

# How much does degree choice matter?\*

Jack Britton (*Institute for Fiscal Studies*)<sup>†</sup>

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

We use a large and novel administrative dataset to investigate returns to different university ‘degrees’ (subject-institution combinations) in the United Kingdom. Conditioning on a rich set of background characteristics, we find substantial variation in returns across degrees with similar selectivity levels, suggesting students’ degree choices matter a lot for later-life earnings. Returns increase with university selectivity much more at the top of the selectivity distribution than further down, and much more for some subjects than others. Returns are poorly correlated with observable degree characteristics other than selectivity, which could have important 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.

<sup>†</sup>Corresponding author, Email: [jack.britton@ifs.org.uk](mailto: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 is the first to exploit an innovative new administrative dataset to look at labour market outcomes at the degree level - that is, the combination 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 for degrees that are similarly selective. This implies degree choice matters much more than some of the previous evidence has suggested. Slightly more than half of the variation in returns is within subject, across institutions. We find only a weak relationship between selectivity and returns through much of the selectivity distribution, but a strong positive relationship at the top end. This suggests 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 commonly available measures of degree quality, such as student satisfaction ratings and early career earnings, are not well related to returns. Such measures might therefore mislead student choice and distort university incentives.

We use 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 roughly 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 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 studies have been able to use. We exploit this to test the likely role of unobservable factors, such as preferences for factors other than the monetary returns to degrees (which Wiswall and Zafar, 2014, have shown to be important drivers of choices, and may also be correlated with subsequent earnings outcomes), in driving our results. We show that our headline findings are robust to a bounding exercise, following Oster (2019), suggesting that unobservable factors are not likely to affect our main conclusions. We also show that the main findings are robust to the exclusion of subsets of our control variables and are also insensitive to reasonable changes in the sample selected or the regression specification that we use. We also show that we would get different conclusions if we were to estimate degree effects as being equal to the sum of subject and institution effects. This highlights the value of our data, as while the additive approach could be estimated using smaller datasets, it is only possible to obtain reliable degree-level estimates using large scale dataset such as ours.

We start by estimating the earnings returns to different subjects and universities. We estimate returns relative to a base case (history for subjects and Sheffield Hallam for universities), within the set of people who attend. We find that medicine, economics and law are amongst the subjects with the highest earnings returns, while social care and creative arts are the subjects with the lowest returns. We find a weak association between university selectivity (as measured by the average prior attainment of people at the university) and returns amongst the bottom half of the selectivity distribution, but a much stronger relationship amongst the top half, suggesting there are 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 then turn to the most novel contribution of the paper and estimate returns at the ‘degree’ level, which is the interaction of institution and subject. We can 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 dramatic variation in raw earnings outcomes across different de-

degrees: the standard deviation of the degree-level fixed effects, without any controls, is 32 percentage points and the 90:10 range is 75 percentage points. These figures drop to 22 and 52 percentage points respectively once we estimate degree-level fixed effects controlling 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 (as measured by the average prior attainment of students at the degree level), the standard deviation in returns is still more than 16 percentage points, increasing to 27 percentage points 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 65% of the variation for the most selective degrees. Combined, these results suggest that degree choice is crucial for subsequent earnings outcomes, right across the selectivity distribution. This holds even holding subject choice fixed - it is not at all uncommon to see differences in returns of 30 percentage points between degrees in the same subject at similarly selective universities. We highlight the value of our dataset by showing that the variation in estimated returns and the returns to selectivity would be understated (and the share of variation that is within-subject overstated) if we were to simply estimate degree returns as the sum of institution and course returns.

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, other than selectivity, existing measures of degree quality are not well correlated with returns once we control for selectivity. 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.<sup>1</sup> We then show that other measures of degree quality including publicly available subject-specific university rankings, completion rates and de-

---

<sup>1</sup>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.

gree performance are all correlated with returns, but this correlation 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, early career earnings and wages in the surrounding area are also 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 inequality, 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. (2021) 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 Section 5. We start by showing heterogeneity in returns across institutions and subjects before turning to our estimates of degree level returns. In the final part of the results section we look at the relationship between degree level returns and selectivity, as well as with other observable characteristics of the degree. Section 6 concludes.

## **2 Literature**

Our work contributes to a substantial academic literature on variation in the returns to different types of higher education, which typically looks at how returns vary by what *or* where people

studied, or both separately, but very rarely the interaction of the two. Altonji et al. (2012) reviews the evidence on the returns to field choices, highlighting that the majority of papers estimating returns assume selection on observables. Walker and Zhu (2011), Walker and Zhu (2018) and Chevalier (2011) all use this approach to study returns in the UK. Our findings are broadly consistent with these papers, although we provide a richer set of estimates as each of those papers is based on survey data with much smaller sample sizes and a more parsimonious set of covariates.

Hastings et al. (2013) and Kirkeboen et al. (2016) exploit discontinuities in entry cutoffs to identify returns to a range of different subjects. The latter argues that returns based on OLS regressions would overstate the causal effects, although Bleemer and Mehta (2022) finds that returns estimates for economics using a regression discontinuity approach are very similar to observational earnings differences. We are therefore slightly cautious about our cross-subject returns estimates, but believe that our results still provide an important contribution, especially given our data on subject-specific prior attainment.

Due to data limitations, the literature on where people study frequently 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). Some papers from outside the UK have been able to estimate returns at the institution level (Cunha and Miller, 2014; Kirkeboen et al., 2016; Mountjoy and Hickman, 2020; Chetty et al., 2020), typically identifying returns based on OLS estimation with rich background characteristics. Some of these papers 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, via 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 and find their results to be consistent with those obtained using OLS conditioning on rich observables, without con-

trolling 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 (although this is in Chile, which is quite different institutional setting to our own). Dillon and Smith (2020) make a similar argument in a recent paper that focuses on match effects in higher education in the United States. Our finding that publicly available metrics that likely influence the choices of students in the UK are not well correlated with earnings (once we control for selectivity) supports our own argument that selection on observables is a credible assumption.

The evidence on the relationship between returns and university selectivity is mixed. Dale and Krueger (2002, 2014) and Mountjoy and Hickman (2020), which all control for the application sets of students in the United States, and Kirkeboen et al. (2016), which exploits discontinuities at entry thresholds in Norway (where university selectivity is not likely to be as important as the UK), all suggest a weak relationship. However, the UK evidence consistently finds a strong relationship between university selectivity and returns (for example, Walker and Zhu, 2018), as does Cunha and Miller (2014), which uses a similar approach to Mountjoy and Hickman (2020) and finds a strong relationship for universities in Texas. Several other 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). We find a weak association between selectivity and returns throughout 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.<sup>2</sup>

However, our main contribution to the literature is to investigate returns at the degree level (namely, the combination of subject and institution). Hastings et al. (2013) is the only other paper to have done this, and therefore the only paper which has been able to look at how returns to selectivity vary by subject. Like our paper, they find that institution selectivity matters much more

---

<sup>2</sup>The discrepancy between our 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), which has been shown to be an issue (Foote and Stange, 2022). This is dramatically less important in the UK, where very few students work abroad after graduation - Britton et al. (2019) estimate that between one and ten years from graduation, only around 1.5% of each student cohort is living outside the UK.

in some subject areas than others. We show that for subjects with high average returns, selectivity matters a lot, while for subjects with low returns it does not. This is an important result as it suggests that the focus on getting students into the best institution possible regardless of subject is misplaced. We extend upon Hastings et al. (2013) by also highlighting the variation in returns to different degrees, which reveals the extent to which degree choices can potentially impact later-life outcomes.<sup>3</sup> Further, we also look more carefully at the relationship between observable degree-level characteristics and returns.<sup>4</sup> This enables us to consider the relevance of information on degree quality that is available to students and regulators before data on earnings at age 30 become available. Our finding that the measures of degree quality that we consider are not related to returns once we control for selectivity is important, as we know that there is substantial variation in returns even for degrees with similar selectivity. The result also adds weight to our findings as if students make choices based on this information (and there is some evidence that they do; Gibbons et al., 2015), then students are not likely to be selecting into courses based on earnings potential, reducing concerns about selection on unobservables driving our results. Finally, in addition to influencing student choices, this information can also affect the regulator’s ratings of teaching quality, and the priorities and incentives of universities.

### **3 Data and institutional background**

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, giving detail on each of the component parts of LEO, alongside describing the institutional background.

---

<sup>3</sup>More generally, 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.

<sup>4</sup>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 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) took their age 16 exams between 2002 and 2007; (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. 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 exam files as well as a census of English schools. In England, students take national, externally marked examinations at age 11, 16, 17 and 18, and we have all four 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 at the end of secondary school, which is in the summer of the school year people turn 16.<sup>5</sup> 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 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, adding together scores from each of the exams taken.<sup>6</sup> The age 17 and 18 assessment data are based primarily on scores in AS (age 17) and A-level (age 18) exams, which are usually taken one and two years after GCSEs. Students typically study three or four subjects and received grades from A-F. Again, we observe the subjects taken

---

<sup>5</sup>The school year in England runs from September 1 to August 31. People whose birthday is in July or August will take their GCSE exams just before turning 16.

<sup>6</sup>For this, 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.

and the grades achieved. Students could also take vocational qualifications such as in hospitality or retail during this period alongside or instead of A-levels. We also observe grades these courses in the data.

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.<sup>7</sup> Following several previous papers (e.g. Chowdry et al., 2013; Campbell et al., 2021), we combine these multiple measures into one continuous index of socio economic status (SES) at age 16 using principal components analysis.<sup>8</sup> The approximately 7% of pupils who attend private secondary schools are missing the School Census data (but we do observe their exam records).<sup>9</sup> We keep this group in the analysis and include missing dummies for any missing School Census information.

### 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 graduated from their degree.<sup>10</sup>

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

---

<sup>7</sup>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).

<sup>8</sup>Campbell et al. (2021) show that this measure of SES aligns well with an alternative measure of parental SES from a linked data source.

<sup>9</sup>Some students who also attended a private primary school have no age 11 exam records, but these students do all have age 16, 17 and age 18 exam records.

<sup>10</sup>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.

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.

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'.<sup>11</sup> To help summarise our findings, we sometimes further group these subjects in three groups: LEM (Law, Economics and Management/Business), STEM (Science, Technology, Engineering and Mathematics), and Other, which consists of other social sciences, arts and humanities subjects.

Individuals attend one of more than 100 UK universities which provide undergraduate degrees. We are interested in the relationship between returns and university (and indeed, degree) selectivity. To determine selectivity, we take the average GCSE scores of the pupils admitted. This is not a perfect reflection of how universities actually select students in practice, as this is often based on A-Level grades, and sometimes other factors, like musical ability, but it results in a ranking of institutions that look very sensible to people familiar with the system (with the Universities of Oxford and Cambridge clearly at the top, for example).<sup>12</sup>

For some of our analysis we also classify universities into five broader groups based on a combination of existing well-known groupings and their selectivity. We have the "Russell Group", which is a self-selected consortium of 24 relatively high-status universities that are amongst the most selective institutions in the country.<sup>13</sup> From this group, we pull out the four most selective universities (namely, the Universities of Oxford and Cambridge, the London School of Economics,

---

<sup>11</sup>For a complete list of these, see <https://www.hesa.ac.uk/innovation/hecos>.

<sup>12</sup>The average selectivity of the different universities (and subjects) can be seen in Section A2 in the Appendix.

<sup>13</sup>This is clearly captured by our measure of selectivity, as seen in Figure A1 in the Appendix.

and Imperial College London) and refer to those universities as the “Elite Russell Group”. We then have the group of “Old universities” which includes the 31 non-Russell Group institutions which existed before 1992, when there was a large expansion in the number of universities in England. The remaining universities are non-traditional universities, such as arts 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 (‘Other, more selective’ and ‘Other, 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. 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,<sup>14</sup> we only make use of the data from 2013/14 onwards. This has the additional advantage of avoiding the immediate labour market impact of the 2008 recession. The tax data only includes information 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 (DWP). They hard link the tax and DWP datasets via individual’s unique National Insurance Number (which is equivalent to the US Social Security Number), while they employ fuzzy matching between National Insurance records and the school records using first name, surname, date of birth, postcode and gender.<sup>15</sup>

The earliest cohort for which we have individual level school records took age 16 exams in the summer of 2002. Almost all of these individuals were born between September 1, 1985 and August 31, 1986, meaning they were aged 30 midway through the final tax year for which we have earnings records (2016/17).<sup>16</sup>

---

<sup>14</sup>By age 30, around 10% of individuals in our sample have self-employment income.

<sup>15</sup>This is necessary as National Insurance Numbers are not included in the NPD records. As discussed in Appendix Section A1, match rates are around 95% for the overall population, and more than 99% for graduates.

<sup>16</sup>Students born between September 1, 1985 and August 31, 1986 were all 16 on August 31, 2002, and therefore 30

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 Sample summary

Table 1 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. Around 20,000 men attended one of the four Elite Russell group universities compared to around 17,000 women.<sup>17</sup>

Table 1: Number of students by university and subject groups

	Women		Men	
	N	Share	N	Share
<i>University group</i>				
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
<i>Subject group</i>				
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

Notes: 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'.

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). A little under 20% of students took LEM courses, while around 35% took STEM courses and just under half of students

on August 31, 2016. This is approximately midway through the 2016/17 tax year which runs from April to April. Individuals who sat their age 16 exams in 2002 but skipped a year in school or were held back a year might be born outside of these dates, but in practice this is the case for less than 1% of our sample.

<sup>17</sup>One reason for this is that two of 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.

studied Other subjects, which consist of arts, humanities and social science degrees (excluding economics). Women were much more likely to study the Other subjects (54% vs 40%) and are less likely than men to study LEM (16% vs 21%) and STEM degrees (30% vs 39%). Basic descriptives of our sample and summaries of raw earnings outcomes by subject and institution are provided Appendix Section A3.

## 4 Earnings model

We estimate the earnings returns to studying different subjects, attending different universities, and finally, studying different degrees. All of our estimates are relative to a base case within higher education. That is, we focus on relative returns to different types of higher education, conditional on people attending.<sup>18</sup> Our identification strategy relies on the assumption that, conditional on attending, selection into different higher education routes is based on observable characteristics. The basic premise follows much of the returns to education literature (for example, Blundell et al., 2000) by estimating one of the following regression models:

$$\ln(y_{it}) = \alpha + X_i' \gamma + \sum_j \beta_j^S S_{ji} + \sum_j \beta_j^U U_{ji} + \epsilon_{it} \quad (1)$$

$$\ln(y_{it}) = \alpha^D + X_i' \gamma^D + \sum_j \beta_j^D D_{ji} + \epsilon_{it}^D \quad (2)$$

where  $S_{ji}$  is an indicator for the subject,  $U_{ji}$  for the university, or  $D_{ji}$  for the 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  $t$ .<sup>19</sup>

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. That is, we assume:

$$\text{cov}(S_{ji}, \epsilon_{it} | X_{it}) = \text{cov}(U_{ji}, \epsilon_{it} | X_{it}) = 0 \quad \forall j \quad (3)$$

<sup>18</sup>This decision affects the interpretation of our findings. By focusing on relative returns, our estimates apply for those who chose to go to higher education, but they may not be reliable 'out-of-sample' predictions for non-attendees.

<sup>19</sup>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.

and

$$\text{cov}(D_{ji}, \epsilon_{it}^D | X_{it}) = 0 \quad \forall j \quad (4)$$

which says that conditional on the control variables  $X$  there is no correlation between the earnings residual and the type of higher education studied. The challenge in estimating the earnings returns by type of higher education is therefore to account for all the differences between individuals that might affect both their decision of where to enrol and what to study and their earnings prospects.

#### 4.1 Pooled earnings model

We start by documenting the regression specification that we use, which extends the model given by equation 2. The oldest cohort in our sample, the 2002 GCSE cohort, are last observed 14 years after their GCSE exams, in the 2016/17 tax year, our last year of data. For our headline estimates we report returns 14 years after GCSEs in order to allow for growth in returns with age as much as possible while keeping our estimates within sample. As described above, we refer to this as ‘age 30’ (and similarly for other years) for simplicity as almost everybody in the cohort will be aged 30 midway through the tax year.

To avoid relying solely on observations from one cohort of students, we include several cohorts of students and 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  $c \in \{2002, \dots, 2007\}$  at time  $t \in \{-5, \dots, 0\}$ , 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), for our estimation of degree returns we model log real earnings as follows:<sup>20</sup>

$$\begin{aligned} \ln(y_{ict}) = & X_i' \gamma + \rho I(\text{age}_{start} > 18) + \omega_1 t + \omega_2 t^2 + \sum_{c=2003}^{2007} \delta_c C_c \\ & + \sum_j \beta_j D_{ji} + \sum_j \beta_{1j} (D_{ji} t) + \sum_j \beta_{2j} (D_{ji} t^2) + \epsilon_{ict} \end{aligned} \quad (5)$$

---

<sup>20</sup>This is a straightforward extension of equation 2 above. We extend upon equation 1 in the equivalent way.

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 ( $C_c$ ) based on GCSE year (with 2002 the omitted category), the treatment of interest ( $D_i$ ; namely, which degree was studied), a treatment-specific quadratic trend in age ( $D_i f(t)$ ) and a random component ( $\epsilon_{ict}$ ).

We exclude individuals still in education or with earnings below £1,000.<sup>21</sup> We further Winsorize 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.<sup>22</sup> 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 5.<sup>23</sup> 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. As mentioned above, when estimating equation 5, we only include individuals who attended higher education and estimate returns relative to a base case.<sup>24</sup>

## 4.2 Control variables

We are able to control for a rich set of covariates due to the detailed 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, this includes gender and, for all children who attended a state secondary school (about 93% of each cohort), 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

---

<sup>21</sup>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.

<sup>22</sup>We solely observe annual earnings in the tax records, and do not observe hours worked.

<sup>23</sup>This means for our headline estimates, the coefficient of interest is  $\beta_j$ . 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.

<sup>24</sup>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.



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), 17 (AS-level) and 18 (A-level) as described earlier, as well as number of subjects taken and subject mix.<sup>25</sup> 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.<sup>26</sup> 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. 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 Universities 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 build sequential common support, and depend on functional form assumptions for identification of returns to elite institutions compared to attending the least selective universities.<sup>27</sup>

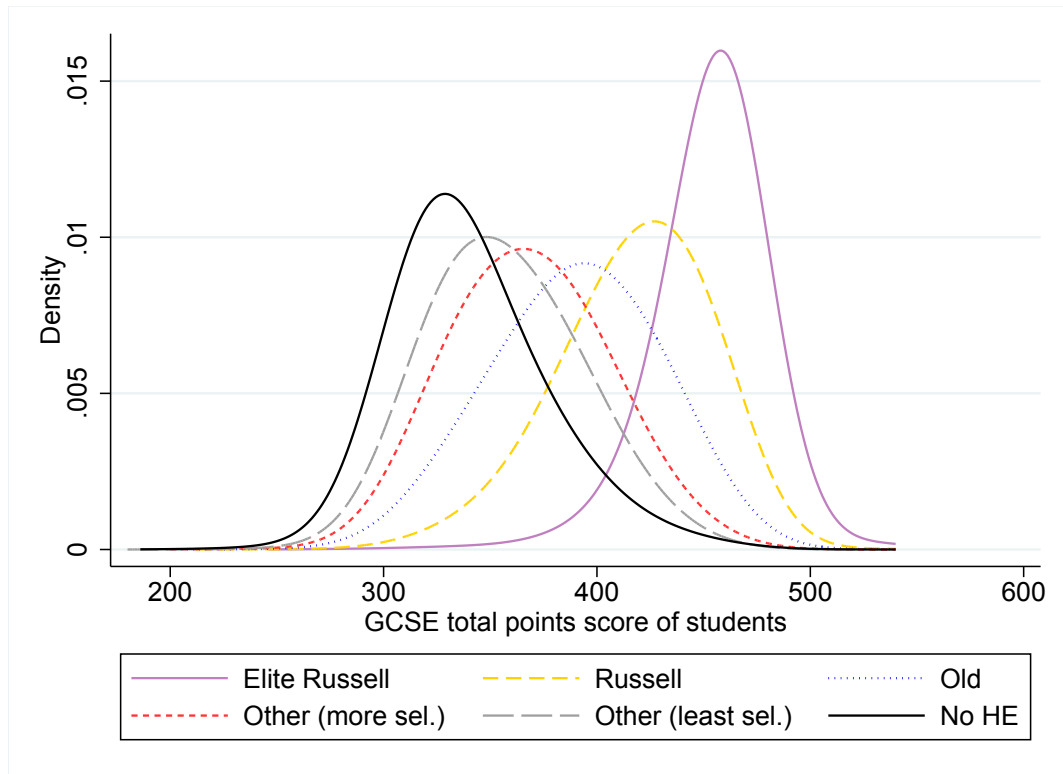
---

<sup>25</sup>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 qualification above GCSE level; a cubic in total A-level point scores; 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.

<sup>26</sup>We are unable to control for non-cognitive skills. However Buchmueller and Walker (2020) estimate a 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.

<sup>27</sup>A very similar argument to this is made in Hoxby (2018). We test the robustness of our results to alternative

Figure 1: Distribution of GCSE points score by university group



Notes: 2004 GCSE cohort only. No HE consists of individuals who did not take an undergraduate degree, but passed their end-of-school (age 16) exams. This is provided as a reference point only, as we do not include these individuals in the estimation. Means and variances for each group are provided in Appendix Table A4.

## 5 Returns by subject, institution and degree

This section presents our main results. We start by showing how returns vary across different subjects and universities before turning to returns by degree. Finally, we consider how well other indicators of university quality are correlated with returns.

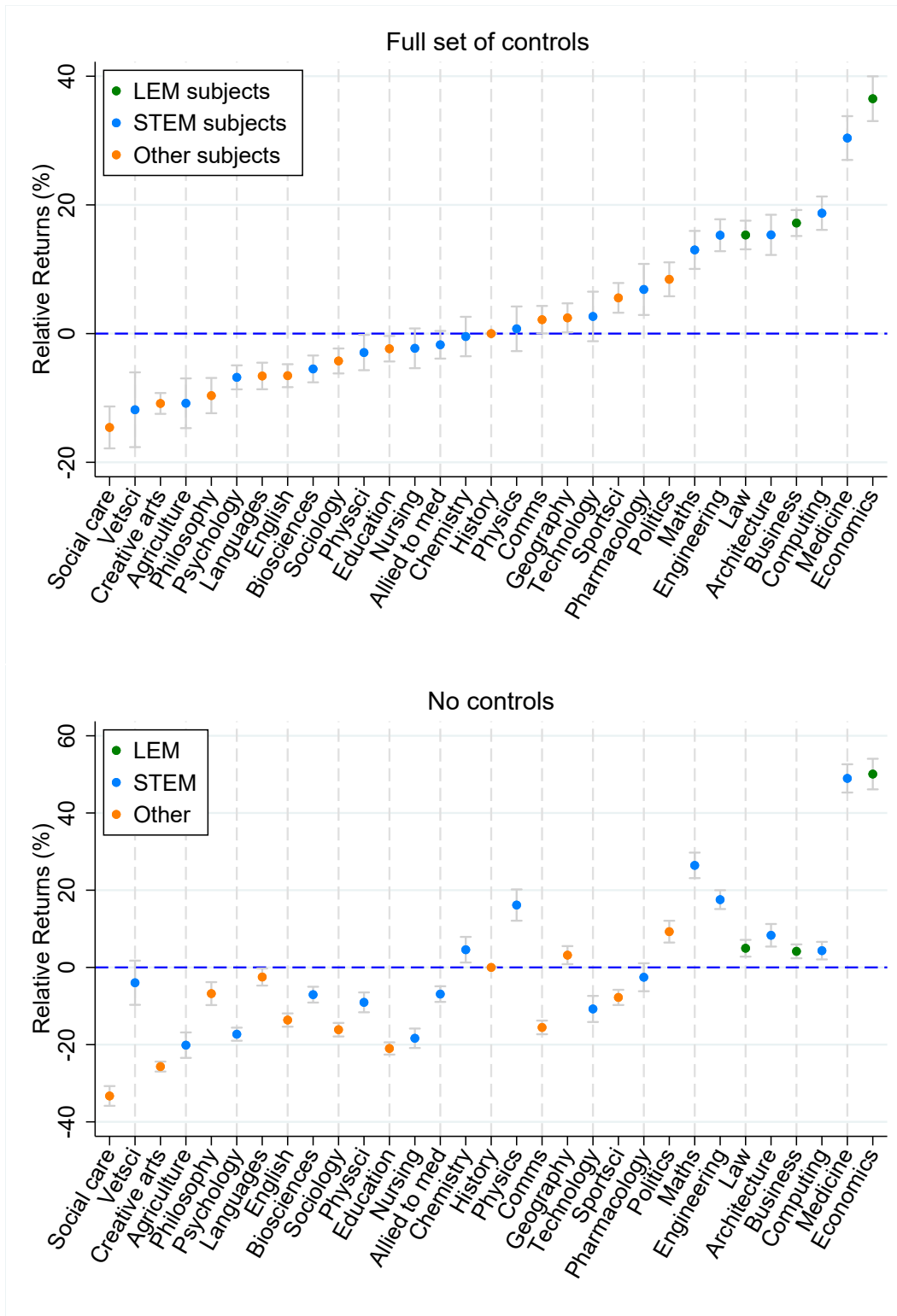
### 5.1 Subject returns

We start by looking at returns by subject in Figure 2, relative to the base case subject of history. The top panel shows the estimates from our final regression model, with the full set of controls (with the point estimates converted into percentage terms).<sup>28</sup>

functional form assumptions.

<sup>28</sup>These estimates are extremely similar when we look at men and women separately. See Figure A5 in the Appendix.

Figure 2: Estimated returns at age 30 by subject



Notes: Estimates have been converted into percentages and are relative to studying History. 95% confidence intervals are shown with standard errors clustered at the individual level.

There is significant variation in returns across subjects. Economics and medicine are the best

performing subjects, being associated with earnings boosts of around 36% and around 30% relative to history, respectively. We also 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 relatively poorly.

There is an interesting pattern in the returns across our different broad subject groups. In general, the three LEM subjects do very well, the 'Other' subjects (which mostly consist of arts, languages and humanities) tend to do relatively poorly, while the returns for STEM subjects are 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.

The bottom panel shows the subject estimates without any control variables.<sup>29</sup> Overall, the controls make quite a big difference to the distribution of subject fixed effects. 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. We also see fairly large upward shifts in relative returns estimates for some of the lowest earning subjects, particularly social care, creative arts, communications and education. Notably, the returns for law, business and computing increase considerably, from around 5% to between 15 and 20% .

## 5.2 University returns

Figure 3 then shows our estimates of the relative returns to different higher education institutions.<sup>30</sup> The institution fixed effects are all shown relative to the base case of Sheffield Hallam University, which is a large, mid-ranking institution. Again, the estimates have been converted into percentage terms from log points, and now institutions are sorted on their selectivity rank (which is discussed in more detail in Section 3). The 'unconditional' estimates on the left include

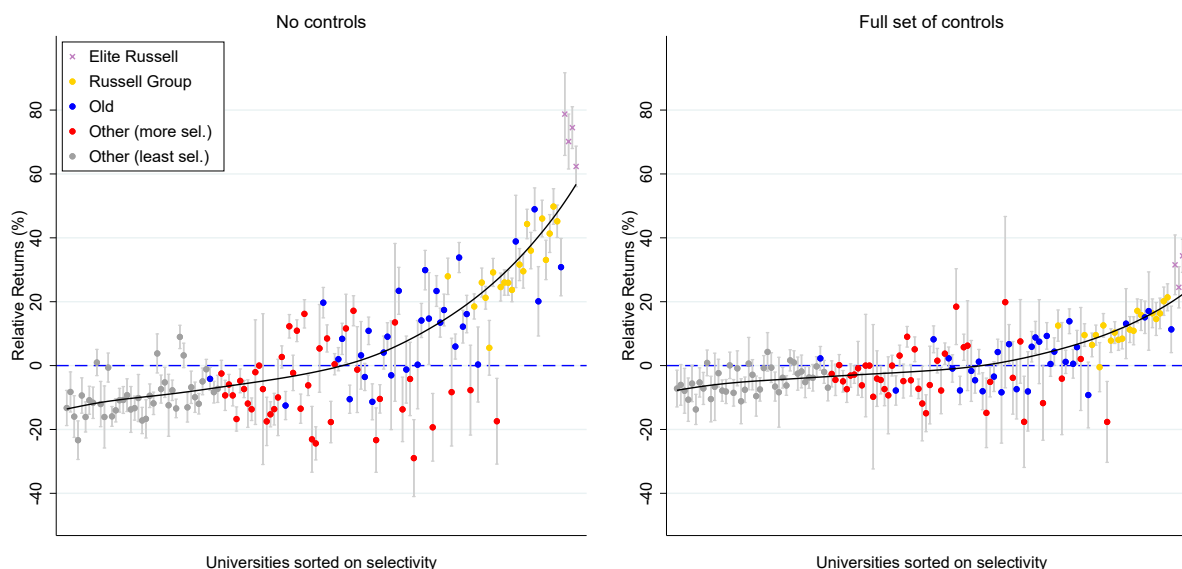
---

<sup>29</sup>All of the point estimates for this and for subsequent results throughout the section are provided in an Online Appendix.

<sup>30</sup>To clarify, these estimates are from the same regression model as the subject estimates in the previous subsection. That is, we control for institution for the subject estimates and for subject for the institution estimates.

only the university fixed effects in the model, while the ‘conditional’ estimates on the right include all of the background controls include the full set of background and attainment controls plus secondary school fixed effects and subject controls.<sup>31</sup>

Figure 3: Estimated returns at age 30 by university



Notes: 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 the conditional model, the relationship between returns and selectivity is quite weak for universities in the bottom two-thirds of the selectivity distribution, but it is clearly 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

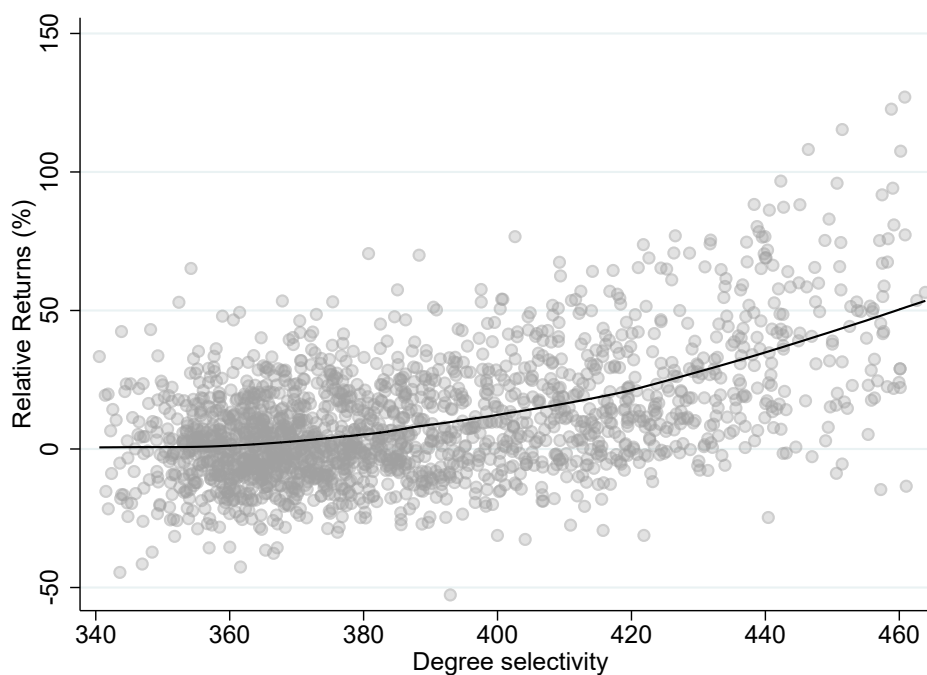
<sup>31</sup>Again, the results are extremely similar when split by gender. See Figure A6 in the Appendix.

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.

### 5.3 Returns by degree

We now turn to our estimates of returns at the degree (that is, the combination of subject and institution) level. Figure 4 shows the returns estimates for more than 1,900 degrees plotted against the selectivity of those degrees (the returns are estimated relative a base case of history at Sheffield Hallam).<sup>32</sup> There is substantial variation in these returns estimates: the standard deviation of the returns estimates is 22 ppts and the 90:10 range is 52 ppts (dropping from 32 ppts and 75 ppts, respectively, without any controls).

Figure 4: Course returns against selectivity



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

<sup>32</sup> Again, all individual returns estimates can be found in the Online Appendix. The number of returns estimates are lower than the number of institutions multiplied by the number of subjects, partly because not all subjects are offered at all universities, and partly because we do not show estimates that are based on small sample sizes. 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’ for any degrees.

The best performing degrees are commonly in law and economics at 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.

More generally, we see that average returns increase considerably as we move from the least 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 towards the top of the selectivity distribution. However, there is also substantial variation around this figure - even amongst the most selective degrees, relative returns range from more than 100% to below zero.

Table 2: Summary of degree estimates

	Selectivity Quintile (Least to Most Selective)				
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
<i>Main results</i>					
Selectivity slope	-0.06	0.22	0.27	0.45	0.85
Standard deviation	15.93	14.77	16.18	19.88	27.04
Share within subject	0.49	0.50	0.60	0.54	0.65
Number of Degrees	386	386	386	386	385

Notes: Table shows the slope of a linear regression of returns on selectivity (average GCSE score of student intake); the standard deviation of returns; and the share of the variation that is within subject, across five selectivity bands. The selectivity bands range from 341-362 (Q1), 362-372 (Q2), 372-388 (Q3), 388-414 (Q4) and 414-464 (Q5).

The positive relationship between returns and selectivity, and the variation in returns, are documented in Table 2 which reports the slope coefficient from a linear regression of returns on selectivity within selectivity bands. We see that this increases from -0.06 (meaning a 100 point increase in GCSE points is associated with a 6 percentage point decrease in returns) in the least selective band of degrees, to 0.85 in the most selective band (meaning a 100 point increase in GCSE points is associated with an 85 percentage point increase in returns).<sup>33</sup>

However, selectivity by no means explains all the variation in returns. Table 2 also highlights

<sup>33</sup>An increase of 100 GCSE points means going from the lowest ranked degree to roughly the 95th percentile degree.

the considerable variation in returns within bands of similarly selective degrees. The standard deviation of returns amongst the least selective band is around 16 percentage points, and this doubles to around 27 percentage points for the most selective band of degrees. As confirmed by Figure 4, 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.

Finally, Table 2 reports the share of the variation in returns that is *within* subject (noting that not all subjects are represented equally across the selectivity bands). Throughout the selectivity distribution, at least half of the variation is within subject, across institutions for all selectivity bands, rising to around two-thirds of the variation for the most selective degrees.<sup>34</sup>

## 5.4 Robustness

Table 3 repeats Table 2 above with a number of alternative approaches in order to test the robustness of our findings. We start by testing the strongest assumption behind our results that selection into different degrees is primarily based on observable characteristics. The concern here is that there are factors that we do not control for that drive selection into certain degrees and are also correlated with subsequent earnings outcomes. In the first robustness check in the table we show how our estimates are affected when we replace the point estimate with the bounded estimate *for each degree*, drawing on the bounding technique developed in Oster (2019).<sup>35</sup> This technique allows us to approximate how much unobserved factors could plausibly shift our individual returns estimates, based on the extent to which observed factors shift them. The exercise requires us to set two parameters, namely  $\Pi$  and  $\delta$ . The former ( $\Pi$ ) informs our assumption about the theoretical maximum  $R^2$  (denoted  $R_{max}$ ) we would obtain if we were able to control for all the relevant variables, such that  $R_{max} = \Pi \tilde{R}$ , where  $\tilde{R}$  is the  $R^2$  in the final model. The latter ( $\delta$ ) is the ratio of the importance of unobserved and observed variables in driving selection. We follow recommendations from Oster (2019) and use  $\Pi = 1.3$  and  $\delta = 1$ . Our estimates are not unaffected

<sup>34</sup>We calculate this by simply considering the within subject versus the between subject variation in the degree estimates. We also did a Shapley-Owen decomposition of the degree level estimates and got very similar results.

<sup>35</sup>We therefore attempted to estimate Oster bounds for all 2000 degree estimates. This was computationally challenging, and we needed to make some simplifications. Most notably, we were unable to estimate them using the full model and instead estimate them for only the 2002 GCSE cohort. We also had to omit school effects from the estimation, and finally, we found that the estimation frequently crashed for degrees with larger sample sizes and we therefore only estimate bounds for around 90% of degrees.



by the bounding exercise - the returns to selectivity are clearly lower, for example, but we still take this exercise as a positive endorsement of our key results given the stringent nature of this test.

Table 3: Summary of degree estimates, alternative specifications

	Selectivity Quintile (Least to Most Selective)				
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
<i>Main Results</i>					
Selectivity slope	-0.06	0.22	0.27	0.45	0.85
Standard deviation	15.93	14.77	16.18	19.88	27.04
Share within subject	0.49	0.50	0.60	0.54	0.65
<i>Oster Bounds</i>					
Selectivity slope	0.04	0.13	-0.09	0.19	0.54
Standard deviation	16.89	16.15	15.05	18.01	22.99
Share within subject	0.25	0.30	0.39	0.35	0.60
<i>Cross-sectional Model</i>					
Selectivity slope	0.02	0.38	0.28	0.45	0.78
Standard deviation	14.93	14.22	15.05	19.16	25.94
Share within subject	0.41	0.39	0.52	0.45	0.62
<i>Excl. School FE and Background</i>					
Selectivity slope	-0.09	0.18	0.28	0.47	0.90
Standard deviation	16.89	15.53	17.35	20.84	28.22
Share within subject	0.53	0.54	0.64	0.57	0.64
<i>Semi-parametric Model</i>					
Selectivity slope	-0.06	0.22	0.28	0.46	0.82
Standard deviation	15.98	14.80	16.22	19.98	26.89
Share within subject	0.50	0.49	0.60	0.54	0.65
<i>Shrinkage</i>					
Selectivity slope	-0.10	0.16	0.20	0.35	0.65
Standard deviation	11.43	11.04	12.25	15.28	20.61
Share within subject	0.48	0.51	0.62	0.56	0.65
<i>Excl. Dropouts</i>					
Selectivity slope	-0.23	0.08	0.24	0.46	0.82
Standard deviation	16.05	15.74	16.99	20.43	26.96
Share within subject	0.50	0.54	0.60	0.55	0.66
<i>Additive Effects</i>					
Selectivity slope	0.08	0.06	0.46	0.31	0.50
Standard deviation	11.99	11.15	12.16	12.68	14.08
Share within subject	0.87	0.87	0.80	0.78	0.70
Number of courses	386	386	386	386	385

Notes: See Table 2. Alternative specifications are explained in the text.

In particular, we still see large increases in the returns to selectivity moving up the selectivity distribution as well as large increases in the variation in returns. The share of the variation that is within subject is also similar at the top end, although it is now smaller at the bottom end. In Ap-

pendix Section A4, we show the bounded estimates graphically while also providing equivalent figures for the subject and institution estimates from the previous subsections.<sup>36</sup>

Next, we consider the extent to which our results are robust to alternative specifications. First, we show the overall conclusions are unchanged when we estimate the model on just one cohort of data (“Cross-sectional”), which shows that our results are not driven by the panel model and the functional form assumptions underlying that model. Second, we find that the results are very similar when we exclude school fixed effects and background controls other than prior attainment from the estimation. Third, we get almost identical results when we control much more flexibly for prior attainment (“Semi-parametric”) by including ventile dummies for GCSE and A level points scores. Fourth, we find that we get quite similar results when we account for estimation error by constructing shrunk Empirical Bayes estimates (see Appendix A5 for more detail on how we construct these estimates).

The table also shows that our results are not driven by the decision to include university dropouts within our estimation sample, although we note that the returns to selectivity are lower amongst the least selective degrees, showing that it is returns towards the bottom end of the selectivity distribution that are most affected by this decision. Finally, we show that our conclusions *are* affected when we do not estimate individual degree-level returns and instead estimate degree returns as the sum of institution and subject effects. The additive estimates understate the variation in returns amongst most selective degrees to a large extent, and systematically overstate the importance of subject relative to institution effects. They also do not so clearly fit with the pattern of the returns to selectivity being much higher at the top of the selectivity distribution (returns are similar in quintile 3 to quintile 5, for example). We believe this highlights the value of our data, as estimating individual degree effects is much more demanding in terms of sample sizes.

## 5.5 Within subject variation in returns

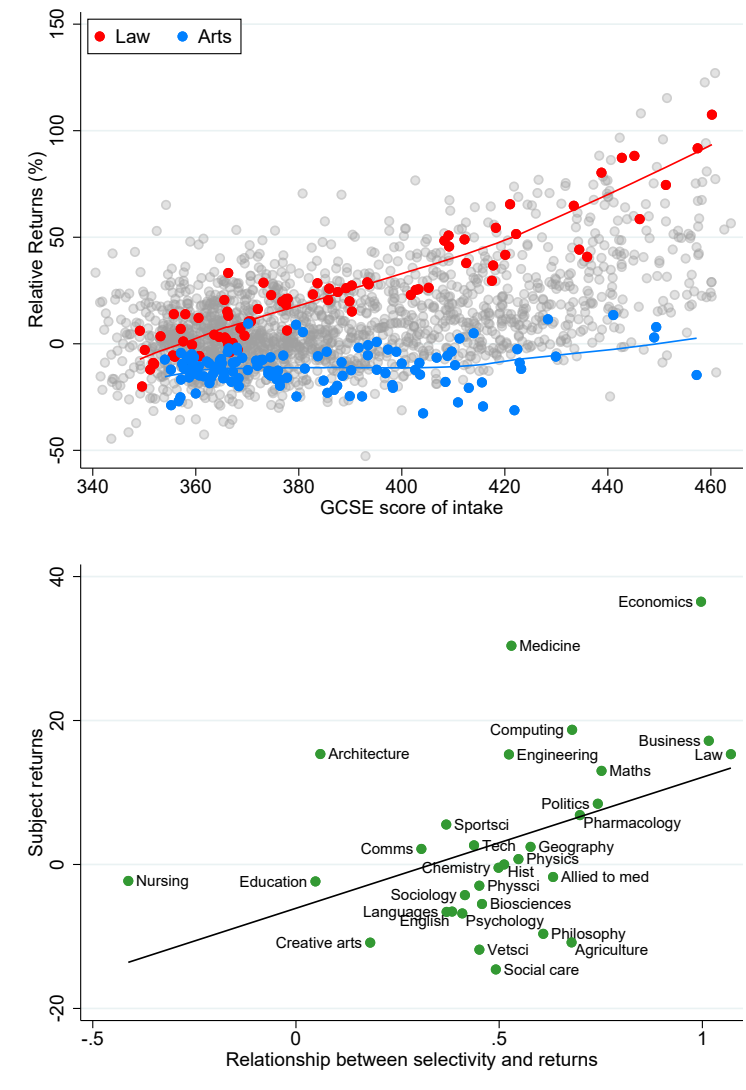
As described above, more than half of the variation in degree returns is within subject, across institution. The top panel of Figure 5 highlights individual institution estimates for two specific

---

<sup>36</sup>We also show, in Appendix Figure A8, that we require  $\delta$  to be around 2 before the pattern of subject returns is lost. This would require that unobservables are twice as important as observables for driving selection, which we think is implausible given the richness of the controls we are able to include in the model.

subjects, law and creative arts. 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 lead to at least a 30 percentage point difference in returns. This holds true throughout the selectivity distribution.

Figure 5: Relationship of returns and selectivity at the subject level



Notes: Top panel shows degree level estimates plotted against average GCSE scores of intake. Bottom panel plots subject returns estimates from the previous section on the within-subject returns to selectivity, using a linear regression of returns on selectivity.

Figure 5 also shows the relationship between returns and selectivity. For law, the relationship between returns and selectivity is strong, while for arts there is a much weaker relationship. The

bottom panel of Figure 5 shows the returns to selectivity for all subjects (now based on a separate linear regression of returns on selectivity for each subject), plotted against the individual subject returns estimates from Section 5.1. Perhaps unsurprisingly, there is a positive correlation between the returns to attending a more selective university within subject and the overall subject returns.<sup>37</sup>

## 5.6 Correlates with degree returns

Our findings so far have suggested that degree choices can make a substantial difference to earnings outcomes at age 30. Further, there remains substantial variation in returns, even when comparing degrees that are similarly selective. Here, we consider whether other degree characteristics are predictive of returns outcomes, again within subject area. In addition to showing the correlation between selectivity and returns for each subject, Table 4 shows the relationship between returns and the following set of degree characteristics:<sup>38</sup>

- **League table rating:** we use data from the Times Good University Guide from 2006 (which is around the time our cohorts were starting university) on the subject-specific league table rankings of universities. The ranking is based on a score out of 100 for each university in each subject that combine several characteristics of the degree, including student-staff ratios and research intensity, and we use this overall rating as a degree characteristic. Gibbons et al. (2015) highlights the importance of league tables in driving institution choices of prospective students.
- **Student satisfaction:** we use data from the National Survey of Students in 2010 (the earliest year available) on the share of students reporting that overall they were satisfied with their university course. This metric has recently been included as an input into the government's Teaching Excellence Framework, a measure of teaching quality.
- **Age 22 returns:** we estimate early-career returns from the LEO data using a similar methodology as for our main estimates, but at a different age. This is usually in the first year after

---

<sup>37</sup>We also looked at the returns to selectivity by subject estimates using Oster Bounds and we get a similar pattern. The correlation between the set of estimates on the returns to selectivity by subject using our headline estimates and using the Oster estimates is 0.85.

<sup>38</sup>Summary statistics for each of these measures, and information on how they correlate with one another, are provided in Appendix Section A6.

graduating from university.<sup>39</sup> We report this because early-career earnings outcomes are often used as a measure of degree quality.

- **Completion rate:** we calculate completion rates from the LEO data based on the number of students from the 2002-2007 GCSE cohorts who completed their degree before before 2016 (the end of our university records), as a share of the total number of students from those cohorts who started a degree before the age of 21. This is another metric that is used as a measure of degree quality.
- **First class degree rate:** we calculate the number of students who graduated with ‘First class honours’ in the 2002-2007 GCSE cohorts, as a share of the total number of students from these cohorts who graduated from a degree they started before the age of 21. First class honours is the highest degree classification. This rate varies across different subjects and universities (see Britton et al., 2022, for a detailed summary of this).
- **Local wages:** we use a measure of average wages for all workers in the travel to work area<sup>40</sup> the university is located from the 2016 Annual Survey of Hours and Earnings (aligning with the year of our returns estimates). This measure is intended to capture the strength of the universities’ local labour market.

Subjects are sorted by their overall returns estimates from Section 5.1, with the highest returning subjects at the top. Column (1) of Table 4 repeats the result from above that returns at age 30 are generally very highly correlated with selectivity, although this varies a lot by subject. We then see that league table rankings, completion rates and first class degree rates are all well correlated with returns. Again, the correlations are generally much stronger for the higher-returning subjects at the top than the lower returning subjects at the bottom. 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 with student satisfaction ratings and early career (age 22) returns are much noisier and generally weaker. In fact we even see negative correlations between student satisfaction and returns for economics, maths and sociology, suggesting students studying towards these

---

<sup>39</sup>Over 85% of the graduates we observe at age 22 have entered the labour market by this age.

<sup>40</sup>These are similar to Commuting Zones in the United States.

degrees do not value or appreciate things that are well correlated with their subsequent labour market success (for example, students might dislike being taught difficult things, but those things might be valued in the labour market).

Table 4: Correlates with age 30 degree returns

	(1) Selectivity	(2) League Table	(3) Student Sat.	(4) Age 22 Returns	(5) Complete Rate	(6) First Rate	(7) Local Wages
Economics	.792*** (.000)	.623*** (.000)	-.080 (.568)	.731*** (.000)	.658*** (.000)	.550*** (.000)	.087 (.519)
Medicine	.088 (.640)	.003 (.992)	.212 (.457)		.455* (.010)	-.123 (.511)	-.187 (.313)
Business	.888*** (.000)	.561*** (.000)	.242* (.020)	.549*** (.000)	.772*** (.000)	.456*** (.000)	.159 (.101)
Architecture	-.056 (.722)	.257 (.158)	.388* (.038)	.281 (.144)	.214 (.169)	.180 (.247)	-.033 (.834)
Law	.928*** (.000)	.891*** (.000)	.071 (.560)	-.338** (.003)	.789*** (.000)	.601*** (.000)	.184 (.107)
Engineering	.521*** (.000)	.247 (.051)	.351** (.005)	.016 (.882)	.499*** (.000)	.269* (.020)	-.019 (.871)
Maths	.696*** (.000)	.471*** (.000)	-.057 (.661)	.536*** (.000)	.590*** (.000)	.430** (.001)	.351** (.007)
Comms	.264* (.015)	.291* (.028)	.286* (.025)	.161 (.169)	.264* (.015)	.147 (.178)	.003 (.980)
History	.610*** (.000)	.240* (.028)	.016 (.885)	.200 (.091)	.557*** (.000)	.517*** (.000)	.311** (.003)
Allied to med	.625*** (.000)	.354** (.001)	.296** (.007)	.058 (.604)	.434*** (.000)	.555*** (.000)	.170 (.132)
Education	-.081 (.521)	-.177 (.243)	.049 (.692)	.077 (.499)	.104 (.409)	.020 (.874)	.194 (.122)
Phyisci	.541*** (.000)	.213 (.123)	.028 (.819)	.187 (.212)	.369** (.008)	.323* (.022)	.305* (.032)
Sociology	.489*** (.000)	.419*** (.000)	-.087 (.424)	-.172 (.092)	.501*** (.000)	.115 (.305)	.153 (.169)
Biosciences	.575*** (.000)	.441*** (.000)	.243* (.042)	.343** (.004)	.500*** (.000)	.449*** (.000)	.225 (.058)
Languages	.541*** (.000)	.387** (.004)	.227 (.081)	-.044 (.769)	.453*** (.000)	.334** (.005)	.113 (.357)
Creative arts	.184* (.047)	.139 (.266)	.089 (.431)	.247** (.007)	.145 (.120)	.081 (.385)	-.126 (.181)

Notes: Raw correlations reported, with p-values in the parenthesis (for the null of zero correlation). \* indicates  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . Descriptions of each of the variables are given in the text. Only subjects for which we could obtain league table rankings and student satisfaction scores are shown. Subjects are sorted on their average returns (highest to lowest).

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.<sup>41</sup>

Table 5: Correlates with age 30 degree returns, controlling for selectivity

	(1) League Table	(2) Student Sat.	(3) Age 22 Returns	(4) Complete Rate	(5) First Rate	(6) Local Wages
Economics	-.050 (.738)	.003 (.970)	.286* (.031)	-.327 (.092)	.058 (.595)	.067 (.421)
Medicine	.001 (.998)	.357 (.237)		.457* (.013)	-.186 (.364)	-.249 (.213)
Business	.039 (.517)	.070 (.162)	.098 (.080)	.072 (.394)	.064 (.208)	.056 (.218)
Architecture	.352 (.096)	.471* (.021)	.370 (.064)	.292 (.091)	.249 (.151)	-.039 (.805)
Law	.081 (.568)	-.020 (.667)	-.019 (.669)	-.036 (.669)	.056 (.300)	.083 (.053)
Engineering	-.374 (.062)	.125 (.335)	-.009 (.919)	.214 (.243)	-.078 (.544)	-.048 (.634)
Maths	.125 (.547)	-.262* (.011)	.292* (.011)	.079 (.633)	.074 (.522)	.261** (.006)
Comms	.238 (.335)	.241 (.058)	.191 (.080)	.150 (.364)	.010 (.936)	.007 (.945)
History	-.213 (.107)	.007 (.943)	.163 (.070)	.111 (.522)	-.211 (.302)	.141 (.115)
Allied to med	.024 (.846)	.059 (.578)	-.005 (.954)	-.300 (.063)	.212 (.102)	.129 (.146)
Education	-.255 (.308)	.070 (.587)	.056 (.628)	.373* (.048)	.107 (.498)	.186 (.146)
Physsci	-.448* (.035)	-.232 (.057)	.128 (.345)	.070 (.650)	-.099 (.562)	.102 (.448)
Sociology	.281 (.063)	-.210* (.041)	-.113 (.226)	.306 (.077)	-.436** (.001)	.113 (.251)
Biosciences	.113 (.522)	-.040 (.732)	.196 (.056)	.103 (.542)	.137 (.287)	.147 (.138)
Languages	.052 (.769)	.119 (.308)	.018 (.863)	-.031 (.877)	-.019 (.888)	.105 (.315)
Creative arts	.077 (.668)	.011 (.928)	.272** (.002)	.003 (.981)	-.069 (.575)	-.164 (.081)

Notes: 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.

<sup>41</sup>In Appendix Table A7 we show that the correlations increase rapidly with age (as would be expected). Age 24 returns are much more highly correlated with age 30 returns than age 22 returns are.

Finally, we see that the correlation between returns and wages in the local labour market is generally weak. It is around 0.3 for maths, history and physical sciences, but lower in other subject areas. This is perhaps surprising, but in fact, by age 27 only around one third of graduates from these cohorts are still working within the travel to work area their university is located. This figure is lower still for Russell Group universities outside of London (often dipping below 10%), despite returns for those universities being high. Furthermore, a very large share of these graduates are also from that area originally, and we control for this in the regressions.<sup>42</sup>

In Table 5 we then look at the correlation of each of these measures with returns at age 30 once we control for selectivity. We are interested in this because we know that there is substantial variation in returns even amongst degrees with relatively similar selectivity levels. 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.<sup>43</sup> 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.<sup>44</sup>

## 6 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. We find very substantial variation in returns at the degree level even for similarly selective degrees. Around half of the variation

---

<sup>42</sup>These facts draw upon recent work (Britton et al., 2021) that explored geographical mobility of graduates in much more detail.

<sup>43</sup>In practice, due to practical data issues, many of the measures we use here would not actually have been available to the cohorts we are interested in at the point they entered university. However, earlier measures are likely to be *less* well correlated with actual outcomes, as older measures are likely to reflect lagged performance of the university on these measures rather than performance of the university while the students were actually enrolled. We are therefore confident that our observation that student information is not well correlated with returns is likely to hold.

<sup>44</sup>On the other hand, as described above, it adds weight to our claims that our results are not driven by selection on unobservable factors.



in returns is *within* subject, across universities. Returns increase with selectivity but only in some subjects. Our results suggest that degree choice matters a lot for earnings outcomes at age 30. We show these findings are robust to the empirical specification used, the exact sample of students included, and to unobserved selection. We would expect variation in returns to only increase further with age.

While degree choice matters a lot, once we control for a simple measure of degree selectivity, we find that 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 choose to target factors such as student satisfaction or first class degree shares when those things might not be beneficial in the long term.<sup>45</sup>

One policy implication from our work is that information on earnings outcomes of past graduates should be made more readily available to prospective higher education students. 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. Another policy implication could be for 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 redistribution. First, there is a long lag between changes to university practice and changes to earnings returns. Second, it is also possible that the returns do not reflect university productivity and are instead a product of peers, labour market signalling or both. Third, a university degree may have important positive impacts that might not be reflected in our earnings returns estimates. Indeed, an additional policy implication of our work could be for government to consider tax breaks or public sector pay increases in areas that have high social value but are associated with degrees that result in low returns, such as social care.

---

<sup>45</sup> 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.

## 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.
- Bleemer, Zachary and Aashish Mehta**, “Will studying economics make you rich? A regression discontinuity analysis of the returns to college major,” *American Economic Journal: Applied Economics*, 2022, 14 (2), 1–22.
- Blundell, Richard, 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, Ian Walker, Ben Waltmann, and Yu Zhu**, “How much does it pay to get good grades at university?,” 2022.
- , **Laura van der Erve, Ben Waltmann, and Xiaowei Xu**, “London calling? Higher education, geographical mobility and early-career earnings: Research report: September 2021,” 2021.
- , **Neil Shephard, and Anna Vignoles**, “A comparison of sample survey measures of earnings of English graduates with administrative data,” *Journal for the Royal Statistical Society, Series A*, 2019.

- 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 Gill Wyness**, "Matching in the dark? Inequalities in student to degree match," *Journal of Labor Economics*, 2021.
- 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.
- Foote, Andrew and Kevin M Stange**, “Attrition from Administrative Data: Problems and Solutions with an Application to Postsecondary Education,” Technical Report, National Bureau of Economic Research 2022.
- 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.

**Oster, Emily**, “Unobservable selection and coefficient stability: Theory and evidence,” *Journal of Business & Economic Statistics*, 2019, 37 (2), 187–204.

**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 —**, “University selectivity and the relative returns to higher education: Evidence from the UK,” *Labour Economics*, 2018.

**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**, “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 stated 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

Notes: 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.<sup>46</sup> The proportion of individuals who do not match to the 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.<sup>47</sup> Among those who attend university - the main focus of this paper - the match rate is much higher, at more than 99%.

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<sup>48</sup> 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.<sup>49</sup> We can see that this group represents around 56% of all students with linked HMRC records.

<sup>46</sup>This step was done separately by the Department for Work and Pensions before we had access to the data.

<sup>47</sup>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.

<sup>48</sup>Less than 10% of those without five good GCSEs start an undergraduate degree by age 21.

<sup>49</sup>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.

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.<sup>50</sup> 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 one-third of the overall cohort.<sup>51</sup>

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

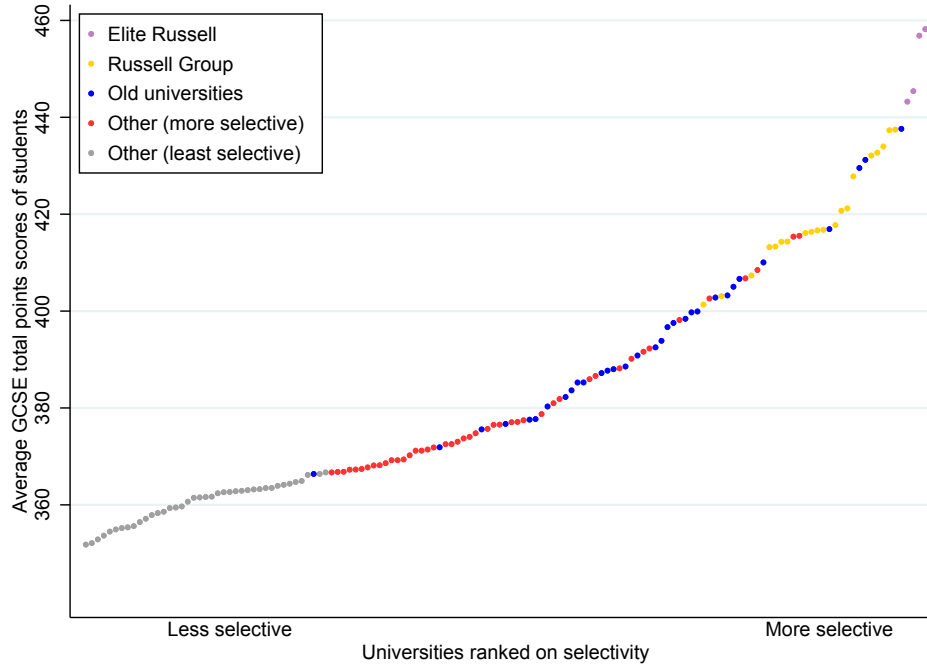
Notes: Column 1 is taken from Column 4 of Table A1. Columns 2-4 sum to Column 1. PT indicates part-time, PG indicates postgraduate.

<sup>50</sup>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.

<sup>51</sup>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 and university selectivity

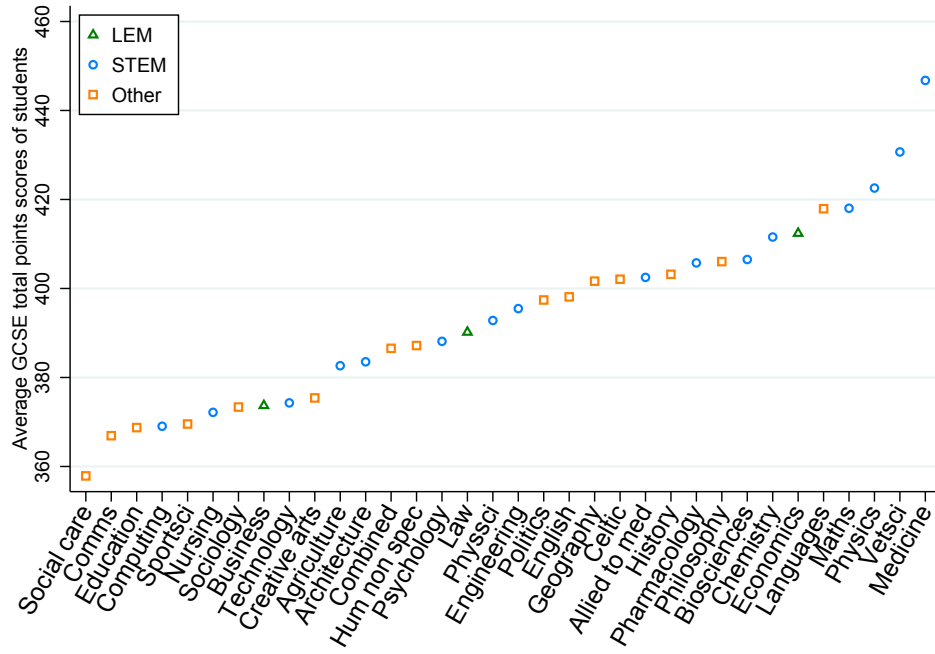
Figure A1: Selectivity by university at age 30



Notes: 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 A2: Selectivity by subject at age 30



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

### A3 Additional descriptives

Table A3: Background characteristics by attainment group

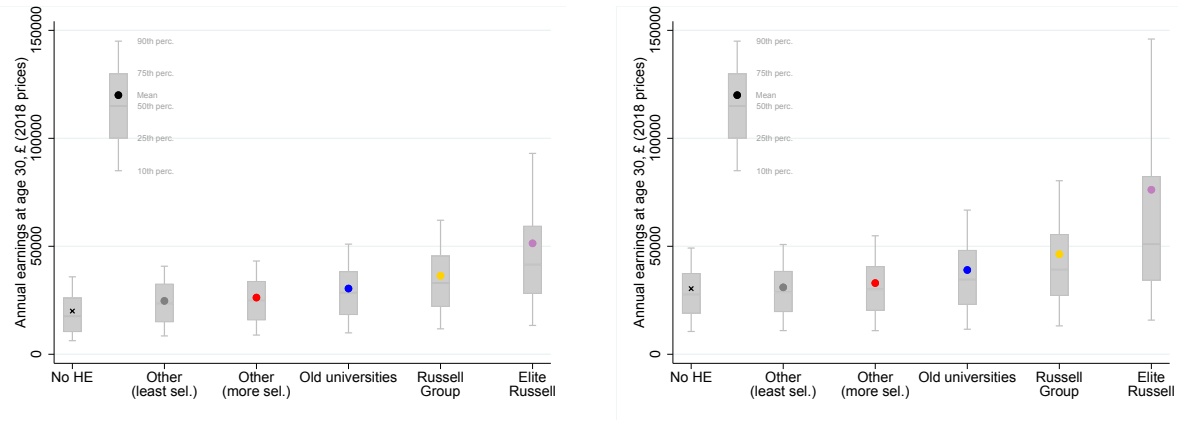
	Women		Men	
	No UG	UG	No UG	UG
<i>Background</i>				
State school	0.95	0.85	0.94	0.83
<i>of which:</i>				
SES Q1 (richest)	0.23	0.33	0.25	0.35
SES Q2	0.25	0.25	0.26	0.25
SES Q3	0.22	0.20	0.22	0.19
SES Q4	0.18	0.13	0.16	0.12
SES Q5 (poorest)	0.13	0.09	0.11	0.08
FSM	0.07	0.05	0.06	0.04
EAL	0.04	0.10	0.04	0.10
SEN	0.03	0.02	0.05	0.03
<i>Ethnicity</i>				
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
<i>Attainment</i>				
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

Notes: 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.

Table A4: GCSE points scores by university group

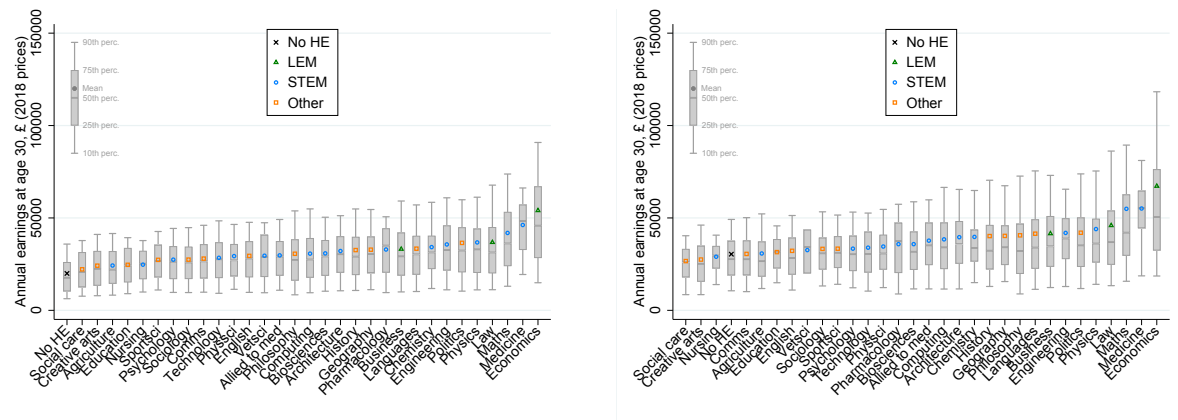
University group	Mean	SD
Elite Russell	453.4	21.75
Russell Group	417.4	33.02
Old universities	389.8	37.02
Other (more selective)	368.2	34.37
Other (less selective)	356.8	33.26
No HE	340.4	33.12

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



Notes: 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 A4: Real earnings by degree subject at age 30 - Women (left) and men (right)

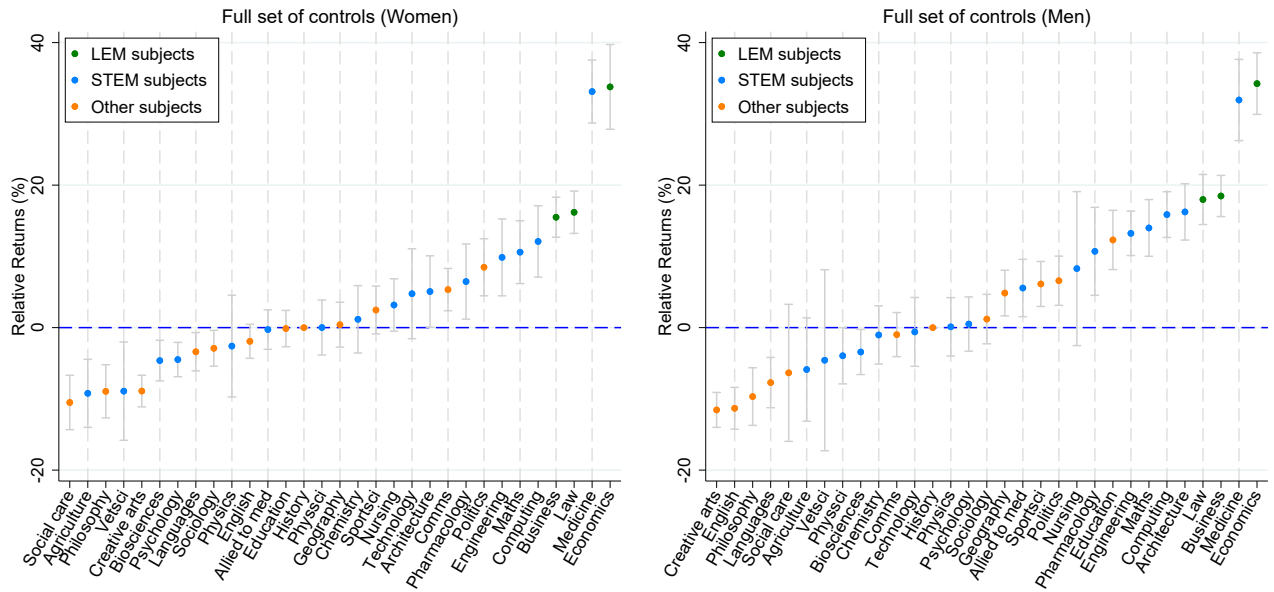


Notes: 2002 GCSE cohort in 2016/17, conditioning on earnings being above £1,000. No HE consists of individuals who did not take an undergraduate degree, but passed their age 16 exams (obtaining at least 5 A\*-C GCSEs).

## A4 Additional results

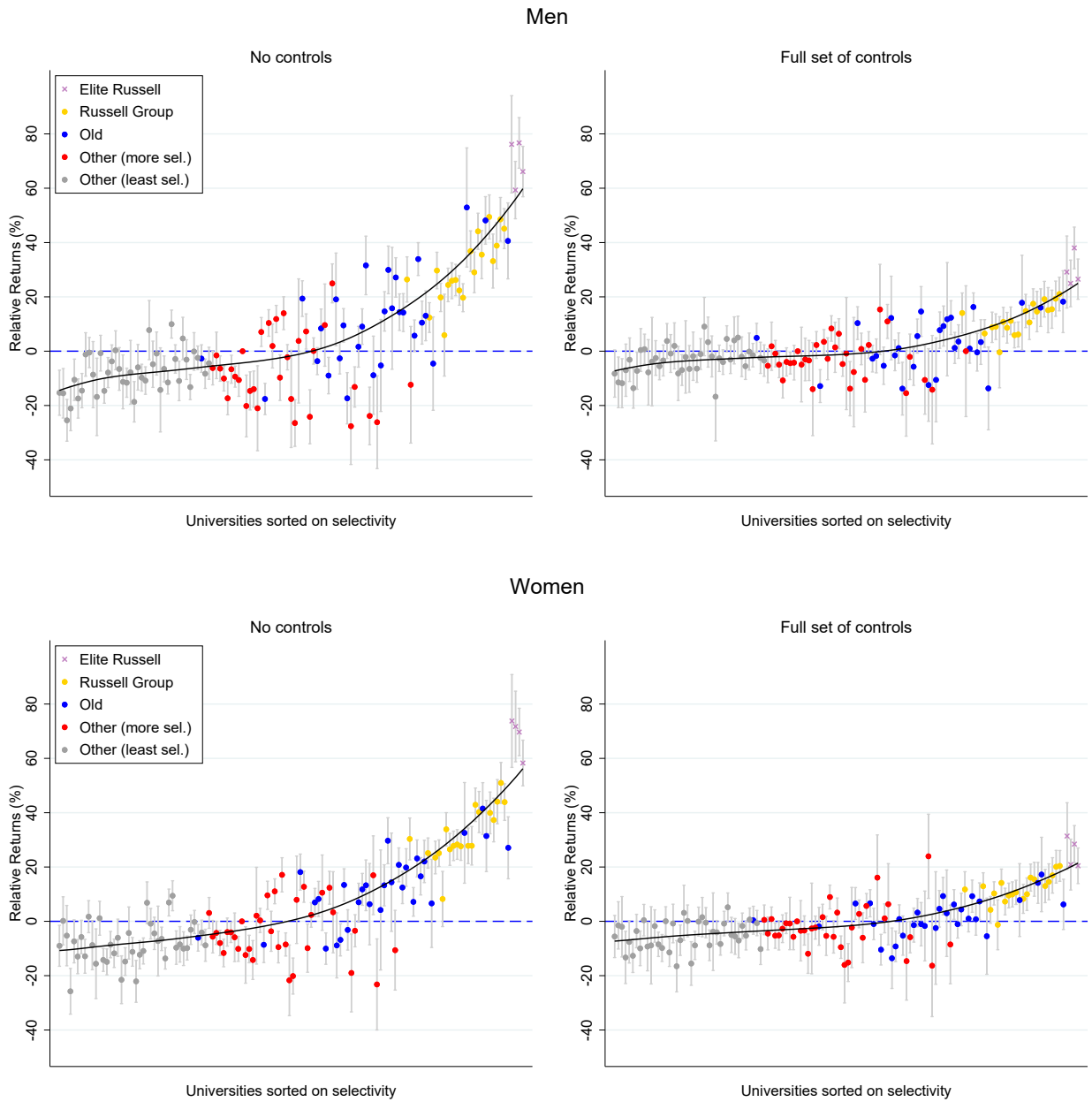
### A4.1 Returns estimates split by gender

Figure A5: Estimated returns at age 30 by subject



Notes: Equivalent to Figure 2, split by gender. Correlation of returns by gender is 0.91.

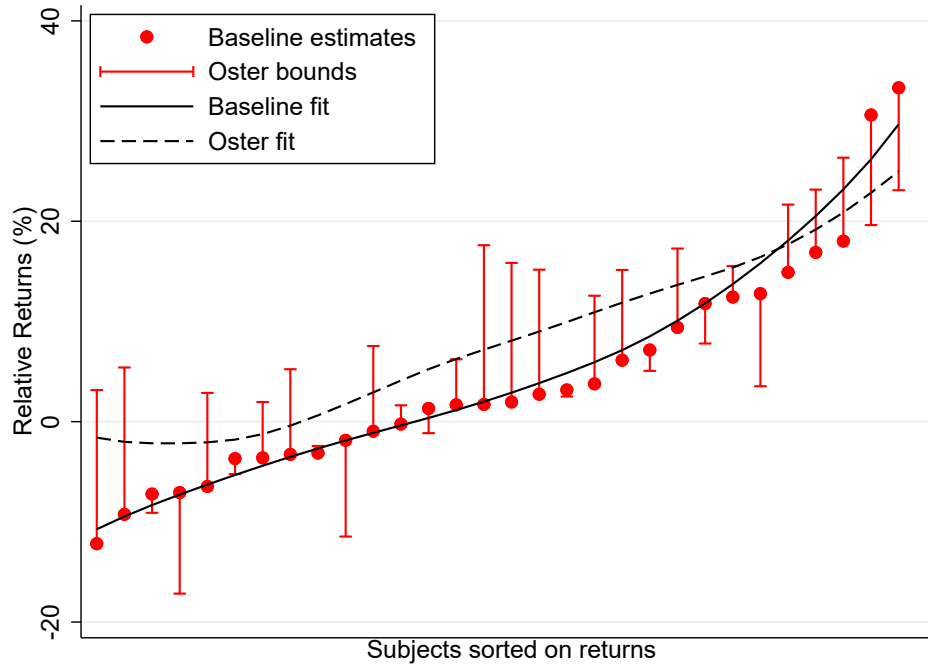
Figure A6: Estimated returns at age 30 by institution



Notes: Equivalent to Figure 3, split by gender.

## A4.2 Bounding

Figure A7: Subject returns, testing for robustness to selection on unobservables



Notes: We use  $\delta=1$  and  $R_{max} = 1.3 * \bar{R}$ .

Figure A8: Subject returns, testing Oster bound sensitivity to different choices of delta

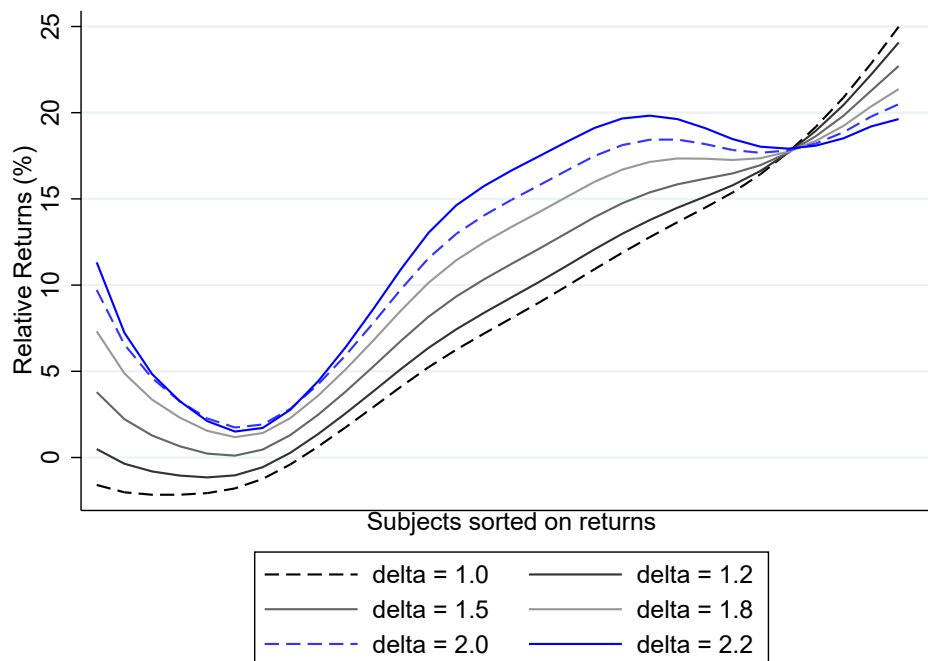
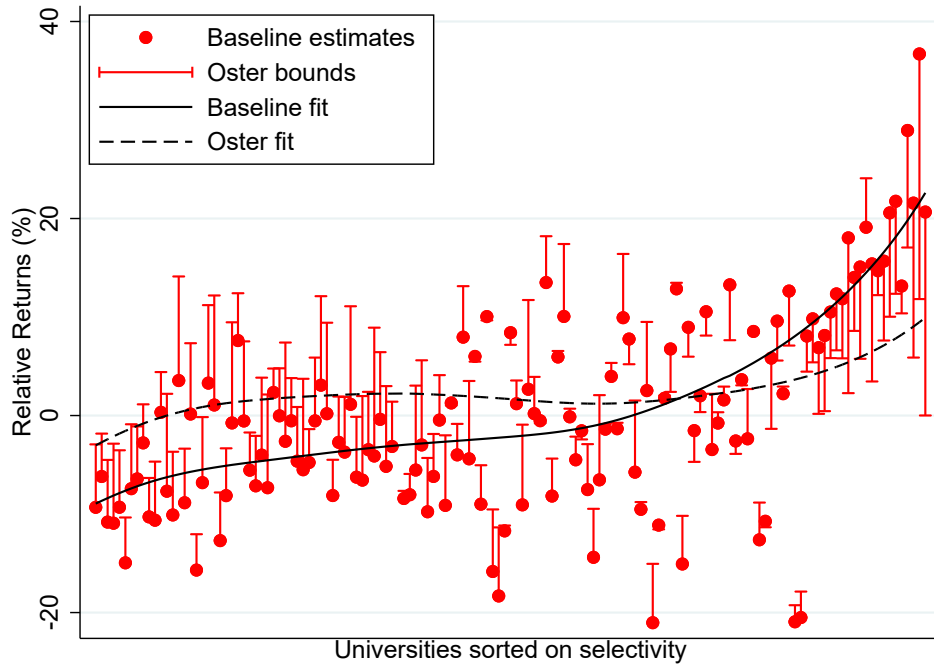
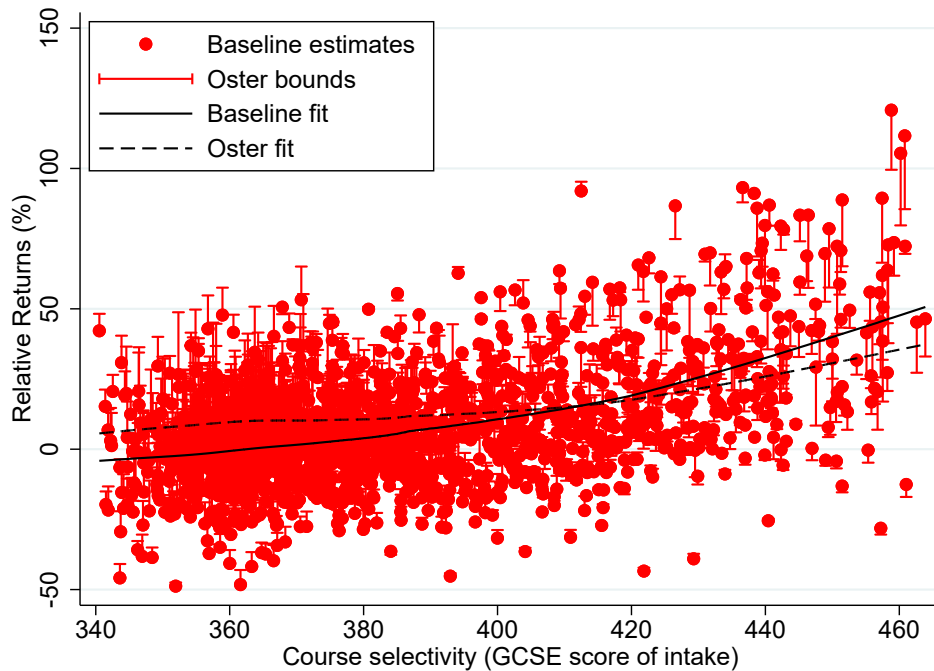


Figure A9: University returns, testing for robustness to selection on unobservables



Notes: We use  $\delta=1$  and  $R_{max} = 1.3 * \bar{R}$ .

Figure A10: Degree returns, testing for robustness to selection on unobservables



Notes: We use  $\delta=1$  and  $R_{max} = 1.3 * \bar{R}$ .

## A5 Empirical Bayes shrinkage

Following the approach taken in Mountjoy and Hickman (2020), we account for estimation error using Empirical Bayes (EB) estimates (Robbins, 1956), where we shrink returns estimates for each degree to the mean return  $\bar{\beta}$  in proportion to their imprecision. We assume a normal distribution of the true returns  $\beta_j$  and the estimation error  $\epsilon_j$ , which for each degree  $j$  gives us the shrinkage factor as  $F_j = \frac{\sigma_{\bar{\beta}}^2}{\sigma_{\bar{\beta}}^2 + \sigma_{\epsilon_j}^2}$ , where  $\sigma_{\epsilon_j}^2$  is estimated as the square of the standard error of  $\hat{\beta}_j$ , our estimated returns for degree  $j$ . We estimate  $\sigma_{\bar{\beta}}^2$  as  $\sqrt{\sigma_{\hat{\beta}}^2 - \sigma_{\epsilon}^2}$ , where  $\sigma_{\hat{\beta}}^2$  is the sample variance of the returns estimates  $\beta_j$ , weighted by the number of students in each degree  $j$  and  $\sigma_{\epsilon}^2$  is estimated as the average of the squared standard errors of the returns estimates  $\beta_j$ , similarly weighted by the number of students in each degree  $j$ . The EB estimates are then constructed as:  $\beta_j^{EB} = F_j \hat{\beta}_j + (1 - F_j) \bar{\beta}$ .

## A6 Quality measures

### A6.1 Summarising the quality measures

Table A5: Quality measures, summary stats

	N	Mean	SD	Min	Max
Selectivity	1155	387.74	28.67	341.42	464.04
League Table	840	79.68	10.10	22.50	100.00
Student Satisfaction	899	76.22	5.95	52.00	95.00
Age 22 Returns	941	17.85	23.77	-39.74	115.76
Completion Rate	1155	0.88	0.08	0.50	1.00
First Class Rate	1155	0.12	0.07	0.00	0.40
Local Wages	1153	23823.96	3592.72	17378.00	30985.00

Notes: Summary stats are at the degree level, excluding subjects that are not included in Tables 4 and 5 (N is the number of degrees).



Table A6: Quality measures, summary stats

	Select.	League Table	Student Sat.	Age 22 Returns	Complete Rate	First Rate	Local Wages
Selectivity	1.000	0.724	0.280	0.107	0.772	0.479	0.025
League Table	0.724	1.000	0.169	0.058	0.679	0.274	0.001
Student Sat.	0.280	0.169	1.000	-0.040	0.286	0.147	-0.276
Age 22 Returns	0.107	0.058	-0.040	1.000	0.046	0.293	0.101
Completion Rate	0.772	0.679	0.286	0.046	1.000	0.200	-0.034
First Class Rate	0.479	0.274	0.147	0.293	0.200	1.000	0.050
Local Wages	0.025	0.001	-0.276	0.101	-0.034	0.050	1.000

Notes: Summary stats are at the degree level, excluding subjects that are not included in Tables 4 and 5.

## A6.2 Correlation of returns by age

Table A7: Correlation between age 30 degree returns and returns at earlier ages

	(1) Age 22 Earnings	(2) Age 24 Earnings	(3) Age 26 Earnings	(4) Age 28 Earnings
Economics	.731*** (.000)	.690*** (.000)	.757*** (.000)	.895*** (.000)
Medicine		.188 (.379)	.106 (.572)	.522** (.003)
Business	.549*** (.000)	.834*** (.000)	.872*** (.000)	.904*** (.000)
Architecture	.281 (.144)	.633*** (.000)	.628*** (.000)	.832*** (.000)
Law	-.338** (.003)	.761*** (.000)	.870*** (.000)	.940*** (.000)
Engineering	.016 (.882)	.449*** (.000)	.464*** (.000)	.698*** (.000)
Maths	.536*** (.000)	.738*** (.000)	.740*** (.000)	.758*** (.000)
Comms	.161 (.169)	.391*** (.000)	.445*** (.000)	.702*** (.000)
History	.200 (.091)	.652*** (.000)	.693*** (.000)	.839*** (.000)
Allied to med	.058 (.604)	.230* (.025)	.591*** (.000)	.811*** (.000)
Education	.077 (.499)	.548*** (.000)	.552*** (.000)	.782*** (.000)
Physsci	.187 (.212)	.602*** (.000)	.718*** (.000)	.773*** (.000)
Sociology	-.172 (.092)	.324** (.003)	.543*** (.000)	.675*** (.000)
Biosciences	.343** (.004)	.520*** (.000)	.580*** (.000)	.789*** (.000)
Languages	-.044 (.769)	.323** (.007)	.424*** (.000)	.673*** (.000)
Creative arts	.247** (.007)	.451*** (.000)	.429*** (.000)	.699*** (.000)

Notes: Raw correlations reported, with p-values in the parenthesis (for the null of zero correlation). \* indicates  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .