Earnings, financial and non-pecuniary factors in university attendance

Oliver Cassagneau-Francis*

22nd February 2023

Abstract

Why do some people choose to attend university, and enjoy state-subsidised benefits, while others do not? We shed new light on this key issue by comparing and quantifying the roles of earnings and non-pecuniary factors in the educational decisions of young people in the UK, exploiting information on young people's beliefs about the advantages and disadvantages of university. We also investigate changes in these factors over time, and their implications for social mobility. We specify a model of educational choice, explicitly including expectations about earnings, financial, and non-pecuniary factors. Our estimation strategy exploits panel survey data on young people's expectations about key outcomes both at, and after, university, linked to their realised outcomes. Income maximisation, despite its prevalent role in the literature, is only part of the story: other factors are at least as important as earnings in determining whether someone goes to university. Non-pecuniary factors also play an important role both the SES-gap in educational attainment, and the huge growth in degree attainment between the 1980s and 2010s.

Keywords: Higher education; Earnings; Psychic costs; Wage premium; Educational choice.

JEL codes: E24; I23; I26; J24

*UCL; email: o.cassagneau-francis@ucl.ac.uk. I am grateful to my advisors Ghazala Azmat and Jean-Marc Robin for their invaluable guidance and support throughout this project. I thank Zsofia Barany, Johannes Boehm, Jeanne Commault, Pierre-Philippe Combes, Nicolo Dalvit, Michele Fioretti, Eric French, George-Levi Gayle, Kerstin Holzheu, Guy Laroque, Hugo Lhuillier, Michela Tincani, and numerous seminar participants at Sciences Po for helpful comments and discussions. I thank the Centre for Longitudinal Studies at UCL for administering the survey and, along with the UK Data Service, providing access to the data. The majority of the work for this paper was carried out while I was a PhD student at Sciences Po, Paris. The usual disclaimer applies.

1 Introduction

University graduates enjoy higher wages, better health, and are more likely to report feeling happy with their lives than their less-educated peers (Heckman, Humphries, and Veramendi, 2018; Oreopoulos and Petronijevic, 2013; Oreopoulos and Salvanes, 2011). These benefits are often heavily subsidised, with governments in developed countries spending around 1% of their countries' GDP on higher education (OECD, 2018). Therefore, understanding the determinants of young people's decision to attend university is key, not only to better understand the direct effects of educational policies, but also for wider issues such as inequality and to identify the beneficiaries of public spending. Traditionally, economists have focused on higher wages, pecuniary costs, and financial frictions to explain this decision—a narrative of comparative advantage and credit constraints. However, more recent work suggests this narrative is missing an important part, as "[t]he evidence against strict income maximization is overwhelming" (Heckman, Lochner, and Todd, 2006, p. 436).

A growing literature has attempted estimating these "psychic costs", as the non-earnings factors are commonly termed. Authors have typically used a residual term to capture these factors, relying upon data containing information on family background, earnings, and educational choices (Cunha and Heckman, 2007). However, some of these same authors have highlighted issues with relying on a residual catch-all term, with Heckman et al. (2006, p. 436) remarking that "explanations based on psychic costs are intrinsically unsatisfactory [as o]ne can rationalize any economic choice data by an appeal to psychic costs."

In this paper, we study the role of both pecuniary and non-pecuniary factors (or "psychic costs") in the decision to attend university, with the aim of exploiting data which combines young people's subjective beliefs about the (pecuniary and non-pecuniary) aspects of university, with information on their later outcomes and educational choices. Our main analysis addresses the following question: what is the relative importance of wages versus other non-pecuniary factors in the decision to attend university? We then extend our analysis to study heterogeneity in educational decisions across different socio-economic groups and over time.

Our data is from a longitudinal study of young people in the UK, which follows a representative sample of students born in 1989 or 1990. The cohort members were surveyed annually between the ages of 13 and 18 — a period in which they were in compulsory education (up to 16), and then either transitioning to work, or on to further and higher education. They were contacted and surveyed again at age 25. The dataset contains information on: (i) young people's beliefs about (the advantages and disadvantages of) university obtained prior to their decision to attend; (ii) their educational and career

choices; (iii) and their later wages.

Given that one of our chief aims is to quantify the non-pecuniary factors in the decision to attend university, these subjective beliefs about university, which include many non-pecuniary aspects, are central to our analysis. Examples of these non-pecuniary factors are: the effort required to gain a place at university; aspects of life at university (social life, studying, leaving home, stress, etc); and aspects of life after university (access to better jobs, graduate "identity", debt). These beliefs are recorded in the form of open-ended questions about the subjective advantages and disadvantages of attending university, and were obtained from a representative sample of young people. As far as we know, we are the first to analyse educational decisions using data containing detailed information, including elicited beliefs about non-pecuniary factors, from a representative sample including data on realised outcomes after university. Previous work has relied on small, selected samples (Boneva and Rauh, 2020), or did not have access to information on young people's beliefs (D'Haultfoeuille and Maurel, 2013).

To guide our empirical analysis, we specify a parsimonious model of educational choice in the spirit of Roy (1951) explicitly including both earnings and other (chiefly non-pecuniary) factors. The structure of the model allows us to quantify and compare the relative contributions of different factors in the decision to attend university, exploiting choice data, subjective beliefs about life at and after university, and realised earnings. We map observed (realised) earnings into potential (expected) earnings as the mean of realised earnings at age 25 conditional on a set of observed characteristics at age 16. The model is then straightforward to estimate using standard techniques from the discrete-choice literature (Mcfadden, 2001). Having estimated our model, we are able to combine estimated preferences with (observed and estimated) expectations to obtain distributions of the relative contributions of earnings and other factors to the decision to attend university. These distributions are: (i) students' expected earnings premium, and; (ii) students' (observed) expected "other factors premium", from attending university. We rescale both distributions of premiums so that they are expressed as a percentage change in wages, allowing a direct comparison.

Our results highlight the important role for non-pecuniary factors as a determinant of higher education attendance. Although the distributions of earnings and of other factors share similar shapes and locations — bell-shaped, with slightly positive means — the dispersion of the other factors distribution is twice as high. This much larger dispersion, along with the similar shapes and locations of the distributions near zero, suggest that the chief determinant of whether or not someone decides to go to university is their expectations about factors other than their future earnings. To underline this conclusion, we study the effects of varying values of these factors, performing the same counterfactual exercise as D'Haultfoeuille and Maurel (2013). This involves fixing the values of young

people's pecuniary and non-pecuniary factors at certain percentiles of the distribution, and recalculating the proportion who get a university degree under these counterfactual values. Assigning everyone in the sample an "other factors premium" equal to the 10th percentile results in 35% obtain a university degree; assigning them values equal to the 90th percentile results in 86% graduating — a change of over 50*pp*. Repeating the same exercise with earnings (assigning different values of the earnings premium to everyone) results in 50% (10th percentile) and 73% (90th percentile) of people attending and graduating university, a much smaller variation.

Next, we split the sample into three groups by socio-economic status (SES) measured by parental earnings at sixteen,[1] to investigate the role of earnings and other factors in the socio-economic gap in university attainment. We recalculate the distributions of earnings and other factors premiums for each of the three SES groups. The distribution of the expected graduate earnings premium is remarkably stable across the three groups, with means ranging from 7 log-points (low SES) to 9 log-points (high SES), and dispersion slightly decreasing with parental income. For other factors, there is much more variation across SES: the low-SES mean is 3 log-points, while the high-SES mean is 15 log-points. The socio-economic gap in university attainment is almost entirely driven by factors other than earnings.

Finally, we re-estimate the model on data from an earlier cohort born in 1970. Comparing the distributions of earnings and other factors from the earlier with the later cohort allows us to assess their role in the huge growth in higher education seen over this period. The distribution of the expected graduate-wage premium remained quite stable over this period, with its mean and dispersion only decreasing slightly. The distrubution of the other factors premium, however, changed drastically, shifting right so that the strongly negative mean of the 1970 cohort became slightly positive for the 1990 cohort. The variance of other factors premium distribution also increased slightly. Taken together these results suggest the increase in degree attainment—which went from 29.9% in the 1970 sample, to 62.2% in the 1990—was entirely driven by changes in the other factors premium.

1.1 Related literature

This paper joins a long tradition of studying educational and occupational decisions in economics and social science. Arguably this tradition in economics begain with Roy's seminal 1951 paper on occupational choice. Roy models (and their extensions) have since been applied to educational choice, a literature which includes an important series of papers by Cunha, Heckman and coauthors (Cunha, Heckman, and Navarro, 2004; Cunha and

[1] These correspond to the bottom quintile of parental earnings in the sample (low SES), the middle-three quintiles (middle SES), and the top quintile (high SES).

Heckman, 2007; Heckman et al., 2006). These and more recent papers highlight the importance of non-pecuniary factors (often called "psychic costs") in explaining educational choices, both at the intensive (e.g. major choices in Wiswall and Zafar (2015) and institutional choices in Delavande and Zafar (2019)) and extensive margins (D'Haultfoeuille and Maurel, 2013; Boneva and Rauh, 2020). Arcidiacono, Hotz, Maurel, and Romano (2020) show the importance of non-pecuniary factors in occupational choice.

Our contributions to this literature are the following. We exploit information on realised earnings and choices, linked to young people's subjective beliefs about the advantages and disadvantages of attending university, including about the non-pecuniary aspects. A growing literature studies young people's choices by eliciting expecations about future earnings from students, but in general these do not elicit expectations about non-pecuniary factors (Manski, 1993; Dominitz and Manski, 1996; Arcidiacono, Hotz, and Kang, 2012; Arcidiacono et al., 2020).[2] For much of this important work realised outcomes are not (yet) available.[3] In addition, most of the prior work using elicited expectations has often used smaller, selected samples, either from a single US college (Arcidiacono et al., 2020) or self-selected survey respondents (Boneva and Rauh, 2020). Our data is from a large, representative sample.

There is also a growing recent literature on the differences across socio-economic groups in education attainment and choices, work to which this paper is closely related. The persistence of the educational attainment gap between more- and less-advantaged students, even during a period of huge expansion in higher education, is documented by Blanden and Machin (2004). More recently, differences across social groups in terms of subject and institution have been highlighted as a key driver of differences in labour market outcomes (Britton, van der Erve, Belfield, Dearden, Vignoles, Dickson, Zhu, Walker, Sibieta, and Buscha, 2021). The drivers of these differences are also beginning to be explored. Boneva and Rauh (2020) study the role of students' beliefs about pecuniary and non-pecuniary outcomes in the decision to attend university in England. Anders (2012) and Anders and Micklewright (2015) investigate how young people's expectations about applying to university evolve differently between the ages of 14 and 17 according to their socio-economic group, using the same cohort study that we analyse in this paper. We contribute to this literature by comparing and quantifying earnings and other factors in the decision to attend university across socio-economic groups.

Finally, through our analysis of the evolution of factors in the decision to attend university over two decades, our work is related to the literature studying the recent growth in university attendance in the UK. Higher education in England has seen substantial changes

[2]Boneva and Rauh (2020) is a notable exception.

[3]Outcomes are beginning to become available for some surveys which elicited expectations (see for example Arcidiacono et al. (2020) and Gong, Stinebrickner, and Stinebrickner (2019)).

in recent decades, undergoing substantial growth and an overhaul of its funding system. Growth in attainment has been much steeper in the UK than in the US. The proportion of UK (US) BAs in a given cohort at age 30 increased from less than 10% (25%) for those born in 1950, to nearly 40% (35%) for those born in 1985 (Blundell, Green, and Jin, 2021). Alongside this rapid growth in attainment, the graduate wage premium has remained flat in the UK, while it has been steadily increasing in the US (Blundell et al., 2021). However, this recent growth in higher education in the UK did not occur equally across socio-economic groups, with the children of richer parents disproportionately benefitting from the expansion (Blanden and Machin, 2004). Walker and Zhu (2008) study the impact of this expansion on the graduate wage premium, finding no change for men, and a slight increase for women. Such rapid growth in higher education, over a period of increased fees and stagnant returns, raises questions about what drove so many more people to attend university — questions we shed new light upon in this paper.

Outline. The rest of the paper is organised as follows. Section 2 describes the data we use in this paper, and presents some initial analysis. Section 3 describes the model we estimate to obtain our main results, with our estimation stratgy in section 3.2. Our main results are in section 4, followed by analysis by socio-economic group (section 4.2) and over time (section 4.3). Section 5 concludes.

2 Context and data

The data used in this paper come from *Next Steps*, a British cohort study which follows a representative sample of 15,770 people born in England in 1989 or 1990 (IOE Centre for Longitudinal Studies, 2018). These young people were able to leave school at age 16, in 2006, and those who went on to higher education would have entered university at age 18 or 19, in 2008 or 2009. They made the decision to apply to university at age 16 or 17, after deciding whether to stay in full-time education after 16. These students would face tuition fees of around 3,000 GBP per year, though there are extensive government-provided loans and grants available to cover both tuition fees and living costs. A detailed discussion of the application process, and the UK system of tuition fees, loans and grants is in appendix A. Typical university degrees in the UK are 3-year bachelor's degrees, with a few subjects offering longer courses as standard.[4] Therefore, the majority of young people who choose to attend university will have graduated by the time they are 25.

Two important periods for this paper are just before young people apply to university (age 16 or 17), and when the majority have entered the labour market including those who attend university (age 25). We have data on the *Next Steps* cohort members at both these

[4]Engineering is often a 4-year combined bachelor's and master's degree, while an undergraduate medicine degree takes at least 5 years.

Table 1: Description of selected variables

Variable	Description
Earnings	Usual weekly earnings in GBP reported by the CM if employed at age 25 (wave 8).
Education	An indicator variable for whether the CM reports having an UG degree at age 25 (wave 8).
# of A-levels	The number of A-levels the CM reported taking at age 16 / 17 (wave 4).
Parental income	Total annual parental income when CMs were 16/17 (wave 4). The data is recorded in 12 bins.
Subjective beliefs	Harmonised, open-ended responses to questions about the advantages and disadvantages of attending university (wave 4).

key stages. Data collection involved annual face-to-face interviews between 2004 and 2010 (waves 1–7), plus a further round of interviews in 2015 (wave 8).[5] In our analysis we use information on: schooling, family background, and subjective beliefs about university and the future at age 16 (before applying to university); and on earnings and qualifications at age twenty-five (after entry to the labour market). A selection of these variables are described in table 1. Importantly, we have a direct measure of students' beliefs about the future. We supplement the data from Next Steps with data from the earlier British Cohort Study (BCS) to analyse changes in factors across time. The BCS is a similar study to Next Steps, following nearly 17,000 people born in the UK in a single week in April 1970.

Data description. Table 2 presents summary statistics from waves 4 (CMs aged 16) and 8 (CMs aged 25) of Next Steps, for all cohort members in our sample, and then split by whether they hold a degree at 25. Only those with a minimum of 5 GCSEs at A*-C or equivalent were asked about their subjective beliefs about university, information vital to the analysis in this paper.[6] The young people not asked about their expectations are not included in the analysis. We also drop young people for whom we do not observe a wage at age 25 (and those who reported wages above the 99th or below the 1st percentiles), who did not respond in either wave 4 or wave 8, who did not answer question about their perception of their ability, or who did not provide information on their qualifications at age 25. Although dropping students without 5 GCSEs at A*-C or equivalent reduces the size of our sample significantly, those omitted are likely students who would have found it very difficult to attend university. They are an important group to study, but their

[5]The study is ongoing and the cohort members will be interviewed again in 2021, with plans to make the data available by 2023.

[6]These are referred to as "high-achieving" students in the survey documentation (IOE Centre for Longitudinal Studies, 2008). Blundell et al. (2021) consider grade C at GCSE as the UK equivalent to high-school graduation in the US.

Table 2: Descriptive statistics

	All	$D=0$	$D=1$
N	3,469	1,311	2,158
Female	0.57	0.56	0.57
Ethnicity			
White	0.73	0.81	0.68
South Asian	0.16	0.10	0.20
Black	0.04	0.03	0.05
Other	0.07	0.06	0.07
Main parent's SOC			
SOC 1–3	0.42	0.35	0.46
SOC 4–5	0.22	0.24	0.21
SOC 6–9	0.36	0.41	0.33
Self-assessed ability†	0.25	0.08	0.34
# A-levels	3.74	3.43	3.92
Degree	0.63	0.00	1.00
Earnings at 25 (weekly in GBP)			
Mean	451.90	416.58	473.01
Std. dev.	197.00	192.84	196.44

Notes: †A composite measure combining students' response to asked a series of questions about their perceived abilities in maths, science, and english. All information except earnings and qualification was recorded at age 16 or 17, in waves 4 and 5.

omission is not fatal to the current analysis.

2.1 Subjective beliefs about university

Information on students' beliefs about their future potential life at and after university is a key feature of this paper. These subjective beliefs were collected as open-ended responses to two questions, one about the advantages and one about the disadvantages of university.[7] The CMs could mention as many or as few advantages (disadvantages) as they wished. The interviewer noted down their interviewee's responses, and similar responses across individuals were then grouped into the harmonised responses we use in our analysis. These harmonised responses are listed in tables 3 (advantages) and 4 (disadvantages), grouped into the following categories: career (non-pecuniary); earnings; financial / debt; social life / environment; education; and time. These tables also display the proportion of young people who mentioned each response, overall and separately for graduates and non-graduates. The final column is the difference in proportion of graduates ($D=1$) and non-graduates ($D=0$) who mentioned each response, expressed in percentage points (pp).

[7]The exact wording of the question(s) was: "What do you think the advantages (disadvantages), if any, might be for someone of going to university to study for a degree?" (IOE Centre for Longitudinal Studies, 2008).

Figure 1 also shows the overall numbers of young people who mentioned each recorded advantage and disadvantage of attending university, ordered by number of mentions, rather than in categories. Focusing first on the reported advantages in figure 1a, access to "better opportunities" and to "better jobs" were the two most common advantages of a university degree mentioned by respondents. In close third was "more qualifications", with getting a "well-paid job" in fourth place. An enjoyable "social life" rounds out the top five most popular advantages, with "learning", "personal development", and "gain life skills" also popular responses. Although some of the responses are arguably linked to higher pay, there are many that are not, for example "social life" and "personal development". In addition, the presence of "well-paid job" as a specific response suggests other career-related responses are capturing broader notions than pay alone.[8] Turning to the disadvantages in figure 1b, the three responses mentioned most often are all financial concerns: "get into debt", "costs (general)", and "too expensive" — concerns which arguably still "pecuniary". However, many of the disadvantages mentioned reflect fully non-pecuniary aspects of a person's career ("no job guarantee"), or life at university ("heavy workload", "leave home"). Together these responses provide information on students' beliefs about the pecuniary, financial, and non-pecuniary aspects of attending university.

Returning to tables 3 and 4 we can take a higher level view, with the advantages and disadvantages displayed in broad categories. We can also see which categories were the most mentioned by young people. Focusing first on the advantages in table 3, the two most mentioned advantages are related to the (non-pecuniary) aspects of a person's future career. This is the most mentioned category, with over 66% of the sample mentioning something to do with their career as an advantage. Education is the next most popular category, followed by earnings, personal development, and then social life, which is still mentioned by nearly 17% of our sample.

2.2 Beliefs and university attendance

We also compare the answers of young people who go on to attend university ("graduates") and those who do not ("non-graduates"). The third and fourth columns of tables 3 and 4 show the proportion of respondents who mentioned each advantage or disadvantage of university separately for graduates ($D = 1$) and non-graduates ($D = 0$). The final column in these tables shows the difference in proportion between graduates and non-graduates mentioning each aspect of university. Generally the advantages of university were more likely to be mentioned by graduates (signalled by a positive difference in the final column of table 3), with all but one of the top-8 advantages being mentioned by a

[8] While it is difficult to know exactly what respondents meant by "better opportunities" and "better jobs", the presence of another (harmonised) response specifically referring to "well-paid job" suggests these might refer to a broader notion of quality than is captured by pay alone.

Table 3: Students' subjective beliefs about university (advantages)

	Prop. mentioning (%)			
Response (harmonised)	All	$D=0$	$D=1$	Diff. (pp)
Career (non-pecuniary)	66.27	58.52	70.91	12.39
Will lead to a good / better job (than would otherwise get)	32.37	29.01	34.32	5.31
Gives someone better opportunities in life	33.12	27.66	36.36	8.71
Is essential for the career they want to go into	3.32	3.80	3.02	-0.78
Shows that you have certain skills	1.82	1.86	1.78	-0.09
To delay entering work / more time to decide on a career	0.63	0.54	0.67	0.13
Earnings				
Will lead to a well paid job	21.33	19.29	22.51	3.22
Social life / environment	16.18	13.95	17.51	3.56
The social life / lifestyle / meeting new people / it's fun	15.22	13.23	16.38	3.16
To leave home / get away from the area	2.02	1.76	2.17	0.42
Education	34.01	35.31	33.23	-2.08
To carry on learning / I am good at / interested in my chosen subject	8.65	7.07	9.57	2.50
Get more / better / higher qualifications	26.38	28.77	24.90	-3.87
Personal development	17.68	13.52	20.17	6.65
Makes someone independent / maturity / personal development	8.76	6.04	10.36	4.33
Gives you more confidence	0.95	0.48	1.22	0.75
People will respect me more	0.35	0.34	0.35	0.01
Leads to a better life / good life (general)	2.13	1.89	2.27	0.38
Prepare you for life / gain life skills	7.52	6.33	8.25	1.92

Notes: Students with at least 5 GCSEs at A*–C or equivalent were asked these questions ($N = 3,469$). Where there are values in the right hand columns for the categories (in bold), these are the proportion of young people who mention at least one of the advantages in that category.

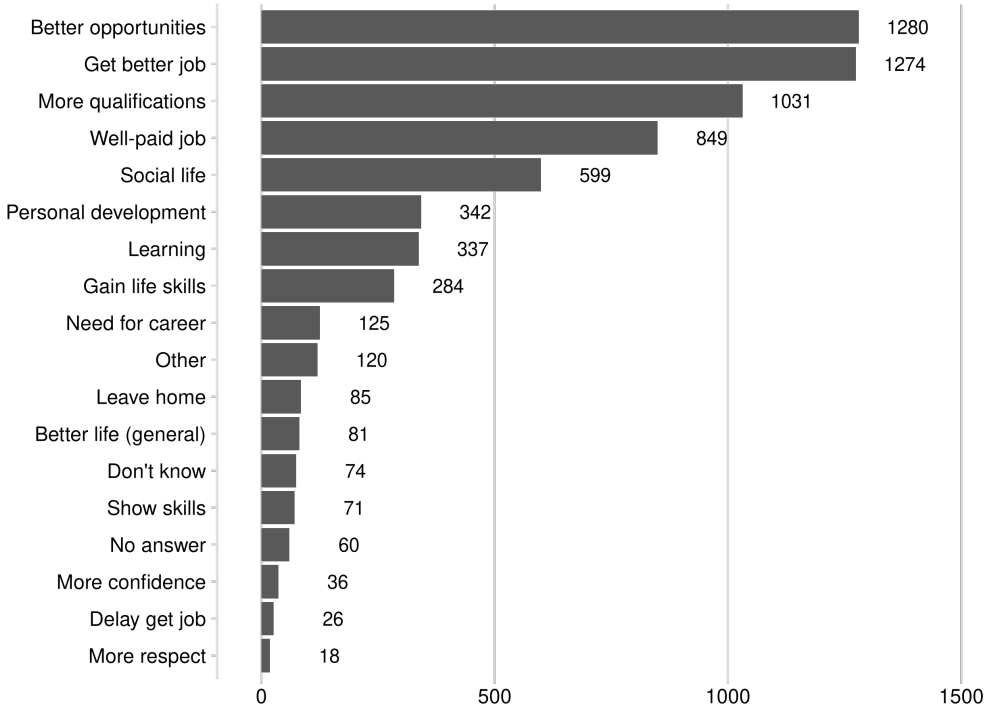
Table 4: Students' subjective beliefs about university (disadvantages)

	Prop. mentioning (%)			
Response (harmonised)	All	$D=0$	$D=1$	Diff. (*pp*)
Career (non-pecuniary)	14.79	14.52	14.95	0.43
No guarantee of a good job at the end	6.89	6.48	7.14	0.66
Don't need to go to university for the job someone may want	0.49	0.62	0.41	-0.20
Get less work experience	1.96	2.23	1.79	-0.44
Financial / debt	73.67	69.62	76.09	6.47
Now				
It is expensive	9.74	8.76	10.33	1.56
Not becoming financially independent	0.98	0.89	1.02	0.13
Not being able to start earning money / start work	6.34	6.54	6.21	-0.33
Costs (general / non specific)	23.18	21.14	24.38	3.24
Tuition fees / Accommodation costs / Living expenses	3.89	2.69	4.58	1.89
Future				
Getting into debt/have to borrow money	37.36	36.36	37.90	1.54
Social life / environment	9.37	9.75	9.14	-0.61
Leaving home / family / friends	8.16	8.90	7.72	-1.18
Stress	1.38	1.00	1.60	0.60
Education				
The workload can be hard / doubts about ability to finish course	6.00	5.37	6.38	1.00
Time	8.19	9.53	7.38	-2.15
Takes a long time	7.09	8.27	6.39	-1.88
Waste of time (general / non-specific)	1.18	1.34	1.09	-0.25

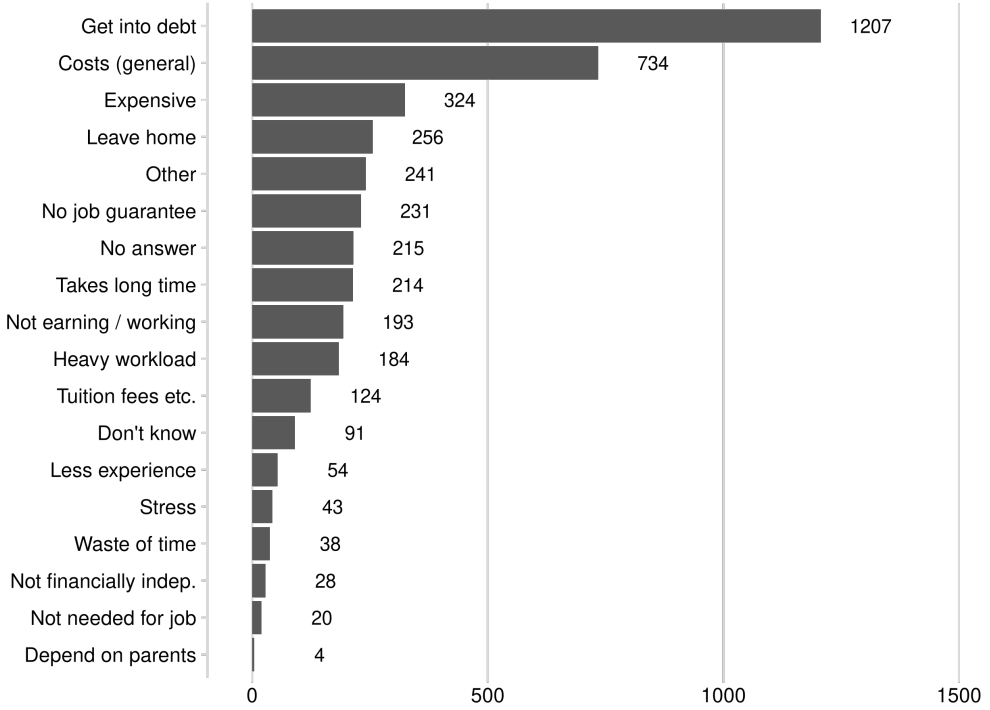
Notes: Only young people with at least 5 GCSEs at A*–C or equivalent were asked these questions ($N = 3,469$). Where there are values in the right hand columns for the categories (in bold), these are the proportion of young people asked who mention at least one of the disadvantages in that category.

Figure 1: Proportion of students who mentioned specific advantages and disadvantages about going to university

(a) Harmonised *advantages*



(b) Harmonised *disadvantages*



Notes: Only students with at least 5 GCSEs at A*–C or equivalent were asked these questions ($N = 3,469$).

higher proportion of graduates.

The disadvantages are more balanced in terms of who is more likely to mention them, with the majority of differences in the final column of table 4 smaller than 2*pp* in magnitude. Still, somewhat surprisingly, the top 3 disadvantages are all more likely to be mentioned by graduates (and are all related to financial aspects of attending university). Our initial analysis suggests those who go on to university are more likely to believe it will be beneficial for their career (and earnings), for their personal development, and for their social lives. Interestingly, future graduates also seem to worry more about the potential financial downsides of attending university — perhaps suggesting that these are not a major barrier to university attendance in the UK.

As a first step towards comparing and quantifying the factors that potential students consider when deciding whether to attend university, we estimate a probability model to assess the predictive content in their reported beliefs. We started by estimating logit models with an indicator for holding a degree at age 25 the dependent variable, and all recorded advantages and disadvantages as (binary) indepedent variables. Estimates of key parameters are in appendix B, table B2.[9] Many of the estimates are sizeable, but they are not very precisely estimated, as evidenced by the large standard errors. There are also a large number of estimates, which combined with their (im)precision, makes this model difficult to interpret.

To address these issues, we estimate a similar model using indicators for mentioning any response in each of the broader categories in tables 3 and 4 in place of the harmonised responses. The estimated parameters from this model are more straightforward to interpret. These are presented in table 5. The estimates in column (1) were obtained by regressing an indicator for holding a university degree at age 25 on indicators for mentioning any response from each of the broad categories of responses, and in column (2) we also include the following background characteristics as covariates: ethnicity, gender, A-levels, parental income, and a self-assessed ability measure.. The signs next to the categories in parentheses signal whether the responses in that category are described as advantages (+) or disadvantages (−).

The results of this exercise reflect our initial findings, with the categories with the biggest gaps between mentions by graduates and non-graduates in tables 3 and 4 having the largest coefficients. Importantly, even when we add a range of background controls in column (2) many of the coefficients remain sizable and statistically different from zero. This suggests that these questions are capturing variation in beliefs across individuals

[9] As these are qualitative survey responses, they are coded as indicator variables and are relative to a reference category, which is those who did not mention the corresponding advantage or disadvantage when surveyed. Also included are a range of background characteristics (ethnicity, gender, A-levels and parental income) for which we do not report the parameter estimates.

which impact their decision to attend university. The most important factors in the decision to attend university continue to be related to career (advantage), earnings (advantage), personal development (advantage), and the time it takes to get a degree (disadvantage).

The initial analysis presented in this section has highlighted the different factors young people consider when applying to university, and we have made a start at comparing their relative importance. However, although one of the responses we include in the model is about earnings, missing from our analysis so far is a proper measure of wages. Including earnings in our model will further benefit our analysis in (at least) two ways: (i) it will provide a better measure of the pecuniary benefits of attending university; (ii) comparing the contributions of other factors in the decision to wages will anchor these contributions to an interpretable metric.

3 Empirical framework

In this section we present a framework designed to allow us to compare and quantify the contributions of different factors in the decision to attend university, exploiting information on realised earnings, observed choices, and subjective beliefs. We build upon the analysis of the previous section by introducing a simple model which allows us to combine these different sources of data answer our research question: what is the relative importance of wages and non-pecuniary factors in the decision to attend university?

We start by introducing the key objects of the model and describing the behaviour of young people in our setup, along with some key assumptions we make to ensure our model is identified.

Utility of university or work. An individual's utility from choosing university ($s=1$), or work ($s=0$) is a linear combination of these different factors

$$U_{s,i} = \alpha Y_{s,i} + \theta'_{s,i}\gamma + Z'_{16,i}\delta_s + \epsilon_{s,i} \tag{1}$$

where Y_s represents the pecuniary factors (the logarithm of log weekly earnings in our application), θ_s is a vector of non-pecuniary (non-earnings) factors on which we have information, Z_{16} contains individual characteristics that may impact the non-pecuniary costs / benefits of university, and ϵ_s is a mean-zero random-utility term, all conditional on choice s.

Decision to attend university. At the time young people make their decision, they do not know the value that many of these outcomes will take, and so form expectations

Table 5: Logit estimates (response categories)

Dependent variable:	Degree	
	(1)	(2)
Earnings	0.354^{***}	0.294^{***}
	(0.094)	(0.101)
Career (+)	0.658^{***}	0.510^{***}
	(0.081)	(0.088)
Career (−)	0.013	−0.054
	(0.102)	(0.110)
Financial (−)	0.169^{**}	0.271^{***}
	(0.084)	(0.094)
Social life (+)	0.272^{***}	0.153
	(0.103)	(0.110)
Social life (−)	-0.250^{**}	−0.121
	(0.127)	(0.138)
Education (+)	0.129	0.106
	(0.081)	(0.087)
Education (−)	0.197	0.354^{**}
	(0.156)	(0.169)
Personal development (+)	0.617^{***}	0.517^{***}
	(0.102)	(0.108)
Time (−)	-0.281^{**}	-0.253^{*}
	(0.129)	(0.141)
Constant	-0.282^{***}	-0.775^{***}
	(0.102)	(0.223)
Observations	3,469	3,469
Log Likelihood	−2,237.408	−2,003.386
Akaike Inf. Crit.	4,496.816	4,088.772

Notes: *p<0.1; **p<0.05; ***p<0.01

Column (1) contains estimates from a logistic regression of an idndicator for holding a university degree at age 25 on indicators for mentioning each of the broad categories of responses in tables 3 and 4. Column (2) contains estimates for a regression also including the following background characteristics: ethnicity, gender, A-levels, parental income, and a self-assessed ability measure.

about their utility under each choice, based on their information set, $\mathcal{I}_i$:

$$\mathbb{E}[U_{s,i}|\mathcal{I}_i] = \mathbb{E}[\alpha Y_{s,i} + \theta'_{s,i}\gamma + Z'_{16,i}\delta_s + \epsilon_{s,i}|\mathcal{I}_i] \tag{2}$$

Individuals compare their expected utility of attending university, $U_1^{\mathcal{I}}(\equiv \mathbb{E}[U_1|\mathcal{I}_i])$, to that of working, $U_0^{\mathcal{I}}$, and choose the option with the higher expected utility. Therefore,

$$S \equiv \mathbb{1}\{U_1^{\mathcal{I}} - U_0^{\mathcal{I}} > 0\}. \tag{3}$$

This can be rewritten as the difference between expected (pecuniary) outcomes, and expected "costs" of attending university, in the spirit of Roy (1951).

$$S \equiv \begin{cases} 1, & \text{if} \quad \alpha(Y_1^{\mathcal{I}} - Y_0^{\mathcal{I}}) + (\theta_1^{\mathcal{I}} - \theta_0^{\mathcal{I}})'\gamma + Z'_{16}(\delta_1 - \delta_0) + \epsilon_1^{\mathcal{I}} - \epsilon_0^{\mathcal{I}} > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

This formulation leads naturally to an expression for the probability of attending university, conditional on expected earnings ($Y_s^{\mathcal{I}}$) and observed non-pecuniary factors (θ, Z_{16})

$$\Pr(S = 1|\mathcal{I}) = \Pr\left(\alpha(Y_1^{\mathcal{I}} - Y_0^{\mathcal{I}}) + (\theta_1^{\mathcal{I}} - \theta_0^{\mathcal{I}})'\gamma + Z'_{16}(\delta_1 - \delta_0) > \epsilon_0^{\mathcal{I}} - \epsilon_1^{\mathcal{I}}\right) \tag{5}$$

A chief aim of this paper is to estimate the relative importance of the pecuniary and non-pecuniary factors in the decision; i.e. how important is $\alpha(Y_1^{\mathcal{I}} - Y_0^{\mathcal{I}})$ versus $[(\theta_1^{\mathcal{I}} - \theta_0^{\mathcal{I}})'\gamma + Z'_{16}(\delta_1 - \delta_0)]$ when evaluating this conditional probability.

Expectations about (future) earnings. We need to specify exactly how young people form expectations about Y_s: what is in their information set, and how they use this information to form their expectations. We make the following assumptions:[10] (i) young people know the *true* process generating future incomes, but they only possess very limited information about the future — their information set reflects their current observable characteristics, i.e. X_{16}; (ii) young people only consider their earnings at age 25 (or these are a sufficient statistic for what they consider) when deciding whether to go to university.

Put differently, they are very good at predicting mean *realised* earnings among their peers conditional on X_{16}, but they are not very good at predicting their own *future*

[10]The current gold standard are elicited expectations about earnings, following the advice of Manski (1993). However, these are rare, especially in large representative samples. Our assumption that young people know the true income process is standard in economics. For example, Cunha and Heckman (2007) assume this, and develop a method for testing the contents of young people's information sets. Willis and Rosen (1979) and D'Haultfoeuille and Maurel (2013) also assume a similar model for expected earnings, though they allow for an unobserved component in the young people's information set, i.e. $\mathbb{E}[Y_s|X_{16}, \eta_1, \eta_0]$, where η_s are not observed by the econometrician. We assume that we observe all the information that young people use to form their expectations about earnings. We plan to compare our model (and subsequent results) with other models of expectations / information sets in future work.

characteristics (or they do not know how earnings depend on their future characteristics). Then $Y_s^{\mathcal{I}} \equiv \mathbb{E}[Y_s \mid X_{16}]$. Under these assumptions earnings expectations, $Y_s^{\mathcal{I}}$ are identified from *realised* earnings, and the students' characteristics at 16. As we only observe either Y_1 or Y_0 for each individual, these assumptions allow us to estimate expected wages $Y_1^{\mathcal{I}}$ and $Y_0^{\mathcal{I}}$ for young person, and hence an expected graduate wage premium, $Y_1^{\mathcal{I}} - Y_0^{\mathcal{I}}$.

Expectations about other (non-pecuniary) factors. We use the harmonised responses to open-ended questions about the advantages and disadvantages of going to university, discussed in detail in section 2, to measure the expected other factors premium, $\theta_1^{\mathcal{I}} - \theta_0^{\mathcal{I}}$. Limited somewhat by the nature of these questions, we assume that individuals either believe there to be no difference in this factor whether they go to university or not, or they believe there will be a difference, which is fixed to be of constant size across all individuals who hold this belief. Therefore for each factor mentioned by *any* student, the component of $\theta_{1,i}^{\mathcal{I}} - \theta_{0,i}^{\mathcal{I}}$ takes one value (normalised to 1) if mentioned by student i, and another value (normalised to 0) if not mentioned. The parameter γ_j on the j-th component of $\theta_{1,i}^{\mathcal{I}} - \theta_{0,i}^{\mathcal{I}}$ then reflects (average) preferences for this aspect of university. We also allow the non-pecuniary factors to vary with individual characteristics, captured by the vector Z_{16}.

3.1 Identifying the parameters in the utility function

Recall the probability of attending university, conditional on expectations about earnings and other factors, in the model:

$$\Pr\left(S = 1 | Y_1^{\mathcal{I}} - Y_0^{\mathcal{I}}, \theta_1^{\mathcal{I}} - \theta_0^{\mathcal{I}}, Z_{16}\right) = \Pr\left(\alpha(Y_1^{\mathcal{I}} - Y_0^{\mathcal{I}}) + (\theta_1^{\mathcal{I}} - \theta_0^{\mathcal{I}})'\gamma + Z_{16}'\delta > \epsilon_0^{\mathcal{I}} - \epsilon_1^{\mathcal{I}}\right). \quad (6)$$

Identification of α, $\delta \equiv \delta_1 - \delta_0$ and γ then requires assumptions on the distribution of the random-utility terms, ϵ_1 and ϵ_0. A standard assumption in the discrete-choice literature is that these follow a type-I extreme-value distribution, meaning their difference follows a logistic distribution: $(\epsilon_0 - \epsilon_1) \sim \text{Logit}$. The parameters α, δ and γ capture the relative contribution of earnings and observed other factors to young people's utility, and hence in their decision to attend university. These parameters are only identified up to a scale normalisation.

3.2 Estimation

Having laid out the assumptions we make to identify our model, we now describe our estimation strategy.

Expected graduate-wage premium, $Y_1^{\mathcal{I}} - Y_0^{\mathcal{I}}$. Under our model for young people's expectations, the expected earnings we need are $\mathbb{E}[Y_s|X_{16}]$. Given X_{16} we use OLS to estimate this conditional expectation.[11] We estimate the simplest linear conditional expectation for each level of education, with no interactions. We then use the estimated coefficients, $\hat{\beta}_{s,16}$, to obtain estimates $\widehat{Y}_s^{\mathcal{I}} = X_{16}'\hat{\beta}_{s,16}$. The estimated expected graduate-wage premium is simply $\widehat{Y}_1^{\mathcal{I}} - \widehat{Y}_0^{\mathcal{I}} = X_{16}'(\hat{\beta}_{1,16} - \hat{\beta}_{0,16})$. We include the following covariates in X_{16}: parents' occupations, parents' education level, a measure of parental income, the number of A-levels a student is taking, gender, and whether high pay is important to them.

The parameters of the utility function, α, δ and γ. We estimate the parameters of the utility function using logistic regression. To avoid perfect multicollinearity when estimating equation (6), X_{16} must not be a subset of Z_{16}. An alternative would be to transform log-wages by some (non-linear) function. We choose to exclude beliefs relating to earnings from Z_{16}.

Distributions of earnings and other factors. The aim of this paper is to compare and to quantify the roles of earnings and non-pecuniary factors in the decision to attend university. To do this, we estimate comparable distributions of the different factors using the following strategy: (i) obtain estimates $\hat{\alpha}$, $\hat{\gamma}$, and $\widehat{Y}_1^{\mathcal{I}} - \widehat{Y}_0^{\mathcal{I}}$; (ii) recombine these estimates with the data $(X_{16}, \theta_1^{\mathcal{I}} - \theta_0^{\mathcal{I}}, Z_{16})$, to calculate the "contribution" of each (type of) factor; (iii) transform these utility values to be equivalent to a difference in log-earnings; (iv) use a kernel-density estimator to plot the empirical distributions.

4 Results

In this section we present and discuss the results of estimating the model described in section 3. We first present results for the full sample, and then study variation across different socio-economic groups and over time.

[11] By using OLS we do not control for selection. Young people base their decision on their expected graduate and non-graduate earnings, $Y_s^{\mathcal{I}}$. Our assumption about how they form these expectations means that we have access to the same information they do—therefore, there is not selection on unobservables. As mentioned in a previous footnote, we plan to test different models of assumptions in future work. We did attempt to estimate different wage equations allowing for selection, which we present in the appendix, figure C1. When assuming normal errors following Heckman (1979) we found that for the majority of CMs their graduate earnings premium $\widehat{Y}_1^{\mathcal{I}} - \widehat{Y}_0^{\mathcal{I}}$ was negative, suggesting that the normality assumption is problematic. Estimating a two-stage model with a more flexible first-stage (using Coppejans (2001) mixture-of-distributions (MOD) estimator) resulted in an almost identical observed graduate wage premium distribution.

Table 6: Summary statistics for the earnings and other factors premium distributions

	Mean	$q_{.1}$	$q_{.25}$	$q_{.5}$	$q_{.75}$	$q_{.9}$
Total	0.14	-0.15	-0.01	0.14	0.30	0.43
Earnings	0.08	-0.05	0.01	0.08	0.15	0.21
Non-earnings	0.06	-0.23	-0.09	0.06	0.21	0.36
Career	0.03	0.00	0.00	0.05	0.05	0.05
Financial	0.05	0.00	0.00	0.06	0.06	0.07
Education	0.01	-0.01	0.00	0.00	0.00	0.06
Personal dev.	0.02	0.00	0.00	0.00	0.00	0.06
Social life	0.00	0.00	0.00	0.00	0.00	0.02
Other	0.18	-0.09	0.04	0.17	0.31	0.45

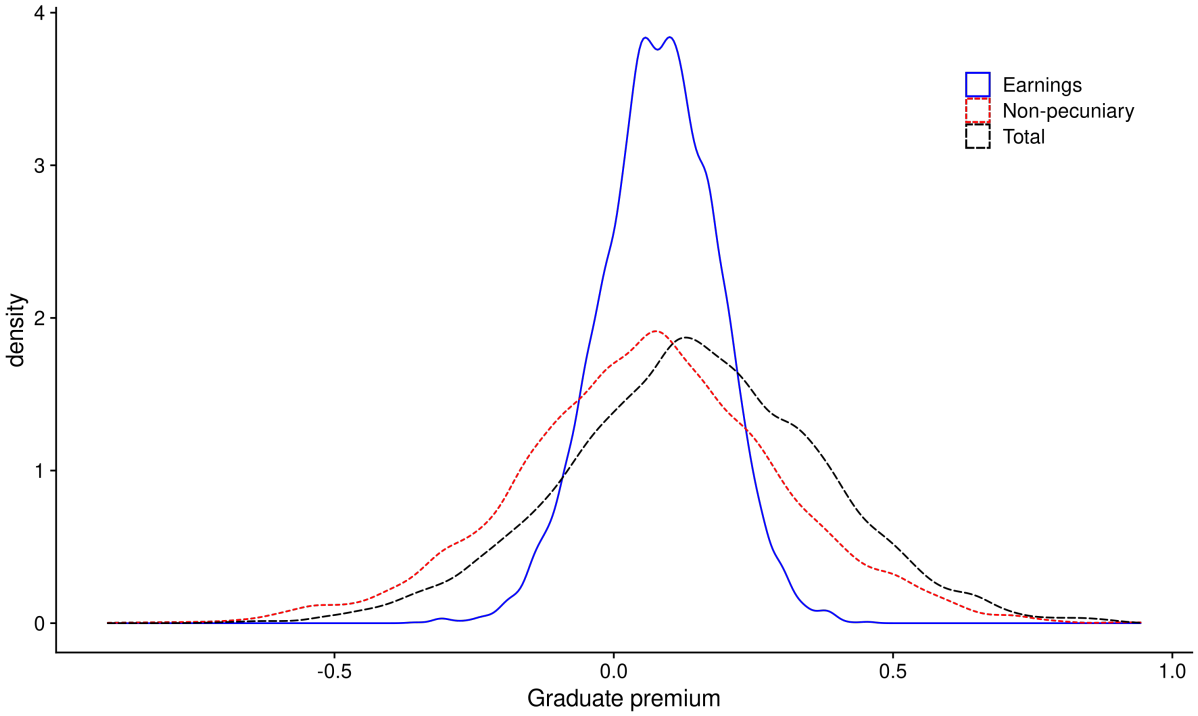
Notes: The values for "Time" are omitted as they do not vary between the 1st and 9th deciles. The values are in units equivalent to a difference in log-wages.

4.1 Main results

Kernel density estimates of the distributions of earnings premiums (blue solid line) and other factors premiums (red dashed line) are presented in figure 2. Table 6 presents further statistics on these distributions: their mean, first and last deciles, and quartiles. The locations of the two distributions are remarkably similar, evidenced by their similar means: 8 log-points (earnings) and 6 log-points (non-earnings). However, the dispersion of the other factors premium is much higher. This difference is visible both in figure 2, and by comparing the interquartile ranges in table 6. The interquartile range is 14 log-points for the earnings premium, while it is over twice as large at 30 log-points for the non-earnings factors. The same is true of the interdecile range (26 versus 59 log-points), and the standard deviation (10 versus 24 log-points).

We can also compare the number of young people who have positive values of each factor. Among graduates, 80% have a positive value of pecuniary factors and 73% of non-pecuniary factors. Therefore, 20% of graduates still attend university despite expecting negative pecuniary returns. For non-graduates, 76% expect positive pecuniary returns, though only 42% expect positive non-pecuniary returns. Therefore, it appears to be chiefly the influence of these non-earnings factors that determines whether a young person decides to attend university, a role reflected in the similarity between the distribution of all factors in the decision (black long-dashed line, figure 2) and the other factors (red dashed line, figure 2).

Figure 2: Distributions of earnings and other factors premiums in the decision to go to university



Notes: The values of the factors are estimated as described in 3.2. The distributions are then estimated (and plotted) with the kernel density estimator in the `R` package `ggplot2`, using the default Gaussian kernel and bandwidth (Wickham, 2016).

4.1.1 Counterfactual exercise

In order to further highlight the importance of non-earnings factors, we perform the following counterfactual exercise, borrowed from D'Haultfoeuille and Maurel (2013). We calculate the predicted probabilities of university attendance under different (fixed across the sample) values of the pecuniary and non-pecuniary factors. The results of this exercise are in table 7. If everyone in the sample had other factors equal to the 10th percentile value, only 35% of people would attend university—over 27*pp* fewer than did actually attend. Moreover, assigning everyone other factors equal to the 90th percentile results in over 86% of people attending university, an increase of over 20*pp*. Conversely, varying the expected graduate-wage premium between the 10th and 90th percentiles has a much smaller effect on university attendance. Over 50% of young people would still attend if they expected a pecuniary premium equal to the 10th percentile, while 73% would attend if we fix all expectations about the pecuniary gains at the 90th percentile. This emphasises the key role non-pecuniary factors play in the decision to attend university.

4.1.2 Decomposing non-pecuniary factors.

So far our analysis has considered all non-earnings factors together. However, using information on young people's beliefs, we can attempt to decompose these non-pecuniary

Table 7: University attendance under counterfactual factor values

Counterfactual	Earnings	Other	University (%)
Data	0.08	0.06	62.2
Earnings			
10th percentile	-0.053	-	50.8
25th percentile	0.014	-	56.8
75th percentile	0.152	-	68.4
90th percentile	0.210	-	72.7
Other			
10th percentile	-	-0.233	34.7
25th percentile	-	-0.087	49.4
75th percentile	-	0.212	77.2
90th percentile	-	0.357	86.1

Notes: The "units" of earnings and other factors are equivalent to a difference in log-wages. University is the fraction who attend under the counterfactual distribution. The values in row "Data" are: the median values of earnings and other factors premiums.

factors. We do so by calculating the portion of non-pecuniary returns attributable to variation in a given aspect of university. For example, the following aspects of university mentioned by students are related to their career: get better job, better opportunities, need for career, show skills, delay get job, not earning / working, no job guarantee, not needed for job, and less experience. Therefore, we can calculate the expected values of non-pecuniary career-related returns using these variables and their estimated coefficients.

The summary statistics for the distributions of non-pecuniary returns associated to career, financial, educational, personal development, social life, and other are in table 6. The allocation of responses to each category is detailed in table B1, and also corresponds to the categorisation in tables 3 and 4. The category "time" is omitted from table 6 as the values of this factor did not vary between the 10th and 90th percentiles. The variation in "other" is due to variation in non-pecuniary factors captured by the included background characteristics. Although some varation in non-pecuniary returns is associated with the recorded beliefs, the majority is due to these background characteristics. Given the limited variation in the indicator variables we rely upon to measure beliefs, it is likely that the variation attributed to each category in table 6 represents a lower bound for the true contributions of these aspects to the non-pecuniary returns to university.

Nevertheless, we can still say something about the contributions to non-pecuniary returns for some of these aspects of (life at and after) university. For both career-related and financial non-pecuniary returns, the 25th percentile value is zero, while the median is 5 log points. This variation is similar to that of pecuniary returns, which range from 1 log-point at the 25th percentile to 8 log-points at the median. Therefore, despite failing

to explain the majority of variation in non-pecuniary returns, the portion of these returns which is explained by variation in beliefs is still sizable compared to expected pecuniary returns.

4.1.3 Limitations

The analysis in this paper has a number of limitations, which we will begin to discuss here. First, we rely on strong assumptions about how young people form expectations about the pecuniary returns to university, both in terms of what measure of realised earnings they base they use to form these expectations, and on what information they include in their information set when forming these expectations. It is unclear exactly how these assumptions are likely to have impacted our results.

For example, if young people are less myopic than we assume, and consider their lifetime earnings, then their expectations will also depend on the relative growth rates of graduate vs non-graduate wages. Then, if graduate wage growth is higher than non-graduate, our results would *underestimate* the expected pecuniary returns to university. However, we do not have full realised lifetime earnings for this cohort as they are still at the start of their careers, so would have to make additional assumptions on how they form expectations about this growth rate. Regarding the contents of young people's information sets, it is possible we do not observe all the information that young people use to form their expectations about the pecuniary and non-pecuniary returns. Recent work has developed methods to allow unobserved components in the expected pecuniary returns (D'Haultfoeuille and Maurel, 2013), and to determine the contents of young people's information sets (Cunha et al., 2004). We plan to test different models of expectations in future work, including allowing for unobserved heterogeneity in both pecuniary and non-pecuniary returns.

Finally, although we have started to decompose the non-pecuniary returns into meaningful components, the majority of these returns are still unattributed. Therefore, in future work it will be important to use (where available), and collect, more detailed data on young people's expectations and beliefs about the non-pecuniary aspects of university.

4.2 Results by socio-economic status (SES)

In this section we present the distributions of earnings and other factors premiums *conditional on socio-economic status*. We use parental earnings at age sixteen as a measure of socio-economic status (SES). Comparing the factor distributions across SES allows us to quantify the relative contributions of earnings and other factors to the SES-gap in university attendance (see table 2). The SES-gap in education in the UK is a name for the finding, documented by many researchers, that young people from less advantaged

backgrounds are much less likely to attend university. Even in our sample, which includes only higher ability students, the SES gap in university attendance is 15 percentage points (*pp*), with 60% of those in the low and middle SES groups (bottom and middle three quintiles of parental income) attending university, compared with 75% in the top SES group (top quintile).

Figure 3 shows the distributions of earnings (left column) and other (right column) factors, for those with parents in the bottom 20% (top row), middle 60% (middle row), and top 20% (bottom row) of the earnings distribution. Focusing first on earnings (left column, figure 3), the distributions of factors across the three groups are similarly located, though the means are slightly increasing in parental income (table 3g). For other factors (right column, figure 3), the distributions across the three groups clearly occupy different locations, and their means are strongly increasing in parental income. The mean other factors in the bottom and middle SES groups are slightly positive while they are strongly positive for the top SES group. The SES-gap in educational attainment is mostly driven by other factors in our analysis.

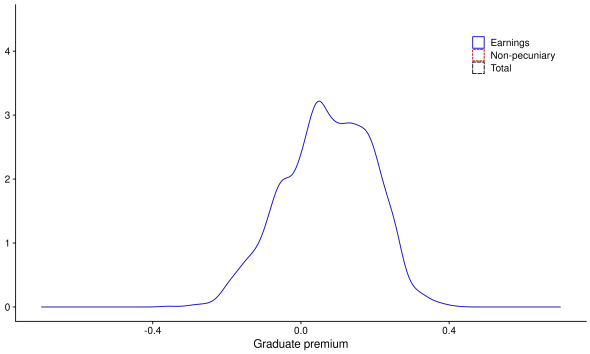
These findings are broadly inline with recent work by Boneva and Rauh (2020) who find that a large part of the gap between high and low SES students is due to differences in other factors premium. They also find a similarly sized role for wage premium, for which we find a much smaller role. There are a number of differences between our analyses that could explain this discrepancy. First, as we include only those individuals asked specifically about their beliefs in our sample, we are forced to focus on higher ability students. Second, Boneva and Rauh (2020) directly elicit expectations about earnings, while we estimate these expectations from realised earnings. Third, our definitions of SES are quite different, as Boneva and Rauh (2020) do not have detailed information on the parents of their sample members, and so define SES based on parental education, while we use parental income. Nonetheless, our findings add to the growing evidence that differences in (beliefs about) the non-pecuniary aspects of university across different groups are key drivers of differences in educational attainment across these groups.

4.3 Changes over two decades

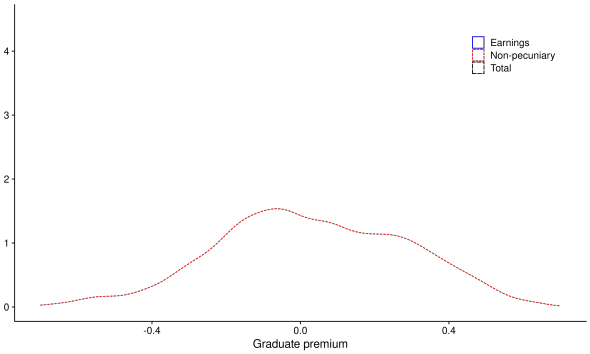
In an effort to shed light on what drove more and more people to attend university in England in recent decades, despite apparent stagnation in the wage returns, we re-estimate our model on data from a cohort born in 1970. These trends are displayed in figure 4, reproduced from Blundell, Green, and Jin (2022). In figure 4a the trend in proportion of those aged 30 with a first degree (BA) is plotted for cohorts born between 1950 and 1985, for the UK (blue line) and US. The level of educational attainment grew much faster for the UK over this period. In figure 4b the ratio of median BA to high school wages is plotted for the same period, which is remarkably flat. Taken together, these facts suggest

Figure 3: Comparing factor distributions by parental income (SES)

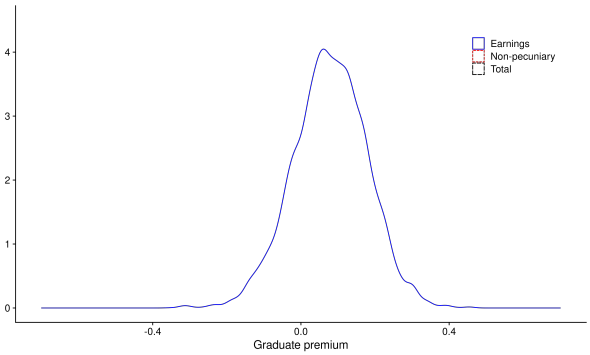
(a) Earnings, bottom 20%



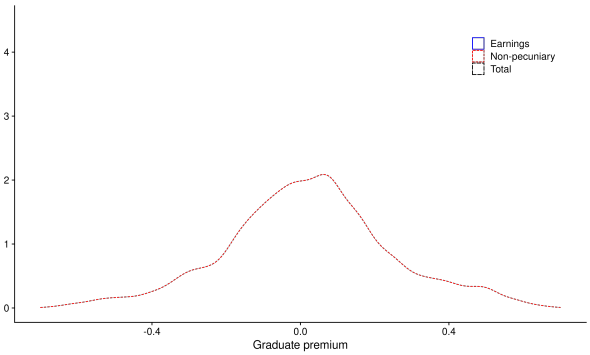
(b) Other, bottom 20%



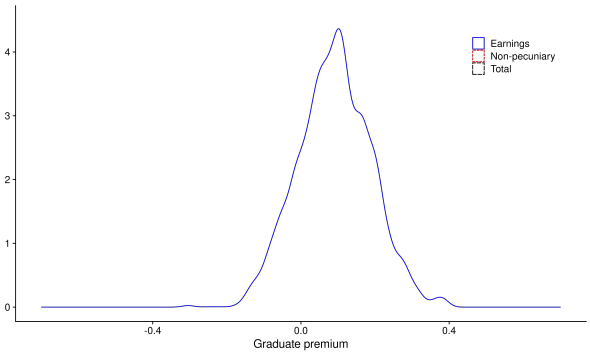
(c) Earnings, middle 60%



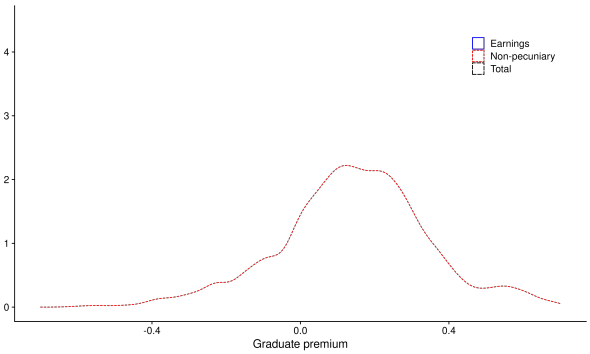
(d) Other, middle 60%



(e) Earnings, top 20%



(f) Other, top 20%



(g) Summary statistics for the distributions in panels (a)–(f)

	Mean	$q_{.1}$	$q_{.25}$	$q_{.5}$	$q_{.75}$	$q_{.9}$
Earnings						
Bottom 20%	0.07	-0.09	-0.01	0.07	0.16	0.22
Middle 60%	0.08	-0.05	0.01	0.08	0.15	0.20
Top 20%	0.09	-0.04	0.03	0.10	0.16	0.21
Other						
Bottom 20%	0.03	-0.29	-0.15	0.02	0.23	0.37
Middle 60%	0.02	-0.26	-0.11	0.02	0.15	0.32
Top 20%	0.15	-0.11	0.03	0.15	0.27	0.39

Notes: The values of the factors are estimated as described in 3.2. The distributions are then estimated (and plotted) with the kernel density estimator in the `R` package `ggplot2`, using the default Gaussian kernel and bandwidth (Wickham, 2016). The mean and standard deviations in panel (g) are in %Δ wage equivalent.

Table 8: Comparison of premiums between cohorts

Cohort:	BCS (1970)	LSYPE (1990)	Change 1970–1990
Degree	29.9%	62.6%	32.7*pp*
Earnings			
10th percentile	0.082	-0.053	-0.140
25th percentile	0.124	0.014	-0.115
50th percentile	0.175	0.083	-0.096
75th percentile	0.231	0.152	-0.080
90th percentile	0.281	0.210	-0.073
Other			
10th percentile	-0.639	-0.233	0.483
25th percentile	-0.500	-0.087	0.492
50th percentile	-0.334	0.065	0.481
75th percentile	-0.150	0.212	0.451
90th percentile	0.003	0.357	0.440

Notes: The percentiles are expressed in units equivalent to the difference in log wages.

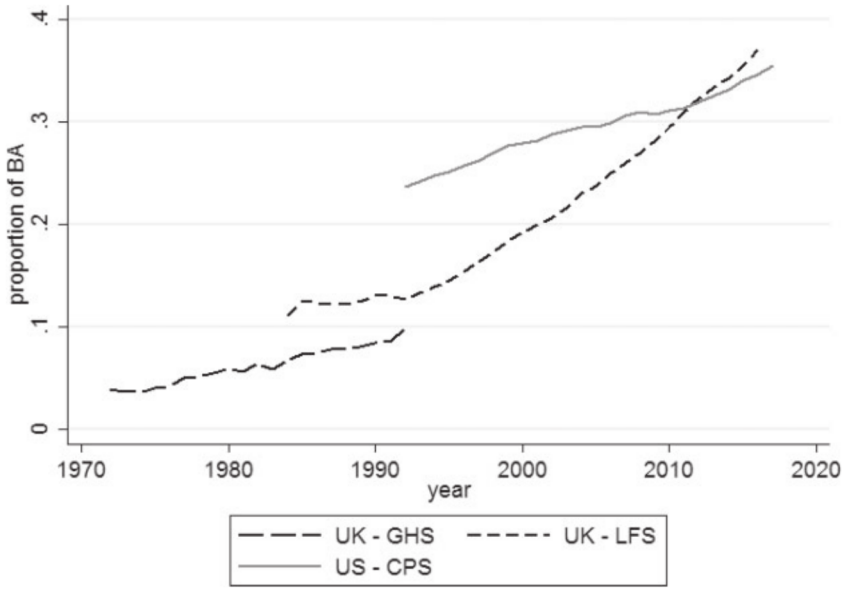
that there must have been a large increase in (expected) non-pecuniary returns to explain the growth in higher education over this period. We test this hypothesis, comparing cohorts born in 1970 and 1990.

The data for the earlier cohort is from the British Cohort Study 1970 (BCS 1970), a similar study to *Next Steps* which follows all 16,000 people born in the UK in a single week in April 1970. The aims of the BCS 1970 are very similar to *Next Steps*, and therefore we have very similar information on the cohort members. In particular, they were interviewed at age 16, when they were asked a series of questions about their expectations for the future, and we also have information on their family background at this point. They were then interviewed again 10 years later at age 26, with their earnings and their qualifications the key information we use from that wave. Having very similar data from two cohorts born 20 years apart allows us to directly compare the distributions of the earnings and other factors premiums we estimate for these two cohorts. We estimate the model on the earlier cohort following the procedure described in section 3.2.

Figure 5 presents the estimated distributions of earnings and other factors premiums for the two cohorts. The mean graduate-wage premium *decreased* on average between those born in 1970 (solid blue line) and 1990 (dashed blue line). Meanwhile, the other factors premium *increased* significantly on average over this period. Key percentiles of these distributions and their differences are in table 8, and we can see that the median earnings premium fell by 9.6 log-points, while the median non-pecuniary factors premium increased by 48.1 log-points. Recall, the units of these premiums are equivalent to a difference in log-wages. In the 1970 cohort, the median cohort member believed their wages would

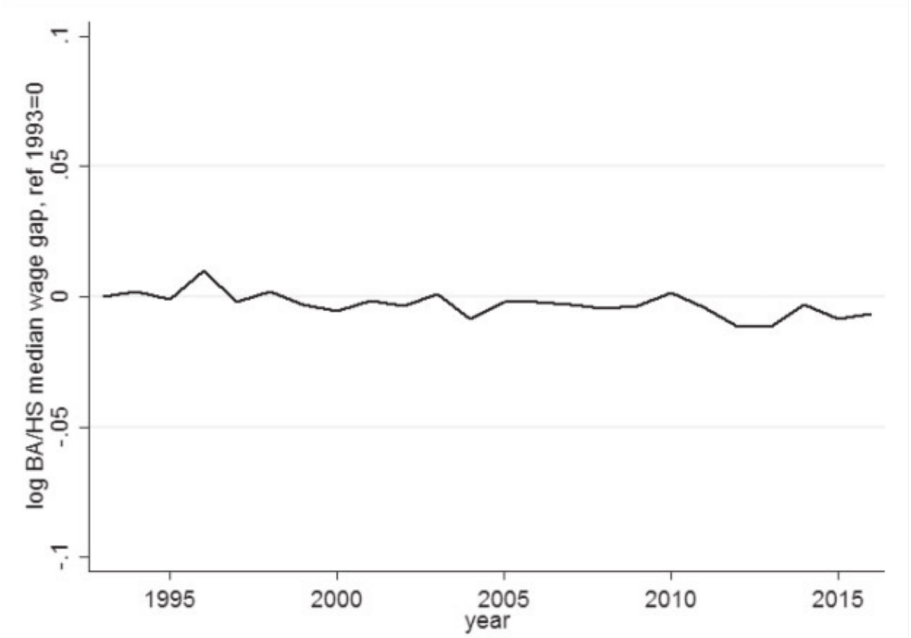
Figure 4: Higher education and wages in the UK vs the US in recent decades

(a) Proportion of people with a BA or higher education by cohort, UK and US



Source: Reproduced from Blundell et al. (2022). Those authors' calculation from the U.K. Labour Force Survey, the U.K. General Household Survey, and the U.S. Current Population Survey.
Notes: BA refers to individuals who have a bachelors or higher degree. Blundell et al. (2022) aggregate each dataset to the level of year and 5-year age band, and regress the BA proportion on year dummies and age-band dummies. The proportion BA numbers are year effects from these regressions plus the level in 1992 for the 30–34 age band.

(b) Ratio of BA median wage to that of high-school graduates 1993–2016, U.K.

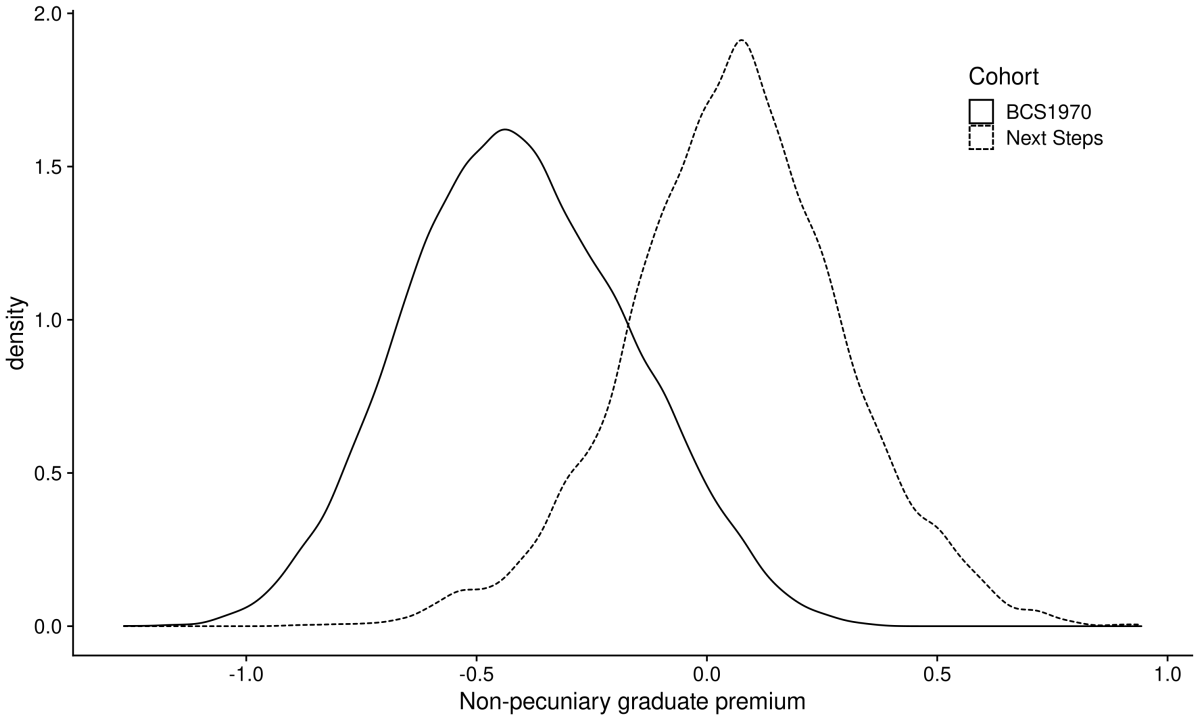


Source: Reproduced from Blundell et al. (2022)
Notes: Wage is hourly. The sample is 20–59 year olds in LFS 1993–2016. BA refers to individuals who have a bachelors or higher degree. Blundell et al. (2022) aggregate LFS to the level of year and 5-year age groups, and regress the log BA to HS median wage ratio on year dummies and age-band dummies. The figure plots the estimated year effects normalized to zero in 1993.

Figure 5: Changes in distributions of factors between cohorts (1970–1990)

(a) Earnings

(b) Other



Source: The data are from the BCS1970 and Next Steps surveys.

Notes: The values of the factors are estimated as described in 3.2. Only young people with at least 5 GCSEs at A*–C or equivalent are included in the sample. The distributions are then estimated (and plotted) with the kernel density estimator in the `R` package `ggplot2`, using the default Gaussian kernel and bandwidth (Wickham, 2016).

be over 16% higher if they attended university, while the median 1990 cohort member believed university would increase their wages by 9.2%. However, the median 1970 cohort member perceived significant non-earnings "costs" to attending university, equivalent to 28.8% of their earnings. These costs had become negative (i.e. benefits) for the later cohort, who perceived non-earnings benefits of attending university equivalent to 6.3% of their earnings.

Therefore, in our framework the large increase in higher education attainment between the two cohorts (see table 8) was entirely driven by an increase in expectations about non-earnings factors. This finding is inline with the evidence, presented at the beginning of this section, that the pecuniary returns to a degree have remained remarkably constant over a period of significant growth in higher education in the UK.

5 Conclusion

In this paper we specify and estimate a model of educational choice, that specifically includes expectations about earnings and other, financial and non-pecuniary, factors. We exploit data on a cohort born at the end of the 1980s which features data on realised earnings and expectations about the non-pecuniary costs and benefits of going to university. Our findings add support to the notion that individuals are not strict income maximisers when they make educational choices. We find that the non-pecuniary premium is able to explain most of the variation across individuals that causes some people to attend university and others to not, with the graduate-earnings premium playing only a minor role. Splitting the sample by parental income (a measure of socio-economic status), we find that differences in factors other than earnings across socio-economic groups are chiefly responsible for the "SES gap" in educational attainment. Finally, comparing the roles of pecuniary and other factors in educational decisions across a period of significant growth in higher education attainment and increased financial costs, we find that the expected graduate premium fell slightly, suggesting increases in the value of non-pecuniary factors drove the expansion in attainment.

Although further work is required to address the limitations of our analysis, our results suggest that a better understanding of the (expected) non-pecuniary costs and benefits of university is vital. Of particular importance is the apparent difference between the most-advantaged young people and their less-advantaged peers in terms of their expectations about the non-pecuniary returns to university. The socio-economic status gap in educational attainment is a barrier to social mobility and contributes to inequality: both in terms of income, and other outcomes known to be related to education such as health.

In future work, we plan to extend the analysis in this paper in a number of directions. First, by using different models of expectations we can test the sensitivity of our find-

ing to our assumptions about how young people form their expectations. Comparing these estimated expectations under different models with the gold-standard of elicited expectations will be an important part of this work. Second, although we have started to decompose the so-called "psychic costs" of university into meaningful components, the limitations of our data mean there is still lots of work to do in this area. This will require both careful analyis of existing data, including from more recent cohort studies in the UK, as well as original data collection.

References

Anders, J. (2012): "The Link between Household Income, University Applications and University Attendance*," *Fiscal Studies*, 33, 185–210.

Anders, J. and J. Micklewright (2015): "Teenagers' Expectations of Applying to University: How do they Change?" *Education Sciences*, 5, 281–305, number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

Arcidiacono, P., V. J. Hotz, and S. Kang (2012): "Modeling college major choices using elicited measures of expectations and counterfactuals," *Journal of Econometrics*, 166, 3–16.

Arcidiacono, P., V. J. Hotz, A. Maurel, and T. Romano (2020): "Ex Ante Returns and Occupational Choice," *Journal of Political Economy*, 128, 4475–4522.

Blanden, J. and S. Machin (2004): "Educational Inequality and the Expansion of UK Higher Education," *Scottish Journal of Political Economy*, 51, 230–249.

Blundell, R., D. A. Green, and W. Jin (2021): "The UK as a Technological Follower: Higher Education Expansion, Technological Adoption, and the Labour Market," *Review of Economic Studies*, 81.

——— (2022): "The U.K. as a Technological Follower: Higher Education Expansion and the College Wage Premium," *The Review of Economic Studies*, 89, 142–180.

Boneva, T. and C. Rauh (2020): "Socio-Economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns," *HCEO Working Paper*, 83.

Britton, J., L. van der Erve, C. Belfield, L. Dearden, A. Vignoles, M. Dickson, Y. Zhu, I. Walker, L. Sibieta, and F. Buscha (2021): "How much does degree choice matter?" Tech. rep., The IFS.

Coppejans, M. (2001): "Estimation of the binary response model using a mixture of distributions estimator (MOD)," *Journal of Econometrics*, 102, 231–269.

Crawford, C. and W. M. Jin (2014): "Payback time? Student debt and loan repayments: what will the 2012 reforms mean for graduates?" Tech. rep., Institute for Fiscal Studies.

Cunha, F. and J. J. Heckman (2007): "Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling," *Labour Economics*, 14, 870–893.

Cunha, F., J. J. Heckman, and S. Navarro (2004): "Separating uncertainty from heterogeneity in life cycle earnings," *Oxford Economic Papers*, 57, 191–261.

Dearden, L., E. Fitzsimons, and A. Goodman (2005): "Higher education funding policy: a guide to the debate," .

Delavande, A. and B. Zafar (2019): "University Choice: The Role of Expected Earnings, Nonpecuniary Outcomes, and Financial Constraints," *Journal of Political Economy*, 127, 2343–2393, publisher: University of Chicago.

D'Haultfoeuille, X. and A. Maurel (2013): "Inference on an extended Roy model, with an application to schooling decisions in France," *Journal of Econometrics*, 174, 95–106.

Dominitz, J. and C. F. Manski (1996): "Eliciting Student Expectations of the Returns to Schooling," *The Journal of Human Resources*, 31, 1–26.

Gong, Y., T. Stinebrickner, and R. Stinebrickner (2019): "Uncertainty about future income: Initial beliefs and resolution during college," *Quantitative Economics*, 10, 607–641, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/QE954.

Heckman, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2018): "Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking," *Journal of Political Economy*, 126, S197–S246, publisher: The University of Chicago Press.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006): "Chapter 7 Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," in *Handbook of the Economics of Education*, Elsevier, vol. 1, 307–458.

IOE Centre for Longitudinal Studies (2008): "Next Steps: Wave 4 Documentation," .

——— (2018): "Next Steps: Sweeps 1–8. 2004–2016. [data collection]." .

Manski, C. F. (1993): "Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?" *NBER*, 43–60.

Mcfadden, D. (2001): "Economic Choices," *THE AMERICAN ECONOMIC REVIEW*, 91, 53.

OECD (2018): *Equity in Education: Breaking Down Barriers to Social Mobility*, PISA, OECD.

Oreopoulos, P. and U. Petronijevic (2013): "Making College Worth It: A Review of the Returns to Higher Education," *The Future of Children*, 23, 41–65.

Oreopoulos, P. and K. G. Salvanes (2011): "Priceless: The Nonpecuniary Benefits of Schooling," *Journal of Economic Perspectives*, 25, 159–184.

Roy, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–146.

Walker, I. and Y. Zhu (2008): "The College Wage Premium and the Expansion of Higher Education in the UK," *The Scandinavian Journal of Economics*, 110, 695–709.

Wickham, H. (2016): *ggplot2: Elegant Graphics for Data Analysis*, Springer-Verlag New York.

Willis, R. J. and S. Rosen (1979): "Education and Self-Selection," *Journal of Political Economy*, 87, S7–S36.

Wiswall, M. and B. Zafar (2015): "Determinants of College Major Choice: Identification using an Information Experiment," *The Review of Economic Studies*, 82, 791–824.

Woodin, T., G. McCulloch, and S. Cowan (2013): *Secondary Education and the Raising of the School-Leaving Age: Coming of Age?*, Secondary Education in a Changing World, Palgrave Macmillan US.

A Institutional context of HE in England

In this section we discuss the organisation of higher education in England. Schooling is compulsory up to the age of sixteen in the UK, and has been since 1972 (Woodin, McCulloch, and Cowan, 2013). Figure A1 presents the time-line of decisions and exams that students (generally) must take to secure a place at university. Two key decisions are: the application to continue on to further education ("sixth form") in the final year of secondary school; and the university application in the final year of sixth form. The main data source follows individuals through secondary school and beyond, from the age of 14 until 19. However, in this paper we will focus exclusively on the decision to attend university and treat the outcome of the decision to continue to sixth form as a predetermined characteristic. Estimating a dynamic discrete-choice model to exploit more of the data is an interesting avenue we hope to explore in future work.

University application process. The UK university application system is quite unique in many ways, and is worthy of study in its own right. Students apply through a centralised system, the "Universities and Colleges Admissions Service" (UCAS)[12] in the autumn of their final year of sixth form. Students can apply for up to five places, where each "place" is a *university-subject pair*. The application consists of a personal statement written by the student, predicted A-levels grades from their teachers, and past national-exam results. These are common across all applications, so students cannot tailor their personal statement to different subjects or institutions.[13] Students then receive *conditional* offers or are rejected from each place they applied, and must select two of their offers: a first choice and a back-up option. The offers made to students in sixth form are (almost exclusively) conditional on their future grades, so for example may require a student sitting 3 A-levels to achieve AAB, with one A in chemistry. The back-up option allows the student to aim high with their first choice, and still have a place somewhere if they fail to achieve those grades. Students sit their A-levels knowing their required grades for each place, and are automatically accepted at their first choice if they achieve the required grade, at their second if they miss the requirement for their first choice, and nowhere if they do not meet either requirement.[14]

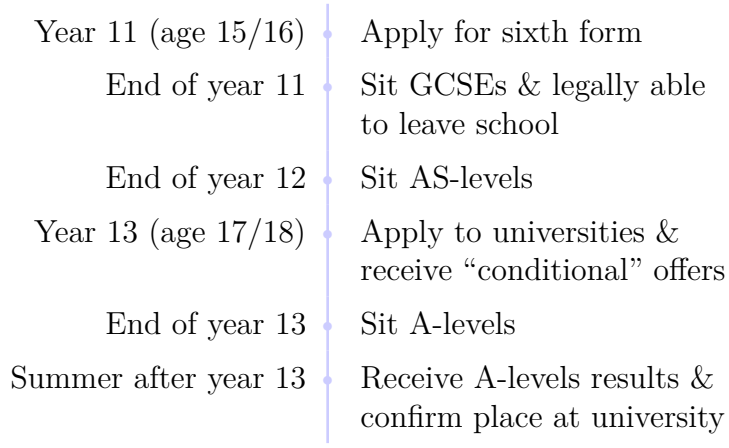
The funding of higher education. Universities in the UK are privately run, but receive state funding *and* are regulated by government over the fees they can charge their students. Tuition fees were first introduced for UK students at UK universities in 1998. Prior to this, universities could not charge fees for tuition. There was also a

[12] Universities and colleges are different entities in the UK, and the names are not used interchangeably, unlike in the US.

[13] This is an implicit barrier which stops people applying to vastly different subjects.

[14] There is a mechanism to allocate students who missed their offers on both their first- and second-choices to places at universities who remain unfilled called "Clearing".

Figure A1: Timeline of educational decisions (1990 cohort)



system of grants and loans in place to cover living costs. In 1998 a means-tested fee was introduced, with the students from the most privileged backgrounds paying £1,000 per year in tuition fees. The poorest students were entitled to a 25% reduction. The situation changed again in 2006, with the introduction of so-called "top-up" fees, which could be set by each university up to a maximum of £3,000.[15] Alongside these fees, the government introduced a comprehensive system of loans and grants to cover both tuition fees and living costs ("maintenance"). Grants and some loans were means-tested, but all students could borrow the full fee, plus some extra for maintenance. The repayment schedule of the loans was made income contingent, meaning that no repayments were required until a graduate earned over a threshold amount, and repayments were set at a percentage of all earnings over this threshold. Therefore, not only does attending university affect the earnings that someone might expect to receive, but their (expected) future earnings will affect how much they expect to pay for their degree, a key feature to capture in the model.

Tuition fees, student loans and maintenance grants The funding of higher education in the UK has changed frequently in recent years moving from a model of direct government funding prior to 1998, to a model with increasingly higher tuition fees alongside a system of government-subsidised loans and grants (see Table 2.1 in Crawford and Jin (2014) for a summary of some of these changes). The majority of the individuals in the main cohort we use left sixth-form in 2007, so they would have experienced the system under reforms that came into force in 2006, henceforth the "2006 reforms". The key features of the system under the 2006 reforms are summarised in table A1.

Student debt levels on graduation. Dearden, Fitzsimons, and Goodman (2005) calculate expected debt levels for a student entering university in 2006/7 (i.e. under the

[15]This maximum fee is set currently at slightly over £9,000, though the increase occurred after the relevant period for the analysis in this paper (in 2012).

Table A1: Details of fees, loans and grants available under the 2006 reforms

Measures	Details
Tuition fees	• Set by university, up to £3,000 p.a. • payable by ALL students
Grants	• Means-tested up to £2,700 p.a. • Tapered to zero at £33,560.
Loans	
Fees	• Equal to fees charged by university. • Available to ALL students.
Maintenance	• £3,555 p.a. if household income <£26,000. • Loan increases from £3,555 p.a. incrementally • Up to £4,405 p.a. if family income between £26,000 and £33,560. • Tapered down to £3,305 at £44,000.
Repayment	• 9% of income above £15,950 (threshold rises with inflation). • State-subsidised loans, zero-real interest rate. • Debt forgiven after 25 years.

Source: Crawford and Jin (2014)

Table A2: Expected debt on graduation (maximum loans under 2006 reforms)

Parental income	Debt on graduation	Share in sample
Low (<£15,970 p.a.)	£19,340	0.20
Middle (∼£25k p.a.)	£19,340	0.09
Upper middle (∼£30k p.a.)	£21,440	0.22
High (>£44k p.a.)	£18,670	0.31
Missing income info.	-	0.18

Source: Dearden et al. (2005) (debt figures) and author's calculations.

first year of the 2006 reforms). Their calculated expected debts are in table A2, along with the share of individuals in each category in the Next Steps cohort. The information in tables A1 and A2 show that although the sticker price of education in the UK was quite high, loans were available to all suggesting credit constraints are not an issue in the UK context. In addition the (maximum) debt burden faced by students appears to be relatively constant across socio-economic groups (though of course the psychological effects of this debt may still vary).

B Data appendix

B.1 Next Steps

Other information collected in wave four

In addition to the data on expectations collected in wave 4, I also use information on family background and schooling up to age sixteen. I use detailed information on parental earnings to estimate a measure of socio-economic status (SES), based on the quintiles of parental earnings (I also use an alternative definition based on means-tested grant eligibility, again calculated from parental earnings). I include information on parents' occupations, ethnicity, education, and income in the model, as well as (limited) information on ability[16] (number of A-levels being taken), and gender. Table 2 presents descriptive statistics for these variables.

B.1.1 Wave eight (age twenty-five)

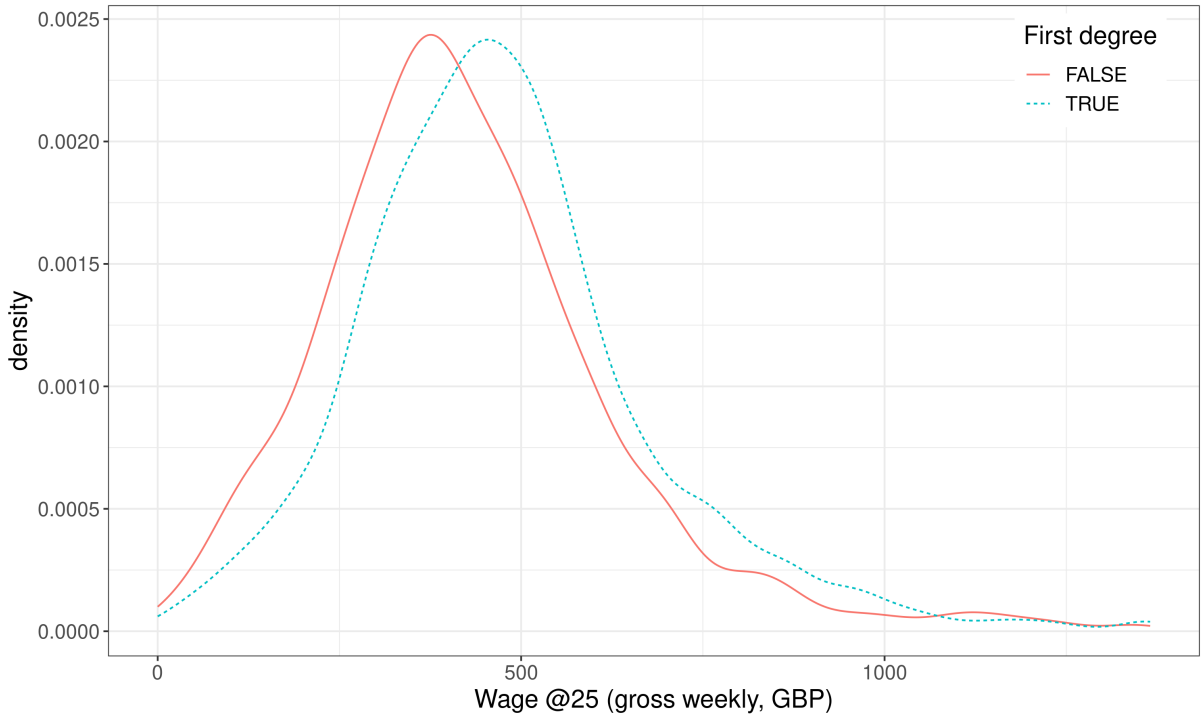
The other key wave of Next Steps for my analysis is the eighth, when the cohort members are aged 25. At this point the majority are working (or at least have worked at some point), and most of those who attend university have completed their degrees.

Degree attainment. The cohort members are asked about any qualifications they have achieved since the last interview (wave seven, five years previously), including whether they hold an undergraduate degree. Table 2 shows information on the proportion of cohort members who hold a degree at 25, including the proportion who attended a member of the Russell Group (a "club" of prestigious research universities in the UK). I also break down degree attainment by SES group (parental income quintiles) in table 2. All these statistics are shown for all respondents to waves 4 and 8, and for the subsample who answered questions about university. Nearly 70% of the analysis subsample hold a degree by the time they are 25, though there is still substantial variation across socio-economic groups. The rate of BAs at 25 among those from the most advantaged backgrounds is 75%, compared with 60% for those from the least advantaged. That the socio-economic attainment gap persists among these "high-achieving" students suggests the issue runs deeper than performance at school.

Wages. As the majority of the cohort members are in work at age 25, a focus of wave 8 is on their careers, occupations and other labour market outcomes. In particular they are asked to provide detailed information about their wages. Figure B1 shows the distribution of weekly wages in the sample, conditional on degree attainment. The conditional

[16]The survey is linked to an administrative education dataset, the National Pupil Database (NPD), so there is much more detailed information on the students' (academic) abilities potentially available. I am currently waiting for access to this linked dataset.

Figure B1: Distribution of weekly wages at age 25, by degree attainment



Notes: The distributions are estimated (and plotted) using the `density` option in the `R` package `ggplot2` (Wickham, 2016), using the default setting of a Gaussian kernel density estimator. Analysis subsample ($N = 4,640$).

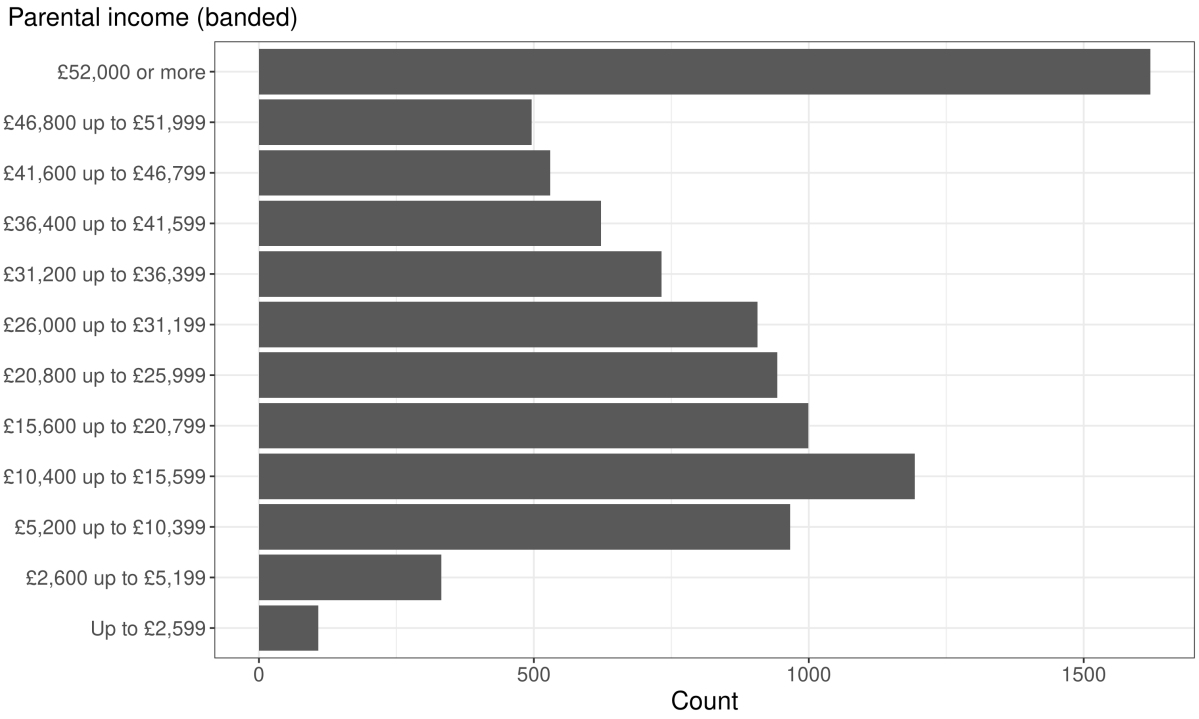
distributions look very similar, with the distribution corresponding to holders of an undergraduate (first) degree shifted slightly to the right. The mean and variance of these distributions are in table 2. However, such analysis does not reveal expected, nor counterfactual, wages: i.e. what graduates (expect they) would earn had they not gone to university, and vice versa. For that we need the model and assumptions detailed in section 3.

B.2 Additional information on Next Steps

Next Steps started in 2004 when the members were in secondary school aged 13 or 14. They were then interviewed annually for the next six years, until aged 18 or 19 (waves 1–7). A further round of interviews (wave 8) was conducted in 2016 when the members were aged 25 or 26, and another is planned for 2021. For consistency with the BCS data, we will focus on the data collected at age 16 (or thereabouts, wave 4) and at 25 (wave 8).

Parental income. Next Steps records information on member's family background in waves 1–7. Though data on parental income was collected in wave 4, it was recorded in 12 bins, with the top (and most populous) bin starting at £52,000 p.a. (see figure B2). More detail was collected in wave 1—over 30 bins, plus further information for some top-coded families—as well as continuous data on parents' salaries for some families (see figures B3 and B4).

Figure B2: Parental income in Next Steps (wave 4)

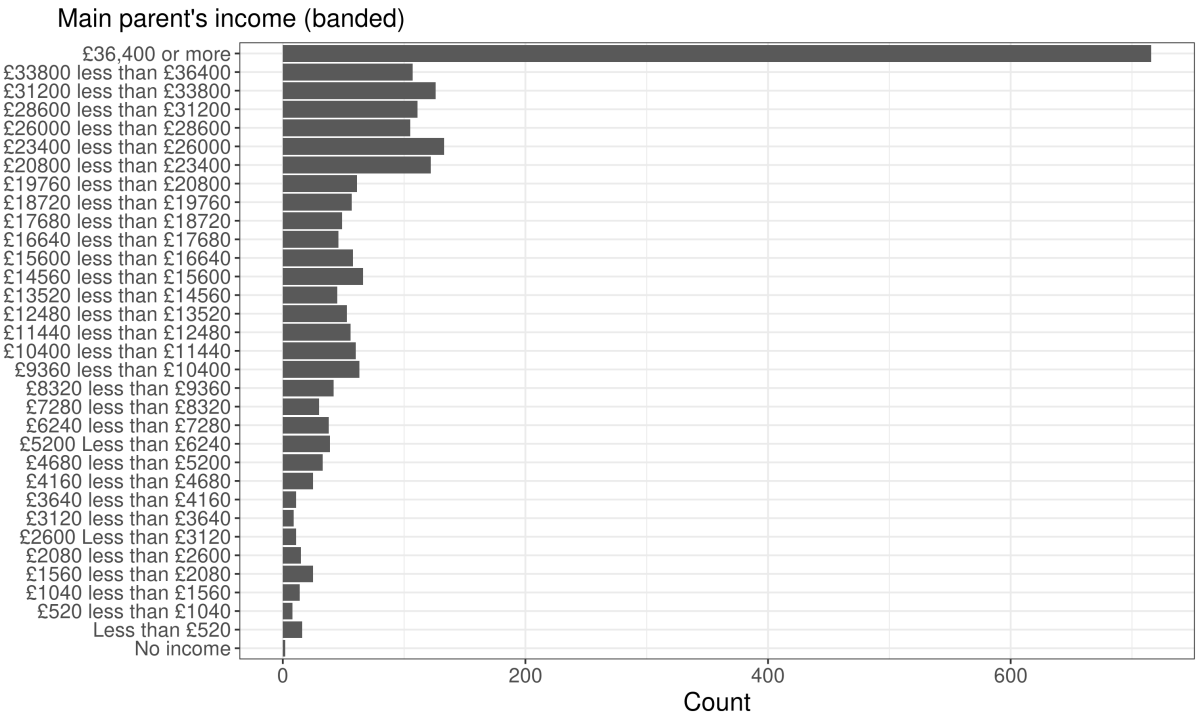


Source: LSYPE wave 4 (CLS, 2018).

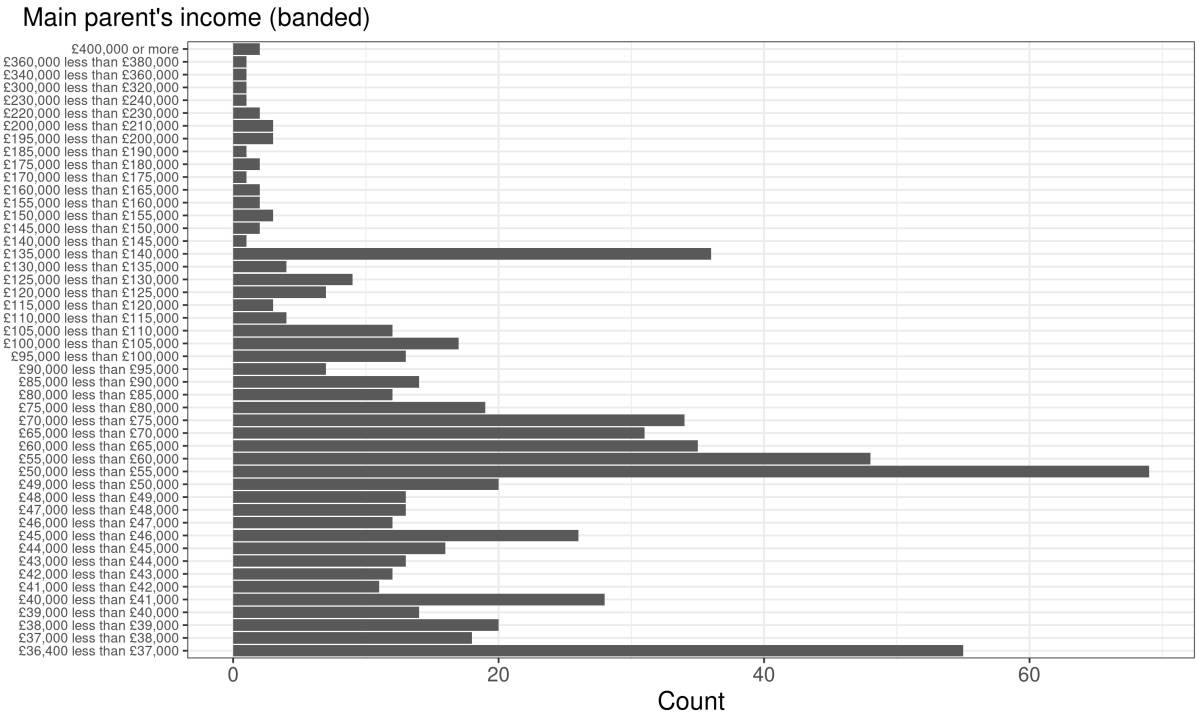
Undergraduate degree. Figure B5 shows the proportion of individuals in Next Steps who hold a degree at 25, broken down by gender.

Figure B3: Main parent's income in Next Steps (wave 1)

(a) Banded



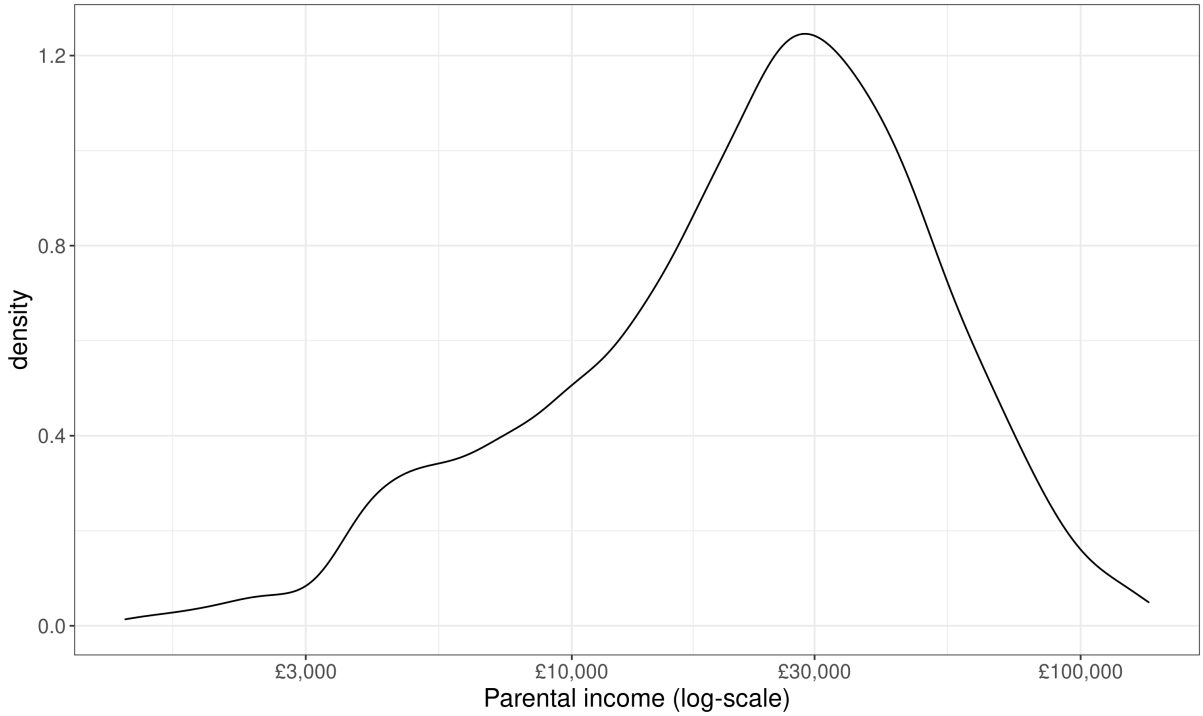
(b) Top-code ($> £36,400$) detail



Source: LSYPE wave 1 (CLS, 2018).

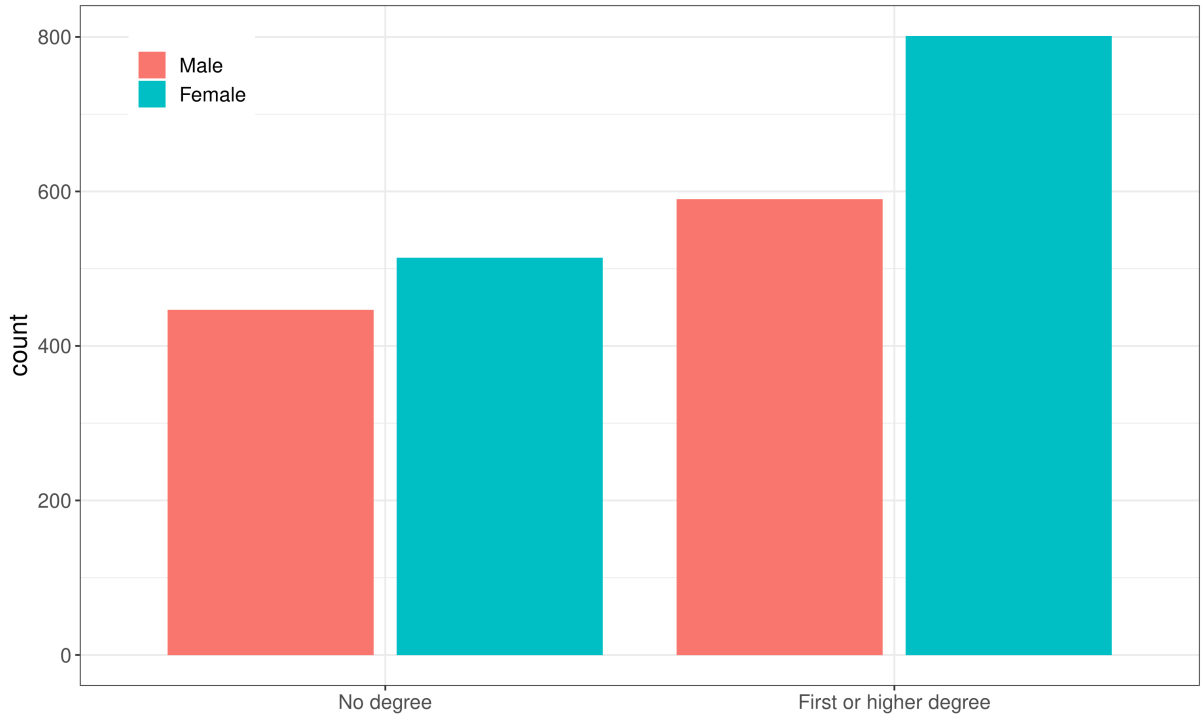
Notes: The top panel (a) shows all recorded earnings for main parents in wave 1. Panel (b) shows a detailed breakdown of the top band from panel (a).

Figure B4: Parental income in Next Steps (wave 1, density)



Source: LSYPE wave 1 (CLS, 2018).
Notes: This plot shows the density of (log-)annual earnings, calculated using the default kernel density estimator of the `geom_density()` function in the `ggplot2 R` package (Wickham, 2016).

Figure B5: Undergraduate degree at 25 by gender (LSYPE)



Source: LSYPE wave 8 (CLS, 2018).

Table B1: The advantages (+) and disadvantages (−) of going to university

Response (harmonised)	+/−
Career	
Will lead to a good/better job (than would otherwise get)	+
Will lead to a well paid job	+
Gives someone better opportunities in life	+
Is essential for the career they want to go into	+
Shows that you have certain skills	+
To delay entering work/ more time to decide on a career	+
Not being able to start earning money/start work	−
No guarantee of a good job at the end	−
Don't need to go to university for the job someone may want	−
Get less work experience	−
Financial / debt	
Now	
It is expensive	−
Not becoming financially independent	−
Not being able to start earning money/start work	−
Costs (general/non specific)	−
Tuition fees/Accommodation costs/Living expenses	−
Future	
Will lead to a well paid job	+
Getting into debt/have to borrow money	−
Social life / environment	
The social life/ lifestyle / meeting new people / it's fun	+
To leave home/ get away from the area	+
Leaving home/family/friends	−
Stress	−
Education	
To carry on learning / I am good at / interested in my chosen subject	+
Get more qualifications/better/higher qualifications	+
The workload can be hard/ doubts about ability to finish course	−
Personal development	
Makes someone independent/ maturity / personal development	+
Gives you more confidence	+
People will respect me more	+
Leads to a better life/good life (general)	+
Prepare you for life/gain life skills	+
Time	
To delay entering work/ more time to decide on a career	+
Takes a long time	−
Waste of time (general/non-specific)	−

Table B2: Logit estimates (all responses, first-stage)

(a) Advantages

Dependent variable:	Degree
Get better job	0.422***
	(0.099)
Well-paid job	0.266***
	(0.103)
Better opportunities	0.454***
	(0.100)
Need for career	−0.028
	(0.226)
Show skills	0.194
	(0.299)
Delay get job	0.670
	(0.569)
Social life	0.138
	(0.116)
Leave home	0.052
	(0.287)
Learning	0.297**
	(0.149)
More qualifications	0.055
	(0.096)
Personal development	0.621***
	(0.158)
More confidence	1.109**
	(0.513)
More respect	−0.135
	(0.630)
Better life (general)	0.134
	(0.272)
Gain life skills	0.298*
	(0.154)
Other	0.208
	(0.232)
Don't know	−0.012
	(0.296)
No answer	−0.719*
	(0.373)
Observations	3,469
Log Likelihood	−1,985.799
Akaike Inf. Crit.	4,105.599

(b) Disadvantages

Dependent variable:	Degree
Expensive	0.108
	(0.152)
Get into debt	0.141
	(0.112)
Depend on parents	12.457
	(228.304)
Not financially indep.	−0.013
	(0.385)
Not earning / working	−0.092
	(0.169)
Costs (general)	0.200*
	(0.121)
No job guarantee	0.065
	(0.162)
Not needed for job	−1.163**
	(0.544)
Less experience	−0.223
	(0.278)
Heavy workload	0.256
	(0.175)
Leave home	−0.286*
	(0.151)
Takes long time	−0.273*
	(0.155)
Waste of time	−0.272
	(0.369)
Tuition fees etc.	0.400*
	(0.229)
Stress	0.523
	(0.372)
Other	−0.197
	(0.159)
Don't know	−0.134
	(0.244)
No answer	−0.100
	(0.190)
Observations	3,469
Log Likelihood	−1,985.799
Akaike Inf. Crit.	4,105.599

Notes: *p<0.1; **p<0.05; ***p<0.01. The two panels contain estimates from the same regression, which also included the following background characteristics: ethnicity, gender, A-levels, parental income, and a self-assessed ability measure.

C Results

Figure C1: Comparing wage premium distributions with and without selection correction