE scores of the students), many other measures of degree quality, including subject-specific league table rankings of universities, are not at all well correlated with returns. This has important implications, as students are making choices that can have enormous implications for their future outcomes with poor information on which to base those choices. This is likely to drive up the costs of higher education, to damage the productivity of the economy and to increase inequality, as poorer students are likely to be more reliant on publicly available information. It is also likely to create perverse incentives for universities, which may wish to target factors such as student satisfaction or first class degree shares when those things might not be beneficial in the long term.[43]

[42] The findings in this section are robust to the exact specification used to estimate returns. In Appendix Table A7 we show that returns are extremely highly correlated across specifications, with correlations of more than 0.95 for most subjects when we compare our main returns estimates to estimates excluding dropouts, using age 30 only, or with shrinkage applied. In the final column, we show the results are essentially the same when we estimate degree returns for each subject completely separately.

[43] A notable example of this is the dramatic increase in first class degree shares that have occurred at UK universities in recent years as competition for domestic students has increased following the removal of student number caps.

One potential solution to this could be to make information on the earnings outcomes of students more readily available when prospective students are making their higher education choices. In the UK this is increasingly plausible given the data linkage created for this work, and other countries may wish to develop similar data sources. A more extreme solution would be for the government to use the returns estimates to protect or boost funding where returns are high and restrict it where they are not. However, there are a few reasons why caution should be exercised before using degree level returns estimates to justify funding cuts. First, there is a long lag between changes to university practice and changes to earnings returns. The current estimates are based on people who started university between 10 and 15 years ago, and we have seen that looking at early earnings outcomes can be misleading. Second, a university degree may have important positive impacts that might not be reflected in our earnings returns estimates. Third, it is also possible that the returns do not reflect university productivity and are instead a product of peers, labour market signalling or both. Understanding what drives the very large differences in returns is an important topic for future research.

Future research should also look in more detail at what drives the higher returns of certain universities within given subject areas. This could include investigations of the specific practices of the successful universities, such as the style of teaching and the content included. Other channels to explore would be the location of students after graduation, and the occupations they enter. Future work could also investigate the signalling component of these returns by surveying employers about how they value degrees from certain universities amongst job applicants. Such a study would help to highlight any labour market biases that might need to be addressed, and would also promote practices that are associated with good outcomes of students that could potentially boost teaching quality throughout higher education.

References

Altonji, Joseph G, Erica Blom, and Costas Meghir, "Heterogeneity in human capital investments: High school curriculum, college major, and careers," *Annu. Rev. Econ.*, 2012, *4* (1), 185–223.

Andrews, Rodney J, Scott A Imberman, and Michael F Lovenheim, "Risky Business? The Effect of Majoring in Business on Earnings and Educational Attainment," Technical Report, National Bureau of Economic Research 2017.

Anelli, Massimo, "The returns to elite university education: A quasi-experimental analysis," *Journal of the European Economic Association*, 2020, *18* (6), 2824–2868.

Black, Dan A and Jeffrey A Smith, "Estimating the returns to college quality with multiple proxies for quality," *Journal of labor Economics*, 2006, *24* (3), 701–728.

Blundell, Richard, David Green, and Wenchao Jin, "The UK wage premium puzzle: how did a large increase in university graduates leave the education premium unchanged," *Institute for Fiscal Studies Working Paper*, 2016, *16* (1).

_ , Lorraine Dearden, Alissa Goodman, and Howard Reed, "The returns to higher education in Britain: evidence from a British cohort," *The Economic Journal*, 2000, *110* (461), 82–99.

_ , _ , and Barbara Sianesi, "Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2005, *168* (3), 473–512.

Britton, Jack, Lorraine Dearden, Neil Shephard, and Anna Vignoles, "How English domiciled graduate earnings vary with gender, institution attended, subject and socio-economic background," Technical Report, Institute for Fiscal Studies WP1606 2016.

Broecke, Stijn, "University selectivity and earnings: Evidence from UK data on applications and admissions to university," *Economics of Education Review*, 2012, *31* (3), 96–107.

Buchmueller, Gerda and Ian Walker, "The Graduate Wage and Earnings Premia and the Role of Non-Cognitive Skills," *IZA Working paper 13248*, 2020.

Campbell, Stuart, Lindsey Macmillan, Richard Murphy, and Wyness, "Inequalities in Student to Course Match: Evidence from Linked Administrative Data," Technical Report, Centre for Economic Performance, LSE 2019.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan, "Mobility report cards: The role of colleges in intergenerational mobility," Technical Report, National Bureau of Economic Research 2017.

_ , _ , _ , _ , **and** _ , "Income Segregation and Intergenerational Mobility Across Colleges in the United States," *Quarterly Journal of Economics*, 2020.

Chevalier, Arnaud, "Subject choice and earnings of UK graduates," *Economics of Education Review*, 2011, *30* (6), 1187–1201.

_ **and Gavan Conlon**, "Does it pay to attend a prestigious university?," 2003.

Chowdry, Haroon, Claire Crawford, Lorraine Dearden, Alissa Goodman, and Anna Vignoles, "Widening participation in higher education: analysis using linked administrative data," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2013, *176* (2), 431–457.

Cunha, Jesse M and Trey Miller, "Measuring value-added in higher education: Possibilities and limitations in the use of administrative data," *Economics of Education Review*, 2014, *42*, 64–77.

Dale, Stacy B and Alan B Krueger, "Estimating the effects of college characteristics over the career using administrative earnings data," *Journal of human resources*, 2014, *49* (2), 323–358.

Dale, Stacy Berg and Alan B Krueger, "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables," *The Quarterly Journal of Economics*, 2002, *117* (4), 1491–1527.

Dillon, Eleanor Wiske and Jeffrey Andrew Smith, "The consequences of academic match between students and colleges," *Journal of Human Resources*, 2020, *55* (3), 767–808.

Gibbons, Stephen, Eric Neumayer, and Richard Perkins, "Student satisfaction, league tables and university applications: evidence from Britain," *Economics of Education Review*, 2015, *48*, 148–164.

Hastings, Justine, Christopher A Neilson, and Seth D Zimmerman, "The effects of earnings disclosure on college enrollment decisions," Technical Report, National Bureau of Economic Research 2018.

Hastings, Justine S, Christopher A Neilson, and Seth D Zimmerman, "Are some degrees worth more than others? Evidence from college admission cutoffs in Chile," Technical Report, National Bureau of Economic Research 2013.

Hoekstra, Mark, "The effect of attending the flagship state university on earnings: A discontinuity-based approach," *The Review of Economics and Statistics*, 2009, *91* (4), 717–724.

Hoxby, Caroline M, "The productivity of US postsecondary institutions," in "Productivity in Higher Education," University of Chicago Press, 2018.

Hussain, Iftikhar, Sandra McNally, and Shqiponja Telhaj, "University quality and graduate wages in the UK," Technical Report, IZA Discussion Papers 2009.

Kirkeboen, Lars J, Edwin Leuven, and Magne Mogstad, "Field of study, earnings, and self-selection," *The Quarterly Journal of Economics*, 2016, *131* (3), 1057–1111.

Mountjoy, Jack and Brent Hickman, "The Returns to College (s): Estimating Value-Added and Match Effects in Higher Education," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2020, (2020-08).

OECD, "Education at a Glance 2014: OECD Indicators," 2014.

O'Leary, Nigel C and Peter J Sloane, "The return to a university education in Great Britain," *National Institute Economic Review*, 2005, *193* (1), 75–89.

Ost, Ben, Weixiang Pan, and Douglas Webber, "The returns to college persistence for marginal students: Regression discontinuity evidence from university dismissal policies," *Journal of Labor Economics*, 2018, *36* (3), 779–805.

Saavedra, Juan E, "The returns to college quality: A regression discontinuity analysis," 2008.

Walker, Ian and Yu Zhu, "Differences by degree: Evidence of the net financial rates of return to undergraduate study for England and Wales," *Economics of Education Review*, 2011, *30* (6), 1177–1186.

__ **and** __ , "The impact of university degrees on the lifecycle of earnings: some further analysis," 2013.

__ **and** __ , "University selectivity and the relative returns to higher education: Evidence from the UK," *Labour Economics*, 2018.

Webber, Douglas A, "Are college costs worth it? How ability, major, and debt affect the returns to schooling," *Economics of Education Review*, 2016, *53*, 296–310.

Wiswall, Matthew and Basit Zafar, "Determinants of college major choice: Identification using an information experiment," *The Review of Economic Studies*, 2014, *82* (2), 791–824.

Zimmerman, Seth D, "The returns to college admission for academically marginal students," *Journal of Labor Economics*, 2014, *32* (4), 711–754.

__ , "Elite colleges and upward mobility to top jobs and top incomes," *American Economic Review*, 2019, *109* (1), 1–47.

Appendix

A1 Sample selection

Table A1 provides details of the LEO dataset, by GCSE cohort (based on the year these exams were taken, as discussed above). The first column shows all individuals with an age 16 GCSE record in the NPD who attended school in England.

In column 2 we drop some people who appear in the baseline sample whom we cannot use for our analysis. This is around 10% of the overall population and primarily consists of people with statemented special educational needs who were unable to take the examinations, people who are in the records but were not in Year 11 at school (for example, people who took some GCSE examinations early or did some retakes) and people with lots of missing background data or exam records. This leaves us with a 'usable sample' of between 520,000 and 600,000 individuals per cohort.

Table A1: LEO sample by GCSE year

	Population (1)	Non-missing NPD (2)	Linked (3)	Passed age 16 exams (4)
2002	589,663	521,153	486,717	279,409
2003	621,929	566,279	531,139	296,365
2004	644,873	601,000	569,854	312,579
2005	644,345	601,300	572,970	320,643
2006	653,971	589,383	568,392	325,581
2007	662,225	598,641	577,184	332,322
Total	3,817,006	3,477,756	3,306,256	1,866,899

Note: Column 1 is the full sample of English domiciled pupils in the NPD. Column 2 excludes people with incomplete school records. Column 3 shows the number of those individuals who can be matched to the HMRC tax records. Column 4 shows the number of individuals who passed their age 16 exams (obtained at least five A*-C GCSE grades).

In column 3 we document the match rate to the HMRC tax data. Across the six cohorts around 95% of individuals are linked to the tax data, with match rates going up slightly across cohorts. Individuals never matching to the tax data means that there is never a record of them in the 11 years of tax or benefits data, or - more likely - because matching to the tax records was not possible due to incorrect or missing information.[44] The proportion of individuals who do not match to the

[44] This step was done separately by the Department for Work and Pensions before we had access to the data.

tax data is approximately twice as large for women as it is for men, suggesting that women are more likely to never be in contact with the tax authorities. Aside from this gender difference, we essentially treat these people as missing at random in our analysis.[45]

Finally, column 4 shows the number of people who passed their age 16 exams, as defined by obtaining at least five A*-C grades in GCSE exams. This level of attainment is a near-universal prerequisite for entry to university[46] and we will therefore focus on this group in our analysis, as we only want to include individuals who conceivably had the option of going to higher education in our control sample.[47] We can see that this group represents around 56% of all students with linked HMRC records.

Table A2 shows how the final sample given in column 4 of Table A1 breaks down. Column 2 shows that around a third of those who passed their age 16 exams do not start an undergraduate degree. In column 3, we show the individuals who enter university as mature or part-time students.[48] We define mature students as anyone entering their first undergraduate degree more than three years after leaving school at age 18, while part-time status is a variable we observe in the HESA dataset. Combined, this group is about 6% of the individuals who passed their age 16 exams, and we exclude it from our analysis entirely. The primary reason for this is that we only observe earnings data up to age 30, which limits the number of years mature and part-time students with linked NPD records can possibly have been in the labour market after graduation (for example, someone who started a three-year degree at age 25 would only have had one or two years of labour market experience as a graduate by age 30). The focus of our paper is therefore on the impact of graduating from a full-time university degree started soon after leaving school, which is by far the most common route for obtaining an undergraduate degree. Finally, column 4 shows the individuals with high GCSEs whom we observe doing standard undergraduate degrees in UK universities. This is close to 60% of those passing their age 16 exams, and roughly

[45]In practice these people are more likely to be deprived or from an independent school. However it is a very small share of the overall population and therefore unlikely to affect our conclusions.

[46]Less than 10% of those without five good GCSEs start an undergraduate degree by age 21.

[47]This is less restrictive than Blundell et al. (2005) and Walker and Zhu (2018), who use individuals with at least one A-level as a control group. We take this decision because during our sample period, more than 10% of individuals who attend HE did not take any A levels or other KS5 qualifications.

[48]We also include a very small number of individuals who start their degrees before age 17 in this column, or for whom we only observe a postgraduate qualification. We think it is most likely that the latter individuals have taken an undergraduate qualification abroad and should therefore be excluded from the analysis.

one-third of the overall cohort.[49]

Table A2: LEO sample by GCSE year

	Baseline (1)	No UG (2)	PT/Mature/PG (3)	UG sample (4)
2002	279,409	98,524	20,091	160,794
2003	296,365	102,790	20,483	173,092
2004	312,579	110,091	21,255	181,233
2005	320,643	114,130	19,691	186,822
2006	325,581	110,938	18,093	196,550
2007	332,322	113,112	15,446	203,764
Total	1,866,899	649,585	115,059	1,102,255

Note: Column 1 is taken from Column 4 of Table A1. Columns 2-4 sum to Column 1. PT indicates part-time, PG indicates postgraduate.

[49] Although it is commonly cited that around half of people go to university, only around one-third of these cohorts start a 'standard' undergraduate degree within three years of leaving school.

A2 Subject groups definition

Table A3: Subjects included in each subject group

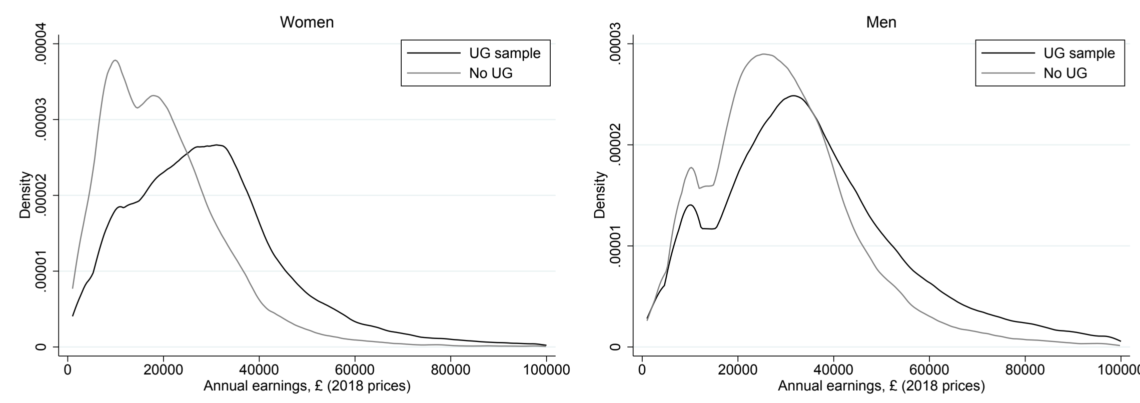
Subject	**Subject group**	**CAH2 code and description**
Agriculture	STEM	(CAH06-01) agriculture, food and related studies
Allied to med	STEM	(CAH02-03) subjects allied to medicine not otherwise specified
Architecture	STEM	(CAH13-01) architecture, building and planning
Biosciences	STEM	(CAH03-01) biosciences
Business	LEM	(CAH17-01) business and management
Chemistry	STEM	(CAH07-02) chemistry
Comms	Other	(CAH18-01) communications and media
Computing	STEM	(CAH11-01) computing
Creative arts	Other	(CAH21-01) creative arts and design
Economics	LEM	(CAH15-02) economics
Education	Other	(CAH22-01) education and teaching
Engineering	STEM	(CAH10-01) engineering
English	Other	(CAH19-01) English studies
Geography	STEM	(CAH12-01) geographical and environmental studies
History	Other	(CAH20-01) history and archaeology
Languages	Other	(CAH19-03) languages, linguistics and classics
Law	LEM	(CAH16-01) law
Maths	STEM	(CAH09-01) mathematical sciences
Medicine	STEM	(CAH01-01) medicine and dentistry
Nursing	STEM	(CAH02-01) nursing
Pharmacology	STEM	(CAH02-02) pharmacology, toxicology and pharmacy
Philosophy	Other	(CAH20-02) philosophy and religious studies
Physics	STEM	(CAH07-01) physics and astronomy
Physsci	STEM	(CAH07-03) physical, material and forensic sciences
Politics	Other	(CAH15-03) politics
Psychology	STEM	(CAH04-01) psychology
Social care	Other	(CAH15-04) health and social care
Sociology	Other	(CAH15-01) sociology, social policy and anthropology
Sportsci	STEM	(CAH03-02) sport and exercise sciences
Technology	STEM	(CAH10-02) technology
Vetsci	STEM	(CAH05-01) veterinary sciences

Note: For sample size reasons we do not include individuals studying: (CAH08-01) general and others in sciences; (CAH14-01); humanities and liberal arts (non-specific); (CAH19-02) Celtic studies; (CAH23-01) combined and general studies. See https://www.hesa.ac.uk/innovation/hecos for more information about the CAH2 subject mapping.

A3 Earnings descriptives

Figure A1 shows the earnings of men and women at age 30 for those who did and did not go to higher education, for those with earnings between £1,000 and £100,000 in the given tax year. We see that men earn more than women and that those who attended HE earn more than those who did not, particularly for women.[50] Average annual taxable earnings of those earning more than £1,000 (including those earning above £100,000) for male graduates are £41,000 versus £31,000 for non-graduates, while the equivalent figures for women are £31,000 and £20,000. The medians are around £2,000 below the mean in all cases except for graduate men where the difference is closer to £6,000. This is due to the very long right hand tail of earnings for graduate men (not shown in the figure).

Figure A1: Real earnings distributions by education level and gender at age 30



Note: Includes the 2002 GCSE cohort in 2016/17, roughly age 30, in the range £1,000 - £100,000. No HE consists of individuals who did not take an undergraduate degree, but passed their age 16 exams (obtaining at least 5 A*-C GCSEs). The tax data includes PAYE and SA earnings.

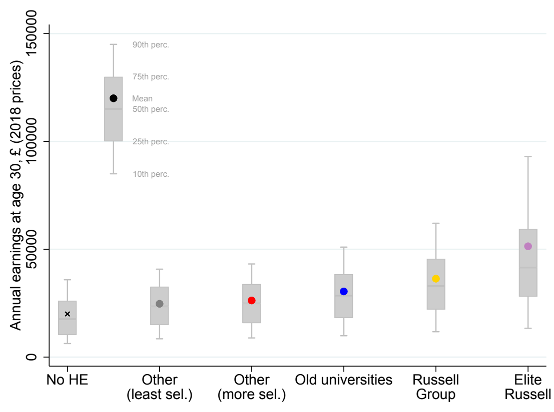
We turn to consider how earnings vary among those who attended HE. Figure A2 shows the earnings distribution of individuals in the different university groupings for men and women separately. We see that average earnings increase with the selectivity of the institutions, with a significant jump for the Elite Russell Group. For men average earnings of those from the most selective universities are more than £75,000, while at the lower end, the earnings distribution for the least selective universities is very similar to that for those who did not attend higher education. For women the differences between the least selective universities and individuals who did not

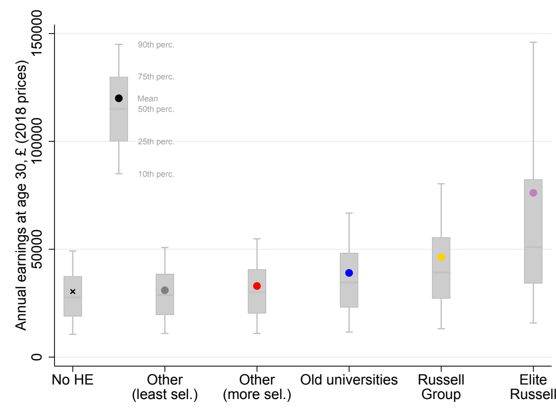
[50]There is also a clear spike in the distribution at around £10,000. This is due to bunching at the income tax and national insurance contribution thresholds.

attend HE are larger.

The figure also highlights various points in the earnings distribution, showing that the variance of earnings increases dramatically with institution selectivity. There is also a significant right hand tail for men from the more selective institutions, with the mean more than £25,000 higher than the median for the Elite Russell Group. Excluding those earning below £1,000, just under 10% of men who attended the Elite Russell Group earn more than £150,000 per year.

Figure A2: Real earnings by HEI type at age 30 - Women (left) and men (right)





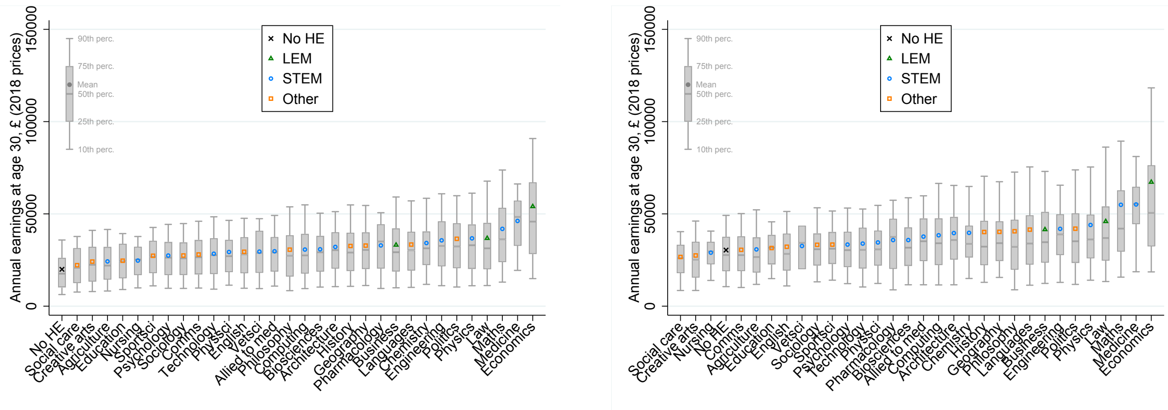
Note: 2002 GCSE cohort in 2016/17, conditioning on earnings being above £1,000. No HE consists of individuals who did not take an undergraduate degree, but passed their age 16 exams (obtaining at least 5 A*-C GCSEs).

Figure A3 then shows the equivalent distributions by individual subject,[51] with the broader subject groups also highlighted. Economics, maths, medicine and law are the subjects with the highest earnings at age 30 for both gender, while social care and creative arts and nursing have the lowest earnings.

All LEM subjects have high earnings, while earnings of STEM graduates are a bit more mixed. Maths, medicine, physics and engineering graduates all have high earnings, but nursing, agriculture, veterinary sciences and psychology graduates do not. We see a similar pattern among 'Other' subjects, with politics, languages, history and geography doing reasonably well, but lower earnings for the remaining subjects. We also note that that the spread of earnings is typically lower in the subjects that feed heavily into public sector careers (and centrally regulated pay scales), such as nursing, education and medicine.

[51] For Veterinary Sciences we drop the tail due to insufficient sample sizes to stay within data disclosure rules.

Figure A3: Real earnings by degree subject at age 30 - Women (left) and men (right)



Note: 2002 GCSE cohort in 2016/17, conditioning on earnings being above £1,000. No HE consists of individuals who did not take an undergraduate degree, but passed their age 16 exams (obtaining at least 5 A*-C GCSEs).

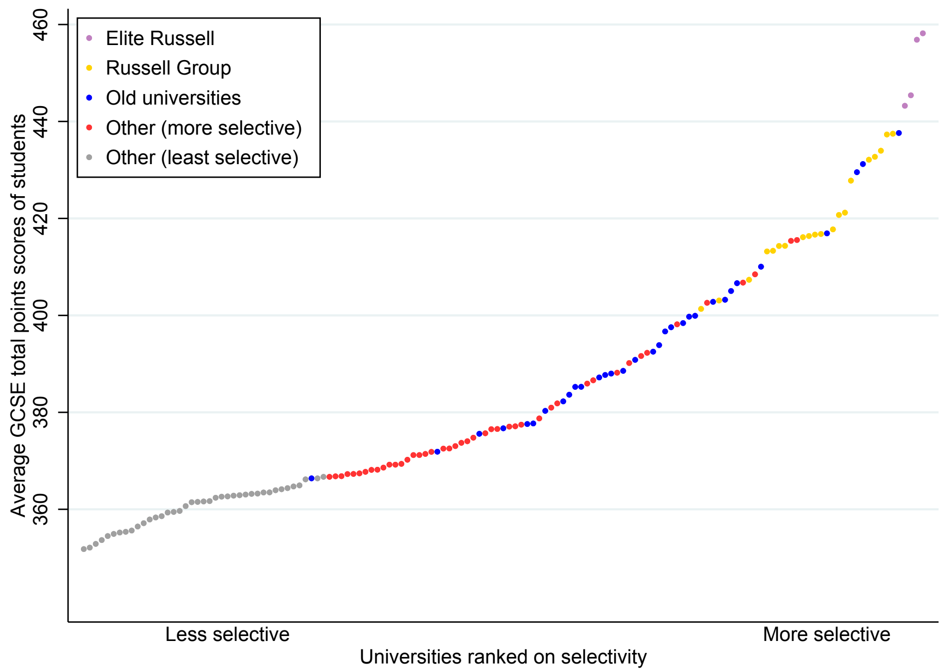
A4 Selectivity

As we described in Section 3.3, the higher education system in the UK is highly selective, meaning the highest status universities take the students with the highest prior attainment. We display this feature visually in Figure A4, which plots the average GCSE points scores of the students of the different universities.[52] On the y axis 6 points is one grade higher in one exam - 100 points is therefore a substantial difference of around 17 grades across all GCSEs taken (students typically take around ten). We see that the 'Elite Russell' group is by far the most selective group, followed by the rest of the Russell Group, although there is some overlap with some of the old universities and more selective other institutions. The least selective other institutions are all at the bottom by construction, as we defined the Other group of universities based on GCSE scores.[53] There are also large differences in average prior attainment for people doing different subjects - see Figure A5.

[52] We show total GCSE points, based on the following: A* = 58 points, A = 52 points, B = 46 points, C = 40 points, D = 34 points, E = 28 points, F = 22 points and G = 16 points (there is also an 'Unclassified', U grade which is worth 0 points). The total GCSE points score is then obtained by adding up the points for the different subjects.

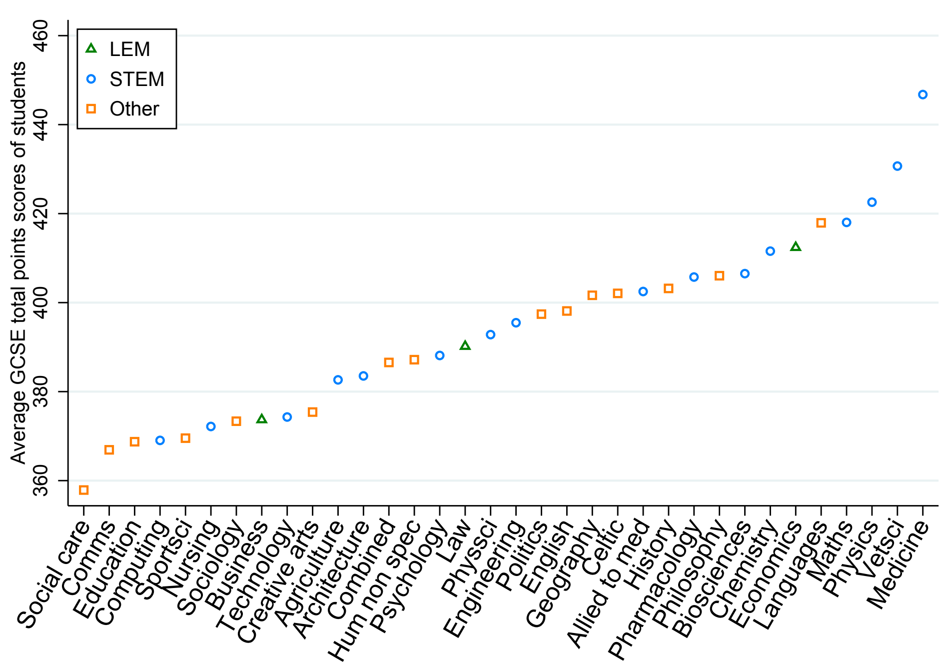
[53] It is important to note that we refer to the universities' selectivity taking into account age 16 test scores only. In fact, several of the universities here will be selective on other metrics such as music ability or arts portfolios. We do not account for that here.

Figure A4: Selectivity by university at age 30



Note: Selectivity is based on the average total GCSE points scores of each institutions' full-time, non-mature students from the 2004-2007 GCSE cohorts.

Figure A5: Selectivity by subject at age 30

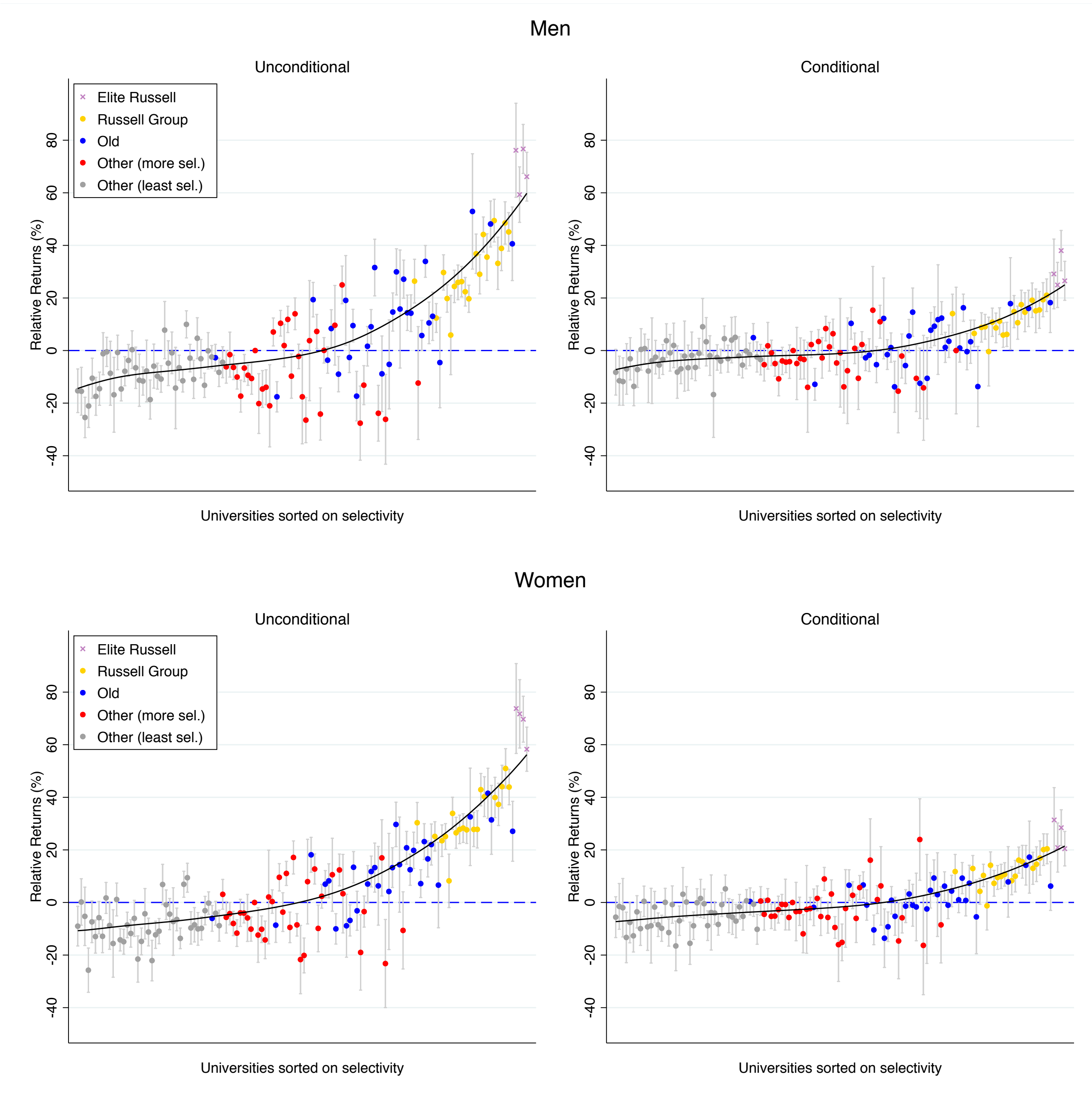


Note: Selectivity is based on the average total GCSE points scores of each subjects' full-time, non-mature students from the 2004-2007 GCSE cohorts.

A5 Additional results

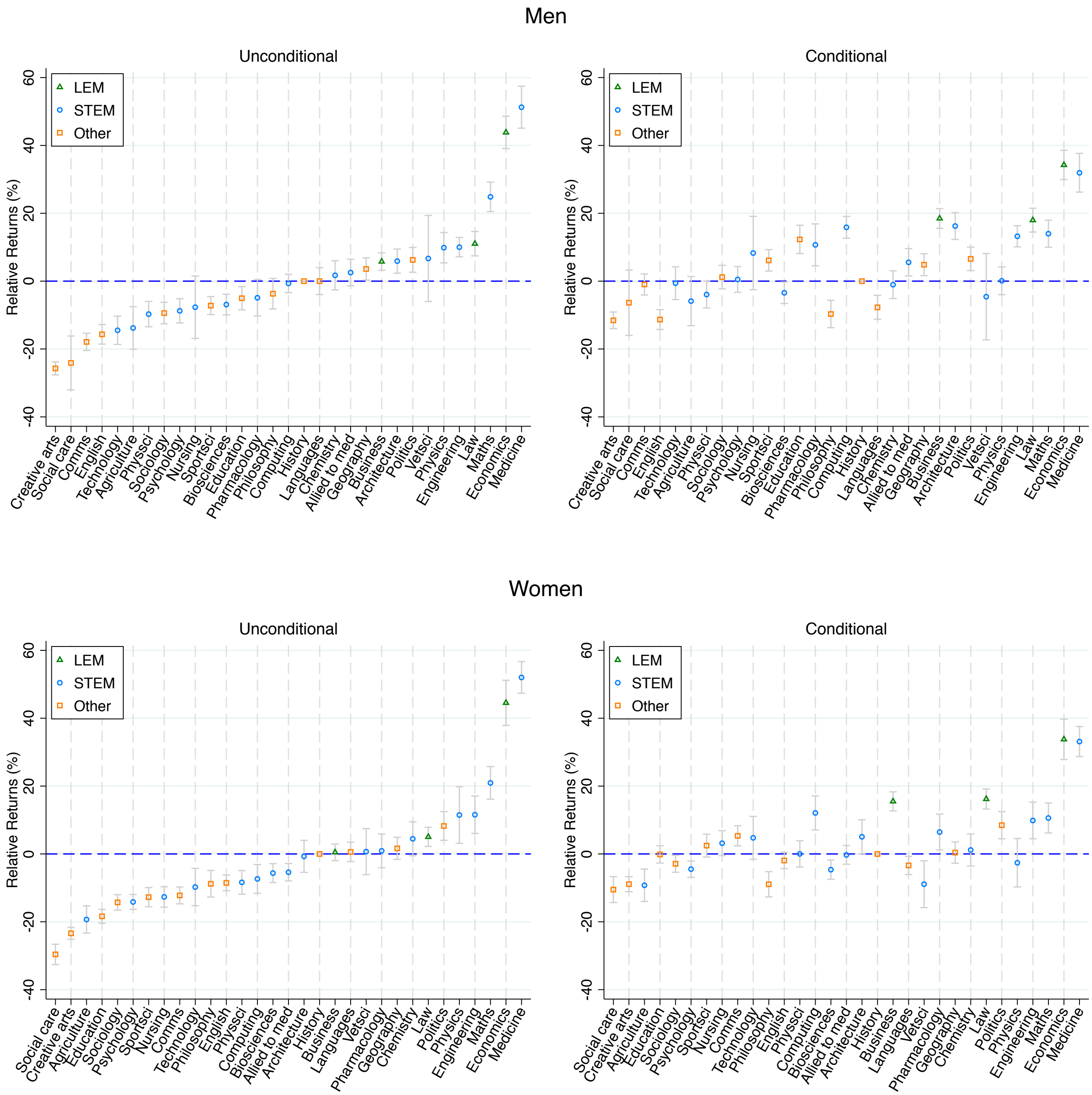
A5.1 Relative returns estimates by gender

Figure A6: Estimated returns at age 30 by institution



Note: Equivalent to Figure 2, split by gender.

Figure A7: Estimated returns at age 30 by subject



Note: Equivalent to Figure 3, split by gender.

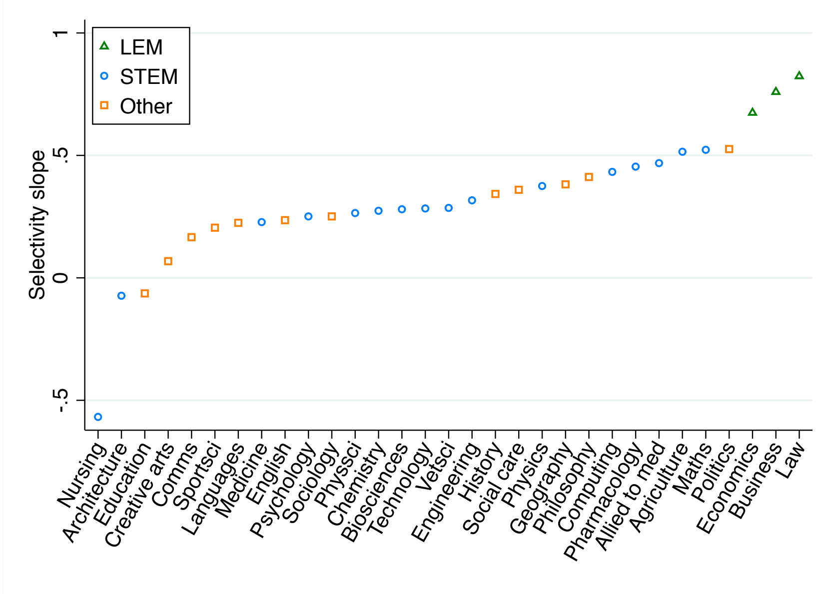
A5.2 Additional degree returns results

Table A4: Best and worst performing degrees

University	Subject	Relative returns (%)
University of Cambridge	Economics	127.0
Oxford University	Business	122.7
University of Cambridge	Computing	115.3
University College London	Economics	108.1
University of Cambridge	Law	107.5
University of St Andrews	Economics	96.7
University of Warwick	Economics	95.9
Oxford University	Economics	94.1
Oxford University	Law	91.7
University of Aberdeen	Medicine	88.3
School of Oriental and African Studies	Philosophy	-52.7
Roehampton University	Social care	-44.5
University of Gloucestershire	Social care	-42.6
University of St Mark & St John	Social care	-41.5
University of Central Lancashire	Philosophy	-37.7
University of Wolverhampton	Politics	-37.2
University of Worcester	Allied to med	-36.5
Roehampton University	Allied to med	-35.7
University of Glamorgan	Psychology	-35.6
London Metropolitan University	Politics	-35.4

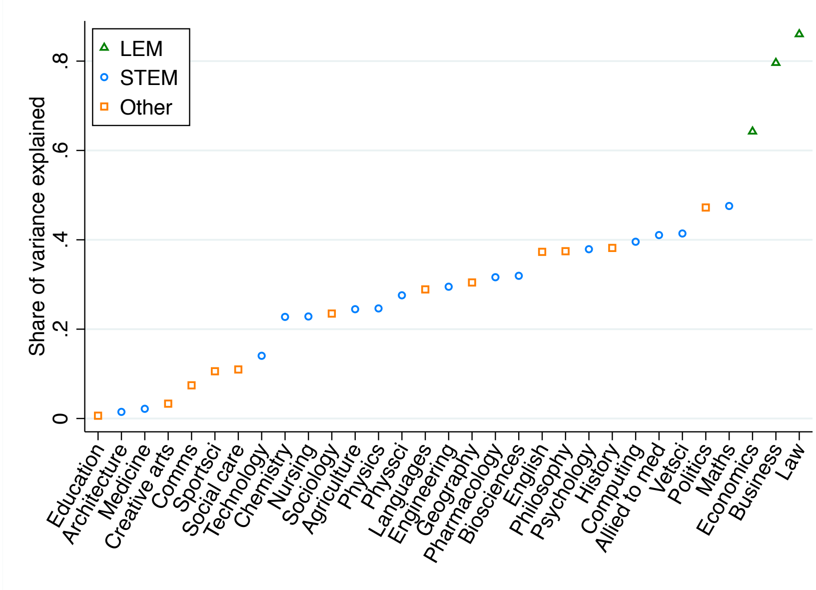
Note: Selected estimates of relative returns (in percentage points) from Figure 4. Returns are relative to History at Sheffield Hallam University.

Figure A8: Returns-selectivity relationship by subject



Notes: Figure shows, for each subject, the slope of a degree-level regression of earnings returns on average GCSE score of the degree intake.

Figure A9: Goodness-of-fit by subject



Notes: Figure shows, for each subject, the R^2 of a course-level regression of earnings returns on average GCSE score of the course intake. Subjects are coloured according to their broad subject group (LEM, STEM, Other). Slope coefficients of these regressions are shown in Appendix Figure A8.

Table A5: Correlations with main institution returns of alternative specifications

	Excl. dropouts	Cross-sectional	Shrinkage	Within uni group
Elite Russell	0.998	0.997	0.997	0.983
Old universities	0.979	0.975	0.984	0.947
Other (more selective)	0.980	0.956	0.938	0.981
Other (least selective)	0.956	0.941	0.982	0.930

Note: Column (1) shows the correlation of our main institution returns with returns estimated on a sample which excludes individuals who did not graduate from their degree. Column (2) shows the correlation with institution returns estimated at age 30, on a cross-sectional sample only. Column (3) shows the correlation with the institution returns after shrinkage has been applied, where we shrink degrees returns to the average degree return. Column (4) shows the correlation with institution returns estimated within subsamples for each university group.

Table A6: Correlations of selectivity with returns for alternative specifications

	Main results	Excl. school FEs	Excl. school FEs and background
LEM			
Business	0.892	0.869	0.890
Economics	0.801	0.838	0.822
Law	0.927	0.921	0.925
STEM			
Agriculture	0.495	0.496	0.491
Allied to med	0.641	0.676	0.656
Architecture	-0.122	-0.102	-0.112
Biosciences	0.565	0.639	0.588
Chemistry	0.477	0.590	0.519
Computing	0.629	0.626	0.620
Engineering	0.543	0.559	0.558
Maths	0.690	0.713	0.707
Medicine	0.147	0.205	0.176
Nursing	-0.478	-0.440	-0.487
Pharmacology	0.562	0.577	0.553
Physics	0.496	0.516	0.507
Physsci	0.525	0.619	0.570
Psychology	0.616	0.649	0.639
Technology	0.375	0.370	0.360
Vetsci	0.644	0.728	0.696
Other			
Comms	0.273	0.285	0.288
Creative arts	0.182	0.264	0.230
Education	-0.079	-0.035	-0.074
English	0.611	0.651	0.633
Geography	0.552	0.556	0.550
History	0.618	0.637	0.628
Languages	0.538	0.555	0.554
Philosophy	0.612	0.650	0.642
Politics	0.687	0.716	0.702
Social care	0.331	0.445	0.359
Sociology	0.485	0.553	0.516
Sportsci	0.325	0.307	0.322

Note: Column (1) shows the correlation of selectivity with our main degree returns. Column (2) shows the correlation of selectivity with degree returns estimated when excluding school FEs from the controls. Column (3) shows the correlation of selectivity with degree returns estimated when both school FEs and individual level background characteristics are excluded from the controls.

Table A7: Correlations with main degree returns of alternative specifications

	Excl. dropouts	Cross-sectional	Shrinkage	Within subject
LEM				
Business	0.985	0.976	0.983	0.990
Economics	0.983	0.962	0.974	0.924
Law	0.990	0.984	0.997	0.982
STEM				
Agriculture	0.987	0.973	0.996	0.947
Allied to med	0.985	0.955	0.985	0.977
Architecture	0.934	0.916	0.985	0.949
Biosciences	0.960	0.925	0.989	0.943
Chemistry	0.963	0.901	0.992	0.855
Computing	0.958	0.938	0.988	0.964
Engineering	0.934	0.944	0.982	0.966
Maths	0.949	0.914	0.986	0.871
Medicine	0.973	0.914	0.925	0.860
Nursing	0.936	0.899	0.991	0.885
Pharmacology	0.957	0.792	0.993	0.805
Physics	0.953	0.944	0.993	0.886
Physsci	0.915	0.881	0.976	0.942
Psychology	0.932	0.932	0.989	0.957
Technology	0.951	0.933	0.990	0.874
Vetsci	0.980	0.993	0.996	0.888
Other				
Comms	0.935	0.905	0.986	0.962
Creative arts	0.954	0.903	0.955	0.907
Education	0.968	0.946	0.976	0.965
English	0.960	0.892	0.986	0.960
Geography	0.983	0.963	0.987	0.978
History	0.979	0.943	0.993	0.936
Languages	0.919	0.894	0.980	0.911
Philosophy	0.970	0.942	0.986	0.952
Politics	0.972	0.960	0.985	0.982
Social care	0.906	0.932	0.968	0.886
Sociology	0.932	0.961	0.985	0.953
Sportsci	0.944	0.882	0.995	0.952

Note: Column (1) shows the correlation of our main degree returns with returns estimated on a sample which excludes individuals who did not graduate from their degree. Column (2) shows the correlation with degree returns estimated at age 30, on a cross-sectional sample only. Column (3) shows the correlation with the degrees returns after shrinkage has been applied, where we shrink degrees returns to the average degree return. Column (4) shows the correlation with course returns estimated within subsamples for each subject.