Journal of Econometrics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Academic and non-academic investments at university: The role of expectations, preferences and constraints[☆]

Adeline Delavande [a,*], Emilia Del Bono [b], Angus Holford [b]

[a] University of Technology Sydney, Australia and Nova School of Business and Economics, Portugal
[b] University of Essex, UK

ARTICLE INFO

Article history:
Received 29 March 2019
Received in revised form 31 January 2020
Accepted 16 March 2020
Available online xxxx

JEL classification:
I23
I24
D84
C81

Keywords:
Subjective expectations
Choice set heterogeneity
Higher education
Ethnic gap

ABSTRACT

This paper estimates a discrete choice model of time allocation decisions made by university students. We consider investments in academic and non-academic activities, such as job placements or volunteering. Identification is achieved using data collected through a recent survey of UK university students on subjective expectations about the returns to these activities, and the enjoyment students derive from them. Unobserved heterogeneity in the choice set is addressed using a sufficient set logit method. The analysis reveals significant ethnic differences in the level of investments, expected academic and labour market returns, and enjoyment of academic and non-academic activities. Simulations suggest that existing constraints play an important role in explaining ethnic gaps in investments.

1. Introduction

Increasing social mobility is a priority for many governments and widening participation into higher education is one possible pathway to do so. However, conditional on going to university, students from disadvantaged family backgrounds and ethnic minorities tend to have worse outcomes than their counterparts: they are more likely to drop-out, they graduate with a lower final mark, they take longer to find a job, have lower earnings once in work, and accumulate less wealth (Bailey and Dynarski, 2011; Meschede et al., 2017; Smith and Naylor, 2001a,b; Zwysen and Longhi, 2018). Controlling for past educational outcomes reduces but typically does not eliminate these inequalities (Arcidiacono and Koedel, 2014; Crawford, 2014). While other dimensions of 'college readiness' are also likely to play a role (Bound et al., 2010), this suggests that differences in outcomes may be driven in part by different choices made while at university.

Students' effort at university is one of the primary inputs into the production function of higher education achievement. Existing evidence suggests that academic investments such as study hours and attendance positively affect marks (Romer, 1993; Schmidt, 1983; Stinebrickner and Stinebrickner, 2004, 2008a). Non-academic (also called extra-curricular)

☆ Acknowledgements: We thank participants at the University of Tilburg, Rotterdam University, University of Bristol, Interdisciplinary Center (IDC) Herzliya, 2018 Workshop on Subjective Expectations, 2018 Misoc Workshop on Higher Education, 2018 Royal Holloway workshop on Education, 2019 Sydney Empirical Micro workshop. We acknowledge funding from the Economic and Social Research Council grants ES/M008622/1, ES/S012486/1 and ES/L009153/1.

* Correspondence to: Department of Economics, UTS Business School, City Campus, Haymarket, PO Box 123, Broadway NSW 2007, Australia.
E-mail address: Adeline.Delavande@uts.edu.au (A. Delavande).

https://doi.org/10.1016/j.jeconom.2020.03.019
0304-4076/© 2020 Published by Elsevier B.V.

investments, such as engaging in competitive sports or undertaking a work placement, are likely to be an important signal to employers about non-cognitive and social skills, with some (thin) evidence that they may impact positively on later labour market outcomes (Lechner and Downward, 2017; Persico et al., 2004; Saniter and Siedler, 2014).

Yet, we still know relatively little about how university students spend their time, and even less on what determines their allocation choices, i.e. whether these choices are mainly driven by expectations about future outcomes, preferences or constraints.[1] The primary reason for this gap in knowledge is lack of data. In this paper, we provide evidence on academic and non-academic investments made by university students and how they vary by social and ethnic background, and use unique data on subjective expectations and preferences to understand the determinants of students' choices.

Our data is sampled from an entire cohort of undergraduates studying at a UK university. We focus on two groups that have received most attention in the recent UK policy debate due to their lower higher education participation and/or academic achievement: students from low socio-economic backgrounds, and Black, Asian and Minority Ethnic (BAME) students (regardless of socio-economic status).[2] Our analysis compares the choices of low socio-economic status (SES) white British students and non-white British students with those of high-SES white British students. A significant percentage of students in our sample come from British ethnic minority (about 30%) and disadvantaged family backgrounds (about 15%) and this gives us the opportunity to investigate social and ethnic differences in time allocation.

We observe significant differences in outcomes and investments in our data, and these differences are mainly noticeable along the ethnic dimension. Among those enrolled in our study in their third year, 26.8% of white British high-SES students were on track for a getting a GPA above 70%, compared to 19.5% of the white British low-SES and only 13.5% of non-white British. Students devote on average 20 hours per week to academic investments (attendance and study) in the second and the third year of their courses. However, non-white British and white British students from high-SES choose to allocate their time differently: non-white British attend lectures and classes significantly less but compensate by spending more time in private study. Low and high-SES white British tend to have similar time allocation by contrast. Students spend also an average of 11 hours per week in non-academic activities (e.g. paid work, paid and unpaid internships, volunteering, having a leadership position in a university society, and competitive sport). Here we see again that although non-white British engage in these activities for a similar amount of hours than white high-SES British, the activities they undertake are different, resulting in the former group being less likely to have accumulated experience that is relevant to their field of study and desired career.

Our objective is to investigate the relative role of preferences, subjective expectations and constraints in the decision to allocate time across different types of activities. Making inference on the decision-making process based on choice data alone is challenging since observed choices might be consistent with several combinations of expectations and preferences (Manski, 1993, 2004). We address this identification issue by using new data on (*i*) subjective expectations about the academic returns to time spent in academic and non-academic activities; (*ii*) subjective expectations about the pecuniary and non-pecuniary labour market returns to academic achievement and non-academic investments; and (*iii*) expected or actual enjoyment of the various activities.

We elicit subjective expectations about the *academic returns* to time spent in various activities by asking students to report their subjective probability of achieving a certain academic outcome (we use a categorical version of the final GPA) conditional on 12 different scenarios where we vary time spent in lectures and classes, private study, and non-academic activities. We then elicit subjective expectations about the *labour market returns* to academic achievement and non-academic investments. Here students are presented with 12 different scenarios in which we vary academic outcomes, work experience relevant to their field of study or desired career, and other extra-curricular experience (e.g. participation in sport or leadership of student clubs).

The data reveal interesting features about students' beliefs regarding their education production function. An hour of attendance is thought to increase the probability of having a "good degree" (i.e. a GPA above 60%) by 3ppt, which is about double the effect of an hour of private study, and almost 7 times larger than the effect of an hour spent on non-academic activities. In relation to labour market outcomes, students perceive high returns both from academic achievement and relevant work experience. There is substantial heterogeneity in beliefs: compared to white high-SES students non-white students tend to have a lower subjective academic return to attendance and study hours, and indicate a lower subjective labour market return to investing in extra-curricular activities. We also gather information about enjoyment of various activities. Here we learn that non-white British enjoy lectures and employment significantly less than white high-SES British. By contrast, there are no significant differences in enjoyment of various activities by SES among the white British students.

[1] With some exceptions (Delaney et al., 2013; Dolton et al., 2003; Grave, 2011; Stinebrickner and Stinebrickner, 2008a), the main evidence about the importance of academic investments – attendance and study hours – comes from small samples of students attending a specific course.

[2] The current focus of many UK higher education policies is to improve access, academic success and labour market outcomes for students from minority ethnic groups and disadvantaged family backgrounds (Connell-Smith and Hubble, 2018; Gaskell and Lingwood, 2019). In 2015 the UK Government set two targets for widening participation in higher education by 2020: to double the proportion of pupils from disadvantaged backgrounds going into higher education; and to increase by 20% the numbers of students from black and minority ethnic backgrounds entering higher education. The Office for Students (England) has recently established a target to eliminate the unexplained gap in degree outcomes between white students and black students by 2024–25. Moreover, a package of measures aiming to tackle ethnic disparities at university was announced by the UK government through the Race Disparity Audit in 2019 (Universities UK, 2019).

Our next step is to combine data on actual choices with data on subjective expectations of the returns to and enjoyment of the activities to estimate a structural discrete choice model of time allocation at university. To do so, we need to address how to take into account constraints that might affect students' choices. Students may be financially constrained and may need to work to fund their studies; they may be unable to secure relevant work experience due to labour demand rationing; they may face time constraints due to commuting or family responsibilities; or they may ignore some choices due to the influence of social norms. This is problematic, as including alternatives that are not in a student's true choice set could lead to inconsistent estimates of the choice model.

In this paper we obtain consistent estimates of the parameters in our model by conditioning on a subset of the true and unobserved choice set (McFadden, 1978). We follow Crawford et al. (2020) who provide some guidance and examples on how to construct such a subset (labelled ''sufficient choice set'') in applications similar to ours. Our key assumption is that individuals of the same gender, ethnicity, SES status and department of study face the same constraints and thus the same choice set. However, students within that group might still have different labour market opportunities. To address this, we make use of unique individual-specific information on internship applications and outcomes. For each individual in our sample, a sufficient choice set is therefore defined by including all choices made by similar students and excluding choices that comprise work relevant experience for students that apply for an internship but fail to secure one.[3]

The estimates reveal that the main drivers of students' investment choices are future earnings, academic achievement and enjoyment of academic investments. We investigate heterogeneity by estimating different structural preference parameters by ethnicity and SES. The main difference that emerges is in relation to the utility from getting a good degree: white British high-SES students have a strong preference for having a good degree, while this preference parameter for non-white British is not significantly different from zero.

We then use the estimated structural preference parameters to gain a better understanding of what drives gaps in investments. Our estimation strategy does not allow us to point-identify individual-specific choice probabilities but we present a lower bound of the predicted probabilities derived under the assumption that students face an unrestricted choice set. With no constraints, the ethnic gaps in all investments are dramatically reduced. By contrast, various counterfactual scenarios in which we attribute white high-SES preferences, beliefs, or enjoyment of activities to non-white students reduces the gap very little, if at all. These results point to an important role of constraints in explaining the observed ethnic differences in time allocation choices.

This paper contributes primarily to the literature on the role of subjective expectations in educational choices. A major focus in this literature has been on the role of future earnings (Arcidiacono, 2004; Beffy et al., 2012; Berger, 1988; Willis and Rosen, 1979). As there is evidence that students tend to be misinformed about the returns to schooling (Betts, 1996; Jensen, 2010; Wiswall and Zafar, 2015), a recent literature has used data on subjective expectations (rather than assumptions on expected earnings) in order to explain students' choices (Arcidiacono et al., 2012; Attanasio and Kaufmann, 2014; Giustinelli, 2016; Wiswall and Zafar, 2015). A smaller group of papers has also looked at the role of expectations about the non-pecuniary benefits of education (Boneva and Rauh, 2017; Delavande and Zafar, 2019; Zafar, 2013) and academic ability (Stinebrickner and Stinebrickner, 2012). The educational choices that have been studied in this context include going to university, choosing among different types of universities, drop-out, and college major.

Our paper innovates in several dimensions. From a substantive point of view, we focus on students' primary inputs into their human capital accumulation, time allocation across different types of investments, and the role of expectations about the returns to academic as well as non-academic uses of time. Acknowledging that students can engage in a wide range of activities (beyond studying and attending lectures) that have employability and labour market returns is an important step forward for our understanding of the production function of academic achievement at university and is motivated by the observation that in an increasingly competitive graduate labour market employers are looking for a wider range of skills and abilities than those signalled through final grades (Association of Graduate Recruiters, 2016; Prospects, 2019).

From a survey design point of view, our module on expectations is novel in that it asks students about their beliefs in relation to future realization of a wide range of outcomes conditional on many different investment choices (12 scenarios). Our application also represents a significant departure from existing work eliciting expectations conditional on discrete alternatives, such as contraceptive method or college major (e.g., Delavande, 2008; Arcidiacono et al., 2019), because time is a continuous variable with a multiplicity of uses (e.g. attendance, study hours and a variety of non-academic activities). This requires careful consideration of the way in which we define discrete values for an underlying continuous variable and set up scenarios that represent alternative allocations. Our results, based on visual graphics to represent the scenarios, are encouraging in that they show that expectations thus elicited exhibit variation that is systematically related to students' choices and can lead to new insights about the drivers of their behaviour.

Finally, from an econometric point of view, our empirical strategy carefully addresses the fact that students may be subject to multiple unobserved constraints. Attention to the way constraints should be modelled is relevant to the

[3] Some of the alternative approaches that deal with choice set heterogeneity include integrating over the distribution of choice sets, which typically requires additional information or assumptions on the choice set formation (Goeree, 2008; Manski, 1977; van Nierop et al., 2010), or using stated choice sets directly elicited from respondents (Ben-Akiva and Boccara, 1995; Stinebrickner and Stinebrickner, 2008b). Another approach is to use hypothetical choice scenarios (Delavande and Zafar, 2019; Louviere et al., 2000; Wiswall and Zafar, 2018). This issue is related to the one faced by those who investigate the role of credit constraints in participation to higher education, where more indirect methods have been used to identify students who are constrained (see Lochner and Monge-Naranjo, 2012, for a review).

expectations literature as constrained individuals may behave as if they do not "act on their expectations", influencing the inference researchers make about the role of expectations in the decision-making process.

This paper is organized as follows. Section 2 outlines a model of investment at university. Section 3 describes our data. We provide descriptive evidence on the academic and non-academic investments at university in Section 4 and on subjective expectations in Section 5. Section 6 addresses potential concerns regarding measurement error and data quality in general. Sections 7 and 8 report estimates from the discrete choice model and corresponding counterfactual scenarios, respectively. Section 9 concludes the paper.

2. A model of academic and non-academic investments at university

In this section, we develop a simple model of investment in academic and non-academic activities at university. A student lives for $T+1$ periods. In period $t=0$, the first period of her life, student i starts higher education. At the end of the period, she leaves university and enters the labour force where she stays until period $t=1,\ldots,T$. Her overall utility depends on consumption, the utility associated with investment I undertaken while at university, academic achievement D, and future job characteristics J, such as being in a job with promotion prospects or in an interesting field of work. We consider academic investments, I_A, such as lecture attendance or study hours, and non-academic investments, I_{NA}, such as doing an internship or volunteering (i.e., $I=\{I_A, I_{NA}\}$).

For tractability, we assume that the utility function is additively separable, linear in academic achievement and job characteristics, and logarithmic in consumption. The lifetime utility of student i is therefore given by:

$$U_i\left(I, c_{it}\right) = u_i\left(I_A, I_{NA}\right) + \theta ln\left(c_{i0}\right) + \rho D + \sum_{t=1}^{T} \beta_i{}^t \left(\theta ln\left(c_{it}\right) + \alpha J\right) + \varepsilon_{iI}, \tag{1}$$

where $u_i\left(I_A, I_{NA}\right)$ is i's net utility associated with a specific combination of academic and non-academic investments, c_{it} is individual i's consumption at time t, θ is the utility value of log consumption, ρ is the utility value of academic achievement, β_i is the individual i's time preference discount factor. The vector α contains the utility value of job characteristics J, and ε_{iI} is a random term which is specific to the individual and the combination of investments chosen, observable to student i at the time investments are made but not to the econometrician. We assume for simplicity that there is no borrowing or lending, so student i will consume her earnings y_{it} at every period $t=1,\ldots,T$. At university ($t=0$), students are assumed to consume out of a parental or governmental allowance y_{i0}.

A key feature of the model is that, at the time of choosing the investments, the student faces uncertainty about academic achievement, the employment probability, the lifetime earnings and the job characteristics associated with each choice. She has beliefs about how academic and non-academic investments influence these outcomes. For simplicity, we assume that academic and non-academic investments influence beliefs about academic achievement, but that only academic achievement and non-academic investments influence the subjective probability for labour market outcomes. In particular, student i has individual-specific beliefs on: (***i***) $P_i(D|I_A, I_{NA})$, i.e. about how academic and non-academic investments influence academic achievement; (***ii***) $P_i(job_t\,|\,D, I_{NA})$, i.e. about how academic achievement and non-academic investments influence the probability of being in employment at time t; and (***iii***) $F_i(y_{it}, J\,|\,D, I_{NA})$, i.e. about how academic achievement and non-academic investments influence future earnings and job characteristics. We accommodate three levels of academic achievement ($d = 1$ to 3), corresponding to a *first class* degree, an *upper second class* degree, or a *lower second class* degree or worse, following the UK undergraduate degree classification system.[4] We denote by y_{ut} the earnings if student i is unemployed at time t and assume there is no uncertainty associated with it.

Student i will choose the combination of investments $\{I_A, I_{NA}\}$ that maximizes her lifetime expected utility:

$$\begin{aligned}
&\max_{I_A, I_{NA}} u_i\left(I_A, I_{NA}\right) + \rho \sum_{d=1}^{3} D \times P_i(D=d|I_A, I_{NA}) + \theta ln\left(y_{i0}\right) \\
&\quad + \sum_{d=1}^{3} P_i(D=d|I_A, I_{NA}) \sum_{t=0}^{T} \beta_i{}^t \left[\begin{array}{c} P_i(job_t\,|\,D=d, I_{NA}) \displaystyle\int \left(\theta ln\left(y_{it}\right) + \alpha J\right) dF_i(y_{it}, J\,|\,D=d, I_{NA}) \\ + \left(1 - P_i(job_t\,|\,D=d, I_{NA})\right) \times \ln(y_{ut}) \end{array}\right] \\
&\quad + \varepsilon_{iI}
\end{aligned} \tag{2}$$

subject to the constraint $\{I_A, I_{NA}\} \in FS_i$, where FS_i is the feasible choice set of student i which captures time, labour demand or other constraints.

The goal of the empirical analysis is to estimate the parameters of the utility function (up to scale) using data on the chosen investments as well as data on the individual-specific expectations. Identification of this model is discussed in Section 7.

4 The classification system is described in more detail in Section 5.1.

3. Data

3.1. The BOOST2018 study

The BOOST2018 Study is a longitudinal survey of undergraduate students who enrolled at one UK university in the academic year 2015/16, and (for the vast majority) completed their degree in 2017/18. Each year students were invited to reply to three on-line surveys, one for each term, and to attend one laboratory session.

The on-line surveys were variable in length but generally took about one hour with the exception of the summer term surveys (waves 4, 8 and 12), which were shorter to allow for the fact that students take their exams at the end of the academic year. Participation in the surveys was incentivized using monetary rewards — between £8 and £20 for on-line surveys and on average £30 for the laboratory sessions. The on-line surveys were designed to collect information on students' academic investments (hours of study), non-academic investments (working for pay, participation in volunteering groups, etc.), their expectations about future academic achievement and labour market outcomes (earnings, probability of employment) as well as, at wave 9, non-pecuniary future job attributes.

The survey data was linked to administrative records held by the university. Specifically, we use here information on the student demographics (gender and age), socio-economic status as measured by parental occupation and the university participation rate in their neighbourhood of domicile, and marks. Ethnicity is self-identified at the time of enrolment at the university. We also obtained access to their timetable of scheduled lectures and classes and weekly records of attendance – administered through a swipe-card electronic system – to derive measures of attendance that are not affected by self-reporting.[5]

The sampling frame comprised all undergraduate students enrolling in the first year of an undergraduate (Bachelor's) course in October 2015. The target population consisted of 2621 students. In order to participate in the study, each student was required to sign a consent form. All students who enrolled in the study received £5 as an incentive. By the end of the Autumn term of the academic year 2015/16, when the participation register was closed, 1978 students had given their consent (about 75% of the target sample). Enrolment re-opened to eligible students at the beginning of the second and third years, resulting in a small number of additional participants ($n = 19$).

Because of the presence of monetary incentives and the advertising campaigns aimed at keeping the study salient to the population, participation to the surveys was consistently high. Between 774 and 1276 students took part in the surveys at different points in time, with higher response rates for the main on-line waves (between 55% and 68%), and lower rates for the laboratory sessions (between 45% and 59%) and summer term surveys (between 52% and 56%) (see Appendix Table A1).

3.2. Students' demographic characteristics

Table 1 shows the demographic characteristics of students in the target population and compares it to the students who ever enrolled or were still participating during the third year of BOOST2018. The target population (column 1) is almost equally composed of male and female students; 90% are 'young', meaning aged 21 or under on entry; and around 69% are of British nationality. The socio-economic characteristics of this population are quite diverse, resulting in a significant overall percentage from minority/non-white British ethnic backgrounds (27%), and from white British low-SES families (15%). We focus on these groups, in comparison to students from white British high-SES backgrounds, as closing the long-standing gaps in degree performance and labour market outcomes between these groups is a high-profile policy objective across Higher Education institutions in the UK (see also footnote 2).

We define ethnicity and socio-economic status using university administrative records. Specifically, we define the non-white British group as those students reporting to be from a non-white etc. Where parental occupation was not available, we classified those domiciled in the top 40% of postcodes (ZIP codes) for Higher Education participation as being of high SES.

Column 2 shows that the sample of students ever enrolled in BOOST2018 is broadly representative of the target population. There is a slightly higher participation of women and Overseas students, but the sample is generally a clear reflection of the demographics, social and academic (i.e. by faculty) diversity of the institution. Attrition does not seem to be a significant problem either, as shown in column 3, which refers to the sample of students still enrolled in BOOST2018 during their third year at university.

We next provide descriptive statistics for our estimation samples. In our empirical analysis, we seek to provide evidence on students' investments during their second and third year at university. This is because marks obtained during the first year do not count toward the overall final mark in the UK and non-academic investments are less common in that year. Our first estimation sample, described in column 4, includes respondents for whom we have information on investments in their second (waves 5 and 7) and third years (waves 9 and 11) as well as expectations about the returns to academic and non-academic investments (wave 9), i.e., students participating in either wave 5 or 7, and in wave 9 of the study. The

5 The measures of attendance obtained using administrative records are not error-free (e.g. the system might not work on some days, or lectures can be cancelled at the last moment so that an absence is falsely recorded), but they are such that it is much more unlikely that errors in measurement are correlated to individual characteristics than it would be the case with self-reports.

second estimation sample, described in column 5, restricts the analysis to those who additionally reply to the last wave (wave 12), which provides information on the outcomes of job placement applications that we use to define respondents' sufficient choice set (see Section 7).

Both estimation samples, based on responses to different combinations of surveys, slightly over-represent non-white British and underrepresent low-SES white British students relative to the underlying target population. They also contain a higher percentage of female students than the population, and this is not unusual in a longitudinal study (see e.g. Lynn and Borkowska, 2018; Department for Education, 2011). Students who continue to participate in the study over the three years are also positively selected on early academic performance. This partly reflects drop-out from university between the first and third years (approximately 13% are no longer enrolled at the start of the third year), and is partly explained by the fact that women have higher achievement than men on average.

Note that in the UK system students choose their field of study prior to enrolling at university. The estimation samples closely reflect enrolment rates by field of study in the population, and there are important gender, social and ethnic differences across subjects. Within both the student population and our main estimation sample, non-white British are less likely to be enrolled in the humanities and more likely to be enrolled in social science compared to white high-SES, and females are more likely to be from white low-SES background (see Appendix Table A2). In the analysis of students' behaviour and expectations we always include field of study fixed-effects and control for gender. Therefore, we do not expect this sorting to cause problems for the interpretation of our results.

4. Students' academic and non-academic investments at university

In this section we discuss our measures of academic and non-academic investments. We consider the average level of investment over the last two years at university during term time, with the exception of non-academic activities undertaken in the summer between the second and the third year (Appendix A1 for the exact question wording and Appendix A2 for details 24 on how we take averages over time). Our analysis is restricted to the sample of students who reply to "wave 5 or 7, and wave 9" of the survey.

4.1. Academic investments

We consider two broad types of academic investments: attendance to classes and lectures, and hours of study. While attendance is derived from administrative records, hours of study are self-reported from the question "Not counting hours spent in class and lectures, how many hours in a typical week during term time do you usually study?".

We focus on three different measures of attendance: (i) average hours per week; (ii) the percentage of scheduled hours (obtained by dividing hours of attendance by the number of hours a student is expected to attend according to the course she is enrolled in); and (iii) high attendance (attending above median percentage of scheduled hours). The last two measures take into consideration the fact that the total scheduled hours vary by department. Note that attendance is not compulsory and officially is left at the discretion of the students. However in practice some courses are partially assessed through coursework or practical assignments that must be completed in scheduled laboratory time, or for which attendance at specific classes is a pre-requisite.[6] Table 2 presents OLS regression coefficients showing the differences in academic investments by demographic and social characteristics of the population. Students attend on average 64% of their scheduled hours, which corresponds to about 6 h weekly. A striking result emerges according to students' background: according to all attendance measures, non-white British attend lectures and classes significantly less than high-SES white British. For example, they have a 9.4ppt lower attendance rate (column 2) and are 20ppt less likely to attend more than the median attendance (column 3). These effects, cumulated over 2 years, are very large. Interestingly, there is no difference in terms of attendance between low-SES and high-SES white British.

We then focus on two measures of private study: (i) hours per week; and (ii) an indicator for high study hours (more than 12.5 h, which is the sample median). Students spend on average 14 h per week in private study. Looking at the difference by students' background reveals again that non-white British stand out: they are 9.9ppt more likely to engage in high study hours compared to high-SES white British (column 5). There is however some within-group heterogeneity as we do not see statistical differences in terms of actual hours per week (column 4). Again, there is no difference in terms of attendance between low-SES and high-SES white British.

Overall, students spend almost 20 h per week in academic activities (column 6), but there is no difference by students' ethnic or social grouping. So while non-white British and high-SES white British spend on average the same number of hours of academic investment, they choose different time allocations, with non-white British attending lectures and classes significantly less but compensating with more study hours outside of classes.[7]

While not the main focus of this paper, Table 2 also reveals that women invest more than men along all dimensions: they have a 5ppt higher attendance rate and are 20ppt more likely to study a high number of hours (which is about twice the gap between non-white and high-SES white British), which results in an additional 3.7 h of academic investment per week overall.

6 In addition, Overseas students must maintain a sufficient level of attendance to comply with the conditions of their study visa.

7 Although not shown here for reasons of space, there is substitution between academic investments and recreational activities. According to information collected through time diaries, students who invest more in academic activities (study and attendance) report fewer hours of recreational time spent alone, while those who invest more in non-academic activities seem to substitute away from recreational time spent with others.

Table 1
Demographic characteristics for the population and relevant samples.

		Population	BOOST participants		Estimation samples	
			Ever enrolled	Enrolled for third year of study	Waves (5 | 7) and 9	Waves (5 | 7) and 9 and 12
		%	%	%	%	%
		(1)	(2)	(3)	(4)	(5)
Sex:						
Female		50.3	52.4***	53.7***	57.2***	59.2***
Demographic group:						
White British high-SES		24.5	24.2	24.0	23.7	24.3
White British low-SES		15.3	15.1	14.7	14.1	13.0**
Non-white British		26.6	29.4***	29.7***	30.8***	30.8***
EU		16.2	15.9	16.4	17.9*	17.6
Overseas		15.2	13.8***	13.6***	12.1***	12.5**
Other/Refused		2.3	1.6***	1.3***	1.2***	1.4*
Age group:						
Young (less than 21 at entry)		90.5	92.0***	92.6***	92.6***	92.8**
Faculty:						
Humanities (6 departments)		28.1	28.0	28.1	28.1	28.0
Science and Health (5 departments)		32.0	29.7	30.4**	32.7	33.3
Social Sciences (5 departments)		40.2	40.2	41.1**	39.5	38.8
Academic performance:						
First year mark	Above 70: 'first class' level	[NA]	18.2	20.2***	22.6***	23.1***
	Above 60: good degree' level	[NA]	52.3	57.5***	60.9***	62.4***
N		2621	1997	1738	1002	773

Note: Faculty of Humanities comprises departments of Art History; History; Interdisciplinary Studies; Law; Literature, Film and Theatre; and Philosophy. Faculty of Science and Health comprises departments of Psychology; Biological Sciences; Computer Science; Health and Human Sciences; Mathematics. Faculty of Social Sciences comprises department so of Sociology; Business; Economics; Government; and Language and Linguistics. Those not completing their first year of study or who drop out are included in the denominator for the proportion achieving a first class or good degree. NA = "Not available". Significance levels for difference in sample proportions relative to population (except for academic performance, which is not available for students who did not enrol in the study, so here we compare to the ever enrolled sample). **Symbols:** * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

Table 2
Academic investments by demographic characteristics.

	Attendance			Study		Total academic investment (hours p.w.)	High attendance and high study
	Hours per week	% of scheduled hours	High	Hours per week	High		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	0.492***	0.050***	0.097***	3.171***	0.201***	3.663***	0.183***
	(0.135)	(0.013)	(0.032)	(0.559)	(0.034)	(0.591)	(0.031)
Demographic group: (Base = White British high-SES)							
White British low-SES	−0.006	−0.008	0.029	−0.640	−0.005	−0.646	0.018
	(0.206)	(0.019)	(0.049)	(0.853)	(0.052)	(0.902)	(0.047)
Non-white British	−0.507***	−0.094***	−0.204***	0.969	0.099**	0.462	−0.051
	(0.172)	(0.016)	(0.041)	(0.712)	(0.044)	(0.753)	(0.039)
EU	0.798***	0.029	0.101**	0.224	0.059	1.022	0.078*
	(0.198)	(0.019)	(0.047)	(0.820)	(0.050)	(0.867)	(0.045)
Overseas	1.073***	0.084***	0.224***	0.472	0.047	1.546	0.121**
	(0.223)	(0.021)	(0.053)	(0.926)	(0.057)	(0.979)	(0.051)
Sample mean:	5.84	63.63	0.567	14.08	0.482	19.92	0.298
N	1002	1002	1002	1002	1002	1002	1002

Note: Observations are from the estimation sample of students participating in wave 5 or 7 & wave 9 (see Table 1, column 4). Coefficients shown are from different OLS regressions of time allocation on demographic and socio-economic characteristics of students. Additional controls are: indicator for ethnicity "other/refused", indicator for mature student, and 16 department of study dummies. Total academic investment is defined by hours (p.w.) of study and hours of attendance; high attendance is defined as 1 if student attends 62% of scheduled hours or more, and zero otherwise; high study is defined as 1 if students reports 12.5 h or private study or more per week. **Symbols:** * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

Table 3
Non-academic investments by demographic characteristics.

	Positive non-academic investments in summer vacation	Non-academic investments: (hours per week)	Has attained			
			Relevant work experience	Extra-curricular experience	Neither	Both
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.147***	2.875***	0.046	0.062**	−0.053**	0.055*
	(0.032)	(0.860)	(0.033)	(0.025)	(0.022)	(0.032)
Demographic group: (Base = White British high-SES)						
White British low-SES	−0.030	0.157	−0.081	0.036	−0.001	−0.045
	(0.050)	(1.313)	(0.050)	(0.038)	(0.034)	(0.049)
Non-white British	−0.093**	0.666	−0.081*	−0.049	0.063**	−0.067
	(0.041)	(1.094)	(0.042)	(0.032)	(0.029)	(0.040)
EU	−0.104**	5.167***	0.062	0.049	−0.032	0.079*
	(0.048)	(1.262)	(0.048)	(0.037)	(0.033)	(0.047)
Overseas	−0.223***	−0.888	0.066	−0.113***	0.115***	0.067
	(0.054)	(1.432)	(0.055)	(0.042)	(0.037)	(0.053)
Sample mean	0.656	11.35	0.331	0.842	0.123	0.296
N	1002	1002	1002	1002	1002	1002

Note: Observations are from the estimation sample of students participating in wave 5 or 7 & wave 9 (see Table 1, column 4). Coefficients shown are from different OLS regressions of time allocation on demographic and socio-economic characteristics of students. Additional controls are: indicator for ethnicity "other/refused", indicator for mature student, and 16 department of study dummies. Dependent variables: Column (1), binary variable equal to one if respondent volunteered or took paid or unpaid work experience during summer vacation between second and third years of study. Column (2), continuous variable comprising sum of hours spent on paid and unpaid work (including internships), volunteering, competitive sport, leadership roles for Student Union, and other qualifying activities for the university's 'Employability' award. Columns (3) and (4), binary variables equal to one if they have accumulated experience relevant to their field of study or their desired career (column 3); or relevant to neither their field of study nor desired career (column 4) – see Appendix A2 for exact definitions of qualifying activities. Column (5), binary variable equal to one if they have not accumulated either of these kinds of experience (Relevant work experience = 0 and Extra-curricular experience = 0). Column (6), binary variable equal to one if one if they have accumulated both these kinds of experience (Relevant work experience = 1 and Extra-curricular experience = 1). **Symbols:** * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

4.2. Non-academic investments

Our measures of non-academic investments are derived from several questions covering the following activities: working for pay; performing a leadership role for the student union, or a sports club or a student society; training for or participating in sporting competitions; volunteering; being in a paid or unpaid internship or in a career placement; and a residual category for engaging in other activities recognized by the university's employability awards (such as being a student representative). Questions related to these activities are asked with reference to the current term, and retrospectively for activities performed during the summer vacation. For employment, internship and volunteering activities, students were also asked to report whether this experience was relevant to their field of study or desired career.

We measure time spent in these activities as *hours per week* during university term-time, and also look at indicator variables for the accumulation of experience (including during summer vacations) that is related (or not) to the field of study or desired career. Specifically, we consider an indicator variable for *relevant work experience* that is equal to 1 if the student has engaged in paid work, internship, or volunteering experience reported as being relevant to their field of study or desired career. We also consider a similar indicator variable for *extra-curricular experience*, which includes activities such as participation in sport competitions, as well as employment, internship or volunteering reported as not being relevant to the student's field of study or desired career (see Appendixes A1 and A2 in the Supplementary Material).

Table 3 shows that students spend just over 11 h per week during term time on average in non-academic investments (column 2), and that two-thirds have engaged in non-academic investment in the summer between the 2nd and 3rd years (column 1). Interestingly, we see again that non-white British and high-SES white British spend on average the same amount of time in non-academic activities during term-time, but that their time allocation differs. In particular, non-white British are 8ppt less likely to have attained relevant work experience by the end of their studies (column 3), and 5ppt less likely to have extra-curricular experience — although the coefficient is imprecisely estimated (column 4). They are also 9ppt less likely to have engaged in any activities during the summer (column 1). As a result, they are 6ppt more likely to have accumulated neither work relevant experience nor extra-curricular experience (column 5). The time allocation of low-SES white British is not statistically different from that of high-SES (all columns of Table 3). We note nevertheless that they are 8ppt less likely to have accumulated relevant work experience (the same magnitude as non-white) although the coefficient is not very precisely estimated.

Further breaking down the hours spent and prevalence in the various individual activities, non-white British are less likely to have internship experience (6.4% versus 10.9% for white high-SES and 11.3% for white low-SES) or paid employment (51% versus 56% for high-SES and 61% for low-SES) than their white counterparts. However, as shown in Appendix Table A3, only the former gap is statistically significant after accounting for sex, field of study and age. Moreover, the fact that we do not see large differences in mean hours suggests that ethnic minority students who do engage in these activities work longer hours.

Table 3 also reveals interesting gender differences. Women spend on average 2.9 h per week more than men in non-academic activities and are 6ppt more likely to have accumulated extra-curricular experience. They are however not significantly more likely to have accumulated relevant work experience than males.

5. Expectations related to academic and non-academic investments

In this section we describe our measures of the students' expectations about the productivity of academic and non-academic investments in terms of academic achievement and labour market outcomes. We also provide analyses of our measures of the enjoyment students derive from spending time in each type of activity. These are the key components of the model introduced in Section 2 and which will be estimated in Section 7. As before, we control for department of study in all our specifications as differences in beliefs by field of study may reflect genuine differences in academic performance or labour market prospects.

5.1. Expectations about returns for academic performance

In order to capture students' subjective beliefs about the returns to academic and non-academic investments in terms of their academic outcomes, we elicited students' expectations about the probability they would achieve a different degree class conditional on a specific time allocation. The degree classification we use is a categorical transformation of marks that is almost universally used by UK universities to distinguish levels of academic performance on the award of an undergraduate degree. Although universities set their own precise conditions, generally a Grade Point Average (out of 100) of 70 or above leads to a *first class* degree, 60–69 is considered an *upper second class*, and 50–59 is a *lower second class*. A *good degree*, comprising a first or upper second, is the threshold widely used by employers as a pre-requisite for applying for graduate-level roles. [8] This classification is known and salient to students.

An important challenge from a survey design perspective is to define time allocation in a way that is refined enough to capture students' actual choice while keeping respondents' burden low. To achieve this, we consider 12 different scenarios using all possible combinations of: (i) two levels of attendance at lectures and classes, set at 6 or 10 h; (ii) two levels of study hours, set at 10 or 15 h; and (iii) three levels of non-academic activities, set at zero, 10, or 20 h. The latter two represent approximately the median and 80th percentile among those with positive hours as measured during the second year of study. These combinations were illustrated as proportions of a notional 60 h week, after accounting for essential activities such as sleep and shopping. Fig. 1 (panel A) shows the way the question was visually presented to the students (see Appendix A1 for more details).

To illustrate the raw data, Fig. 2 plots the mean subjective probabilities of attaining a good degree (first class or upper second) for all 12 scenarios, with 95% confidence intervals. This indicates that students expect better academic outcomes with a larger input of any of the three investment types, other things equal, but also that they foresee trade-offs among these inputs that can lead to very similar expected outcomes from very different time allocations.

We summarize these expectations by calculating the mean expected *returns* to an hour of attendance, study, or non-academic investment using an individual fixed-effect regression pooling all respondents and scenarios. Here the dependent variable is the subjective probability of achieving a certain educational outcome, and we exploit variation in the level of investment generated by the 12 scenarios, i.e. the independent variables are the hours of the three activities in the associated scenario. The results are presented in Table 4, in which a clear hierarchy emerges, in that an additional hour of attendance is perceived to increase the probability of a good degree by 3ppt, an additional hour of study by 1.7ppt and an additional hour of non-academic activities only by 0.4ppt. Corresponding coefficients for getting a first class degree are slightly smaller.[9]

Using the 12 observations per person, we also estimate individual-specific regressions as in Blass et al. (2010) to calculate for each student the expected return to a marginal hour spent on each of these activities, holding constant time spent on the others. Fig. 3 (left panel) plots the distribution of these regression coefficients for getting a good degree. The variance as well as mean of the expected return to attendance is greater than that of study, and in turn, greater than that of non-academic investments.

[8] Data for the academic year 2017/18 indicate that 27% of the population of UK undergraduate students graduated with a first class degree, 46% with an upper second degree, and 21% with a lower second or worse, while 6% of students were unclassified. See https://www.hesa.ac.uk/data-and-analysis/students/outcomes.

[9] With 12 data points per person, we have enough degrees of freedom to estimate coefficients on the interactions between inputs. In an auxiliary regression, attendance, study, and non-academic activities are all found to be complements in increasing the probability of getting a first class degree. For reasons of space, we do not expand our analysis to document heterogeneity in these interactions by demographic group.

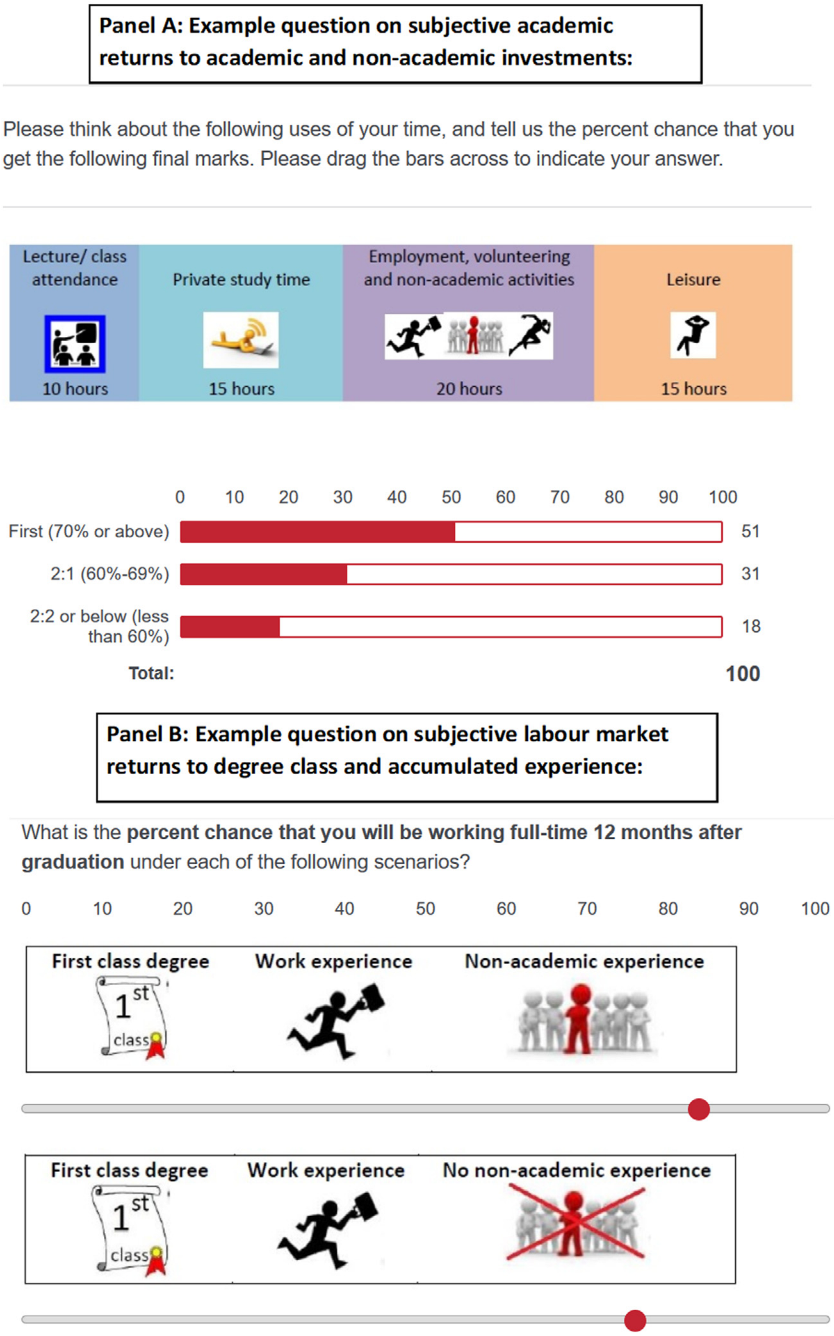

Fig. 1. Question design: Elicitation of subjective returns to academic and non-academic investments. **Note:** Panel A: Expected final degree mark conditional on time allocation. First of twelve scenarios shown. Panel B: Expected probability of working full-time 12 months after graduation conditional on degree class and accumulated experience. First two of twelve scenarios shown.

Finally, in Table 5 we investigate heterogeneity more formally by regressing the individual-specific coefficients representing subjective returns to each activity on the demographic characteristics of the students. Non-white British students differ from their white high-SES colleagues in that they attribute a lower subjective return to both attendance and study hours in relation to getting a first class degree. It is not clear whether these differences stem from different access to information or truly different production functions (by demographic groups or field of study). On the other hand, white British from low and high-SES background hold similar expectations.

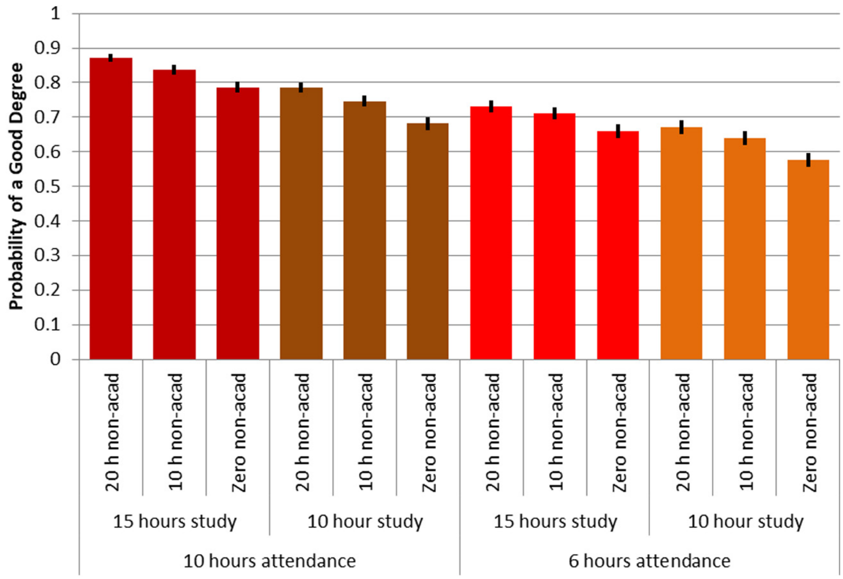

Fig. 2. Expected probability of good degree conditional on academic and non-academic time use. **Note:** Observations are from the estimation sample of students participating in wave 5 or 7 & wave 9 (see Table 1, column 4); N = 1002. Bars indicate sample averages and black line indicates 95% confidence interval.

Table 4
Individual fixed effect models: Effect of marginal hour on percent chance that will obtain given degree class.

	Percent chance of	
	First class degree	Good degree
	(1)	(2)
Attendance	2.485*** (0.060)	2.995*** (0.072)
Study	1.498*** (0.048)	1.650*** (0.057)
Non-academic	0.303*** (0.015)	0.443*** (0.018)
Mean constant term	−10.767*** (0.791)	23.437*** (0.945)
N (observations)	12 023	12 023
N (respondents)	1002	1002

Note: Observations are from the estimation sample of students participating in wave 5 or 7 & wave 9 (see Table 1, column 4). Coefficients shown are from individual fixed-effects regressions of academic achievement indicators on time use (in hours p.w.). **Symbols:** * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

5.2. Expectations about returns for labour market outcomes

Following the model set up in Section 2, we also want to capture expectations about the labour market returns to academic achievement and non-academic investments. We therefore ask students to report their subjective probability to be in different possible labour market states subject to obtaining a certain degree class and having obtained work experience relevant to their field of study or desired career, other extra-curricular experience, neither, or both. We elicit expectations with respect to the following labour market outcomes: percent chance of being in full-time work 12 months after graduation; salary 12 months after graduation and at age 40; percent chance of the job being in a field they like; percent chance of being promoted to a position with a greater level of leadership and responsibility within the next three years, percent chance that this job will be useful for society or the general public (see Appendix A1).

This set-up lends itself more naturally to the use of discrete scenarios but this may still be challenging to respondents as we want to capture *joint* investments. To keep the survey tractable, our design assumes that labour market outcomes do not depend on academic investments once we condition on degree class and non-academic investments. We use 12 scenarios with all possible combinations of: (i) degree classification (first class, upper second, lower second or worse); (ii) two levels of relevant work experience; and (ii) two levels of extra-curricular experience. Fig. 1 (panel B) shows how we use a visual graphic to present the first two of these scenarios.

Table 5
Heterogeneity in academic returns to time use by demographic characteristics: Effect of a marginal hour on percent chance that will obtain given degree class.

For obtaining a:	First class degree			Good degree		
Subjective return to marginal hour of:	Attendance	Study	Non-academic	Attendance	Study	Non-academic
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.458**	0.134	0.043	1.000***	0.398***	0.127***
	(0.227)	(0.128)	(0.039)	(0.288)	(0.151)	(0.047)
Demographic group: (Base = White British high-SES)						
White British low-SES	−0.411	−0.247	0.005	−0.236	−0.153	−0.045
	(0.347)	(0.195)	(0.059)	(0.439)	(0.230)	(0.072)
Non-white British	−0.625**	−0.415**	0.055	−0.347	−0.002	0.072
	(0.290)	(0.163)	(0.050)	(0.366)	(0.192)	(0.060)
EU	−0.220	−0.254	0.054	−0.999**	−0.423*	−0.104
	(0.334)	(0.187)	(0.057)	(0.422)	(0.221)	(0.069)
Overseas	−0.714*	−0.134	0.062	−1.006**	−0.357	−0.053
	(0.377)	(0.212)	(0.065)	(0.477)	(0.250)	(0.078)
N	1002	1002	1002	1002	1002	1002

Note: Observations are from the estimation sample of students participating in wave 5 or 7 & wave 9 (see Table 1, column 4). Coefficients shown are obtained from OLS regressions of the academic returns to time use (as estimated by individual regressions using 12 scenarios per each individual) on demographic and socio-economic characteristics of students. Additional controls are: indicator for ethnicity "other/refused", indicator for mature student, and 16 department of study dummies. **Symbols:** * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

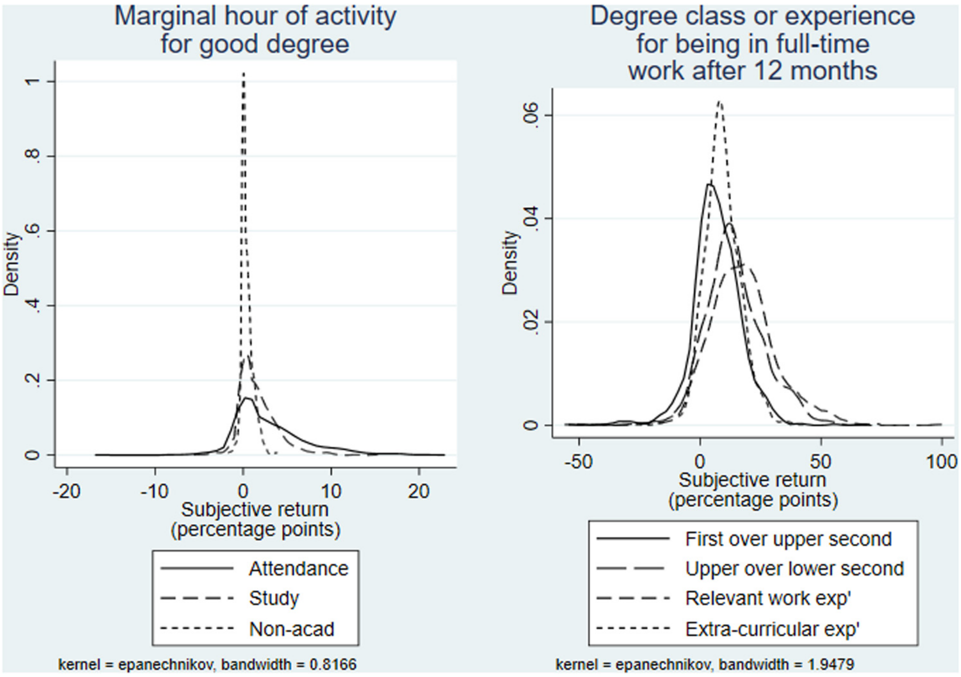

Fig. 3. Distribution of individual subjective returns to time investments for getting a good degree, and degree class and accumulated experience for probability of being in full-time employment. **Note:** Observations are from the estimation sample of students participating in wave 5 or 7 & wave 9 (see Table 1, column 4). Left panel: distribution of the coefficients obtained from individual-specific regressions of academic returns (i.e. getting a good degree) to time use obtained using 12 scenarios per each individual (N = 1002). Right panel: distribution of the coefficients obtained from individual-specific regressions of being in full-time work 12 months after graduation on academic achievement and non-academic investments using 12 scenarios per each individual (N = 948). Return to first class degree is over having an upper second degree, and return to upper second class degree is over having a lower second class degree or worse.

We next estimate the expected labour market returns to degree class and accumulated experience by estimating individual fixed-effect regressions of expected labour market outcomes pooling all respondents and exploiting variation in inputs across the 12 different scenarios. The results are presented in Table 6. There is a strong positive expected return to academic achievement, relevant work experience, and other non-relevant experience on all the outcomes.

Table 6
Individual fixed effect models: Effect of degree class and accumulated experience on expected job attributes.

	In full-time work 12 months after graduation (% chance)	If in full time work				
		12 months after graduation				At age 40
		Salary (£000s)	Promoted in next three years (% chance)	Job is in a field I like (% chance)	Job useful to society (% chance)	Salary (£000s)
	(1)	(2)	(3)	(4)	(5)	(6)
Degree class: (Base category: Lower second class degree or worse)						
First class degree	21.220*** (0.306)	5.080*** (0.449)	18.743*** (0.267)	20.227*** (0.278)	13.496*** (0.250)	18.231*** (1.231)
Upper second class degree	14.262*** (0.306)	3.064*** (0.449)	11.636*** (0.267)	12.363*** (0.278)	7.306*** (0.250)	11.264*** (1.231)
Attained:						
Relevant work experience	17.865*** (0.250)	2.182*** (0.366)	12.371*** (0.218)	11.530*** (0.227)	6.761*** (0.204)	6.323*** (1.005)
Extra-curricular experience	9.217*** (0.250)	1.207*** (0.366)	6.368*** (0.218)	5.902*** (0.227)	4.033*** (0.204)	3.827*** (1.005)
Mean constant term	30.582*** (0.279)	18.659*** (0.410)	39.252*** (0.244)	40.461*** (0.254)	48.278*** (0.228)	46.784*** (1.124)
N respondents	948	948	948	948	948	948

Note: Observations are from the estimation sample of students participating in wave 5 or 7 & wave 9 (see Table 1, column 4). Coefficients shown are obtained from OLS regressions of labour market returns to academic achievement and non-academic investments (as estimated by individual regressions using 12 scenarios per each individual) on demographic and socio-economic characteristics of students. **Symbols:** * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

The employment returns to academic achievements are particularly striking (column 1). Students expect that having an upper-second class, as opposed to a lower second or worse, would increase their probability of being in employment 12 months after graduation by 14.3ppt, and a first class degree by another 7ppt (to 21ppt in total). Having relevant work or extra-curricular experience seems also to matter greatly, increasing employment probabilities by 18 and 9ppt respectively. Across all outcomes, relevant work experience is consistently held to be more important than extra-curricular experience.[10]

Despite similar patterns being observed for all six outcome variables, correlations between expected salaries and both the employment and job attribute probabilities are all positive but relatively low (ρ between 0.06 and 0.15), indicating that students do not expect these outcomes to move together systematically. However, correlations between expected job attribute probabilities are somewhat high, particularly between promotion prospects and working in a field they like (ρ>0.7) suggesting these non-pecuniary outcomes are perceived to be more connected (Appendix Table A4).

We next derive a student-specific measure of these expected returns by estimating coefficients from individual regressions of expected outcomes based on the 12 scenarios. Fig. 3 (right panel) plots the distribution of these returns for the probability of being in full-time work 12 months after graduation, showing that there is heterogeneity in all these coefficients, but that the distribution is more concentrated and closer to zero for extra-curricular experience compared with relevant work experience, and the marginal return of a first class degree compared with an upper second. Table 7 shows regressions of the individual-specific subjective returns to each investment in terms of future employment and earnings on demographic characteristics. Among British students, the only significant difference by ethnicity or SES here is that non-white British tend to believe that there are lower returns to relevant work experience in terms of employment probabilities. Overall, demographic characteristics explain little of the heterogeneity seen in Fig. 3. We obtain similar results when looking at the other labour market outcomes (tables not shown).

5.3. Enjoyment of investments

As well as potentially obtaining a labour market return to their investments, students may experience a utility or disutility in the current period from participating in academic or non-academic activities. To account for this, we asked students to indicate their expected or actual enjoyment in different activities using a zero to 100 scale where "0 means you do or would expect to completely dislike the activity and 100 means you do or would expect to really enjoy the activity" (see Appendix A1 for exact wording).

10 With 12 data points per person, we have the degrees of freedom to assess the perceived complementarity or substitutability of degree class and accumulated experience. In an auxiliary regression for expected probability of being in full-time work, degree class is shown to be complementary to both types of experience, but these experiences substitute for each other. For other labour market outcomes there are few large or significant interaction terms.

Table 7
Heterogeneity in subjective labour market returns to time use and academic achievement.

Return for	Being in full-time work 12 months after graduation (per cent chance)				Salary if in full-time work 12 months after graduation (£000s)			
of	First class	Upper second	Relevant work experience	Extra-curricular experience	First class	Upper second	Relevant work experience	Extra-curricular Experience
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	−2.048***	2.049**	1.979**	1.586***	0.363	−1.082	0.785	0.315
	(0.733)	(0.909)	(0.984)	(0.541)	(1.379)	(1.362)	(0.812)	(0.541)
Demographic group: (Base = White British high-SES)								
White British low-SES	−1.617	1.630	−1.932	−0.073	1.020	−1.463	−0.577	0.298
	(1.107)	(1.372)	(1.485)	(0.817)	(2.081)	(2.055)	(1.226)	(0.817)
Non-white British	0.107	−1.667	−2.549**	−0.301	0.105	−0.128	0.074	−0.166
	(0.928)	(1.150)	(1.244)	(0.685)	(1.744)	(1.722)	(1.027)	(0.684)
EU	−0.065	−2.086	−0.081	−0.199	−0.266	0.862	0.422	−0.082
	(1.075)	(1.332)	(1.442)	(0.793)	(2.021)	(1.996)	(1.190)	(0.793)
Overseas	0.724	−3.395**	−5.649***	−2.706***	3.781	−3.490	0.464	0.235
	(1.237)	(1.533)	(1.659)	(0.913)	(2.326)	(2.297)	(1.370)	(0.913)
N	948	948	948	948	948	948	948	948

Note: Observations taken from main estimation sample of those taking wave 5 or 7, plus wave 9. Exclude 54 cases for whom at least one subjective return parameter across all six job attributes is not identified due to item non-response. Coefficients shown are obtained from OLS regressions of labour market returns to academic achievement and non-academic investments (as estimated by individual regressions using 12 scenarios per each individual) on demographic and socio-economic characteristics of students. Return to first class degree is over having an upper second, and return to Upper second class degree is over having a Lower second class degree or worse. Additional controls are: indicator for ethnicity “other/refused”, indicator for mature student, and 16 department of study dummies. **Symbols:** * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

Table 8
Enjoyment of investment activities by demographic characteristics.

	Study	Classes	Lectures	SU leadership	Competitive sport	Employment	Paid internships	Unpaid internships	Volunteering
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	−1.156	−5.184***	−3.045*	−0.945	−13.128***	−1.958	1.632	−1.166	5.795***
	(1.526)	(1.703)	(1.626)	(2.151)	(2.288)	(1.910)	(1.968)	(2.003)	(1.963)
Demographic group: (Base = White British high-SES)									
White British low-SES	0.764	0.752	2.416	2.317	−3.225	−1.213	4.021	5.605*	4.832
	(2.316)	(2.584)	(2.467)	(3.264)	(3.472)	(2.899)	(2.987)	(3.040)	(2.978)
Non-white British	1.058	−2.515	−4.216**	3.431	−0.477	−6.453***	−1.122	1.671	2.720
	(1.942)	(2.167)	(2.069)	(2.737)	(2.911)	(2.431)	(2.505)	(2.550)	(2.498)
EU	10.729***	9.586***	7.417***	8.400***	1.953	0.144	8.242***	1.622	3.939
	(2.230)	(2.489)	(2.376)	(3.143)	(3.343)	(2.792)	(2.876)	(2.928)	(2.868)
Overseas	7.877***	9.985***	9.651***	7.748**	4.149	1.353	6.778**	1.852	11.020***
	(2.529)	(2.822)	(2.694)	(3.565)	(3.792)	(3.166)	(3.262)	(3.321)	(3.253)
N	962	962	962	962	962	962	962	962	962
Mean response	52.72	53.58	55.26	43.56	48.67	59.52	64.56	39.59	46.90
(Standard deviation)	(21.65)	(24.46)	(23.45)	(30.09)	(32.76)	(26.74)	(27.90)	(28.17)	(27.89)

Note: Observations are from the estimation sample of students participating in wave 5 or 7 & wave 9 (see Table 1, column 4). We exclude 40 additional observations due to item non-response. Coefficients shown are obtained from OLS regressions of enjoyment of different activities on demographic and socio-economic characteristics of students. Additional controls are: indicator for ethnicity “other/refused”, indicator for mature student, and 16 department of study dummies. **Symbols:** * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

Sample means and standard deviations, together with the regression coefficients indicating the difference in the mean enjoyment of these activities by demographic characteristics, other things equal, are shown in Table 8. Non-white British students clearly derive less enjoyment from lectures and (not significantly) classes, but also from working for pay. Low-SES white British report similar enjoyment as high-SES white British for most activities, although they report a higher enjoyment for unpaid internships. Females stand out by reporting lower enjoyment from attendance and participation in competitive sports, while they seem to derive more utility from volunteering.

6. Measurement error and data quality considerations

It is important we consider whether issues of measurement error or – more generally – data quality may confound our interpretation. It would be particularly relevant for our analysis of ethnic and SES differences if measurement error and data quality were to vary significantly along these dimensions.

Table 9
Measurement and response error in conditional expectations.

	Proportion of conditional probabilities at a multiple of 10		Proportion of degenerate probability distributions for degree class	Same conditional probability given for all 12 scenarios		Outliers in expected salary after 12 months		Lower degree class dominates	Expected minus realized final degree mark, conditional on own investments
	Degree class	Job attributes		Work full time	Field they like	Raw data	Cleaned data		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
White British Low SES	0.009	−0.012	−0.004	0.002	−0.025	−0.032	−0.010	0.003	−3.745
	(0.026)	(0.023)	(0.021)	(0.012)	(0.018)	(0.021)	(0.016)	(0.009)	(3.240)
Non-white British	−0.034	−0.038**	0.017	0.000	−0.007	0.011	0.020	0.009	−1.011
	(0.021)	(0.019)	(0.017)	(0.009)	(0.015)	(0.017)	(0.013)	(0.007)	(2.646)
EU	−0.008	−0.016	−0.010	−0.007	−0.030*	−0.004	0.000	−0.004	−0.482
	(0.024)	(0.021)	0.020	(0.011)	(0.017)	(0.020)	(0.015)	(0.008)	(3.272)
Overseas	0.028	−0.013	0.013	−0.005	−0.039**	0.014	−0.009	−0.004	1.213
	(0.027)	(0.024)	(0.022)	(0.012)	(0.019)	(0.022)	(0.017)	(0.009)	(8.968)
Sample mean	0.335	0.257	0.063	0.012	0.031	0.039	0.022	0.007	−1.55
Observations	1002	1002	1002	998	994	963	963	1002	718

Note: Observations are from the estimation sample of students participating in wave 5 or 7 & wave 9 (see Table 1, column 4). We exclude additional observations due to item non-response (columns 4-7) and those who have not yet completed their degree due to a placement or a year abroad (column 9). Coefficients shown are obtained from indicators for specific data features on demographic and socio-economic characteristics of students. No additional covariates. **Symbols:** * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

In order to investigate *measurement error* in self-reported time use in our sample, we take advantage of the fact that we also have elicited self-reported attendance (measured as "what proportion of classes/lectures do you attend") and compare the self-reported data with those obtained from administrative records. The latter are not necessarily error-free, as discussed in Section 3.1, but they are collected in such a way that it is unlikely that errors in measurement are correlated to individual characteristics. Examining the data we see that self-reported attendance is 11ppt higher than documented in the administrative records. However, the correlation between these two measures averaged out over the second and the third year of study is very high at 0.722, and is not significantly different across demographic groups (0.780 for white high-SES, 0.691 for white low-SES and 0.712 for non-white British; see also Appendix Figure A1). Moreover, in a regression similar to that in Table 2, where we use self-reported attendance as the dependent variable, we find that non-white British attend 10.2ppt less than white high-SES British, a difference similar to the 9.3ppt gap observed when using administrative records. The gap between high and low-SES white British is once again both statistically insignificant and quantitatively negligible (0.5ppt, versus 0.8 with administrative records).

As far as the data on expectations is concerned, we look at various indicators of data quality – including rounding, outliers, the propensity to report the same answers to all scenarios, as well as reporting negative returns for higher academic achievement, etc. – and assess how these vary by ethnicity and SES. The main findings are summarized in Table 9 and are discussed below.

We first analyse the issue of *rounding* (columns 1 and 2). All our surveys are administered on-line and we use sliders to elicit subjective expectations of degree classes conditional on time use, and job attributes conditional on degree class and experience, using grid lines at 10 percentage point intervals, at which we might also expect heaping as is common when eliciting probabilistic expectations (e.g., Manski and Molinari, 2010). Indeed, 34% of reported conditional expected degree class probabilities, and 26% of expected job attribute probabilities were multiples of 10 (including zero and 100), versus the 11% we would expect from a uniform distribution. We do not find significant differences in rounding to multiples of 10 by SES status among white British, while non-white British are significantly less likely to round their answers than white high-SES British. The difference is however rather small (less 4 percentage points) and so unlikely to affect our inference.

In column 3 we look at the probability distributions. Assigning a probability of 100% to attaining a range of marks (e.g. a certain degree class) acknowledges no uncertainty in academic performance (outside that range) and results in a *degenerate probability distribution*. It also speeds up the process of answering these questions, as it requires the respondent to move across only one bar instead of three (see Fig. 1, panel A). We look at the proportion of degenerate probabilities reported by an individual across the 12 scenarios in which we vary the time allocations. Interestingly, there are no differences along the ethnic and SES dimension in this respect.[11]

[11] This is also the case conditional on submitting at least one degenerate distribution. About 14.5% of the sample assigns 100% probability to one outcome for at least one of the scenarios.

Although a respondent may genuinely expect *no variation in outcomes across all scenarios*, providing the same probability for all 12 scenarios for either academic or labour market outcomes is also a potential indicator of inattention. It is reassuring that this occurs very rarely; for example only 1.2% of respondents report the same probability of working full-time across all scenarios related to labour market outcomes and 3.1% report the same probability for working in a field they like. There are no significant differences among British students in these response patterns (columns 4 and 5).

Answers to the expected salary questions are sometimes particularly problematic as adding or forgetting a zero might be an easy typing mistake. A small proportion of students (3.9%) input expected salaries with large *outliers*: i.e., a scenario-specific salary more than 10 times larger or smaller than the average of the 12 scenarios in that battery. We inspected these cases individually. Where this clearly reflected either (i) a change in scaling due to answering the first scenario in pounds and the remaining eleven in thousands, or (ii) a typographical error with a trailing zero omitted or added, we re-scale the relevant items to correspond to the order of magnitude of the remaining items. This leaves 2.2% with some 'real' outliers. Again, there is no significant difference by SES or ethnicity in the prevalence of this problem (columns 6 and 7).

It is very unlikely that students believe they would experience worse outcomes on all dimensions with a higher degree class, holding all other investments at the same levels. We investigate how prevalent this is in our data. Importantly, only 0.7% of the sample expects to attain a lower degree class to dominate a higher one, meaning that they expect a *negative return to higher academic achievement* for all outcome variables, and there is no statistical difference by ethnic and social background for this pattern of expectations (column 8).

Students were also asked to report their expected final degree marks conditional on the 12 time allocation scenarios. We take the expected marks reported for the scenario which reflects the student's own contemporaneous investments, and subtract the realized final degree marks to compute the *prediction error* in marks. Students are very accurate on average, with a mean predication error of −1.6 marks out of 100 (though this is not a statistically significant error at any conventional level). We find no significant differences in the prediction error between any demographic groups (column 9). Finally, we consider the *correlation of expectations over time*. Conditional expected salaries, employment probabilities, and job attributes were all collected at both waves 9 and 11. While beliefs may be revised over time, we investigate the stability of answers as a marker of data quality. Within-person correlations of wave 9 and wave 11 answers across the 12 scenarios averaged 0.71 and 0.68 for the 12 month and age 40 salaries, 0.65 for employment probability, and 0.43 to 0.59 for job attribute probabilities. This indicates that these expectations are generally stable over time. There were no large differences in these correlations by ethnicity or SES.

Overall, this investigation gives us confidence in the quality of the expectations data, with no indication of significant differences in reporting by ethnicity and SES background.

7. A discrete choice model of academic and non-academic investments

We now seek to understand the factors driving students' investments. We first discuss how we define the investment allocations and the identification of the model outlined in Section 2, before presenting the estimation results.

7.1. Investment allocations

In our empirical analysis, an investment allocation is defined by: (i) time devoted to lectures and classes, hours of study, and non-academic activities during the academic year; (ii) engagement in non-academic activities outside the academic year, i.e. in the summer between the second and third years; and (iii) the type of non-academic activities a student engages in (i.e. whether they are relevant to the field of study or to desired career or not).

To map as closely as possible information on expectations about outcomes to the actual choices students make, we discretize their time use (see Appendix A2 for details). A choice is therefore a vector of: summer activities (yes/no), attendance (high/low), study hours (high/low), non-academic hours (high/medium/low), relevant work experience (yes/no); extra-curricular experience (yes/no). This leads to 96 different combinations. These are cross-tabulated in Appendix Table A6, with combinations of different levels of summer experience, attendance, study, and non-academic time use shown in 24 rows, and accumulated types of non-academic experience shown in four columns. The percentage of the estimation sample choosing each combination is shown in each cell. Twenty of the cells are empty by construction, meaning students can effectively choose a combination of investments from 76 different options.

Inspecting individual cells, only 5.3% of the sample have made high investments in all dimensions (top left: summer activities, high levels for attendance, study and non-academic activities during the academic year, and both relevant and non-relevant experience), while 4.6% have invested very little (bottom right: no summer activities, and low levels of investments during the academic year). Comparing column totals shows that the most common category for accumulated experience is extra-curricular experience only (54.1% of the sample). Combining information from some of the rows, we see that the most common combination of term-time investments is high attendance, low study hours and low non-academic hours (12.3% of the sample), while 67.5% of the students have engaged in summer activities.

7.2. Identification and empirical specification

7.2.1. Assumptions on the utility function

We estimate the parameters of the utility function described in Section 2 using data on actual investment choices, as well as data on enjoyment and expectations. For tractability, we make the following assumptions:

Taste for academic achievement: We assume that students get utility ρ if they obtain at a good degree, i.e. a final average greater than 60. This is because having a good degree is extremely salient in the UK context.

Individual-specific discount factor We use an individual-specific discount factor $\beta_i = \frac{1}{1+r_i}$, where r_i is the Annual Effective Rate of interest at which respondent i becomes prepared to wait 5 months for a higher cash payment than a fixed amount after 1 month in an incentivized task as part of wave 5 of BOOST2018 (see Appendix A2 for details). The median discount factor for our estimation sample is 0.824. Using an individual-specific discount factor is important as differences in investment allocation by group could be in part due to differences in time preferences. In our data, non-white British students tend to discount the future less than their white counterparts (see Appendix Table A5).

Expected earnings: We assume that all students have the same period zero consumption (see discussion in Section 8 about the availability of maintenance loans and grants). We also assume students' working life takes place between ages 22 and 60. Students were asked to provide their employment probability and expected earnings conditional on degree classification and accumulated experience 12 months after graduation and at age 40. We use these data to compute year-specific earnings and employment probabilities (see Appendix A2 for details). We normalize unemployment earnings y_{ut} to 1 and use the year-specific log earnings and probability of employment together with the individual-specific discount factor β_i to compute the expected lifetime log earnings conditional on degree classification and non-academic experience as follows:

$$lnY_{d,I_{NA}} = \sum_{t=0}^{T} {\beta_i}^t \left[P_i(job_t \,|\, D = d, I_{NA}) lnE(y_{it}|D = d, I_{NA})\right].$$

Our empirical results are robust to other plausible assumptions, or to simply using expected log earnings at age 40. Note that this derivation implies that there is no uncertainty in earnings at time t conditional on degree classification and accumulated experience, that is, the student knows with certainty her conditional earnings at time t. This is a potential limitation; however the model embeds uncertainty in earnings through the uncertainty of being employed and uncertainty of academic achievement.

Job attributes: As seen in the previous section, students' reported probabilities of their job being in a field they like, having good promotion prospects and being useful to society (all conditional on degree class, accumulated experience and being in full-time work) are highly correlated. We therefore use only the job attribute "the job is in a field you like" in our empirical analysis, as this is the attribute that student reported elsewhere to be the most important for what constitutes a good job.[12] Students are assumed to enjoy utility α if they have a job in a field they like.

Enjoyment of investments: We incorporate choice-specific utility to capture the enjoyment net of effort associated with the different time use component (attendance, study and non-academic hours) using the following functional form: $u_i\,(I_C) = \gamma_I H_{I_C} E_I$, where E_I is the elicited enjoyment (re-scaled from 0 to 1) associated with engaging in activity I, H_{I_C} is the number of hours of activity I in choice C and γ_I is an activity-specific preference parameter to be estimated that translates the enjoyment measure into util. This specification requires measures of enjoyment for (i) attendance, (ii) study, (iii) term-time non-academic activities, and (iv) summer activities. We use the enjoyment data described in Section 5.3 for this purpose (see Appendix A2 for details).[13] Note that the mapping between the true latent individual preferences and the reported enjoyment scale is a priori individual-specific. For tractability, we first assume that the mapping is homogeneous in the sample and then relax this assumption by allowing γ_I to vary by social and ethnic background.

Choice-specific dummies: We include a choice-specific dummy μ_C to capture unobserved choice-specific factor enjoyed by the students.

Under those assumptions, we can rewrite student's i lifetime expected utility associated with the investment choice C as follows:

$$\begin{aligned} U_i\,(C) = & \sum_{a=\{At,St,NA,Su\}} \gamma_a H_{a_C} E_a + \rho P_i(good\_degree|C) + \theta\,\ln(y_{i0}) \\ & + \theta \sum_{d=1}^{3} P_i(D = d|C) lnY_{d,C} + \alpha \beta_i P_i\,(J|C) + \mu_C + \varepsilon_{iC}, \end{aligned} \tag{3}$$

[12] At wave 7, 76% of N = 978 respondents in the main estimation sample rated their job being in a field they like as "Very" or "Extremely" important, compared with 57% for it having promotion prospects and 35% for it being useful to society/others.

[13] Note that the enjoyment of employment may include the utility derived from the extra earnings as those are not explicitly embedded in the time $t = 0$ consumption. For example, students report on average a higher enjoyment for paid internship compared to unpaid internship, suggesting that the reported enjoyment also take pecuniary factors into consideration.

where $\{At, St, NA, Su\}$ denotes {Attendance, Study, Non-Academic activities, Summer activities}, H_{a_C} is the hours spent on activity a in choice C (set to 1 for summer activities), E_a is the elicited enjoyment of activity a, $P_i(good\_degree|C)$ is the probability of having a good degree conditional on choice C, $P_i(D = d|C)$ is the probability of degree class d conditional on choice C, Y_C is the expected log life earning conditional on degree class d and choice C and $P_i\,(J|C)$ is the probability that the student works in a field she likes conditional on choice C.

7.2.2. Identification and sufficient set logit estimation

Under the assumption that the random terms e_{iC} are independent for every individual i and choice C and that they have a Type I extreme value distribution, the difference $(\varepsilon_{iC} - \varepsilon_{iK})\,, C \neq K$, is distributed as a logistic. Conditional on the students' expectations and enjoyment for each choice in her choice set S_i, the probability Π_{iC} that student i chooses C is therefore given by:

$$\Pi_{iC} = P\left(V_{iC} + \varepsilon_{iC} > V_{iK} + \varepsilon_{iK}, \text{all}C, K \in S_i, C \neq K\right) = \frac{\exp(V_{iC})}{\sum_{K \in S_i} \exp(V_{iK})}, \tag{4}$$

where V_{iC} is the expected utility maximized in Eq. (3), net of ε_{iC}. Eq. (4) is a multinomial logistic regression that can be estimated by maximum likelihood. The preference parameters of interest, $(\gamma, \rho, \theta, \alpha)$, are identified up to scale from the variation in expectations and enjoyment across individuals and choices.[14]

Eq. (4) assumes that we have information on the individual-specific choice set S_i. As mentioned in Section 7.1, the largest possible choice set comprises 76 options. However, some students may be constrained and unable to choose some of these options. For example, they may apply for an internship or volunteering but not secure one. This is problematic, as including alternatives that are not in a student's true choice set typically leads to a violation of the Independence of Irrelevant Alternative (IIA) assumption and to inconsistent estimators (McFadden, 1978). The likely inconsistency is greater if the decision-maker would have preferred an alternative mistakenly included in the choice set.

To deal with this issue, we follow McFadden (1978) who show that estimating a multinomial logit model using a *true* subset of the unobserved choice set yields consistent estimators (see also the recent survey by Crawford et al., 2020). Intuitively, one can estimate the preference parameters based only on the variation in the characteristics of alternatives that are in the subset, rather than on the full (unobserved) choice set. Crawford et al. (2020) call these subsets *sufficient sets* and a multinomial logit model estimated using these sufficient sets a *sufficient set logit*.

Crawford et al. (2020) also provide some guidance on how to select a sufficient choice set. In a cross-sectional setting, the sufficient choice set of an individual may be defined as the choice she made together with the choices made by individuals of the *same type*. In our application, we assume that the *sufficient choice set* of student i is the set of choices available to individuals of the same gender, ethnic, SES, nationality, and department.[15] Note that white high-SES students have a larger sufficient choice set on average with this definition than white low-SES students and non-white students (12 alternatives versus 9.1 and 9.0 respectively for the other groups), which could capture the fact they may be facing fewer constraints than their counterparts or that they make less homogeneous choices.

In addition, we take advantage of the fact that we have individual-specific information on internship applications and outcomes.[16] For all respondents who applied to a work relevant placement but did not secure one, we exclude any choices that include work relevant experience from their sufficient choice set. This is because we think they would have liked to choose an alternative with relevant work experience (since they have applied), but they have been constrained by labour demand rationing and unable to do so. The resulting sufficient choice set, SS_i, is what we use for our estimation. We therefore estimate equation (4) but use SS_i as individual i's choice set, instead of S_i. With this new restriction, the average size of the sufficient choices set comprises 11.4 alternatives for white high-SES students, 9 for white low-SES students and 8.5 for non-white students.

7.3. Baseline model estimates

Column 1 of Table 10 presents the multinomial logit estimates. Our analysis is restricted to the sample of students who reply to "wave 5 or 7 & wave 9 & wave 12" of the survey. This is because it is only in wave 12 that information on applications for relevant work experience is collected. As shown in Table 1, the profile of this sample only differs in minor ways from our main estimation sample.

In terms of future outcomes, there is a positive and precisely estimated coefficient associated with expected lifetime log earnings and the probability of having a good degree, suggesting that both are relevant in affecting students' investment

14 Note that we use expectations elicited in the third year of university to explain choices made over the second and third years. One concern might be that students are learning or that they report beliefs that rationalize their choices. Previous research in the context of educational choices of US students has found little evidence of students shifting their beliefs about expected outcomes in favour of the options they had chosen (Zafar, 2011; Arcidiacono et al., 2012). Moreover, the estimated parameters are comparable if we use year 3 investments instead of investments made over the last two years at university, suggesting that the effect of (potential) learning is minimal.

15 We combine very small departments within a larger department in a related field.

16 Respondents report whether they having applied to an internship at wave 12. Failure is equal to one if a respondent has applied to an internship at least once but has not secured relevant work experience.

Table 10
Baseline model estimates and Willingness to Pay.

	Coefficient estimates	WTP for 1 percentage point increase in probability or enjoyment
	(1)	(2)
Expected log life earnings	0.018***	
	(0.007)	
Prob of good degree	0.564*	0.311
	(0.304)	(0.219)
Prob of job in field student likes	0.090	0.050
	(0.740)	(0.414)
Utility of attendance	0.472***	0.260**
	(0.109)	(0.113)
Utility of study	0.456***	0.252**
	(0.080)	(0.101)
Utility of non-academic activities	0.047**	0.026*
	(0.023)	(0.016)
Utility of summer activity	0.231	0.127
	(0.398)	(0.225)
Number of observations	6639	
Number of cases	693	

Note: Observations are from the estimation sample of students participating in wave 5 or 7 & wave 9 & wave 12 (see Table 1, column 5). Model specification and estimation described in Section 7. The model contains 59 choice-specific constant terms. **Symbols:** * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

decisions. The probability that a job will be in a field they like has a coefficient much smaller in magnitude than the probability of having a good degree, and is imprecisely estimated. This suggests that the non-pecuniary attributes of a job have only a limited impact on students' investments at university.[17] Outside of the summer activities, the coefficients associated with the enjoyment variables are all positive and precisely estimated, confirming that students are more likely to engage in activities they enjoy. The coefficients associated with attendance enjoyment and study enjoyment are very similar, and ten times larger than the non-academic experience enjoyment. Overall, this suggests that students take into consideration academic achievement, future earnings, as well as enjoyment of academic activities when making their investment decisions.

To gain insight into the magnitude of the estimated parameters, the second column of Table 10 shows the willingness to pay (WTP) implied by the model estimates. Let w denote the percentage of lifetime earnings that would make the student indifferent between having a good degree with probability P_2 instead of probability P_1, other things being equal. Based on utility specification in Eq. (2),

$$\rho P_1 + \theta \ln(Y) = \rho P_2 + \theta \ln((1+w)Y)$$

The WTP, w, calculated as a percentage of lifetime earnings, is then $\exp\left(\rho \frac{P_1 - P_2}{\theta}\right) - 1$. Increasing the probability of having a good degree by 1ppt, that is $P_2 = P_1 + 0.01$, yields a WTP of 0.311. That is, students are on average willing to give up 31% of their discounted lifetime earnings to increase their probability of a good degree by 1ppt. This estimate is very large and implies that students gain significant utility from doing well academically. Similarly large WTPs have been found for university-related non-pecuniary outcomes in other contexts such as parental approval or university ideology (Delavande and Zafar, 2019).

Similarly, let ω denote the percentage of lifetime earnings that would make the student indifferent between enjoying 1 h of attendance with enjoyment E_{A1} instead of enjoyment E_{A2}, other things being equal. The WTP, ω, calculated as a percentage of lifetime earnings, is then $\exp\left(\gamma_A \frac{E_{A1} - E_{A2}}{\theta}\right) - 1$. Table 10 shows that a decrease of 1ppt in their enjoyment for attendance or study leads to a similar utility loss as giving up about 26% of their lifetime earnings. The magnitude is 2.6% of lifetime earnings for non-academic activities. Again, these estimates are very large and suggest that the net utility of engaging in the activities is a very important factor in the decision to invest at university.

It is plausible that the expected returns of each investment allocation are correlated with unobserved preferences for the various activities (Wiswall and Zafar, 2015). We expect our enjoyment measures and investment allocation fixed-effect to deal with this in part, but acknowledge that some aspects of enjoyment (such as preferences complementarity across various activity types) are not fully captured by our measures. We investigate how sensitive the coefficients associated with expected returns are to the inclusion of the enjoyment measures (Appendix Table A7) and find that the estimated preferences for future earnings and for academic achievement are remarkably stable. If unobserved preferences

[17] Including the other expected non-pecuniary job attributes inflates the coefficient on "a field they like" and results in negative and imprecisely estimated coefficients on "promotion" and "job being useful for society".

Table 11
Model estimates with heterogeneous preferences.

	Estimated coefficient on attributes from single estimated model allowing for heterogeneous coefficients						
	Expected log life earnings	Prob of good degree	Prob of job in field student likes	Utility of attendance	Utility of study	Utility of non-academic activities	Utility of summer activity
For each demographic group:							
White high-SES	0.022	1.418**	−0.890	0.378***	0.447***	0.037	0.384
(N = 168)	(0.017)	(0.586)	(1.578)	(0.133)	(0.103)	(0.030)	(0.509)
White low-SES	0.025	0.484	−1.795	0.428***	0.561***	0.029	0.819
(N = 90)	(0.018)	(0.889)	(1.990)	(0.155)	(0.117)	(0.036)	(0.587)
Non-white	0.009	0.162	0.727	0.405***	0.426***	0.044	0.182
(N = 213)	(0.010)	(0.502)	(1.178)	(0.133)	(0.097)	(0.029)	(0.477)
Other	0.033**	0.568	−0.084	0.617***+	0.436***	0.068**	−0.034
(N = 222)	(0.013)	(0.597)	(1.268)	(0.134)	(0.089)	(0.027)	(0.455)
Number of observations				6639			
Number of respondents				693			

Note: Observations are from the estimation sample of students participating in wave 5 or 7 & wave 9 & wave 12 (see Table 1, column 5). All parameters in the table are estimated from a single model allowing for heterogeneous coefficients (see Section 7). **Symbols:** * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$; + $p < 0.1$, ++ $p < 0.05$, +++ $p < 0.01$ represent statistically significance differences from the white British high-SES group coefficients.

for activities and our enjoyment measures are correlated (which is plausible), this stability suggests that these two coefficients are unlikely to be biased due to unobserved preferences. The coefficient measuring the preference for working in a field they like, which is imprecisely estimated throughout, decreases in magnitude once we control for the enjoyment of non-academic activities. The preference parameters for non-pecuniary job amenities may therefore be upward biased if observed and unobserved preferences for activities are positively correlated.

7.4. Heterogeneous preference parameters

The model estimated in the previous section assumed that students from all backgrounds had similar preference parameters for the choices' attributes. We now relax this assumption and allow the preference parameters to be different for white high-SES, white low-SES, non-white and other ethnic groups by using interactions. The results are presented in Table 11. The estimated structural parameters for non-white are not statistically different at conventional levels from those of white high-SES for all choices' attributes, with one important exception: the value attached to academic achievement. The coefficient associated with the probability of having a good degree for non-white is almost ten times smaller than that of white high-SES. The fact that non-white students value academic achievement much less is one of the factors responsible for their lower investments.

It is not entirely clear why non-white students have different preferences for academic achievement. We could speculate that going to university is their main target and academic results matter less if they were predominantly the first generation in their family to go to university, but they are as likely as white high-SES to have a parent with a university degree. Alternatively, it is possible that caring for academic achievement and behaving accordingly on campus is a sign of "acting white" (Austen-Smith and Fryer, 2005; Fryer and Torelli, 2010) or of being "uncool" (Bursztyn et al., 2019) which might be frowned upon among their peers who tend to be homogeneous according to race (e.g., Mayer and Puller, 2008).

8. Understanding the gap in investments

In this section, we use the estimated preference parameters from Table 11 to gain a better understanding about the determinants of gaps in investments between white high-SES and non-white British students, the two groups for whom investment allocations are most different. While the sufficient set logit allows us to obtain point estimates of the preference parameters, we cannot point-identify individual-specific choice probabilities as we do not observe the individual-specific choice set. However, we can provide bounds of the choice probabilities. We first illustrate the role of constraints by showing a lower bound for the choice probabilities derived under the assumption that students can access any alternative and then present other counterfactuals using this unrestricted choice set. All our simulations hence assume no constraints.

Fig. 4 first displays the actual gap in choosing high investments between non-white British and high-SES white British ("Actual choice"). Next, we impose the assumption that students face no constraints (S1: "Unrestricted choices").[18] The

[18] We exclude allocations that no students have chosen as we do not have a choice-specific constant for them. This implies a choice set of 60 alternatives. It is important for this analysis that the model produces accurate predictions of the heterogeneity in investments by social and ethnic groups. Appendix Table A8 presents the actual proportion of students engaging in high investments, and the corresponding predicted probabilities derived using the sufficient choice set, and shows a very good fit indeed.

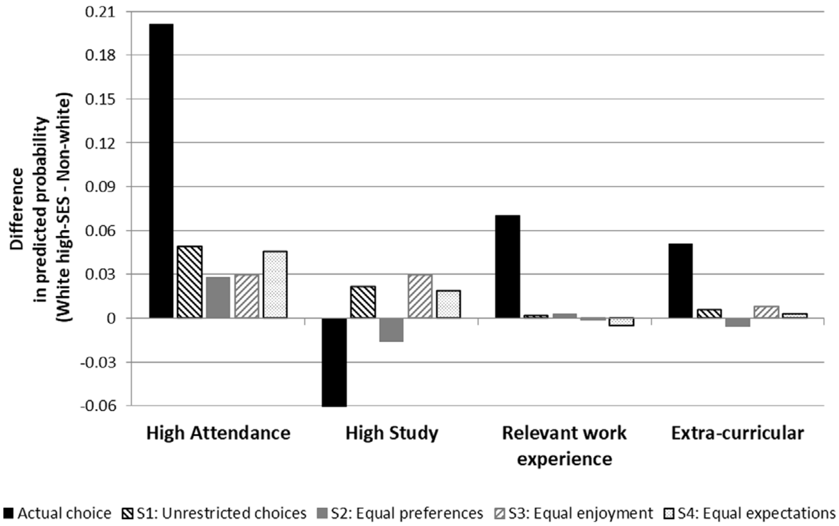

Fig. 4. Simulating the effect of counterfactual scenarios on gaps in investments. **Note:** N = 168 white high-SES, 213 non-white British. Each bar represents the difference in the predicted probability of choosing the indicated level or type of investment under different scenarios. Specification of the counterfactual scenarios as detailed in Section 8. S2, S3 and S4 all assume unrestricted choices.

difference between the actual gap and the predicted gap with no constraints highlights that constraints play a very important role. With the exception of high study (where the actual gap is negative, and unconstrained non-white students would study less than in the constrained case), we see that relaxing the constraints reduces the gap by a factor of 4 for high attendance, and essentially eliminates it for work experience and extra-curricular activities.

We then consider the allocation resulting from the assumption that non-white British students have the same structural preference parameters as white high-SES (S2: "Equal preferences"). This further reduces the gap in attendance by a factor of 2 compared to the previous scenario (S1), and study is moved marginally back in favour of non-white students. The gap in work experience is unchanged at close to zero. Allocating to non-white students the average enjoyment for different activities of white high-SES students (S3: "Equal enjoyment") yields very similar results as in scenario S2. Finally, allocating to non-white students the average expectations about academic and non-academic outcomes of white high-SES (S4: "Equal expectations") slightly widens and closes the attendance and study gaps respectively, though is more effective at reducing the work experience gap than the other two simulations. This exercise clearly shows that the largest reduction in the ethnic gap is achieved by relaxing the constraints that students face, while other factors – such as preferences, enjoyment of activities and beliefs – play a minor role.

A detailed analysis of the constraints students face in making their investment decisions is beyond the scope of this paper, but we take advantage of the richness of our data to explore this a bit further. Table 12 shows the conditional difference by social and ethnic group in variables which (imperfectly) proxy for various constraints. Column 1 investigates financial constraints and shows that self-reported parental annual income is significantly lower for both white British low-SES and non-white British than white British high-SES, by £27,000 and £21,000 on average. However, all British (and EU) students have access to government-backed loans that (i) pay tuition directly to the university and (ii) provide maintenance support directly to students. The latter are means-tested on a sliding scale by parental income, which must be declared to access any maintenance support, and (for this cohort) those from lower-income households were also entitled to some level of non-repayable maintenance grant (up to £3400 per year). Moreover, there is a single application process for both the tuition and maintenance components of financial support, payments are made automatically after the initial application, and repayments of the loan component are only required after graduation on an income-contingent basis. This system of financial support should go a long way to reduce differences in financial constraints across the population of British students. In particular, as we saw earlier, there are no significant differences in the propensity to work for pay between white high-SES, white low-SES, and non-white British students (see Appendix Table A3).

Column 2 focuses on labour demand constraints by looking at the probability to fail to secure relevant work experience. This variable is defined only for those who report having applied to an internship at least once, which is the case for 48% of white high-SES, 38% of white low-SES, and 55% of non-white British.[19] Failure is equal to one if a respondent has applied to an internship at least once but has not secured relevant work experience. We see that non-white British students are 13.8 percentage points more likely than white British high-SES students to fail to obtain relevant work experience, conditional on having applied for it. This may reflect labour market discrimination, or differences in accumulation of or

[19] Plus 76% of EU and 57% for Overseas students.

Table 12
Analysis of constraints by demographic characteristics.

	Combined parental income (£000s)	Failed an application for relevant work experience	Average commuting time (min)	At least one parent known to have a degree
	(1)	(2)	(3)	(4)
Female	−5.976***	0.004	2.499**	−0.038
	(1.793)	(0.056)	(1.116)	(0.033)
Demographic group: (Base = White British high-SES)				
White British low-SES	−26.821***	0.054	0.663	−0.326***
	(2.737)	(0.099)	(1.696)	(0.050)
Non-white British	−20.854***	0.138*	0.631	0.020
	(2.283)	(0.071)	(1.618)	(0.042)
EU	−32.894***	−0.066	−0.391	0.234***
	(2.630)	(0.076)	(1.618)	(0.048)
Overseas	−9.888***	−0.028	−8.546***	0.214***
	(2.972)	(0.091)	(1.859)	(0.054)
N	1002	423	908	1002
Mean response (standard deviation)	38.752 (28.431)	0.532 (0.500)	19.126 (15.696)	0.521 (0.500)

Note: Observations are from the estimation sample of students participating in wave 5 or 7 & wave 9 (see Table 1, column 4). Coefficients shown are from different OLS regressions of the dependent variable on demographic and socio-economic characteristics of students. Additional controls are: indicator for ethnicity "other/refused", indicator for mature student, and 16 department of study dummies. **Symbols:** * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

ability to signal required skills to obtain such positions, or differences in the quality of potential matches yielded through different groups' job search methods.[20]

The third column addresses time constraints. Here we see no significant differences by ethnicity or SES in commuting time from students' residence to campus.[21] Having at least one parent with a university degree may provide information about what it takes to succeed at university. Column 4 shows that low-SES white British students are significantly less likely than high-SES white British students to have at least one parent with a degree (32.6ppt, from an overall sample mean of 52%), but non-white British students are no different from the high-SES white British group in this dimension.

Overall, we think that labour demand constraints, combined with the fact that students believe work experience has large return on the labour market, explain most of the ethnic gap in the accumulation of work and extra-curricular experience. By contrast, our analysis is not able to identify what could explain the existence of constraints on the choice of academic investments, especially in terms of study hours. We speculate that social norms or peer pressure might reduce attendance among non-white British students, and indeed we find that in general they engage more in private study time. The latter may be unobserved by peers or more socially acceptable. This is also consistent with the interpretation we provided for the lower preference non-white students exhibit for academic achievement.

9. Conclusion

An important focus of many higher education policies is to improve access, academic success and labour market outcomes for students from minority groups and disadvantaged family backgrounds. Investments undertaken by students at university are likely to play a major role in determining their academic achievement, employability, occupation and earnings. Using new data from a recent survey of UK university students, we document significant heterogeneity in student's investment decisions along mainly the ethnic dimension, with non-white British students attending lectures and classes significantly less than white high-SES British students, and investing less time accumulating relevant work experience. By contrast, no significant differences in investments are found between low-SES and high-SES white British students.

We use newly collected data on students' expected academic and labour market returns, as well as their non-pecuniary value, together with student's actual investment choices to estimate a structural model of time allocation. The estimates reveal that there are three main drivers of students' time allocation decisions at university: expectations about future earnings, academic achievement, and enjoyment of classes and lectures. Non-pecuniary aspects of future jobs and enjoyment of non-academic activities play a less important role by contrast.

While expectations about future outcomes and enjoyment about the investment matter in informing students' decision-making process, we find that the largest reduction in the ethnic investment gap would be achieved by relaxing

[20] White high-SES students are, for example, 7.6 percentage points more likely than non-white British students to have used "parents and/or relatives" as a source of information about employment opportunities at least once in the past year at wave 9. Non-white British use a range of university-based sources, including the university's careers service, job fairs, lectures and the alumni network significantly more frequently.

[21] Note that almost all Overseas students reside on campus, leading to a much lower average commuting time for this group.

the constraints that students seem to face, as the predicted investments from our model exhibit a smaller ethnic gap under the assumptions that all students can choose from all time allocations. This suggests that offering work experience to all and relaxing the (potential) constraints created by social norms may lead to more balanced investments. We also find that changing expectations or enjoyment of activities of non-white students reduces the gap somewhat. It is however not clear whether some policies can or should be designed to achieve these changes. Non-white British students may have different expectations than white high-SES British students because they have a truly different production function (for example, they may face a glass ceiling in the labour market and as a result experience lower returns on investments). In this case, trying to change their expectations might not be desirable. Instead it might be better to try to increase their enjoyment of lectures and classes. One possibility might be to augment the ethnic diversity of lecturers/teaching assistants, as there is evidence that minority students benefit academically from having same-race teachers at school and university (e.g., Dee, 2004; Fairlie et al., 2014), with some suggestive indication that this effect may be driven in part by increased attendance in the university context (Lusher et al., 2018). By acting as role models, these ethnically diverse lecturers could also increase the value that non-white British students attach to academic achievement.

From a methodological point of view, our analysis highlights that the difficulty in observing individual-specific constraints is particularly relevant to the growing literature estimating discrete choice models using expectations data. Constrained individuals may behave as if expectations are irrelevant to their choice, impairing the inference researchers can make about the role of expectations in decision-making under uncertainty. The solution adopted in this paper, i.e. the sufficient set logit method (Crawford et al., 2020), is one possible way of dealing with this issue. However, to define the sufficient set it is important that researchers collect more information about the different choice environments decision-makers face when making their decisions. This should be an important area of research in the future.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeconom.2020.03.019.

References

Arcidiacono, P., 2004. Ability sorting and the returns to college major. J. Econometrics 121, 343–375. http://dx.doi.org/10.1016/j.jeconom.2003.10.010.

Arcidiacono, P., Aucejo, E.M., Spenner, K., 2012. What happens after enrollment? An analysis of the time path of racial differences in GPA and major choice. IZA J. Labor Econ. 1, http://dx.doi.org/10.1186/2193-8997-1-5.

Arcidiacono, P., Hotz, V.J., Maurel, A., Romano, T., 2019. Ex Ante Returns and Occupational Choice. mimeo.

Arcidiacono, P., Koedel, C., 2014. Race and college success: Evidence from missouri. Amer. Econ. J.: Appl. Econ. 6, 20–57. http://dx.doi.org/10.1257/app.6.3.20.

Association of Graduate Recruiters, 2016. The AGR 2016 Annual Survey. The Association of Graduate Recruiters, London.

Attanasio, O.P., Kaufmann, K.M., 2014. Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender. J. Dev. Econ. 109, 203–216. http://dx.doi.org/10.1016/j.jdeveco.2014.04.003.

Austen-Smith, D., Fryer, R.G., 2005. An economic analysis of acting white. Q. J. Econ. 120, 551–583. http://dx.doi.org/10.1093/qje/120.2.551.

Bailey, M.J., Dynarski, S.M., 2011. Gains and Gaps: Changing Inequality in U.S. College Entry and Completion. NBER Working Papers 17633, National Bureau of Economic Research, Inc, http://dx.doi.org/10.3386/w17633.

Beffy, M., Fougère, D., Maurel, A., 2012. Choosing the field of study in postsecondary education: Do expected earnings matter? Rev. Econ. Stat. 94, 334–347. http://dx.doi.org/10.1162/REST_a_00212.

Ben-Akiva, M., Boccara, B., 1995. Discrete choice models with latent choice sets. Int. J. Res. Mark. 12, 9–24. http://dx.doi.org/10.1016/0167-8116(95)00002-J.

Berger, M.C., 1988. Predicted future earnings and choice of college major. Ind. Labor Relat. Rev. 41, 418–429. http://dx.doi.org/10.2307/2523907.

Betts, J.R., 1996. What do students know about wages? Evidence from a survey of undergraduates. J. Hum. Resour. 31, 27–56. http://dx.doi.org/10.2307/146042.

Blass, A.A., Lach, S., Manski, C.F., 2010. Using elicited choice probabilities to estimate random utility models: Preferences for electricity reliability. Internat. Econom. Rev. 51, 421–440. http://dx.doi.org/10.1111/j.1468-2354.2010.00586.x.

Boneva, T., Rauh, C., 2017. Socio-Economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns. cesifo Working Papers 6756.

Bound, J., Lovenheim, M.F., Turner, S., 2010. Why have college completion rates declined? an analysis of changing student preparation and collegiate resources. Amer. Econ. J.: Appl. Econ. 2, 129–157. http://dx.doi.org/10.1257/app.2.3.129.

Bursztyn, L., Egorov, G., Jensen, R., 2019. Cool to be smart or smart to be cool? Understanding peer pressure in education. Rev. Econom. Stud. 86 (4), 1487–1526. http://dx.doi.org/10.1093/restud/rdy026.

Connell-Smith, A., Hubble, A., 2018. Widening Participation Strategy in Higher Education in England. House of Commons Briefing papers CBP-8204.

Crawford, C., 2014. Socio-Economic Differences in University Outcomes in the UK: Drop-Out, Degree Completion and Degree Class. IFS Working Paper W14/31, http://dx.doi.org/10.1920/wp.ifs.2014.1431.

Crawford, G.S., Griffith, R., Iaria, A., 2020. A survey of preference estimation with unobserved choice set heterogeneity. J. Econometrics http://dx.doi.org/10.1016/j.jeconom.2020.07.024, Forthcoming. Available at https://www.dropbox.com/s/jtb8mzjmnxxvb6q/consideration_sets.pdf?dl=0 (accessed 30th January 2020).

Dee, T.S., 2004. Teachers, race, and student achievement in a randomized experiment. Rev. Econ. Stat. 86, 195–210. http://dx.doi.org/10.1162/003465304323023750.

Delaney, L., Harmon, C., Ryan, M., 2013. The role of noncognitive traits in undergraduate study behaviours. Econ. Educ. Rev. 32, 181–195. http://dx.doi.org/10.1016/j.econedurev.2012.07.009.

Delavande, A., 2008. Pill, patch or shot? Subjective expectations and birth control choice. Internat. Econom. Rev. 49, 999–1042. http://dx.doi.org/10.1111/j.1468-2354.2008.00504.x.

Delavande, A., Zafar, B., 2019. University choice: The role of expected earnings, non-pecuniary outcomes, and financial constraints. J. Political Econ. 127 (5), 2343–2393. http://dx.doi.org/10.1086/701808.

Department for Education, 2011. LSYPE user guide to the datasets: Wave 1 to wave 7.

Dolton, P., Marcenaro, O.D., Navarro, L., 2003. The effective use of student time: A stochastic frontier production function case study. Econ. Educ. Rev. 22, 547–560. http://dx.doi.org/10.1016/S0272-7757(03)00027-X.

Fairlie, R.W., Hoffmann, F., Oreopoulos, P., 2014. A community college instructor like me: Race and ethnicity interactions in the classroom. Amer. Econ. Rev. 104, 2567–2591. http://dx.doi.org/10.1257/aer.104.8.2567.

Fryer, R.G., Torelli, P., 2010. An empirical analysis of acting white. J. Public Econ. 94, 380–396. http://dx.doi.org/10.1016/j.jpubeco.2009.10.011.

Gaskell, S., Lingwood, R., 2019. Widening Opportunity in Higher Education - The Third Phase: Beyond Graduation. Universities UK Report.

Giustinelli, P., 2016. Group decision making with uncertain outcomes: Unpacking child-parent choice of the high school track. Internat. Econom. Rev. 57, 573–602. http://dx.doi.org/10.1111/iere.12168.

Goeree, M.S., 2008. Limited information and advertising in the U.S. personal computer industry. Econometrica 76, 1017–1074. http://dx.doi.org/10.3982/ecta4158.

Grave, B.S., 2011. The effect of student time allocation on academic achievement. Educ. Econ. 19, 291–310. http://dx.doi.org/10.1080/09645292.2011.585794.

Jensen, R., 2010. The (perceived) returns to education and the demand for schooling. Q. J. Econ. 125, 515–548. http://dx.doi.org/10.1162/qjec.2010.125.2.515.

Lechner, M., Downward, P., 2017. Heterogeneous sports participation and labour market outcomes in England. Appl. Econ. 49, 335–348. http://dx.doi.org/10.1080/00036846.2016.1197369.

Lochner, L., Monge-Naranjo, A., 2012. Credit constraints in education. Annu. Rev. Econ. 4, 225–256. http://dx.doi.org/10.1146/annurev-economics-080511-110920.

Louviere, J.J., Hensher, D.A., Swait, J/D., 2000. Stated Choice Methods: Analysis and Applications. Cambridge University Press.

Lusher, L., Campbell, D., Carrell, S., 2018. TAs like me: Racial interactions between graduate teaching assistants and undergraduates. J. Public Econ. 159, 203–224. http://dx.doi.org/10.1016/j.jpubeco.2018.02.005.

Lynn, P., Borkowska, M., 2018. Some Indicators of Sample Representativeness and Attrition Bias for BHPS and Understanding Society 19. Understanding Society Working Paper Series No. 2018-01.

Manski, C.F., 1977. The structure of random utility models. Theory and Decision 8, 229–254. http://dx.doi.org/10.1007/BF00133443.

Manski, C.F., 1993. Identification of endogenous social effects: The reflection problem. Rev. Econom. Stud. 60, 531–542. http://dx.doi.org/10.2307/2298123.

Manski, C.F., 2004. Measuring expectations. Econometrica 72, 1329–1376. http://dx.doi.org/10.1111/j.1468-0262.2004.00537.x.

Manski, C.F., Molinari, F., 2010. Rounding probabilistic expectations in surveys. J. Bus. Econom. Statist. 28 (2), 219–231.

Mayer, A., Puller, S.L., 2008. The old boy (and girl) network: Social network formation on university campuses. J. Public Econ. 92, 329–347. http://dx.doi.org/10.1016/j.jpubeco.2007.09.001.

McFadden, D., 1978. Modeling the choice of residential location. In: Karlqvist, A., Lundqvist, L., Snickars, F., Weibull, J. (Eds.), Spatial Interaction Theory and Planning Models. North-Holland, Amsterdam, pp. 75–96.

Meschede, T., Taylor, J., Mann, A., Shapiro, T., 2017. Family achievements?: How a college degree accumulates wealth for whites and not for blacks. In: Federal Reserve Bank of St. Louis Review, Vol. 99. pp. 121–137. http://dx.doi.org/10.20955/r.2017.121-137.

Persico, N., Postlewaite, A., Silverman, D., 2004. The effect of adolescent experience on labor market outcomes : The case of height. J. Political Econ. 112, 1019–1053. http://dx.doi.org/10.1086/422566.

Prospects, 2019. Skills shortages in the UK 2019/20: An insight into occupational shortages in the UK labour market. downloadable from: https://www.prospects.ac.uk/careers-advice/getting-a-job/skills-shortages-in-the-graduate-job-market.

Romer, D., 1993. Do students go to class? Should they? J. Econ. Perspect. 7, 167–174. http://dx.doi.org/10.1257/jep.7.3.167.

Saniter, N., Siedler, T., 2014. Door Opener Or Waste of Time? The Effects of Student Internships on Labor Market Outcomes. IZA Discussion Paper No: 8141.

Schmidt, R.M., 1983. Who maximizes what? A study in student time allocation. Amer. Econ. Rev. 73, 23–28.

Smith, J., Naylor, R., 2001a. Determinants of degree performance in UK universities: A statistical analysis of the 1993 student cohort. Oxf. Bull. Econ. Stat. 63, 29–60. http://dx.doi.org/10.1111/1468-0084.00208.

Smith, J.P., Naylor, R.A., 2001b. Dropping out of university: A statistical analysis of the probability of withdrawal for UK university students. J. R. Stat. Soc. Ser. A 164, 389–405. http://dx.doi.org/10.1111/1467-985X.00209.

Stinebrickner, R., Stinebrickner, T.R., 2004. Time-use and college outcomes. J. Econometrics 121, 243–269. http://dx.doi.org/10.1016/j.jeconom.2003.10.013.

Stinebrickner, R., Stinebrickner, T.R., 2008a. The causal effect of studying on academic performance the causal effect of studying on academic. B. E. J. Econ. Anal. Policy Front. 8, 14. http://dx.doi.org/10.2202/1935-1682.1868.

Stinebrickner, B.R., Stinebrickner, T., 2008b. The effect of credit constraints on the college drop-out decision : A direct approach using a new panel study. Amer. Econ. Rev. 98, 2163–2184. http://dx.doi.org/10.1257/aer.98.5.2163.

Stinebrickner, T., Stinebrickner, R., 2012. Learning about academic ability and the college dropout decision. J. Labor Econ. 30, 707–748. http://dx.doi.org/10.1086/666525.

Universities UK, 2019. Black, asian and minority ethnic student attainment at UK universities: Closing the gap.

van Nierop, E., Bronnenberg, B., Paap, R., Wedel, M., Franses, P.H., 2010. Retrieving unobserved consideration sets from household panel data. J. Mark. Res. XLVII, 63–74. http://dx.doi.org/10.1509/jmkr.47.1.63.

Willis, R.J., Rosen, S., 1979. Education and self-selection. J. Political Econ. 87, S7–S36.

Wiswall, M., Zafar, B., 2015. Determinants of college major choice: Identification using an information experiment. Rev. Econom. Stud. 791–824. http://dx.doi.org/10.1093/restud/rdu044.

Wiswall, M., Zafar, B., 2018. Preference for the workplace, investment in human capital, and gender. Q. J. Econ. 133, 457–507. http://dx.doi.org/10.1093/qje/qjx035.

Zafar, B., 2011. How do college students form expectations? J. Labor Econ. 29, 301–348. http://dx.doi.org/10.2139/ssrn.1432526.

Zafar, B., 2013. College major choice and the gender gap. J. Hum. Resour. 48, 545–595. http://dx.doi.org/10.3368/jhr.48.3.545.

Zwysen, W., Longhi, S., 2018. Employment and earning differences in the early career of ethnic minority British graduates: The importance of university career, parental background and area characteristics. J. Ethn. Migr. Stud. 44, 154–172. http://dx.doi.org/10.1080/1369183X.2017.1338559.