Socio-Economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns

Teodora Boneva and Christopher Rauh*

October 25, 2017

Job Market Paper
Please click here for the most recent version.

Students from low socio-economic backgrounds are significantly less likely to go to university compared to students from more advantaged backgrounds with similar levels of skills. While traditional models emphasize the role of credit constraints in explaining the socio-economic gap in university attendance, we investigate to what extent the gap can be explained by differences in beliefs about the pecuniary and non-pecuniary benefits of university education. For this purpose, we elicit students' beliefs about the benefits of attending university as well as their intentions to go to university in a sample of 2,540 secondary school students in England. Our analysis proceeds in three steps. First, we document that relative to high socio-economic status students, students with low socio-economic status perceive both the pecuniary and non-pecuniary returns to university to be significantly lower. Second, we derive and estimate a choice model which allows us to investigate the relative importance of different factors. For both low and high socio-economic status students, the perceived non-pecuniary benefits can explain a large share of the variation in intentions to go to university. Amongst the non-pecuniary outcomes, expected job satisfaction, parental approval, and perceptions about social life during the 3-4 years after finishing secondary school are most important. Third, we perform a decomposition analysis and find that 49% of the socio-economic gap can be explained by differences in beliefs across socio-economic groups, and that 37% can be explained by differences in beliefs about the non-pecuniary returns alone.

JEL: I24, I26, J13, J24, J62
Keywords: Higher education, beliefs, socio-economic inequality, intergenerational mobility

* Boneva: University College London (email: t.boneva@ucl.ac.uk). Rauh: University of Montreal (email: christopher.raphael.rauh@umontreal.ca). We thank Akash Raja and Jack Light for their excellent research assistance. We are further grateful to Orazio Attanasio, Kai Barron, Richard Blundell, Mariacristina De Nardi, Armin Falk, Christina Gravert, James Heckman, Stephanie Heger, Simon Jäger, Katja Kaufmann, Emir Kamenica, Alice Kügler, Michael Lovenheim, Kate Orkin, Imran Rasul, Uta Schönberg, Jenny Williams and Basit Zafar, as well as conference and seminar participants at UCL, University of Oxford, University of Mannheim, CESifo, Bonn/briq, ECBE, NHH Bergen, WinE and the New York Fed for their helpful comments and suggestions. Boneva acknowledges support from the British Academy. This study has been approved by the UCL Research Ethics Committee (Project ID number: 9287/002).

1. Introduction

It has been well documented that students from low socio-economic backgrounds are significantly less likely to attend university compared to students from more advantaged backgrounds with similar levels of prior academic achievement.[1] In the UK, students with at least one parent holding a university degree are about 14 percentage points more likely to go to university compared to students with similar levels of skills but less well educated parents.[2] The decision to attend university is a life-changing decision with a large impact on labor market, health, marriage, and crime outcomes.[3] Therefore, it is essential to understand why students with low socio-economic status are less likely to go to university. This open question is of high policy relevance given the low levels of socio-economic mobility in the UK and in many other countries where educational attainment and income are highly correlated across generations (Blanden, Gregg and Macmillan 2007; Black and Devereux 2011; Chetty et al. 2014*a*,*b*, 2017).

While traditional models have emphasized the importance of credit constraints in explaining the socio-economic gap in enrollment (see Keane and Wolpin 2001; Carneiro and Heckman 2002; Gayle, Berridge and Davies 2002; Cunha et al. 2006; Belley and Lochner 2007; Lochner and Monge-Naranjo 2012), it is not well understood why we observe socio-economic differences in university attendance in countries in which grants and loans are available to students from disadvantaged backgrounds.[4] Traditional choice models based on rational expectations about discounted future income streams fail to generate the enrollment gaps observed in the data. Instead, the models need to rely on a residual catch-all-term generally referred to as

[1] See, for example, Machin and Vignoles (2004); Blanden and Machin (2004); Blanden and Gregg (2004); Chowdry et al. (2013) for the UK, Bailey and Dynarski (2011); Chetty et al. (2014*a*,*b*, 2017) for the US, Checchi, Fiorio and Leonardi (2013) for Italy, and Stifterverband (2016) for Germany.

[2] We use data from the British Household Panel Study (BHPS) and the UK Longitudinal Household Survey (UKLHS) to calculate the socio-economic gap in university attendance conditional on a range of cognitive and non-cognitive skills. The results of this analysis are reported in Table D.1 in the Appendix and are robust to the inclusion of cohort fixed effects.

[3] See, for example, Oreopoulos and Salvanes (2011), Oreopoulos and Petronjievic (2013), and Heckman, Humphries and Veramendi (2016) for evidence on the pecuniary and non-pecuniary benefits of university education.

[4] See www.gov.uk/student-finance for an overview of student loans, grants, bursaries and scholarships available in the UK.

'psychic cost' or 'consumption value', which is allowed to vary across groups with different background characteristics.[5] Summarizing the results in the literature, Heckman, Lochner and Todd (2006) note: "*The evidence against strict income maximization is overwhelming. However, explanations based on psychic costs are intrinsically unsatisfactory. One can rationalize any economic choice data by an appeal to psychic costs* [p. 436]". To better understand socio-economic gaps in university enrollment, it is therefore crucial to obtain a better understanding of what 'psychic costs' actually represent, and how these costs vary across different socio-economic groups.

In this paper, we shed light on students' motives to obtain university education and proceed in three steps. First, we elicit students' beliefs about the benefits of university education in a sample of 2,540 secondary school students in England (ages 13-18), and we document how students' beliefs about the pecuniary and non-pecuniary benefits/costs of university education differ by socio-economic status (SES). Here we do not only focus on students' beliefs about their likely labor market outcomes later in life but also on students' perceptions about what their lives are likely to be like during the 3-4 years after they finish secondary school. Second, we estimate a choice model that allows for differences in beliefs and differences in preferences, and we investigate the relative importance of the different factors in students' decisions to go to university. Finally, we investigate to what extent the socio-economic gap in university enrollment can be explained by differences in students' beliefs about the pecuniary and non-pecuniary returns to university education, and discuss which policies might be effective in narrowing the socio-economic gap.

To investigate the role of beliefs in educational investment decisions, it is not possible to rely on choice data alone. Observed choices can be consistent with many different combinations of beliefs and preferences (Manski 2004). To overcome this identification problem, it is important to obtain direct measures of individual beliefs, both about the pecuniary and non-pecuniary benefits of education. For this purpose, we collect novel primary survey data and directly elicit individual intentions to attend university as well as individual beliefs

[5]See, e.g., Carneiro, Hansen and Heckman (2003); Cunha, Heckman and Navarro (2005); Heckman, Lochner and Todd (2006); Cunha, Heckman and Navarro (2006); Cunha and Heckman (2007, 2008); Carneiro, Heckman and Vytlacil (2011).

about the pecuniary and non-pecuniary benefits of university education. To obtain a measure of perceived pecuniary labor market returns, we ask students about their expected labor market income and probability of having a job (at age 30) under different hypothetical scenarios (Dominitz and Manski 1996). The hypothetical scenarios include (i) not going to/not graduating from university, (ii) graduating with low grades in their preferred subject field, and (iii) graduating with high grades in their preferred subject field. To elicit students' beliefs about the non-pecuniary aspects of university education, we use probabilistic questions (Manski 2004). In particular, we ask students to state how they perceive the probability of multiple outcomes under different scenarios that relate to (i) their social life during the 3-4 years after finishing secondary school, (ii) their enjoyment of the study/work tasks during these 3-4 years, (iii) their financial worries during those years, (iv) the approval they obtain from their parents for their decision and (v) their likely job satisfaction at age 30. In order to validate responses, we follow up a subsample of students and find very high test-retest correlations. Moreover, we find very high correlations between stated intentions to attend university and actual application decisions, as well as between stated intentions to choose a specific subject field and actual subject choice. We also show that our survey data is consistent with administrative data on a number of different dimensions.

Our analyses reveal three main findings which contribute to our understanding of what drives socio-economic differences in university attendance. First, relative to high SES students, low SES students do not only perceive the pecuniary benefits of university education to be significantly lower but they also perceive the returns to be lower when it comes to a wide range of different non-pecuniary factors. For example, while both low and high SES students believe it is more likely they will enjoy their future job and obtain parental approval for their choice if they go to university, this difference is markedly more pronounced for high SES students. Second, the estimates of our choice model reveal that both low and high SES students do take pecuniary returns into account. However, when we compare the relative importance of pecuniary and non-pecuniary returns in students' intentions to go to university, we find that for both low and high SES students the perceived non-pecuniary benefits account for a much larger share of the variation. Amongst the non-pecuniary outcomes, expected job

satisfaction, parental approval, and perceptions about social life during the 3-4 years after finishing school play an important role in the decision. We also find evidence that perceived non-pecuniary benefits play a larger relative role for students with low socio-economic status. For example, the perceived probability of enjoying one's future job accounts for 29% of the variation for low SES students, whereas for high SES students it only accounts for 15%. Third, the results reveal that 49% of the socio-economic gap in students' intentions to go to university can be explained by differences in students' beliefs about the pecuniary and non-pecuniary benefits of university education. Out of the explained gap, 25% can be accounted for by differences in beliefs about pecuniary labor market returns, while 75% can be explained by differences in beliefs about the non-pecuniary benefits we elicit.

While traditional policies have mainly focused on relieving credit constraints to increase university enrollment rates among socio-economically disadvantaged groups, our results, by contrast, suggest that alleviating credit constraints will not be sufficient for closing the SES gap in university enrollment. Students from low SES backgrounds perceive the pecuniary and non-pecuniary benefits of university education to be lower, and differences in beliefs about returns explain a large proportion of the socio-economic gap in attendance. To effectively design policies which narrow the socio-economic gap in enrollment, we will need a better understanding of whether low SES students only *perceive* the returns to be significantly lower or whether the returns do indeed differ across socio-economic groups. If low SES students were misinformed about what university life is like or what labor market opportunities university graduates have, then interventions which provide students with information are likely to be effective. For example, campaigns that provide information on the social activities students can engage in whilst at university or the variety of jobs available to university graduates may encourage more students from disadvantaged backgrounds to apply. While this evidence is merely suggestive, we note that low SES students in our sample know significantly fewer people they can ask about university and that perceived returns are increasing in the number of people students can ask about university life. While low SES students may misperceive the returns, it may also well be that the returns are in fact lower for students from disadvantaged backgrounds. For example, if it is indeed true that low SES students experience

more difficulties fitting in with the social environment at universities, as has been suggested by psychologists (e.g. Janke et al. 2017), then universities might want to consider addressing this problem if they are interested in attracting and retaining the best talent.

Our study builds on and contributes to several strands of the literature. First, the study relates to the growing literature which investigates the role of beliefs in human capital investment decisions. While previous work has mainly focused on the role of actual and perceived *pecuniary* returns in explaining educational attainment (e.g. Dominitz and Manski 1996, Jensen 2010, Abramitzky and Lavy 2014, Attanasio and Kaufmann 2014, Kaufmann 2014, Almas et al. 2016), we investigate the role of perceived *non-pecuniary* benefits and examine differences in motives for educational attainment across socio-economic groups.[6] Indeed, we find that perceived non-pecuniary benefits play a very important role for the decision to enroll in university as well as for explaining the socio-economic gap in enrollment.

Second, our paper relates to the literature documenting how additional services and amenities provided by universities influence enrollment. Jacob, McCall and Stange (2013) find that demand-side pressures have pushed US colleges to increase expenditures on consumption amenities, such as student activities, sports, and dormitories. The number of applications received by a specific university has been shown to increase after successful basketball and football seasons (Pope and Pope 2009) and after improvements in a quality-of-life ranking focusing on non-academic amenities (Alter and Reback 2014). While these studies provide indirect evidence that students value non-pecuniary attributes of university life, we directly measure beliefs about a wide range of different pecuniary and non-pecuniary benefits, allowing us to examine how they relate to choices and differ across socio-economic groups.

Third, our study relates to the literature which examines the role of beliefs about pecuniary and non-pecuniary benefits in students' choice of major (Montmarquette, Cannings and Mahseredjian 2002; Arcidiacono 2004; Arcidiacono, Hotz and Kang 2012; Beffy, Fougere and Maurel 2012; Zafar 2013; Arcidiacono et al. 2014; Stinebrickner and Stinebrickner 2014;

[6]Notable recent exceptions include Attanasio and Kaufmann (2017) who investigate the role of perceived marriage market returns and Belfield et al. (2017) who show that there is a strong positive association between perceived enjoyment of university and stated intentions to continue in higher education.

Wiswall and Zafar 2015*a*,*b*; Hastings et al. 2016; Wiswall and Zafar 2017; Baker et al. 2017), high-school track (Giustinelli 2016), and which specific university to attend (Delavande and Zafar 2014). In contrast to these studies, our analysis addresses the question of whether or not to go to university (i.e. the extensive margin), rather than which specific major, high-school track or university to choose. Given the large estimated gains from a college degree (Heckman, Humphries and Veramendi 2016), the decision whether or not to enroll is crucial for many later life outcomes as well as for social mobility, and likely to be even more important than which specific major or university to select.

More broadly, our study contributes to the work that investigates the sources of socio-economic gaps in educational attainment. Previous studies have examined the role of credit constraints (e.g. Keane and Wolpin 2001, Carneiro and Heckman 2002, Gayle, Berridge and Davies 2002, Cunha et al. 2006, Belley and Lochner 2007, Lochner and Monge-Naranjo 2012) and the availability of financial aid (e.g. Angrist et al. 2016), as well as debt aversion (e.g. Field 2009) and incomplete information about the costs of attending university/the availability of financial aid (e.g. Hoxby and Turner 2013). Studies have also found that arguably small differences in the application process can have sizeable impacts on students' application decisions. For example, Bettinger et al. (2012) show that high school students whose parents receive assistance with filling out the financial aid application form are significantly more likely to attend university. Pallais (2015) shows that small changes in the cost of sending standardized test scores to colleges affects students' college application decisions. The results of our study contribute to this literature by drawing attention to the importance of differences in perceived pecuniary and non-pecuniary benefits across socio-economic groups.

This paper is organized as follows. Section 2 derives the choice model and presents the estimation approach. Section 3 provides details on the survey design we use to elicit students' beliefs. Section 4 provides details on the sample and the survey data. Section 5 documents how students with different socio-economic backgrounds perceive the pecuniary and non-pecuniary returns to attending university. Section 6 presents the results of the choice model estimation, while Section 7 provides the results of different policy simulations. Section 8 discusses the implications of our findings and Section 9 concludes.

2. Theory

In the following, we formulate the discrete choice problem students face when deciding whether or not to go to university (Section 2.1), and describe the estimation approach we use to estimate the choice model (Section 2.2).[7]

2.1. The Discrete Choice Problem

Students face the choice of whether to go to university ($j = 1$) or not ($j = 0$). The choice is static yielding n separable binary outcomes, which are denoted as $\{b_n \in \{0,1\}\}_{n=1}^{N}$, and m continuous outcomes, which we denote as $\{c_m\}_{m=1}^{M}$. A binary outcome can be, for example, whether or not the student enjoys his/her social life, while a continuous outcome can be the earnings of the student at age 30.[8] In addition, each individual i has a vector of characteristics, X_i (e.g. socio-economic status). Let $U_i(b, c, X_i)$ be the utility that an individual derives from the vector of binary outcomes, b, the vector of continuous outcomes, c, and individual characteristics, X_i. The individual chooses alternative j so as to maximize her subjective expected utility (SEU). The SEU of alternative j can be written as:

$$SEU_{ij} = \int U_i(b, c, X_i) dP_{ij}(b, c)$$

where $P_{ij}(b, c)$ denotes the probability of the occurrence (b, c) if alternative j is chosen. Assuming that utility is additively separable across the outcomes, we can write the SEU as:[9]

$$(1) \qquad SEU_{ij} = \sum_{n=1}^{N} P_{ij}(b_n = 1) \triangle u_n(X_i) + \sum_{n=1}^{N} u_n(b_n = 0, X_i) + \sum_{m=1}^{M} \gamma_m(X_i) E_{ij}[c_m] + \varepsilon_{ij}$$

where $u_n(b_n, X_i)$ is the utility associated with the binary outcome b_n for an individual with

[7] Similar frameworks have been used to model students' choices in other contexts, e.g. students' choice of high school track (Giustinelli 2016) or college major (Zafar 2013).

[8] Note that we define all non-pecuniary outcomes as binary despite the fact that the nature of these outcomes is also likely to be continuous. The reason for this design choice is that there is no natural metric in which to measure non-pecuniary outcomes, which is why we follow the approach advocated by Manski (2004) and elicit perceived probabilities on a 0-100% scale. We provide a more detailed discussion of the method in Section 3.3.

[9] See Appendix A for a derivation of this expression.

characteristics X_i, $\gamma_m(X_i)$ is a constant for the continuous outcome c_m for an individual with characteristics X_i, and ε_{ij} is a random error term. $\triangle u_n(X_i)$ is defined as $\triangle u_n(X_i) \equiv u_n(b_n = 1, X_i) - u_n(b_n = 0, X_i)$. The utility individual i derives from (b, c) is the same for all individuals with the same characteristics X_i up to the random error term, and is allowed to differ across individuals with different characteristics X_i.

Note that we directly elicit individual beliefs about the probability of the various binary outcomes given the different alternatives j, $P_{ij}(b_n = 1)$. Moreover, there is no need to know the utility an individual derives from an outcome not occurring, $u_n(b_n = 0, X_i)$, because this is assumed to be the same for each alternative and therefore is differenced out in the choice problem. For the continuous outcomes, we elicit the expected value of the different outcomes for each alternative j, $E_{ij}[c_m]$. The parameters to be estimated are the $\triangle u_n(X_i)$'s and the $\gamma_m(X_i)$'s. Intuitively, $\triangle u_n(X_i)$ is the change in utility that results from the occurrence of outcome b_n for an individual with characteristics X_i, while $\gamma_m(X_i)$ is the parameter in the utility function for the continuous outcome c_m for an individual with characteristics X_i.

The individual will choose to go to university, i.e. $j = 1$, if the SEU from going to university exceeds the SEU from not going to university. The probability that $j = 1$ is chosen given the student beliefs, $P_{ij}(b_n = 1)$ and $E_{ij}[c_m]$, is:

$$Pr(j = 1|P_{ij}(b_n = 1), E_{ij}[c_m]) = Pr(SEU_{ij=1} \geq SEU_{ij=0})$$

$$= Pr(\sum_{n=1}^{N} P_{ij=1}(b_n = 1) \triangle u_n(X_i) - \sum_{n=1}^{N} P_{ij=0}(b_n = 1) \triangle u_n(X_i)$$

$$+ \sum_{m=1}^{M} \gamma_m(X_i) E_{ij=1}[c_m] - \sum_{m=1}^{M} \gamma_m(X_i) E_{ij=0}[c_m] \geq \varepsilon_{ij=0} - \varepsilon_{ij=1})$$

2.2. Choice Model Estimation

Under the assumption that the error terms ε_{ij} have a Type-I extreme value distribution, $\varepsilon_{ij=0} - \varepsilon_{ij=1}$ has a standard logistic distribution. Simplifying the notation $P_{ij}(b_n = 1) = P_{ijn}$

and $E_{ij}[c_m] = E_{ijm}$, we can write the probability of $j = 1$ being chosen as:

$$Pr\big(j = 1|P_{ijn}, E_{ijm}\big) = \Lambda\Big(\sum_{n=1}^{N}(P_{ij=1n} - P_{ij=0n}) \triangle u_n(X_i) + \sum_{m=1}^{M} \gamma_m(X_i)(E_{ij=1m} - E_{ij=0m})\Big)$$

where $\Lambda(a) = \frac{1}{1+e^{-a}}$ is the logistic function. Note that only differences in perceived probabilities and expected values enter the decision problem. Moreover, the $\triangle u_n(X_i)$'s and the $\gamma_m(X_i)$'s will only be identified up to a constant.

Given that we collect our survey data *prospectively* rather than retrospectively (see Section 3.1), the main outcome variable we use for our analysis is the perceived probability of going to university, which is measured on a 0-100% scale. Since our outcome variable is not binary, we do not estimate the choice model using a logistic regression. In principle, models in which the response variable appears as a proportion could be estimated using least squares by regressing the logit transformation of the outcome variable (i.e. the log odds ratio) on the explanatory variables:

$$\log(\frac{y_i}{1 - y_i}) = Z_i\beta + \epsilon_i,$$

where y_i is the outcome of interest, Z_i are the explanatory variables and β is the vector of parameters to be estimated. While this estimation strategy has been used in previous work, there are two problems with this approach: first, the log odds ratio cannot be calculated for values that are exactly 0 or 1. This is problematic given that in our survey data we do have respondents who perceive the likelihood of going to university to be exactly 0 or 100%. Second, estimation of the model using least squares does not account for the heteroskedastic nature of ϵ. To estimate the model, we therefore use the approach proposed by Papke and Wooldridge (1996), which overcomes both of these challenges. The estimation strategy makes use of the logit link function (that is, the logit transformation of the response variable) and the Bernoulli distribution. More specifically, the parameters of the model are estimated by quasi-maximum likelihood, in which the following Bernoulli log-likelihood function is maximized:

$$\max_b \sum_{i=1}^{n} l_i(b) = \max_b \sum_{i=1}^{n} \big(y_i \log[\Lambda(Z_i\beta)] + (1 - y_i) \log[1 - \Lambda(Z_i\beta)]\big).$$

The quasi-maximum likelihood estimator (QMLE) of β, obtained from maximizing the quasi-likelihood function, is consistent and $\sqrt{N}$-asymptotically normal, regardless of the distribution of y_i conditional on Z_i. This approach has the advantage that it can handle proportions data in which zeros and ones may appear as well as intermediate values. Moreover, the variance of the Bernoulli distribution must go to zero when the mean goes to either 0 or 1, as in each case the variable is approaching a constant, and the variance is maximized for a variable with a mean of 0.5.

3. Eliciting Student Beliefs

To investigate the role of students' beliefs in their decision to go to university using the choice model described in Section 2, we develop a novel survey tool to elicit students' beliefs about the benefits of university education. Section 3.1 explains the main features of our survey design, while Sections 3.2 and 3.3 present the methodology we use to elicit beliefs about the pecuniary and non-pecuniary benefits, respectively.

3.1. Survey Design

To estimate the choice model described in Section 2, we need to elicit student beliefs about (the likelihood of) potential outcomes if (i) the student goes to university and (ii) the student does not go to university but starts working instead.[10] We elicit these beliefs in a sample of 2,540 secondary school students (ages 13-18) using hypothetical investment scenarios, which are described in detail in Sections 3.2 and 3.3 and summarized in Table 1.[11] Importantly, we choose to ask all questions *prospectively* rather than retrospectively to minimize any potential bias that could arise due to cognitive dissonance or ex-post rationalization (Festinger 1957, Bertrand and Mullainathan 2001, Benitez-Silva et al. 2004, Zafar 2011). Similar to other studies (e.g. Bleemer and Zafar 2015), we use the self-reported likelihood of continuing in full-time education as our main outcome measure because the students have not yet decided

[10]We deliberately chose to make it explicit that the alternative is to start working because we did not want students to think about the possibility of doing a gap year before starting in higher education.

[11]See Appendix B for a full list of all questions.

whether or not to go to university at the time of the survey.[12] More specifically, we ask students to state how likely they think it is that they will go to university if they get the necessary grades at the end of school (0-100%). In the UK, 48% of a given cohort of students continues to higher education (Source: Department for Education Statistical Fiscal Releases 45/2016).

We validate our survey tool in multiple ways: first, we re-survey a subsample of participants two months after the initial survey. This allows us to compute test-retest correlations to assess the reliability of our survey measures. Second, for final year students, we can compare stated intentions to continue in full-time education against actual decisions to apply to university. In the UK, all students who wish to go to university need to apply through a centralized application system (Universities & College Admissions Service – UCAS) by a specific date. We re-survey individuals *after* the application deadline, which allows us to obtain information on whether students actually applied to university or not, and if yes, which subject field they chose to apply to. Third, we can also assess the reliability of the survey by investigating students' responses to questions that are similar to each other but where one of the questions is reverse coded. If students were actively thinking about the questions we would expect to see a negative correlation between these survey items. Finally, we merge our survey data with administrative data on the schools the students currently attend, which allows us to assess whether the mean responses of students within a given school are consistent with the information we can obtain at the school level. For example, we have information on the proportion of students going on to higher education institutions in a given school in the past, and we can correlate that school-level information with students' average beliefs about the likelihood of going to university at that school.

We also note that we took great care to emphasize that the survey is completely anonymous. We did not collect any personal information such as names or addresses, and it was made clear to the students that neither the researchers nor the students' teachers can identify any individual respondent. We matched students across survey waves using a survey ID.

[12]Previous studies have documented that intended educational choices are strong predictors of actual future educational choices (e.g. Jacob and Linkow 2011, Beaman et al. 2012).

3.2. Pecuniary Motives

To elicit student beliefs about the pecuniary returns to education, we present students with different hypothetical scenarios. The use of hypothetical scenarios to elicit beliefs about pecuniary returns to further education was pioneered by Dominitz and Manski (1996), and it has been used successfully in different studies such as Jensen (2010), Kaufmann (2014) and Attanasio and Kaufmann (2014).[13] Our survey design builds on and extends this literature in a number of important ways. First, we condition on the students' preferred field of study, which allows us to examine whether students from different socio-economic groups sort into different subject fields. Second, we allow for differences in beliefs about how well students will do if they choose to go to university. Unlike in the US, dropout rates in the UK are low with only 6% of all students who start a degree failing to qualify. There is, however, significant heterogeneity in how well students do at university. Conditional on starting a degree and graduating, 73% of all student obtain First-class or Upper Second-class honours (2.1), while 27% obtain Lower Second-class honours (2.2) or Third-class honours (Source: Higher Education Statistics Authority).

Our survey design allows students to differ in their beliefs about how well they will do at university, and asks students to imagine scenarios in which they (i) do not go to university/go to university but do not graduate, (ii) go to university and graduate with low grades (Lower Second-class honours or Third-class honours) in their preferred subject field, and (iii) go to university and graduate with high grades (First-class honours or Upper Second-class honours) in their preferred subject field. For each of these three scenarios, we ask students how likely they think it is that they will have a job at age 30 and what their likely earnings will be (conditional on working full-time).[14]

Before we ask students these questions, we ask them to report which field of study they would be most likely to choose if they decided to go to university. More specifically, students are able to choose between five different subject fields: (i) Arts and Humanities (e.g.

[13]In comparison, vaguely worded qualitative questions have been shown to provide little useful information about respondents' expectations (see for example Manski (1990) and Juster (1966)). See Manski (2004) for a review and discussion of different belief elicitation approaches.

[14]We additionally ask students to assume there is no inflation.

languages, history, music, architecture, philosophy), (ii) Life Sciences (e.g. biology, medicine, pharmacy, psychology), (iii) Physical Sciences and Engineering (e.g. maths, computer science, physics, engineering), (iv) Social Sciences (e.g. economics, law, business) and (v) Education. The advantage of this questionnaire design is that we can investigate whether differences in perceived returns across socio-economic groups are driven by differences in subject choice.

We also ask students additional questions that allow us to calculate the perceived probability of being in any of the three different scenarios. In particular, we ask students how likely they think it is that they will graduate if they go to university, $P_i^{graduate}$, and how likely they think it is that they will obtain high grades (a First or a 2.1) conditional on graduating, $P_i^{Hgrades}$. We use this information to calculate the individual perceived pecuniary returns to higher education (conditional on full-time work). More specifically, we take the difference between what student i expects to earn when she goes to university (conditional on working full-time) and what the student expects to earn when she does not go to university, $E[Y_i|j=1] - E[Y_i|j=0]$. While we directly elicit the latter, we calculate the former in the following way:

$$E[Y_i|j=1] = P_i^{graduate}[P_i^{Hgrades} \times Y_i^{Hgrades} + (1 - P_i^{Hgrades}) \times Y_i^{Lgrades}] + (1 - P_i^{graduate}) \times Y_i^{nouni}$$

The expected earnings of an individual who goes to university are a weighted average of the expected earnings if the student graduates with high grades, $Y_i^{Hgrades}$, graduates with low grades, $Y_i^{Lgrades}$, or does not graduate at all, Y_i^{nouni}.

3.3. Non-Pecuniary Motives

To elicit students' beliefs about the non-pecuniary benefits of higher education, we use probabilistic questions. The use of probabilistic questions has the advantage that the responses are interpersonally comparable and more informative than responses on a Likert-scale (see Manski 2004). First, we ask students to report the percent chance that they will enjoy the job they will be doing at age 30 in the three scenarios described above, namely if they (i) do not go to university/do not graduate, (ii) graduate with low grades in their preferred sub-

ject field, and (iii) graduate with high grades in their preferred subject field.[15] Second, we specifically ask students to think about what their lives are likely to be like during the 3-4 years after they leave secondary school (i) if they go to university or (ii) if they start working instead.[16] For example, we ask students the following two questions:

1) *If you go to university, how likely do you think it is that you will meet people with whom you easily get along with? [0-100%]*

2) *If you start working, how likely do you think it is that you will meet people with whom you easily get along with? [0-100%]*

The questions presented in Table 1 relate to different aspects of the students' lives after they finish secondary school. We draw from the economics, psychology and sociology literature to select questions related to their social life, their work/study life, their financial standing as well as the approval they receive from their parents for their decision. For instance, the question regarding parental approval is inspired by the concept introduced to economists by Akerlof and Kranton (2000) that identity, e.g. in terms of being from a working class background, might play an important role in the decision. Relatedly, concerns about losing contact to family and friends could also be linked to identity (similar to 'acting white' as in Austen-Smith and Fryer Jr (2005)) or simply to geographical distance to university, which has been used as an instrumental variable for enrollment (Kane and Rouse 1995). The questions regarding the students' social life are further motivated by recent evidence from case studies which document that students from low SES backgrounds have more difficulties 'fitting in' with the university culture (e.g. Reay, Crozier and Clayton 2010) as well as by evidence on the recent increases in the supply of amenities provided by universities (Jacob, McCall and Stange 2013). We also include questions that relate to the (dis)utility from acquiring new skills, which are inspired by the literature that documents differences in test anxiety across socio-economic groups (Janke et al. 2017). The questions that relate to students' concerns

[15]From these responses, we calculate the perceived probability of enjoying the job at age 30 if the student decides to go/decides not to go to university.

[16]A typical undergraduate degree in the UK takes 3-4 years to be completed.

about their financial standing during the 3-4 years after finishing school are motivated by the literature on credit constraints (e.g. Lochner and Monge-Naranjo 2012).

The specific survey technique we use has been used in the previous literature which examines the importance of non-pecuniary benefits in students' choice of major (e.g. Zafar 2013), students' choice of high school track (e.g. Giustinelli 2016) and their decision of which specific university to attend (e.g. Delavande and Zafar 2014).[17] As is standard in the literature, the subjective probability questions are preceded by a section which explains the use of the 0-100% chance scale and asks respondents to answer some warm-up questions.[18]

Table 1—: Overview of Belief Elicitation Questions

Scenarios	**Outcomes**
At age 30	
(1) If you graduate with high grades	Earnings (conditional on working)
(2) If you graduate with low grades	Have job (0-100%)
(3) If you do not go to university/do not graduate	Enjoy job (0-100%)
During 3-4 years after finishing secondary school	
(1) If you go to university	Enjoy social life and activities (0-100%)
(2) If you work rather than go to university	Meet people you easily get along with (0-100%)
	Lose contact with family and friends (0-100%)
	Feel lonely (0-100%)
	Find your life partner (0-100%)
	Find material/work tasks interesting (0-100%)
	Enjoy studying/work (0-100%)
	Find material hard/workload high(0-100%)
	Feel stressed (0-100%)
	Struggle financially (0-100%)
	Have financial support from your parents (0-100%)
	Have enough money (0-100%)
	Have parental support in your choice (0-100%)

Notes: Students are asked 9 questions regarding potential outcomes at age 30 (3×3) and a total of 26 questions regarding potential outcomes during the 3-4 years after finishing secondary school (2×13).

[17]For example, Zafar (2013) uses the question *'If you were majoring in [X], what do you think is the percent chance that you will enjoy the coursework?'*. Similarly, Giustinelli (2016) asks *'What did you think would be the chances out of 100 that you would enjoy the content of the following curricula, should you enrol in each one of them?'*.

[18]In order to familiarize students with the nature of probabilistic questions, we ask: *'What do you think is the percent chance that it will rain tomorrow?'*

4. Data

In order to examine which beliefs are important for students' decisions to enroll in higher education and to gain a better understanding of what drives the socio-economic gap in enrollment, we collect primary survey data from secondary school students in England. In the following, we describe the characteristics of our sample (Section 4.1) and assess the validity of responses (Section 4.2).

4.1. The Sample

The study sample consists of 2,540 students aged 13-18 from 37 schools in England who are at a critical age as they are about to make the decision of whether or not to go to university. We collected the data with the help of an online survey, which was distributed via the student mailing lists of schools that agreed to participate in the research study (see map in Appendix D).[19] 10.1% of all students who were contacted to participate in the study chose to participate. The main survey data was collected in November 2016 (wave 1). A short follow-up survey was administered in January 2017 (wave 2), for which 319 of the 2,540 students participated. Students were incentivized to participate in the surveys through a prize draw of a voucher worth £100. The median time students needed to complete the survey was 13 minutes. 37% of the students in our sample are male with the average student being in Year 11 (see Table D.2).[20] 759 students in our sample are in their final year of secondary school education (Year 13).

To measure students' socio-economic background, we ask students to report whether or not their parents have gone to university. 55% of students in our sample report that at least one of their parents obtained university education.[21] We define students as low/high SES depending

[19]We did not use any specific selection criteria to select the schools we contacted. The Department for Education provides lists of all secondary schools and sixth form colleges in England. We used these lists of potential schools and contacted the head teachers of a random subset of these schools in no specific order.

[20]Out of the 37 schools that participated, 6 are all-girls and 2 are all-boys schools. We note that we obtain qualitatively similar results when we exclude students attending these schools. The results are presented and discussed in Section 7.

[21]In a representative sample of students in England, 40% of all student have at least one parent with university education (Source: Family Resources Survey).

on whether or not they have at least one parent with a university degree. Put differently, we define those students as low SES who would be first-generation university students if they decided to go to university. To shed more light on whether socio-economic differences in beliefs may be related to differences in access to information, we also ask students about the number of people they know whom they could ask about university life. On average, low SES students know 4.5 people whom they can ask about their experiences at university, while high SES students report knowing 7.2 people (p-value=0.000).[22] 8.8% of all low SES students report they do not know a single person they can ask about university life.

To obtain measures of students' academic ability, we ask all students in Years 12 and 13 to report the grades they achieved on their GCSE English Literature and Maths exams, which are standardized tests students in England take at the end of Year 11. Figure D.2 in the Appendix compares the distribution of grades in our sample to (i) the distribution of grades within the sampled schools and (ii) the distribution of grades in England. The students in our sample have higher grades compared to the average students in the sampled schools as well as the average students in England. While our sample is not representative of the population of students in England, we oversample those students whom we are most interested in: the high-achieving students for whom going to university is actually a realistic option.[23]

4.2. Survey Validation

In order to assess the reliability of survey responses to the question how likely students think it is they will go to university if they get the grades, we re-survey a subsample of all students two months after the initial survey. Whilst it is possible that beliefs can change between the two survey waves, we would not expect major shifts in beliefs for most respondents. The Spearman rank correlation between individual beliefs about the perceived probability

[22]All p-values reported in parentheses are for two-sided t-tests testing differences across socio-economic groups.

[23]There are two different reasons why the students in our sample are not representative of the population of students in England. First, it was the decision of the head teacher whether or not to distribute our survey among students, so the schools included in our sample are not representative of the population of English schools. Table D.3 in the Appendix shows the characteristics of schools in our sample with comparisons to national statistics. Second, participation in the survey was voluntary so the students who decided to participate are a self-selected sample.

of going to university stated in waves 1 and 2 for those students who were not in their final year of post-secondary education is 0.532 (N=202), which is high compared to the test-retest correlation other survey studies find for individual survey items (e.g. Falk et al. 2016).

As explained in Section 3.1, students in the UK who wish to go to university need to apply via a centralized system by a specific date. Given that we re-survey students after this deadline, we can examine whether the stated intentions of final year students correlate with their actual decisions to apply. Panel A of Figure 1 shows the mean perceived likelihood of going to university stated in the initial survey for final year students who chose to apply to university and final year students who did not apply to university (N=117). We can see that stated intentions correlate highly with actual choices. For students who chose to apply to university, the average stated probability in wave 1 is 93% compared to 28% for students who did not apply. The two means are statistically different from each other at the 1% level. We also investigate students' stated subject choices and how those relate to actual application decisions. In wave 1, 27% of all students stated that they would study Arts and Humanities, 27% Life sciences, 21% Physical sciences and engineering, 20% Social sciences and 6% Education.[24] Among the students who applied, 90% did indeed apply to a subject in the subject field they stated was their most preferred in wave 1.[25] As can be seen in Figure D.5, low SES students intend to study similar subject fields as high SES students.

We additionally assess the reliability of our survey tool by investigating the Spearman rank correlations between different survey items, some of which are reverse coded (see Tables D.4 and D.5). As expected, we do find that similar survey items correlate positively (e.g. enjoy social life and meet people with whom you easily get along with), while similar reverse coded survey items correlate negatively (e.g. enjoy social life and feel lonely). We have illustrated

[24]Note that these numbers are comparable to statistics from the Higher Education Statistics Authority which contains information on all students in the UK. Among all students, 22% study Arts & Humanities, 27% study Life Sciences, 20% study Physical sciences and engineering, 24% study Social sciences and 7% study Education.

[25]While attrition between the two surveys is non-random, with students who stated a higher likelihood of going to university being more likely to fill out the follow-up survey (82.6% vs. 77.8%), we find similar results in terms of validity of the survey measure for both socio-economic groups. More specifically, the test-retest correlation is 0.503 for low SES students and 0.478 for high SES students, and for both groups the stated intentions significantly differ by actual application decisions (at the 1% level).

some of the joint distributions of survey responses in Figure D.4. We also note that similar patterns can be found for low and high SES students.

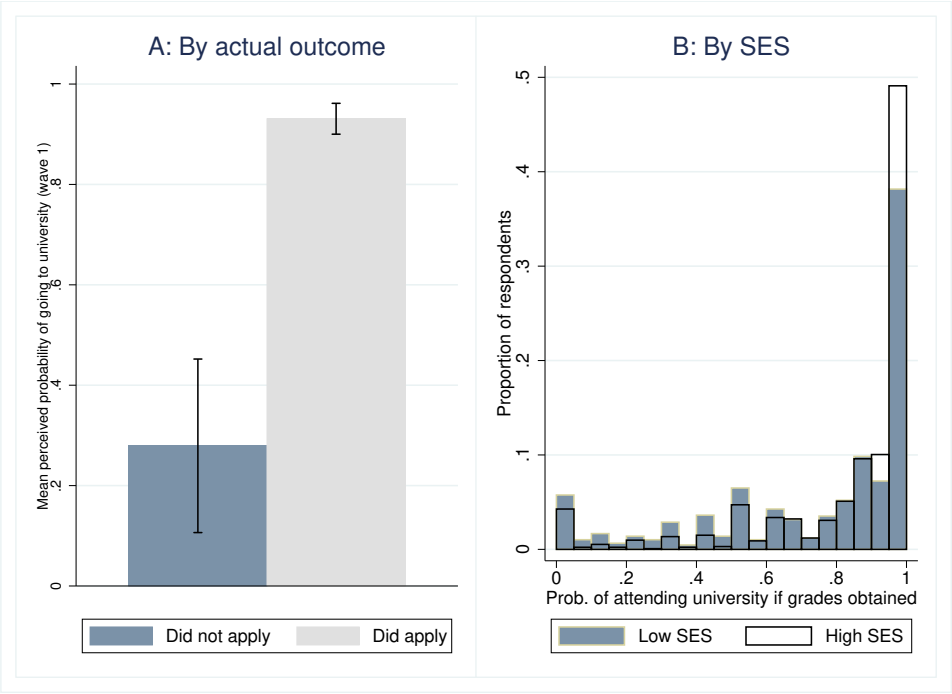
To further investigate the reliability of our response measures, we investigate whether students' mean responses within a given school correlate with school-level information we obtain from administrative data (see Figure D.3). The following findings emerge: first, students who attend schools in which a high proportion of students continue to higher education are also more likely to perceive the likelihood of attending university to be higher; second, students who attend schools with high results on national standardized tests (measured by the average Attainment 8 score, which is based on the students' GCSE results in Year 11) also perceive the probability of obtaining the grades to go to university to be higher. Overall, we conclude that these patterns are consistent and that the reliability of the survey data is likely to be high.

5. Differences in Beliefs across Socio-Economic Groups

In this section, we investigate how students from different socio-economic groups differ in their beliefs. First, we examine students' responses to our main outcome measure, namely the perceived likelihood of going to university conditional on achieving the required entry grades. Panel B of Figure 1 shows the distribution of responses, separately for low and high socio-economic status individuals. The mean perceived probability of going to university is 83% for students who have at least one parent who attended university and 74% for students whose parents have not attended university. The distributions are significantly different from each other at the 1% level. Whilst the histogram shows a significant proportion of students in our sample who are virtually certain that they want to go to university (46% for high SES and 36% for low SES students), there are also a non-trivial proportion of respondents who are either not likely or deem it more unlikely than likely that they will go to university, thereby showing a substantial degree of heterogeneity in beliefs in our sample.

Turning to students' expectations about how well they will do if they decide to go to university, we find that high SES students on average believe there is a 62% chance they will graduate from university with high grades (First or 2.1), a 23% chance they will graduate

Figure 1. : Perceived probability of going to university



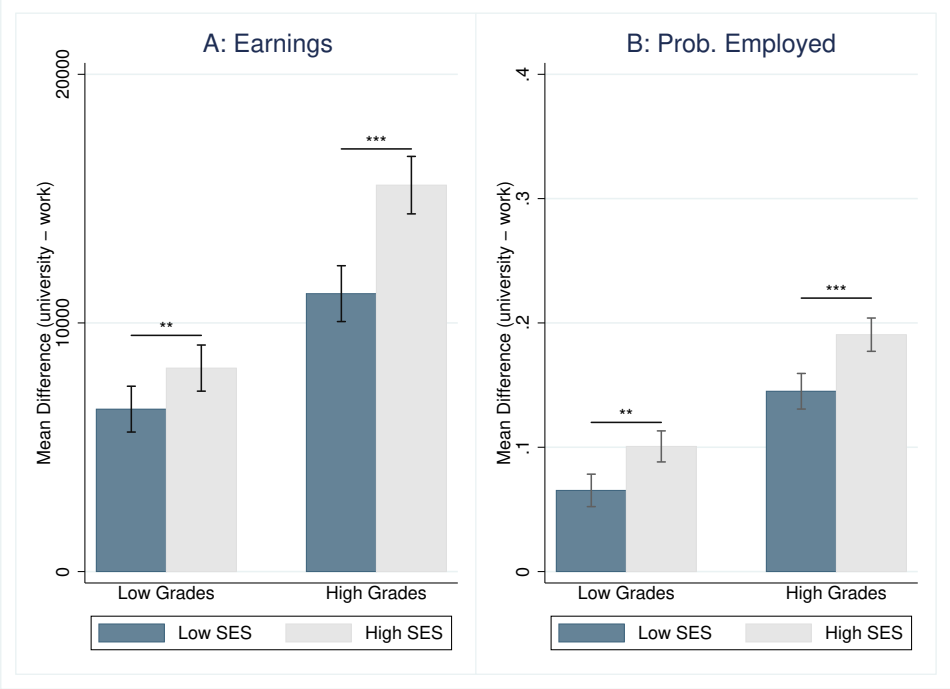
Notes: Panel A shows the probability of going to university stated in wave 1, separately averaged across individuals in their final year who applied to university and those who did not (with 95% confidence intervals). Panel B shows the distribution of stated beliefs of attending university. Bins have a width of 5 percentage points. High SES students are defined as those students who have at least one parent with university education.

from university with low grades, and a 15% chance they will not graduate. Compared to high SES students, low SES students believe they are less likely to graduate and obtain high grades. On average, low SES students believe there is a 55% chance they will graduate with high grades, a 26% chance they will graduate with low grades, and an 18% chance they will not graduate at all.[26]

When we investigate students' expected labor market outcomes for each of these three

[26]Both students' perceptions about the likelihood of graduating as well as students' perceptions about the likelihood of obtaining high grades significantly differs across socio-economic groups at the 1% level (see Table D.8). We further note that administrative data from the Higher Education Statistics Authority reveals that, conditional on starting, 69% of students graduate with high grades (First or 2.1), 25% graduate with low grades and 6% fail to qualify.

Figure 2. : Perceived earnings and probability of employment by SES



Notes: Panel A shows average perceived difference in earnings (cond. on having a job) between (i) graduating from university with low grades and not going to university ('Low Grades') and (ii) graduating from university with high grades and not going to university ('High Grades'). Panel B shows the average perceived difference in the probability of being employed. High SES students are defined as those students who have at least one parent with university education. Stars indicate differences by SES: * p<0.10, ** p<0.05, *** p<0.01.

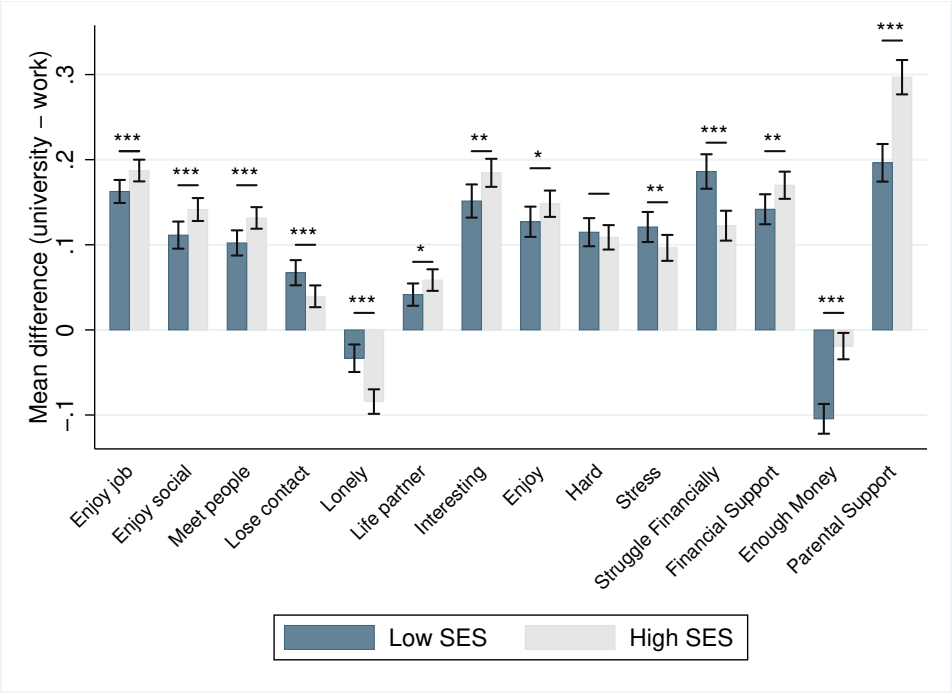
different scenarios, we find that both low and high SES students expect their labor market outcomes to improve with higher levels of achievement, but that this difference is more pronounced for high SES students (see Figure 2). To investigate this question in more detail, we calculate individual perceived returns in terms of labor market outcomes separately for each student (see Section 3.2). More specifically, we calculate the difference between what the student expects to earn at age 30 (conditional on having a job) if she decides to go to university and if she does not decide to go to university, taking into account students' perceptions about how likely it is that they will be in any given terminal node. On average, high SES students believe that their expected earnings will be £12,223 higher if they go to

university compared to if they start working instead (£37,090 vs. £24,867). This difference is significantly lower for low SES students who on average expect the earnings boost from going to university to be £8,242 (£31,208 vs. £22,966). We also find significant differences across socio-economic groups when we investigate the perceived returns to attending university in terms of the probability of being employed. While high SES students expect to be 15 percentage points more likely to have a job if they go to university compared to if they do not go to university (82% vs. 67%), low SES students perceive this difference to be only 10 percentage points (82% vs. 72%). Both of these differences in perceived returns across socio-economic groups are highly significant (at the 1% level).

The average beliefs about potential earnings in our sample are remarkably close to the earnings statistics we obtain from the Labour Force Survey (LFS). On average, students in our sample expect to earn £34,194 if they go to university (cond. on being employed), while they expect to earn £23,912 if they start working instead. Using information on individuals in the LFS aged 25-34 who are employed full-time, we document that the actual realized earnings are £33,642 for individuals who went to university while they are £24,752 for individuals who did not go to university. Whether or not realized outcomes differ across socio-economic groups is a question we discuss in Section 8.

Next we examine students' beliefs about the non-pecuniary benefits of university education and we investigate whether these beliefs differ significantly across socio-economic groups. The results are illustrated in Figure 3. Appendix Table D.7 also presents average beliefs about returns separately by socio-economic status. While both low and high SES students believe it is more likely they will enjoy the job they will be doing at age 30 if they go to university, the perceived return is larger for high SES individuals (p-value=0.010). We also find significant differences in how students with different socio-economic backgrounds perceive the returns in terms of different non-pecuniary outcomes during the 3-4 years after finishing secondary school. On average, students in both socio-economic groups believe that it is more likely they will enjoy their social life, will meet people they easily get along with and will meet their life partner if they go to university. They also think that it is less likely they will feel lonely but also more likely they will lose contact to their family/friends. Notice that there are significant

Figure 3. : Perceived differences in non-pecuniary factors by SES



Notes: The figure shows the average difference in beliefs (university - work) regarding non-pecuniary factors by socio-economic status. High SES students are defined as those students who have at least one parent with university education. Stars indicate differences by SES: * p<0.10, ** p<0.05, *** p<0.01.

differences across the two socio-economic groups in terms of how large these perceived non-pecuniary benefits/costs are. In particular, students with low SES backgrounds perceive the return in terms of the likelihood of enjoying their social life (p-value=0.004), meeting people (p-value=0.003), and meeting their life partner (p-value=0.068) to be significantly lower, while they perceive the cost in terms of feeling lonely (p-value<0.001) and losing contact to their family/friends (p-value=0.005) to be higher.

When it comes to the tasks and workload associated with the different choices, we find that both groups on average find the material/work tasks more interesting if they go to university. They also think they would enjoy studying/working more. At the same time, they believe it is more likely they will find the material hard or the workload too high and that they will feel stressed. Students with low SES perceive the benefit in terms of how interesting

(p-value=0.011) or enjoyable (p-value=0.077) the tasks are to be lower, and they perceive the cost in terms of the likelihood of feeling stressed to be higher (p-value=0.038). In terms of financial factors, students are more likely to think they will struggle financially if they go to university. They also think they will be more likely to be supported financially by their parents and less likely to have enough money to do what they enjoy. Again there are stark differences across socio-economic groups that are likely to reflect the availability of financial resources in the home. Compared to high SES students, low SES students are more likely to think they will struggle more financially if they go to university (p-value<0.001). The difference in the likelihood of receiving financial support from their parents is significantly lower (p-value=0.020), while the difference in terms of having enough money to do what they enjoy is significantly higher (p-value<0.001). Finally, we find large differences across socio-economic groups in terms of the difference in parental approval. Both groups believe that parents will approve of their choice more if they go to university, but this difference is 10 percentage points larger for high SES students (p-value<0.001), which is consistent with theories of identity (e.g. Akerlof and Kranton 2000).[27]

6. Choice Model Estimation

In the following, we first estimate the choice model presented in Section 2 assuming that preferences are homogeneous (Section 6.1). This allows us to investigate the relative importance of the different perceived pecuniary and non-pecuniary factors in the university attendance decision. We then estimate the choice model separately for students with low and high socio-economic status (Section 6.2), which allows us to examine whether low and high SES students differ in their preferences over the different attributes. Put differently, while in Section 5 we document that low and high SES students differ in their beliefs about the returns to university education, here we investigate whether low and high SES students additionally place differential weight on the various perceived returns.

[27]In Table D.7 we additionally report whether the variance in perceived returns differs significantly across socio-economic groups. We find there is substantial variation in perceived returns within both groups and that for some perceived returns (e.g. earnings, probability of having a job) the variance is higher for high SES students.

6.1. Homogeneous Preferences

We estimate the choice model assuming homogeneous preferences and present the quasi-maximum likelihood estimates of the parameters in column 1 of Table 2. The coefficients on expected income and the probability of having a job at age 30 are positive and significant, indicating that perceived pecuniary labor market returns play a significant role in the university education decision.[28] The relative magnitudes of $\{\triangle u_n\}_{n=1}^{N}$ show the importance of the binary outcomes in the choice. The most notable difference in utility levels we find is for enjoying the job one will be doing at age 30, which suggests that non-pecuniary career aspects play an important role for students. The second most important factor is gaining parental approval for the decision of whether or not to go to university; the coefficient is about 72% of the size of the coefficient of enjoying the job one will be doing at age 30. In addition, we also find that different aspects that relate to the students' lives during the 3-4 years after finishing school are significant determinants of the decision to go to university. Amongst the social factors, students' decisions are determined by their perceptions about whether they will enjoy the social life and activities they engage in, whether they will meet people with whom they easily get along with, and whether they will feel lonely but not by whether they think they will find a life partner. What also matters to students is their perception about whether they will find the material/work tasks interesting and whether they will enjoy studying/working. Interestingly, students' choices do not seem to be driven by perceptions about the probability of finding the material hard/workload too high or the probability of feeling stressed. Finally, we also find that it is important to students whether they will receive financial support from their parents and whether they will have enough money during the 3-4 years after they finish school to do what they enjoy. Overall, the results suggest that in addition to thinking about the pecuniary and non-pecuniary career benefits of university education, students also actively consider non-pecuniary benefits/costs that accrue during the time of the university education itself.

To get a better sense of the magnitudes of the estimated parameters, we calculate the

[28] As explained in Section 2, the coefficients are only identified up to a constant and therefore do not have a direct interpretation. We therefore interpret all coefficients relative to each other.

willingness-to-pay for each binary outcome, i.e. we calculate the amount of expected earnings that an individual would be willing to give up in order to experience that outcome. More specifically, the amount that an individual would be willing to forgo in yearly earnings at age 30 for a 1 percentage point change in the probability of the binary outcome n is:

$$WTP_n = \frac{0.01 \times \triangle u_n}{\gamma}.$$

We report the willingness-to-pay (WTP) for each outcome n in column 2 of Table 2. Standard errors are reported in parentheses.[29] Our results reveal that individuals are willing to forgo a substantial amount of earnings for a one percentage point change in the probability of different non-pecuniary outcomes. For example, individuals are willing to give up £1,138 of annual earnings at age 30 for a one percentage point increase in the probability of enjoying the job they will be doing at age 30, and they are willing to forgo £818 for a one percentage point increase in the probability of obtaining parental support in their choice. These results indicate that not only are the non-pecuniary benefits of university education significant determinants of the choice, but they are also economically relevant.

The relative importance of non-pecuniary factors in the university decision is also corroborated in the decomposition analysis we perform.[30] To illustrate the decomposition method, suppose that $Pr(j = 1) = F(Z\beta)$, where Z includes two sets of variables, Z_1 and Z_2. Given the parameter estimates $\hat{\beta}_1$ and $\hat{\beta}_2$, we can calculate the contribution of Z_1 as:

$$M_{Z_1} = ||\overline{Pr(j = 1|\hat{\beta}_1, \hat{\beta}_2)} - \overline{Pr(j = 1|\hat{\beta}_1 = 0, \hat{\beta}_2)}||$$

$$= \sqrt{\left[\sum_{i=1}^{N} \frac{Pr(j = 1|\hat{\beta}_1, \hat{\beta}_2)}{N} - \sum_{i=1}^{N} \frac{Pr(j = 1|\hat{\beta}_1 = 0, \hat{\beta}_2)}{N}\right]^2},$$

where the first term is the average probability of going to university as predicted by the choice model, and the second term is the average predicted probability of going to university

[29] We use the Delta method to calculate the standard errors of this nonlinear combination of estimators.

[30] This decomposition method has also been used by Zafar (2013).

if the set of variables Z_1 were not considered. We can then calculate the *relative* contribution of Z_1 to the choice as $R_{Z_1} = \frac{M_{Z_1}}{M_{Z_1}+M_{Z_2}}$.

For the purpose of the decomposition analysis, we group all variables related to the social life whilst at university (i.e. enjoy social life/activities, meet people, lose contact with family/friends, feel lonely, find life partner), all work/university specific factors (i.e. find material/work interesting, enjoy studying/working, find material hard/workload high, feel stressed), as well as all financial factors (i.e. struggle financially, have financial support from parents, have enough money). We then perform the decomposition analysis and present the results in column 1 of Table 3. Bootstrapped standard errors are reported in parentheses (based on 1,000 repetitions). Our results reveal that the pecuniary labor market returns in terms of expected earnings (conditional on having a job) and the probability of employment account for 12% and 7%, respectively. Both enjoyment of the job and parental support account for 22% each. Factors that are specific to the social life during the 3-4 years after finishing school account for 10%, while work/university specific factors and financial factors account for 14% and 12%, respectively. Overall, we conclude that while pecuniary factors do play a significant role in the choice, the relative importance of non-pecuniary factors is high.

Table 2—: Choice Model Estimation

	All		Low SES		High SES		Difference (p-value)	
	Choice model	WTP	Choice model	WTP	Choice model	WTP	Choice model	WTP
Expected earnings at age 30 ($1000\gamma_1$)	0.014***	-	0.011**	-	0.016***	-	0.42	-
(in £'000s)	(0.003)		(0.005)		(0.005)			
Employed at age 30 (Δu_1)	0.770***	£534.01**	0.699*	£649.69	0.952**	£592.28*	0.64	0.92
	(0.259)	(238.45)	(0.377)	(475.21)	(0.389)	(336.03)		
Enjoy job at age 30 (Δu_2)	1.642***	£1,138.24***	1.912***	£1,776.42*	1.259***	£782.95**	0.19	0.33
	(0.234)	(330.39)	(0.340)	(946.88)	(0.359)	(350.50)		
Enjoy social life and activities (Δu_3)	0.367*	£254.12*	0.476*	£441.77	0.387	£240.99	0.83	0.59
	(0.189)	(144.36)	(0.256)	(308.22)	(0.307)	(206.96)		
Meet people (Δu_4)	0.472**	£327.41**	0.666**	£619.10*	0.381	£236.90	0.49	0.37
	(0.204)	(159.44)	(0.289)	(373.43)	(0.302)	(202.18)		
Lose contact with family/friends (Δu_5)	0.248	£171.82	0.188	£174.79	0.481*	£299.40	0.40	0.66
	(0.167)	(121.43)	(0.229)	(220.59)	(0.260)	(182.81)		
Feel lonely (Δu_6)	-0.369**	-£255.83**	0.190	£176.53	-0.746***	-£464.12**	0.01	0.03
	(0.162)	(124.27)	(0.236)	(226.46)	(0.237)	(187.05)		
Find life partner (Δu_7)	0.117	£81.41	0.275	£255.90	0.002	£1.53	0.48	0.43
	(0.182)	(126.13)	(0.260)	(269.89)	(0.281)	(174.73)		
Find material/work interesting (Δu_8)	0.351**	£243.53**	0.404*	£375.08	0.418*	£260.20*	0.96	0.70
	(0.153)	(114.45)	(0.212)	(250.01)	(0.239)	(158.00)		
Enjoy studying/work (Δu_9)	0.990***	£686.45***	1.040***	£965.97*	0.899***	£558.92**	0.70	0.47
	(0.171)	(203.06)	(0.244)	(502.21)	(0.270)	(246.64)		
Find material hard/workload high (Δu_{10})	0.134	£92.77	-0.276	-£256.38	0.378	£234.79	0.08	0.10
	(0.173)	(121.43)	(0.249)	(246.46)	(0.273)	(174.60)		
Feel stressed (Δu_{11})	0.053	£36.57	0.233	£216.64	-0.141	-£87.40	0.32	0.32
	(0.179)	(124.48)	(0.264)	(260.70)	(0.268)	(165.33)		
Struggle financially (Δu_{12})	0.053	£36.53	0.078	£72.06	0.118	£73.63	0.88	0.99
	(0.133)	(92.84)	(0.184)	(177.71)	(0.212)	(132.48)		
Have financial support from parents (Δu_{13})	0.657***	£455.51***	0.461***	£428.05*	0.928***	£577.25***	0.07	0.65
	(0.125)	(136.85)	(0.178)	(250.57)	(0.187)	(207.54)		
Have enough money (Δu_{14})	-0.355**	-£246.38**	-0.272	-£252.97	-0.313	-£194.79	0.89	0.83
	(0.151)	(119.52)	(0.208)	(255.19)	(0.225)	(153.55)		
Have parental approval (Δu_{15})	1.181***	£818.49***	1.059***	£983.56**	1.255***	£780.32***	0.41	0.71
	(0.114)	(206.79)	(0.158)	(470.81)	(0.176)	(263.49)		
Observations	2540		1077		1334			

Notes: Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Column 1 presents the estimates of the choice model for the whole sample, while column 2 presents the willingness-to-pay calculations. Columns 3-4 present the results for students with low socio-economic status, columns 5-6 present the results for students with high socio-economic status, while columns 7-8 present the p-values for tests of equality of coefficients for the choice model parameter estimates and the WTP estimates between socio-economic groups. For the willingness to pay calculations, standard errors are calculated using the delta method. High SES students are defined as those students who have at least one parent with university education.

Table 3—: Decomposition Analysis

	All	Low SES	High SES	Δ SES
Expected earnings	0.1197	0.0841	0.1354	-0.0513***
	(0.0014)	(0.0020)	(0.0021)	(0.0029)
Employed at age 30	0.0737	0.0632	0.0915	-0.0283***
	(0.0009)	(0.0015)	(0.0014)	(0.0021)
Enjoy job at age 30	0.2218	0.2882	0.1477	0.1405***
	(0.0026)	(0.0070)	(0.0023)	(0.0074)
Social factors	0.1008	0.1231	0.1142	0.0089**
	(0.0012)	(0.0030)	(0.0018)	(0.0035)
Work/university specific factors	0.1421	0.1358	0.1351	0.0007
	(0.0017)	(0.0033)	(0.0021)	(0.0039)
Financial factors	0.1188	0.1225	0.1267	-0.0042
	(0.0014)	(0.0030)	(0.0020)	(0.0036)
Parental approval	0.2231	0.1831	0.2495	-0.0664***
	(0.0026)	(0.0044)	(0.0039)	(0.0059)
Observations	2540	1077	1334	

Notes: * p<0.10, ** p<0.05, *** p<0.01. This table shows the results of the decomposition analysis. Column 1 gives the decomposition results for the full sample, while columns 2 and 3 present the results for low and high socio-economic status students, respectively. '*Social factors*' comprise enjoying the social life/activities, meeting people, feeling lonely, losing contact to family/friends and meeting a life partner. '*Work/university specific factors*' comprise finding the work/university tasks interesting, enjoyable, difficult and stressful. '*Financial factors*' comprise struggling financially, receiving financial support from parents and having enough money. Bootstrapped standard errors calculated from 1,000 repetitions are provided in parentheses as percentages. High SES students are defined as those students who have at least one parent with university education.

6.2. Heterogeneous Preferences by Socio-economic Status

Having documented which factors play a role in students' decisions to attend university when we assume homogeneous preferences, we now turn to the question of whether students from low and high SES backgrounds differ in their preferences over the different attributes. For this purpose, we split the sample by whether or not students have at least one parent who attended university, and we estimate the choice model and perform the decomposition analysis separately for low and high SES students. The results are reported in columns 3-6 of Table 2 and columns 2-3 of Table 3, respectively.

While most of the differences between the estimated preference parameters are statistically insignificant (see columns 7-8 of Table 2), possibly due to a lack of power, some of the point estimates differ substantially across socio-economic groups. For example, while low SES students are willing to forgo £1,776 in earning at age 30 for a 1 percentage point increase in the probability of enjoying their future job, high SES students are only willing to forgo £783 in earnings. The difference, however, is not statistically significant at conventional levels (p=0.19). Using these estimated coefficients, we separately perform the decomposition analysis for low and high SES students (see columns 2-3 of Table 3). Perceived pecuniary labor market returns (i.e. expected earnings and probability of employment) explain a significantly higher proportion of the variance for high SES students than for low SES students. For low SES students, the perceived pecuniary labor market returns account for 15% of the variation, whereas for high SES students they account for 23%. Interestingly, the pattern is reversed when we investigate the proportion of the variance which can be explained by perceived returns in terms of job enjoyment. While for low SES students 29% of the variation can be accounted for by beliefs about the probability of enjoying one's future job, the corresponding number for high SES students is only 15%, indicating that the non-pecuniary labor market benefits of university education are more relevant in the decisions of low SES students. Moreover, we find that parental approval in the choice is significantly more important for high SES than for low SES students (25% vs. 18%). We find no difference in terms of the importance of work/university specific factors and financial factors, and only small differences in terms of the importance of social factors (12% vs. 11%). The results suggest that the motives for educational attainment do differ significantly across socio-economic groups. As a result, informational interventions may have differential effects on students from different backgrounds. If students were misinformed about the earnings of university graduates or the types of jobs available to university graduates, then informational interventions which target students' beliefs about their likely earnings are more likely to be effective for high SES students while informational interventions that inform students about the variety of jobs available to university graduates may be more effective for low SES students, for whom the non-pecuniary labor market benefits play a larger role in the decision.

7. Policy Simulations

There are two main reasons why our model predicts a socio-economic gap in university enrollment. First, students with low socio-economic status hold different beliefs about the pecuniary and non-pecuniary benefits of university education compared to their more advantaged peers (see Section 5). Second, the point estimates of the parameters of the choice model differ across socio-economic groups, which suggests that students of different socio-economic background may place different weight on the different pecuniary/non-pecuniary factors affecting the choice (see Section 6). While beliefs are likely to be malleable and may be affected through targeted interventions, it is likely to be much harder to influence students' preferences. It is therefore important to understand to what extent the socio-economic gap in university enrollment would be reduced if students from different SES backgrounds had the same beliefs and by how much it would be reduced if they had the same preferences.

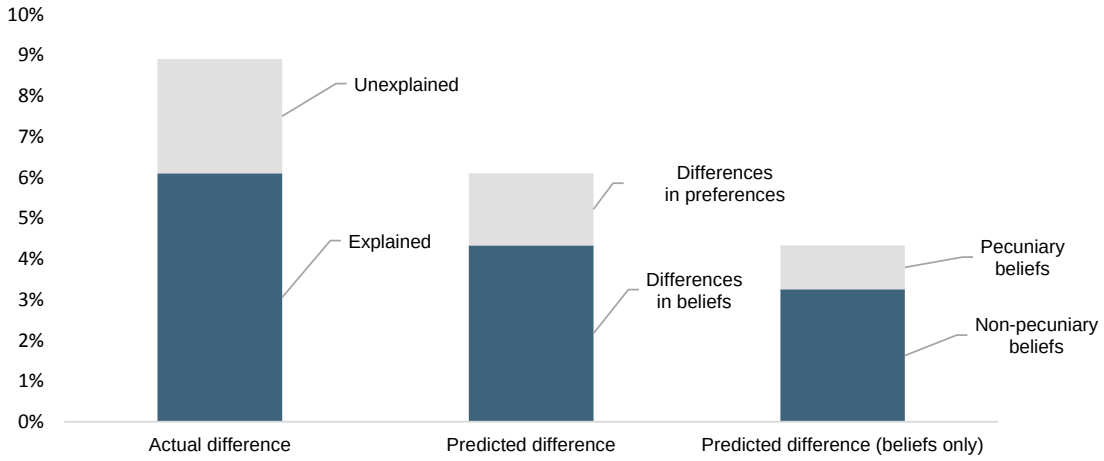
For this purpose, we decompose the predicted socio-economic gap into a composition effect (covariate effect) and a structural effect (coefficients effect). While the former captures the effect of differences in beliefs across low and high SES students, the latter captures the effect of differences in preference parameter estimates across the two groups. Let D_i denote the socio-economic group that individual i belongs to, where $D_i = L$ for low SES students and $D_i = H$ for high SES students. Using the law of iterated expectations, the average outcome gap predicted by the model, $\triangle^{SES}$, can be written as:

$$\begin{aligned}\Delta^{SES} &= E[\Lambda(Z_i\beta^H) \mid D_i = H] - E[\Lambda(Z_i\beta^L) \mid D_i = L] \\ &= \underbrace{E[\Lambda(Z_i\beta^H) \mid D_i = H] - E[\Lambda(Z_i\beta^L) \mid D_i = H]}_{\Delta^{\beta} \text{ (structural effect)}} + \underbrace{E[\Lambda(Z_i\beta^L) \mid D_i = H] - E[\Lambda(Z_i\beta^L) \mid D_i = L]}_{\Delta^{Z} \text{ (composition effect)}},\end{aligned}$$

where β^L and β^H are the preference parameter estimates for low and high SES students, respectively. Using the parameter estimates we obtain for low and high SES students in Section 6, we first document that the socio-economic gap in students' intentions to go to university predicted by our model is 6.2 percentage points, which is 69% of the actual socio-

economic gap in intentions (actual gap: 8.9 percentage points).[31] We then perform the decomposition analysis and find that of the 6.2 percentage point gap predicted by the model, 71% can be explained by socio-economic differences in beliefs (composition effect) while only 29% can be accounted for by differences in preferences (structural effect). Overall, differences in beliefs across socio-economic groups can hence explain 49% of the actual gap in students' intentions to go to university. These results suggest that policy interventions which target students' beliefs have the potential of substantially reducing the socio-economic gap in university enrollment. We illustrate these results in Figure 4.[32]

Figure 4. : Decomposing model predictions compared to actual data



Note: The first column decomposes the actual difference between low and high SES respondents' stated intention to attend university into that which can be predicted by the model and that which is unexplained. The second column decomposes the predicted SES gap into differences in preferences and differences in beliefs. High SES students are defined as those students who have at least one parent with university education.

Given the importance of beliefs in explaining the socio-economic gap, a natural question

[31] We note that 31% of the socio-economic gap remains unexplained. Further research will be needed into which other factors can explain the socio-economic gap in university attendance.

[32] We can also ask the question to what extent the socio-economic gap in intentions to go to university can be explained by differences in beliefs across socio-economic groups if we assume that preferences are homogeneous (see Section 6.1). The predicted socio-economic gap in a model with homogeneous preferences is 5.04 percentage points, so the model can explain 57% of the actual socio-economic gap in students' intentions to go to university.

to ask is whether policy interventions targeting a certain set of beliefs are likely to be more effective in narrowing the SES gap compared to interventions targeting a different set of beliefs. Is an intervention likely to be more effective if it targets students' beliefs about the pecuniary returns of education or if it targets students' beliefs about specific non-pecuniary benefits? To investigate this question, we follow the approach developed in Kaiser (2015).[33] For each perceived return k, we examine by how much the socio-economic gap would be reduced if low SES students had the same distribution of beliefs as high SES students.

The results are presented in Table 4. Several findings emerge regarding which shifts in beliefs would lead to a significant reduction in the socio-economic gap. In particular, if low SES students had the same distribution of perceived returns in terms of earnings (conditional on having a job) as high SES students, the SES gap would be significantly reduced by 0.63 percentage points. Shifting low SES students' beliefs about the returns to university education in terms of job enjoyment would significantly reduce the gap by 0.75 percentage points. Equalizing students' beliefs about the perceived benefits in terms of enjoying one's social life/activities and meeting people one easily gets along with would reduce the gap by 0.26 and 0.34 percentage points, respectively, while equalizing students' beliefs about the probability of finding the material/work tasks interesting and enjoyable would reduce the gap by 0.27 and 0.44 percentage points. We note that equalizing students' beliefs about the probability of struggling financially, having financial support from their parents, or having enough money to do what they enjoy does *not* lead to a significant reduction in the socio-economic gap, indicating that students' perceptions about their finances during the 3-4 years after finishing secondary school do not drive socio-economic differences in enrollment. Overall, out of the socio-economic gap that can be explained by differences in beliefs across socio-economic groups, we find that 25% can be explained by differences in students' beliefs about pecuniary labor market returns, while 75% can be explained by differences in beliefs about the non-pecuniary benefits we elicit, suggesting that socio-economic differences in beliefs

[33]While detailed decompositions are straightforward in linear models (Oaxaca 1973; Blinder 1973), they are more complicated in non-linear models. Relative to previous work (e.g. Fairlie 2005; Yun 2004), the approach developed by Kaiser (2015) has the advantage that it is path independent and that it takes higher order moments of Z_i into account and not only differences in means. More details can be found in the Appendix.

about non-pecuniary returns are important for explaining the socio-economic gap in students' decisions to go to university.[34]

Table D.10 presents the socio-economic gaps and decompositions for the whole sample as well as for different subgroups of the population. Our results are robust to restricting the sample to students who are in their final year, who attend mixed schools and who received A*-B grades in both Mathematics and English Literature in their Year 11 GCSE exams. We also note that the socio-economic gap in intentions to go to university is about twice as large for boys as it is for girls (13.0 percentage points vs. 6.5 percentage points) and that socio-economic differences in beliefs about the non-pecuniary benefits of university education can explain a larger share of the socio-economic gap for boys than for girls (42% for boys and 29% for girls).

Our results suggest that while interventions that target low SES students' beliefs about pecuniary returns (e.g. interventions informing students about the earnings of university graduates or interventions that boost students' confidence in their ability to do well whilst at university) have some potential in reducing the socio-economic gap in attendance, interventions that shift students' beliefs about non-pecuniary benefits may even be more effective. In particular, the results indicate that providing students with information about the variety of jobs available to university graduates or the variety of social activities available at universities may significantly reduce the socio-economic gap in university enrollment. Whether or not such interventions will be effective will depend on whether or not students misperceive the returns to education due to a lack of information, a question we now discuss in the following section.

8. Discussion

The results of our analysis suggest that differences in students' beliefs about the returns to university education can explain a substantial part of the gap in university enrollment between students who would be first generation university students and students whose parents

[34]Note that because students' beliefs about the financial factors do not contribute to the socio-economic gap, the results are not susceptible to whether or not we classify these financial factors as pecuniary or non-pecuniary.

Table 4—: Detailed decomposition of predicted SES gap

SES gap predicted by model	6.150
Contribution of differences in beliefs to predicted gap	4.350***
Expected earnings at age 30 (γ_1)	0.630**
	(0.28)
Employed at age 30 (Δu_1)	0.481
	(0.30)
Enjoy job at age 30 (Δu_2)	0.748***
	(0.28)
Enjoy social life and activities (Δu_3)	0.263*
	(0.16)
Meet people (Δu_4)	0.340*
	(0.17)
Lose contact with family/friends (Δu_5)	-0.087
	(0.11)
Feel lonely (Δu_6)	-0.169
	(0.23)
Find life partner (Δu_7)	0.082
	(0.10)
Find material/work tasks interesting (Δu_8)	0.268*
	(0.15)
Enjoy studying/work (Δu_9)	0.448*
	(0.25)
Find material hard/workload high (Δu_{10})	0.057
	(0.09)
Feel stressed (Δu_{11})	-0.130
	(0.15)
Struggle financially (Δu_{12})	-0.084
	(0.23)
Have financial support from parents (Δu_{13})	0.168
	(0.12)
Have enough money (Δu_{14})	-0.359
	(0.26)
Have parental support in your choice (Δu_{15})	1.693***
	(0.31)

Notes: Of the actual SES gap of 8.9 percentage points, 6.15 percentage points can be predicted by a model allowing for heterogeneous preferences. Of the 6.15 percentage points a total of 4.35 percentage points can be explained by differences in beliefs. The relative contribution of beliefs is further decomposed in rows 3 to 18. Bootstrapped standard errors calculated from 200 repetitions are provided in parentheses.

have attended university. An important question which emerges is to what extent student beliefs are correct. In other words, do students whose parents did not attend university only *perceive* the pecuniary and non-pecuniary returns to university education to be lower, or are these returns actually lower for low SES students? Estimating the causal effects of going to university is arguably a challenging task, and there is is little evidence on how the pecuniary and non-pecuniary returns to education differ across socio-economic groups. We do note, however, that differences in realized outcomes are different across students with different socio-economic background characteristics. While this evidence is solely indicative, it would suggest that actual returns are lower for low SES students if low SES students who decided to go to university were positively selected on gains.

We use data from the UK Labour Force Survey (LFS) 2015 to compare how the realized earnings of individuals differ by education and parental background. The LFS data contain (i) detailed information on individual earnings, (ii) information on whether or not the individual went to university and, if yes, whether or not the individual graduated with high grades (a First or a 2.1), and (iii) information on parental occupations. We use information on the occupation of the main earner in the household when the individual was 14 years old to construct a measure of socio-economic status.[35] Using earnings data on all 25-34 year-old individuals working full-time, we document that high SES university graduates who obtained high grades earn £9,834 more than high SES individuals without university education. In contrast, low SES individuals who graduated from university with high grades only earn £8,017 more than low SES individuals who did not obtain a university degree. These differences may arise because different types of students self-select into university education or because of other differences between the two groups. For example, Crawford et al. (2016) show that low SES students in the UK are significantly less likely to attend high status uni-

[35]More specifically, high SES families are defined as those for whom the main wage earner fell into categories 1, 2 or 3 of the ONS standard occupation classification (which includes (1) Managers, Directors And Senior Officials, (2) Professional Occupations and (3) Associate Professional And Technical Occupations). Individuals from low SES families are categorised as those for whom the main wage earner fell into categories 4 - 9 of the classification (which includes (4) Administrative And Secretarial Occupations, (5) Skilled Trades Occupations, (6) Caring, Leisure And Other Service Occupations, (7) Sales And Customer Service Occupations (8) Process, Plant And Machine Operatives and (9) Elementary Occupations).

versities, a finding that is consistent with evidence from other countries such as the US (see Chetty et al. 2017). It may also be that students from low SES backgrounds have different access to networks, which facilitate finding a job in a high status occupation (see Corak and Piraino 2011).

Regarding the non-pecuniary benefits of university education, recent psychology studies document that first-generation students (i.e. students whose parents did not attend university) often experience difficulties fitting in with the social environment at universities and experience subjective feelings of detachment (e.g. Stephens et al. 2012, Janke et al. 2017), in addition to experiencing financial difficulties and having to work more hours (Walpole 2013).[36] These results are consistent with our finding that low SES students perceive the likelihood of enjoying their social life at university to be significantly lower. To fully understand what deters high achieving low SES students from going to university, it seems crucial to obtain a better understanding of how student experiences at university differ by SES. From a policy perspective, it is also important to understand which interventions may be effective in narrowing the gap in actual student experiences whilst students are at university.

While it may be that student beliefs are rational, it might also be that low SES students hold different beliefs from high SES students because they have less access to information. In Tables D.11 and D.12 we document which individual and local characteristics predict individual beliefs about returns. Not only is parental education significantly related to students' perceptions about the returns to university education, but student beliefs about the pecuniary and non-pecuniary benefits of university education are also increasing in the number of individuals the student knows whom he/she can ask about university. Relating our data to regional data from the 2011 census, we also find that student beliefs are significantly related to local area characteristics. Students in local areas with high unemployment rates perceive the returns to education in terms of the likelihood of having a job at age 30 to be significantly higher. Likewise, students in local areas in which the percentage of individuals with a university degree is higher perceive the returns in terms of earnings (conditional on having a

[36] In our data, we also find that low SES students perceive the likelihood of having to work whilst being at university as significantly higher (71% vs 63%).

job), the likelihood of having a job, and enjoying the job one will be doing to be significantly greater. While this evidence is merely suggestive, it is consistent with the argument that low SES students have limited information about the returns to university education as well as about what university life is likely to be like. Recent studies have shown that providing students with information on population earnings statistics conditional on different educational outcomes can have an impact on students' decisions to pursue further education (e.g. Nguyen 2008, Jensen 2010, Bleemer and Zafar 2015, Kerr et al. 2015, Hastings, Neilson and Zimmerman 2017).[37] More research will be needed on how student beliefs are formed and whether informational interventions that inform students about the pecuniary and especially the non-pecuniary benefits of university education can narrow the socio-economic gap in university enrollment.

9. Conclusion

In this paper, we investigate to what extent the socio-economic gap in students' intentions to go to university can be explained by differences in students' beliefs about the returns to university education. In order to address this question, we collect survey data on students' beliefs about the pecuniary and non-pecuniary benefits of going to university. We also elicit students' intentions to go to university and show that students' intentions are highly correlated with students' actual application decisions. By surveying students prospectively, we minimize biases that can arise due to ex-post rationalization.

We document that students who would belong to the first generation in their family to go to university perceive both the pecuniary as well as the non-pecuniary benefits of university education to be significantly lower. These differences in beliefs can explain 49% of the entire socio-economic gap in students' intentions to go to university. Among the non-pecuniary factors, we find that expected job satisfaction, parental approval and students' perceptions of the quality of their social life during the 3-4 years after they finish school play an important

[37] Alan, Boneva and Ertac (2015) also provide evidence that an educational intervention delivered by teachers in the classroom can affect students' beliefs about the malleability of skills/the productivity of effort and change students' achievements as a result. Recent work also stresses the importance that external rewards for achievement might have for students' beliefs about the returns to educational investments (Sequeira, Spinnewijn and Xu 2016).

role in the choice and in explaining the socio-economic gap. Perceived monetary returns are also found to play a role, but to a lesser extent. Differences in students' beliefs about the non-pecuniary benefits of university education alone can account for 37% of the overall socio-economic gap.

While traditional policies have mainly focused on relieving credit constraints for low socio-economic status students, our results suggest that policies which target students' beliefs about the pecuniary and especially the non-pecuniary benefits of university education may be very effective in narrowing the socio-economic gap. Whether or not informational interventions that target students' beliefs about non-pecuniary returns are effective is an important open question that future research should address with the help of randomized control trials. In addition, it will be equally important to understand to what extent the pecuniary and non-pecuniary returns to university education actually differ across socio-economic groups and how universities can provide better support to low socio-economic status students to attract and retain the best talent.

REFERENCES

Abramitzky, Ran, and Victor Lavy. 2014. "How responsive is investment in schooling to changes in redistributive policies and in returns?" *Econometrica*, 82(4): 1241–1272.

Akerlof, George A, and Rachel E Kranton. 2000. "Economics and identity." *The Quarterly Journal of Economics*, 115(3): 715–753.

Alan, Sule, Teodora Boneva, and Seda Ertac. 2015. "Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit." HCEO Working Paper No. 2015-009.

Almas, Ingvild, Alexander Cappelen, Kjell Salvanes, Erik Sorensen, and Bertil Tungodden. 2016. "What explains the gender gap in college track dropout? Experimental and administrative evidence." *American Economic Review: Papers & Proceedings*, 106(5): 296–302.

Alter, Molly, and Randall Reback. 2014. "True for your school? How changing reputations alter demand for selective US colleges." *Educational Evaluation and Policy Analysis*,

20(10): 1–25.

Angrist, Josh, David Autor, Sally Hudson, and Amanda Pallais. 2016. "Evaluating post-secondary aid: Enrollment, persistence, and projected completion effects." NBER Working Paper 23015.

Arcidiacono, Peter. 2004. "Ability Sorting and the Returns to College Major." *Journal of Econometrics*, 121(1): 343–375.

Arcidiacono, Peter, V Joseph Hotz, and Songman Kang. 2012. "Modeling College Major Choices using Elicited Measures of Expectations and Counterfactuals." *Journal of Econometrics*, 166(1): 3–16.

Arcidiacono, Peter, V Joseph Hotz, Arnaud Maurel, and Teresa Romano. 2014. "Recovering ex ante returns and preferences for occupations using subjective expectations data." National Bureau of Economic Research.

Attanasio, Orazio, and Katja Kaufmann. 2014. "Education Choices and Returns to Schooling: Intra-household Decision Making, Gender and Subjective Expectations." *Journal of Development Economics*, 109: 203–216.

Attanasio, Orazio, and Katja Kaufmann. 2017. "Education choices and returns on the labor and marriage markets: Evidence from data on subjective expectations." Unpublished manuscript.

Austen-Smith, David, and Roland G Fryer Jr. 2005. "An economic analysis of "acting white"." *The Quarterly Journal of Economics*, 120(2): 551–583.

Bailey, Martha, and Susan Dynarski. 2011. "Inequality in Postsecondary Education." In: G. Duncan and R. Murnane, eds., Whither Opportunity: Rising Inequality, Schools, and Children's Life Chances. New York: Russell Sage Foundation.

Baker, Rachel, Eric Bettinger, Brian Jacob, and Ioana Marinescu. 2017. "The effect of labor market information on community college students' major choice." NBER Working Paper 23333.

Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova. 2012. "Female leadership raises aspirations and educational attainment for girls." *Science*, 335(6068): 582–586.

Beffy, Magali, Denis Fougere, and Arnaud Maurel. 2012. "Choosing the field of study in postsecondary education: Do expected earnings matter?" *Review of Economics and Statistics*, 94(1): 334–347.

Belfield, Chris, Teodora Boneva, Christopher Rauh, and Jonathan Shaw. 2017. "What drives enrollment gaps in further education?" Unpublished manuscript.

Belley, Philippe, and Lance Lochner. 2007. "The changing role of family income and ability in determining educational achievement." *Journal of Human Capital*, 1(1): 37–89.

Benitez-Silva, Hugo, Moshe Buchinsky, Hiu Man Chan, Sofia Cheidvasser, and John Rust. 2004. "How large is the bias in self-reported disability?" *Journal of Applied Econometrics*, 19: 649–670.

Bertrand, Marianne, and Sendhil Mullainathan. 2001. "Do people mean what they say? Implications for subjective survey data." *American Economic Review Papers & Proceedings*, 91(2): 67–72.

Bettinger, Eric, Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. "The role of application assistance and information in college decisions: Results from the H&R Block FAFSA Experiment." *Quarterly Journal of Economics*, 127(3): 1205–1242.

Black, Sandra E., and Paul J. Devereux. 2011. "Recent Developments in Intergenerational Mobility." *Handbook of Labor Economics*, 4(b): 1487–1541.

Blanden, Jo, and Paul Gregg. 2004. "Family income and educational attainment: a review of approaches and evidence for Britain." *Oxford Review of Economic Policy*, 20(2): 245–263.

Blanden, Jo, and Stephen Machin. 2004. "Educational inequality and the expansion of UK higher education." *Scottish Journal of Political Economy*, 51(2): 230–249.

Blanden, Jo, Paul Gregg, and Lindsey Macmillan. 2007. "Accounting for intergenerational income persistence: noncognitive skills, ability and education." *The Economic Journal*, 117(519): C43–C60.

Bleemer, Zach, and Basit Zafar. 2015. "Intended College Attendance: Evidence from an Experiment on College Returns and Costs." Working Paper.

Blinder, Alan. 1973. "Wage discrimination: reduced form and structural estimates." *Journal of Human Resources*, 8: 436–55.

Carneiro, Pedro, and James J. Heckman. 2002. "The evidence on credit constraints in post-secondary schooling." *The Economic Journal*, 112: 705–736.

Carneiro, Pedro, James J Heckman, and Edward J Vytlacil. 2011. "Estimating Marginal Returns to Education." *American Economic Review*, 101: 2754–2781.

Carneiro, P, KT Hansen, and JJ Heckman. 2003. "Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice." *International Economic Review*, 44(2): 361–422.

Checchi, Daniele, Carlo V Fiorio, and Marco Leonardi. 2013. "Intergenerational persistence of educational attainment in Italy." *Economics Letters*, 118(1): 229–232.

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

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014*a*. "Where is the land of opportunity? The geography of intergenerational mobility in the United States." *The Quarterly Journal of Economics*, 129(4): 1553–1623.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. 2014*b*. "Is the United States still a land of opportunity? Recent trends in intergenerational mobility." *The American Economic Review*, 104(5): 141–147.

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

Corak, Miles, and Patrizio Piraino. 2011. "The intergenerational transmission of employers." *Journal of Labor Economics*, 29(1): 37–68.

Crawford, Claire, Lorraine Dearden, John Micklewright, and Anna Vignoles. 2016. *Family background and university success: differences in Higher Education access and outcomes in England.*

Cunha, Flavio, and James Heckman. 2008. "A new framework for the analysis of inequality." *Macroeconomic Dynamics*, 12(S2): 315–354.

Cunha, Flavio, and James J Heckman. 2007. "Identifying and estimating the distribu-

tions of ex post and ex ante returns to schooling." *Labour Economics*, 14(6): 870–893.

Cunha, Flavio, James Heckman, and Salvador Navarro. 2005. "Separating uncertainty from heterogeneity in life cycle earnings." *Oxford Economic Papers*, 57(2): 191–261.

Cunha, Flavio, James J Heckman, and Salvador Navarro. 2006. "Counterfactual Analysis of Inequality and Social Mobility." *Mobility and Inequality: Frontiers of Research in Sociology and Economics*, 290–348.

Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy Masterov. 2006. "Interpreting the evidence on life cycle skill formation." *Handbook of the Economics of Education*, 1(12): 697–812.

Delavande, Adeline, and Basit Zafar. 2014. "University choice: the role of expected earnings, non-pecuniary outcomes, and financial constraints." *FRB of New York Staff Report*, 683.

Department for Education Statistical Fiscal Releases, SFR. 45/2016. "Participation Rates in Higher Education." *Academic Years 2006/7 - 2014/15*.

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

Fairlie, Robert. 2005. "An extension of the Blinder-Oaxaca decomposition technique to logit and probit models." *Journal of Economic and Social Measurement*, 30: 305–16.

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde. 2016. "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences." Working Paper.

Festinger, Leon. 1957. *A Theory of Cognitive Dissonance.* Stanford: Stanford University Press.

Field, Erica. 2009. "Educational debt burden and career choice: Evidence from a financial aid experiment at NYU Law School." *American Economic Journal: Applied Economics*.

Gayle, Vernon, Damon Berridge, and Richard Davies. 2002. "Young people's entry into higher education: quantifying influential factors." *Oxford Review of Education*, 28(1): 5–20.

Giustinelli, Pamela. 2016. "Group Decision Making with Uncertain Outcomes: Unpacking

Child-Parent Choice of the High School Track." *International Economic Review*, 57(2).

Hastings, Justine, Christopher Neilson, and Seth Zimmerman. 2017. "The effect of earnings disclosure on college enrollment decisions." NBER Working Paper 21300.

Hastings, Justine, Christopher Neilson, Anely Ramirez, and Seth Zimmerman. 2016. "(Un)informed college and major choice: Evidence from linked survey and administrative data." *Economics of Education Review*, 51: 136–151.

Heckman, James J., John Eric Humphries, and Gregory Veramendi. 2016. "Returns to Education: The Causal Effects of Education on Earnings, Health and Smoking." NBER Working Paper 22291.

Heckman, James J, Lance J Lochner, and Petra E Todd. 2006. "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond." In *Handbook of the Economics of Education*. Vol. 1, , ed. Erik A Hanushek and Finis Welch, 307–458. Elsevier.

Hoxby, Caroline, and Sarah Turner. 2013. "Expanding college opportunities for high-achieving, low income students." *SIEPR Discussion Paper 12-014*.

Jacob, Brian, and Tamara Linkow. 2011. "Educational expectation and attainment." In G. Duncan and R. Mumane, eds., Whither Opportunity? Rising Inequality and the Uncertain Life Changes of Low-Income Children, New York: Russell Sage Press.

Jacob, Brian, Brian McCall, and Kevin M Stange. 2013. "College as Country Club: Do Colleges Cater to Students' Preferences for Consumption?" National Bureau of Economic Research.

Janke, Stefan, Selma C Rudert, Tamara Marksteiner, and Oliver Dickhäuser. 2017. "Knowing one's place: parental educational background influences social identification with academia, test anxiety, and satisfaction with studying at university." *Frontiers in psychology*, 8.

Jensen, Robert. 2010. "The (perceived) returns to education and the demand for schooling." *The Quarterly Journal of Economics*, 125(2): 515–548.

Juster, Thomas. 1966. "Consumer buying intentions and purchase probability: An experiment in survey design." *Journal of the American Statistical Association*, 61: 658–696.

Kaiser, Boris. 2015. "Detailed decompositions in nonlinear models." *Applied Economics Letters*, 22(1): 25–29.

Kane, Thomas J, and Cecilia Elena Rouse. 1995. "Labor-market returns to two-and four-year college." *The American Economic Review*, 85(3): 600–614.

Kaufmann, Katja. 2014. "Understanding the Income Gradient in College Attendance in Mexico: The Role of Heterogeneity in Expected Returns." *Quantitative Economics*, 5(3): 583–630.

Keane, Michael P, and Kenneth I Wolpin. 2001. "The effect of parental transfers and borrowing constraints on educational attainment." *International Economic Review*, 42(4): 1051–1103.

Kerr, Sari Pekkala, Tuomas Pekkarinen, Matti Sarvimaki, and Roope Uusitalo. 2015. "Post-secondary education and information on labor market prospects: a randomized field experiment." IZA Discussion Paper 9372.

Lochner, Lance, and Alexander Monge-Naranjo. 2012. "Credit constraints in education." *Annual Review of Economics*, 4: 225–256.

Machin, Stephen, and Anna Vignoles. 2004. "Educational inequality: The widening socio-economic gap." *Fiscal Studies*, 25(2): 107–128.

Manski, Charles. 1990. "The use of intentions data to predict behavior: A best case analysis." *Journal of the American Statistical Association*, 85: 934–940.

Manski, Charles. 2004. "Measuring Expectations." *Econometrica*, 72(5): 1329–1376.

Montmarquette, Claude, Kathy Cannings, and Sophie Mahseredjian. 2002. "How Do Young People Choose College Majors?" *Economics of Educaion Review*, 21: 543–556.

Nguyen, Trang. 2008. "Information, role models and perceived returns to education: experimental evidence from Madagascar." Working Paper.

Oaxaca, Ronald. 1973. "Male-female wage differentials in urban labor markets." *International Economic Review*, 14: 693–709.

Oreopoulos, Philip, and Kjell G Salvanes. 2011. "Priceless: The nonpecuniary benefits of schooling." *The Journal of Economic Perspectiveso*, 25(1): 159–184.

Oreopoulos, Philip, and Uros Petronjievic. 2013. "Who benefits from college? A review

of research on the returns to higher education." *The Future of Children*, 23(1): 41–65.

Pallais, Amanda. 2015. "Small differences that matter: Mistakes in applying to college." *Journal of Labor Economics*, 33(2): 493–520.

Papke, Leslie, and Jeffrey Wooldridge. 1996. "Econometric methods for fractional response variables with an application to 401(k) plan participation rates." *Journal of Applied Econometrics*, 11: 619–632.

Pope, Devin G, and Jaren C Pope. 2009. "The Impact of College Sports Success on the Quantity and Quality of Student Applications." *Southern Economic Journal*, 75(3): 750–780.

Reay, Diane, Gill Crozier, and John Clayton. 2010. "'Fitting in'or 'standing out': working-class students in UK higher education." *British Educational Research Journal*, 36(1): 107–124.

Sequeira, Sandra, Johannes Spinnewijn, and Guo Xu. 2016. "Rewarding school success and perceived returns to education: evidence from India." *Journal of Economic Behavior and Organization*, 131: 373–392.

Stephens, Nicole, Sarah Townsend, Hazel Markus, and Taylor Phillips. 2012. "A cultural mismatch: Independent cultural norms produce greater increases in cortisol and more negative emotions among first-generation college students." *Journal of Experimental Social Psychology*, 48(6): 1389–1393.

Stifterverband. 2016. "In Zusammenarbeit mit McKinsey. Hochschul-Bildungs-Report 2020. Hochschulbildung für die Arbeitswelt 4.0. Jahresbericht 2016."

Stinebrickner, Ralph, and Todd R Stinebrickner. 2014. "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout." *The Review of Economic Studies*, 81(1): 426–472.

Walpole, Mary Beth. 2013. "Socioeconomic Status and College: How SES Affects College Experiences and Outcomes." *The Review of Higher Education*, 27(1): 45–73.

Wiswall, Matthew, and Basit Zafar. 2015*a*. "Determinants of college major choice: Identification using an information experiment." *The Review of Economic Studies*, 82(2): 791–824.

Wiswall, Matthew, and Basit Zafar. 2015*b*. "How do college students respond to public information about earnings?" *Journal of Human Capital*, 9(2): 117–169.

Wiswall, Matthew, and Basit Zafar. 2017. "Preference for the workplace, human capital and gender." *Quarterly Journal of Economics*.

Yun, Myeong-Su. 2004. "Decomposing differences in the first moment." *Economics Letters*, 82: 275–280.

Zafar, Basit. 2011. "Can subjective expectations data be used in choice models? Evidence on cognitive biases." *Journal of Applied Econometrics*, 26(3): 520–544.

Zafar, Basit. 2013. "College major choice and the gender gap." *Journal of Human Resources*, 48(3): 545–595.

Appendix A: Derivation of the Choice Model

To derive expression (1), we note that, assuming utility is additively separable across the outcomes, we can write $U_i(b, c, X_i)$ as:

$$U_i(b, c, X_i) = \sum_{n=1}^{N} u_n(b_n, X_i) + \sum_{m=1}^{M} \gamma_m(X_i) c_m + \varepsilon_{ij}$$

where $u_n(b_n, X_i)$ is the utility associated with the binary outcome b_n for an individual with characteristics X_i, $\gamma_m(X_i)$ is a constant for the continuous outcome c_m for an individual with characteristics X_i, and ε_{ij} is a random error term. The utility individual i derives from (b, c) is the same for all individuals who have the same characteristics X_i, up to the random error term. The SEU of alternative j can now be re-written as:

$$SEU_{ij} = \sum_{n=1}^{N} \int u_n(b_n, X_i) dP_{ij}(b) + \sum_{m=1}^{M} \gamma_m(X_i) \int c_m dP_{ij}(c) + \varepsilon_{ij}$$

For the binary outcomes, we can further re-write the expression. In particular, we know that $dP_{ij}(b) = P_{ij}(b_n = 1)$ if $b_n = 1$ and $dP_{ij}(b) = P_{ij}(b_n = 0)$ if $b_n = 0$. We can therefore re-write the first expression as:

$$\int u_n(b_n, X_i) dP_{ij}(b) = P_{ij}(b_n = 1) u_n(b_n = 1, X_i) + (1 - P_{ij}(b_n = 1)) u_n(b_n = 0, X_i)$$

$$= P_{ij}(b_n = 1) [\underbrace{u_n(b_n = 1, X_i) - u_n(b_n = 0, X_i)}_{\triangle u_n(X_i)}] + u_n(b_n = 0, X_i)$$

Note that we directly elicit individual beliefs about the probability of the different binary outcomes given the different alternatives j, $P_{ij}(b_n = 1)$. Moreover, there is no need to know the utility an individual derives from an outcome not occurring, $u_n(b_n = 0, X_i)$, because this is assumed to be the same for each alternative and therefore is differenced out in the choice problem. For the continuous outcomes, instead of eliciting the probability

distributions, we elicit the expected value of the different outcomes for each alternative j, $E_{ij}[c_m] = \int c_m dP_{ij}(c)$. Taken together, the subjective expected utility individual i gets from choosing alternative j is:

$$SEU_{ij} = \sum_{n=1}^{N} P_{ij}(b_n = 1) \triangle u_n(X_i) + \sum_{n=1}^{N} u_n(b_n = 0, X_i) + \sum_{m=1}^{M} \gamma_m(X_i) E_{ij}[c_m] + \varepsilon_{ij}$$

Appendix B: Questionnaires

B1. Plans for the future

1) *How likely do you think it is that you will obtain the required A-level grades to go to university?* [0-100 scale where 0 means very unlikely and 100 means very likely]

2) *Assuming you do get the grades in sixth form/college to go to university, how likely do you think it is that you will go to university?* [0-100 scale]

3) *Which field of study would you be most likely to choose if you decided to go to university? We know this is a difficult question. Please select one field of study which you think interests you most, even if you are unsure about which field of study best fits your interests*

   a) Arts and Humanities (e.g. languages, history, music, architecture, philosophy)

   b) Life Sciences (e.g. biology, medicine, pharmacy, psychology)

   c) Physical Sciences and Engineering (e.g. maths, computer science, physics, engineering)

   d) Social Sciences (e.g. economics, law, business)

   e) Education

4) *Imagine you decided not to go to university. In a few sentences, what would be your main reasons for this decision?*

Now imagine you have been admitted to university and you started an undergraduate degree in the field of [insert chosen subject from Question 3 above].

1) *How likely do you think it is that you would complete your degree?* [0-100 scale]

2) *Assuming that you graduate, how likely do you think it is that you would get a First or a 2.1?* [0-100 scale]

B2. Hypothetical Scenarios

Now we would like you to think about the 3-4 years of your life that will come after you leave sixth form/college. Imagine that during those 3-4 years you go to university. What do you think your life during those 3-4 years will be like?

If you go to university, how likely do you think it is that you will... [0-100 scale]

1) ...enjoy the social life and activities you engage in?
2) ...meet people with whom you easily get along with?
3) ...lose contact with your family and current friends?
4) ...feel lonely and not part of a group?
5) ...find the material that is covered interesting?
6) ...enjoy studying for the course?
7) ...find the material too hard and/or work load too high?
8) ...be stressed and anxious about not being able to cope?
9) ...struggle financially?
10) ...receive financial support from your family?
11) ...have enough money to do what you enjoy?
12) ...have support from your parents in your decision to go to university rather than work immediately?
13) ...meet your future partner

Please keep thinking about the 3-4 years of your life that will come after you leave sixth form/college. This time imagine that during those 3-4 years you do not go to university but start working instead. What do you think your life during those 3-4 years will be like?

If you go start working, how likely do you think it is that you will... [0-100 scale]

1) ...enjoy the social life and activities you engage in?
2) ...meet people with whom you easily get along with?
3) ...lose contact with your family and current friends?
4) ...feel lonely and not part of a group?
5) ...find the tasks at work interesting?
6) ...enjoy performing the different tasks at work?
7) ...find the tasks at work too hard and/or work load too high?
8) ...be stressed and anxious about not being able to cope?
9) ...struggle financially?
10) ...receive financial support from your family?
11) ...have enough money to do what you enjoy?
12) ...have support from your parents in your decision to work rather than attend university?
13) ...meet your future partner

Now we would like you to think about what your life might be like when you are 30 years old. What do you think your life will be like if you did enrol in a [insert stated subject choice] degree and graduated with a First or a 2.1? Try to think about the types of jobs that would be available to you and answer the following questions.

1) *How likely do you think it is that you will have a paid job at age 30?* [0-100 scale]

2) *Assuming that you work full-time, how likely do you think it is that you would enjoy the job that you would be doing?* [0-100 scale]

3) *Assuming that you work full-time and that there is no inflation, what do you think your pre-tax earnings are likely to be per year?* [£0, £1,000, £2,000,... £99,000, £100,000, More than £100,000]

Please continue to think about what your life might be like when you are 30 years old. What do you think your life will be like if you did enrol in a [insert stated subject choice] degree and graduated but NOT a First or a 2.1? Try to think about the types of jobs that would be available to you and answer the following questions.

1) *How likely do you think it is that you will have a paid job at age 30?* [0-100 scale]

2) *Assuming that you work full-time, how likely do you think it is that you would enjoy the job that you would be doing?* [0-100 scale]

3) *Assuming that you work full-time and that there is no inflation, what do you think your pre-tax earnings are likely to be per year?* [£0, £1,000, £2,000,... £99,000, £100,000, More than £100,000]

Please continue to think about what your life might be like when you are 30 years old. What do you think your life will be like if you did not go to university or did go to university but did not graduate? Try to think about the types of jobs that would be available to you and answer the following questions.

1) *How likely do you think it is that you will have a paid job at age 30?* [0-100 scale]

2) *Assuming that you work full-time, how likely do you think it is that you would enjoy the job that you would be doing?* [0-100 scale]

3) *Assuming that you work full-time and that there is no inflation, what do you think your pre-tax earnings are likely to be per year?* [£0, £1,000, £2,000,... £99,000, £100,000, More than £100,000]

B3. Follow-up survey (wave 2)

1) *How likely do you think it is that you will obtain the required A-level grades to go to university?* [0-100 scale]

2) *Assuming you do get the grades in sixth form/college to go to university, how likely do you think it is that you will go to university* [0-100 scale]

3) *Did you apply to university?* [Yes, No] (only asked to respondents in Year 13)

    a) *Which field of study did you choose?* (if answer was 'Yes')

        i) Arts and Humanities (e.g. languages, history, music, architecture, philosophy)

        ii) Life Sciences (e.g. biology, medicine, pharmacy, psychology)

        iii) Physical Sciences and Engineering (e.g. maths, computer science, physics, engineering)

        iv) Social Sciences (e.g. economics, law, business)

        v) Education

        vi) Other

Appendix C: Decomposition Method

Using the notation in Kaiser (2015), suppose we have two subgroups of interest, L and H, and let D_i denote the subgroup that individual i belongs to, $D_i \in \{L, H\}$. Let y_i be the outcome of interest and the conditional expectation function (CEF) for y_i be of the form

$$E[y_i \mid Z_i, D_i = g] = \Lambda(Z_i \beta^g),$$

where $g = \{L, H\}$, β^g is a vector of population coefficients, Z_i is a vector of covariates and $\Lambda(\cdot)$ is a logit link function. The average explained outcome gap between the two subgroups, Δ^{SES}, is defined as

$$\Delta^{SES} \equiv E[y_i \mid D_i = H] - E[y_i \mid D_i = L],$$

which using the law of iterated expectations can be written as

$$\begin{aligned}\Delta^{SES} &= E[\Lambda(Z_i\beta^H) \mid D_i = H] - E[\Lambda(Z_i\beta^L) \mid D_i = L] \\ &= \underbrace{E[\Lambda(Z_i\beta^H) \mid D_i = H] - E[\Lambda(Z_i\beta^L) \mid D_i = H]}_{\Delta^{\beta} \text{ (structural effect)}} + \underbrace{E[\Lambda(Z_i\beta^L) \mid D_i = H] - E[\Lambda(Z_i\beta^L) \mid D_i = L]}_{\Delta^{Z} \text{ (composition effect)}},\end{aligned}$$

Δ^{β}, defined as the structural effect, gives the part of the outcome gap for the two subgroups due differing preferences/parameter values, while Δ^{Z}, the composition effect, gives the part that is due to differing beliefs/covariates. The next step is to do a detailed decomposition, whereby Δ^{Z} is decomposed into the contributions from each covariate. More specifically, we want to see how switching the value of covariate k for individual $i \in L$ with that of individual $j \in H$ affects the CEF, holding the parameters β^L fixed. The contribution of covariate k is defined as

$$\Delta_k^Z(Z_i, Z_j) = \frac{\Lambda(Z_j\beta^L) - \Lambda(Z_i\beta^L)}{(Z_j - Z_i)\beta^L}(Z_{jk} - Z_{ik})\beta_k^L$$

In order to obtain an unconditional detailed decomposition, Δ_k^Z, we integrate $\Delta_k^Z(Z_i, Z_j)$

over covariate distributions for groups L and H

$$\Delta_k^Z = \int_u \int_v \Delta_k^Z(u,v) dF_{Z|D=H}(v) dF_{Z|D=L}(u).$$

The sample analogue is given by

$$\hat{\Delta}_k^Z = \frac{1}{N_L} \sum_{i:D_i=L} \frac{1}{N_H} \sum_{j:D_j=H} \Delta_k^Z(Z_i, Z_j)$$

where N_g is the number of individuals in group $g \in \{L, H\}$.

Appendix D: Sample Characteristics and Supplementary Analyses

Table D.1—: University Enrollment Gap

Dependent variable: Attended further education (dummy)			
	(1)	(2)	(3)
Female	0.039***	0.091***	0.077***
	(0.01)	(0.01)	(0.01)
University (parent)	0.189***	0.138***	0.152***
	(0.01)	(0.02)	(0.01)
Controls	No	Yes	Yes
Cohort dummy	No	No	Yes
R-Squared	0.03	0.08	0.23
N	7173	5153	5153

Datasource: BHPS, UKHLS.

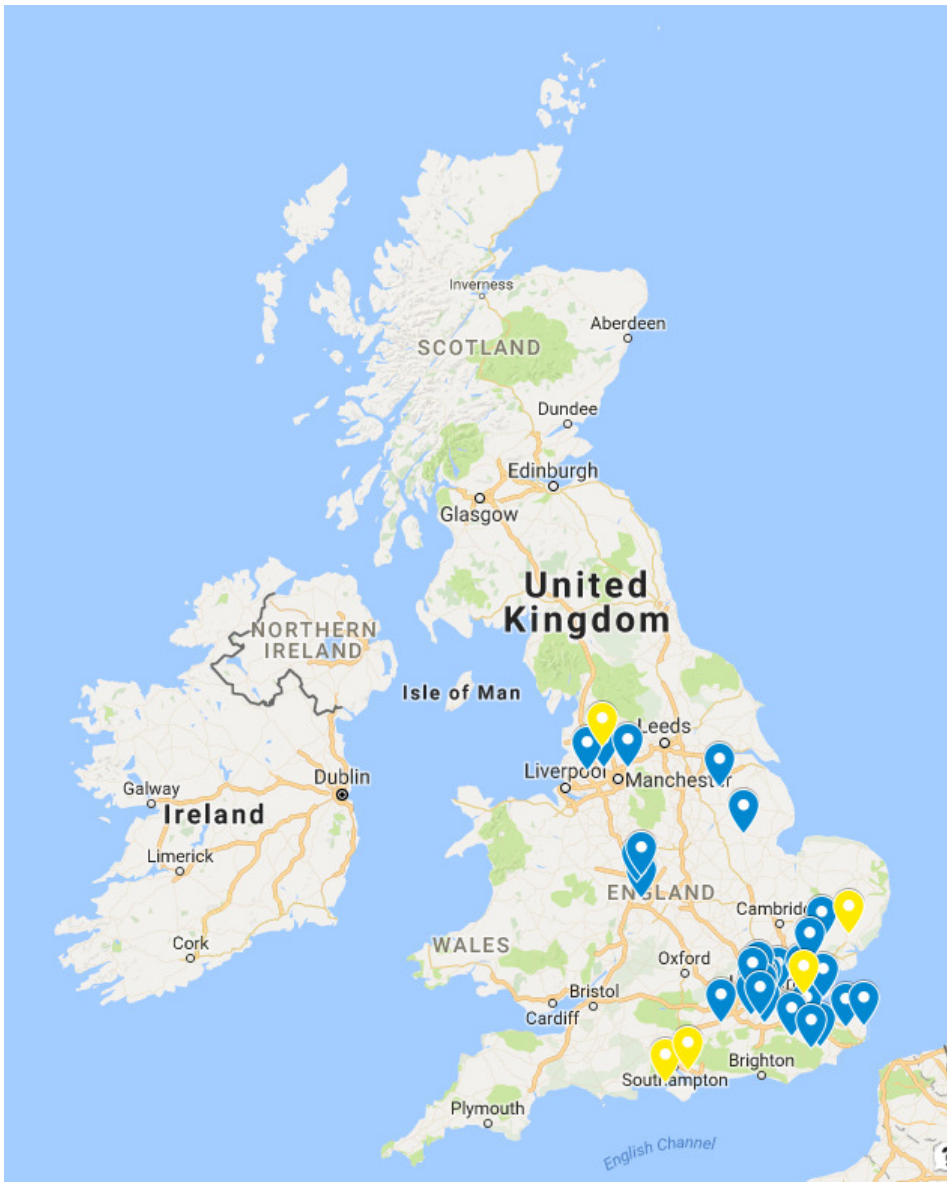
Notes: Estimation technique is OLS. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The sample includes all respondents born after 1980 that are surveyed at least once after the age of 18 residing in England. All regressions include a constant. *University (parent)* is a dummy taking the value one if at least one parent has a university degree. Controls include a constant, scores from each of the Big Five personality traits (openness, agreeableness, conscientiousness, extraversion, neuroticism) and from both cognitive and verbal tests. The *Cohort dummy* includes a dummy for each year of birth.

Table D.2—: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Male respondent	0.371	0.483	0	1	2540
School year	11.223	1.57	9	13	2540
At least one parent has degree	0.553	0.497	0	1	2411
No. of people to ask about university	5.957	3.792	0	11	2515

Notes: 'No. of people to ask about university' refers to the stated number of people the respondent knows whom he/she can ask about university life.

Figure D.1. : Location of Schools in Sample



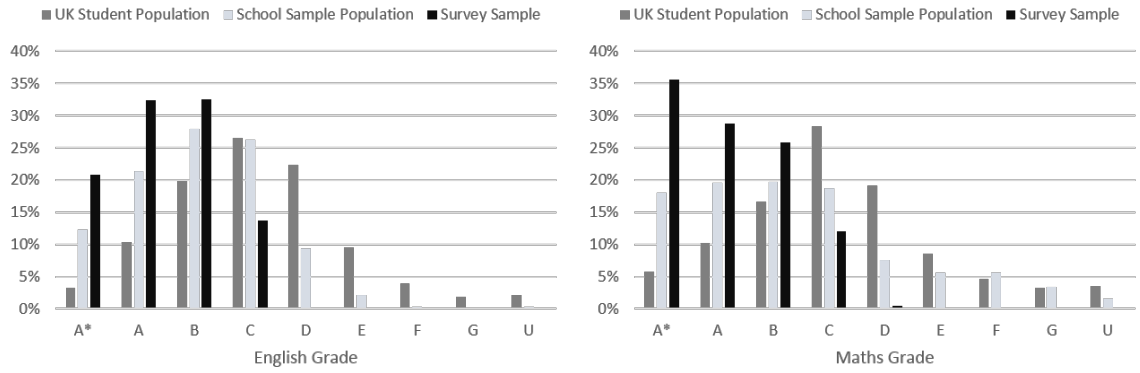
Note: This map shows the location of the schools in our sample. The schools which are marked in blue offer a sixth form (i.e. Years 12 and 13) while the schools marked in yellow do not offer a sixth form.

Table D.3—: School characteristics

School	Sample size			National Statistics				
	Male	Female	Total	Eng not first lang.	Free sch. meals	Attainment 8	GCSE	% university
1	57	9	66	0.6	7.4	66.6	96	70
2	0	97	97	69.4	42.6	54.3	58	63
3	4	7	11	49.8	60.9	-	-	67
4	30	0	30	26.7	10.2	77.4	100	81
5	19	5	24	42	10	72.8	99	81
6	74	8	82	9.5	2.4	74.9	98	77
7	42	56	98	43	29.9	58	76	66
8	57	81	138	2.1	15	51	62	58
9	52	33	85	4.4	45.5	48.7	53	-
10	31	26	57	12.5	37.7	49.8	63	58
11	0	3	3	-	-	70.5	98	-
12	46	45	91	46.4	38.8	46	54	-
13	24	29	53	3.5	20.1	51.1	64	-
14	11	8	19	1.9	24.2	42.2	41	29
15	17	24	41	4.2	37.1	45.7	48	58
16	41	9	50	13.3	6	68	97	61
17	0	117	117	-	-	32.2	-	-
18	23	50	73	5.9	7.4	64	97	76
19	6	9	15	-	-	-	-	-
20	8	15	23	47.6	52.4	55	66	77
21	0	247	247	30.5	29.1	57.5	78	60
22	27	34	61	2.4	19.2	56	76	-
23	25	37	62	-	-	-	-	72
24	7	47	54	7.3	7.3	67.4	98	65
25	0	40	40	9.6	7.1	68.2	99	77
26	12	22	34	4.7	7.9	65.9	94	81
27	27	30	57	4.1	18.4	50.4	55	75
28	17	34	51	2.5	21.2	52.5	67	-
29	6	11	17	1.6	11.1	56.6	75	82
30	32	0	32	44.2	26.4	51.8	64	44
31	52	97	149	-	-	-	-	61
32	5	16	21	65.3	34.3	52.9	57	77
33	144	240	384	-	-	-	-	60
34	11	43	54	4.5	2.9	76.7	100	75
35	0	21	21	-	-	47.4	-	-
36	20	22	42	2.3	29.2	50.5	59	58
37	15	26	41	27.7	29.8	54.6	64	63
School average	25.5	43.2	68.6	19.7	23.1	57.4	75.2	66.9
England average	-	-	-	15.7	29.3	48.5	-	48

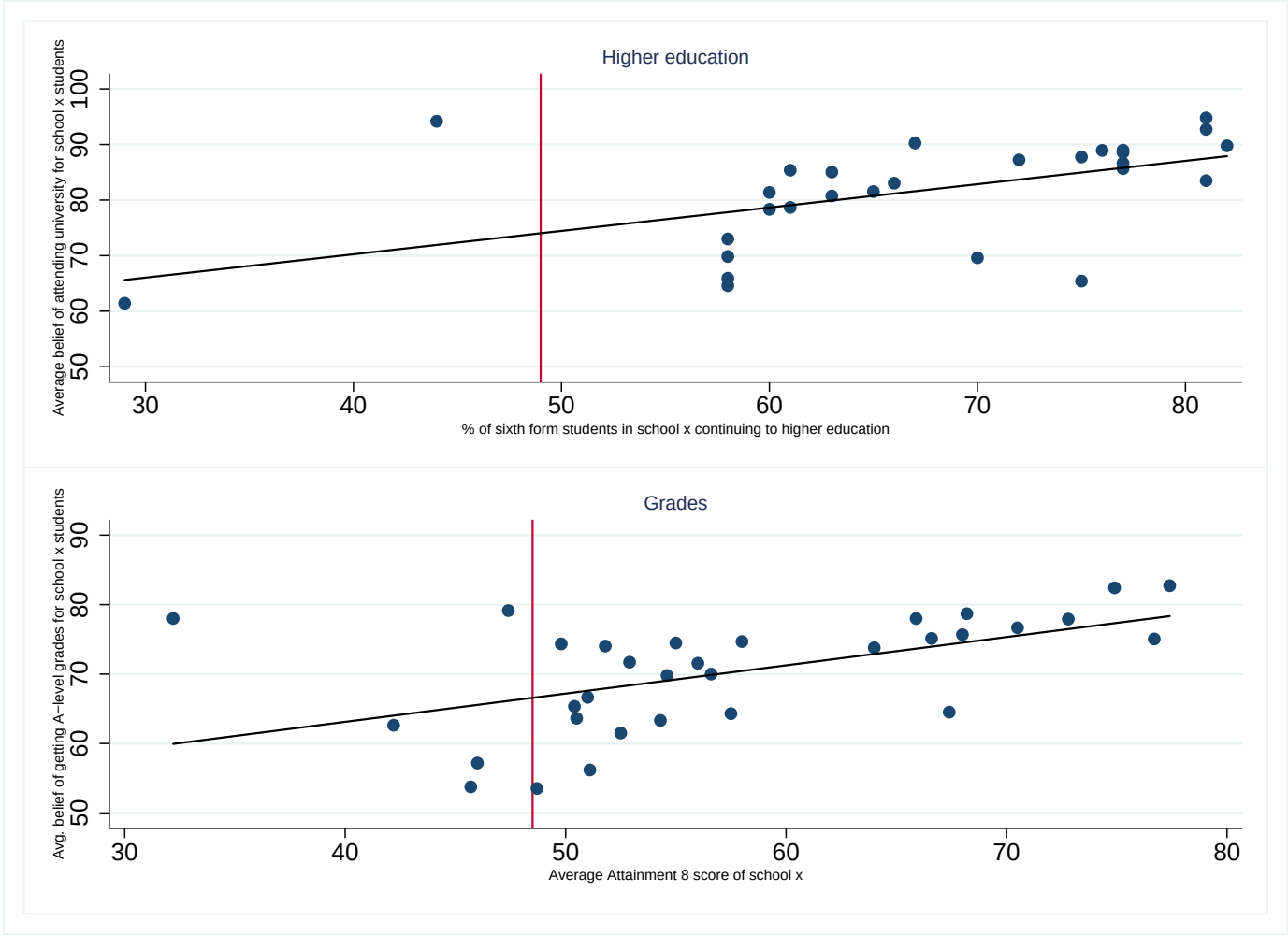
Notes: This table provides information on the 37 schools in our sample. 'Sample size' gives the number of respondents in our sample coming from each school. The data used is for the 2015/16 academic year and is available from https://www.compare-school-performance.service.gov.uk/. 'Eng not first lang.' gives the percentage of pupils at the school for whom English is not a first language. 'Free school meals' gives the percentage of pupils who have been entitled to free school meals at any point in the last 6 years. 'Attainment 8' denotes the average Attainment 8 score per pupil for the school. 'GCSE' gives the percentage of pupils who obtained 5+ A*-C grades at GCSE (including A*-C grades in English and Maths). '% university' denotes the percentage of sixth form students who went onto a UK higher education institution in the year following the completion of their A-level (or equivalent) examinations (this data refers to those who finished Key Stage 5 in the 2015/16 academic year). This data is based on all students at the school, not just the students in our sample. 'Sample average' gives the average of these variables across schools for which the relevant data is available. Where available, 'England average' gives the average across all schools in England, except for '% university' which is the average for England state-funded 16 to 18 schools/colleges. '-' denotes unavailable data on the variable for the specific school. For private/independent schools, data is often not provided. For schools with no sixth form, there is no data on '% university'.

Figure D.2. : Distribution of GCSE grades



Note: The two panels display the distribution of GCSE English Literature and Maths grades, respectively, for (i) the UK student population, (ii) the population of students in the surveyed schools and (iii) the survey sample.

Figure D.3. : Mean Beliefs and School Level Data



Notes: These graphs plot average school level beliefs against actual school level averages. The vertical red line indicates the national average.

Table D.4—: Spearman rank correlations of university beliefs

	Earnings	Have	Enjoy	Social	Meet	Contact	Lonely	Partner	Interest	Enjoy	Hard	Stress	Struggle	Financial	Money	Work	Parental
Expected earnings	1.000																
Have job	0.347***	1.000															
Enjoy job	0.277***	0.524***	1.000														
Enjoy social	0.177***	0.311***	0.278***	1.000													
Meet people	0.164***	0.267***	0.287***	0.593***	1.000												
Lose contact	-0.003	-0.085***	-0.123***	-0.069***	-0.110***	1.000											
Lonely	-0.122***	-0.190***	-0.195***	-0.299***	-0.375***	0.343***	1.000										
Life partner	0.072***	0.067***	0.115***	0.139***	0.197***	-0.010	-0.092***	1.000									
Interesting	0.188***	0.270***	0.361***	0.308***	0.322***	-0.055***	-0.055***	0.089***	1.000								
Enjoy	0.181***	0.276***	0.428***	0.329***	0.317***	-0.086***	-0.141***	0.093***	0.648***	1.000							
Hard	-0.104***	-0.113***	-0.151***	-0.050**	-0.102***	0.199***	0.298***	-0.025	-0.130***	-0.183***	1.000						
Stress	-0.138***	-0.159***	-0.170***	-0.122***	-0.173***	0.198***	0.418***	-0.072***	-0.097***	-0.188***	0.581***	1.000					
Struggle financially	-0.099***	-0.090***	-0.134***	-0.038*	-0.078***	0.233***	0.290***	-0.097***	-0.036*	-0.121***	0.400***	0.442***	1.000				
Financial support	0.080***	0.083***	0.075***	0.115***	0.115***	-0.086***	-0.037*	0.104***	0.110***	0.098***	0.065***	0.058***	-0.080***	1.000			
Enough money	0.129***	0.104***	0.202***	0.160***	0.194***	-0.109***	-0.154***	0.192***	0.154***	0.217***	-0.150***	-0.178***	-0.425***	0.298***	1.000		
Work	0.007	0.125***	0.074***	0.139***	0.091***	0.079***	0.075***	-0.027	0.100***	0.065***	0.157***	0.175***	0.381***	-0.031	-0.149***	1.000	
Parental support	0.163***	0.247***	0.187***	0.220***	0.218***	-0.110***	-0.100***	0.048**	0.282***	0.236***	-0.037*	-0.012	-0.041**	0.263***	0.135***	0.060***	1.000

Notes: * p<0.10, ** p<0.05, *** p<0.01. This table shows Spearman rank correlations for beliefs conditional on going to university. The sample size is 2,540 students. '(Expected) earnings' refers to the expected earnings in pounds at age 30 conditional on having a job. 'Have' denotes the likelihood of having a job at age 30. '(Enjoy) job' denotes the likelihood of enjoying your job at age 30. In terms of social attributes, '(Enjoy) social', 'Meet (people)' and '(Lose) contact' refer to the likelihood of enjoying the social life, meeting people and losing contact with family/friends respectively. 'Lonely' denotes the perceived belief of feeling alone and not part of a group, and '(Life) partner' reflects the belief of meeting your life partner. In terms of university-specific factors, 'Interest(ing)', 'Enjoy', 'Hard' and 'Stress' refer to the likelihood of finding the course interesting, enjoying the course, finding the material hard/workload too high and being stressed respectively. For financial attributes, 'Struggle (financially)', 'Financial (support)', '(Enough) money' and 'Work' denote the perceived likelihood of struggling financially, having financial support from your parents, having enough money to do what you want to do and needing to work whilst at university respectively. 'Parental (support)' is the likelihood of having support from your parents in your decision to go to university rather than work immediately.

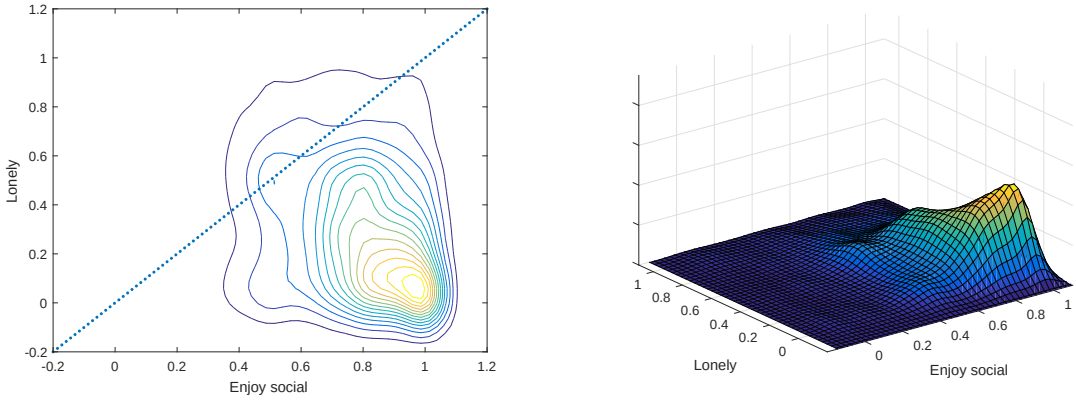
Table D.5—: Spearman rank correlations of work beliefs

	Earnings	Have	Enjoy	Social	Meet	Contact	Lonely	Partner	Interest	Enjoy	Hard	Stress	Struggle	Financial	Money	Parental
Expected earnings	1.000															
Have job	0.390***	1.000														
Enjoy job	0.350***	0.547***	1.000													
Enjoy social	0.136***	0.221***	0.296***	1.000												
Meet people	0.122***	0.180***	0.276***	0.628***	1.000											
Lose contact	-0.035	-0.077***	-0.076***	-0.122***	-0.103***	1.000										
Lonely	-0.059***	-0.110***	-0.202***	-0.306***	-0.353***	0.353***	1.000									
Life partner	0.048	0.053***	0.187***	0.268***	0.293***	0.026	-0.142***	1.000								
Interesting	0.120***	0.144***	0.341***	0.469***	0.469***	-0.043**	-0.238***	0.270***	1.000							
Enjoy	0.109***	0.132***	0.326***	0.456***	0.467***	-0.057***	-0.219***	0.241***	0.762***	1.000						
Hard	-0.076***	-0.102***	0.001	-0.010	0.006	0.235***	0.205***	0.063***	0.129***	0.087***	1.000					
Stress	-0.126***	-0.154***	-0.073***	-0.110***	-0.083***	0.250***	0.342***	-0.007	0.033	0.004	0.630***	1.000				
Struggle financially	-0.115***	-0.152***	-0.088***	-0.092***	-0.079***	0.272***	0.290***	0.003	-0.039**	-0.081***	0.374***	0.455***	1.000			
Financial support	-0.051	-0.055***	0.077***	0.118***	0.125***	0.009	-0.024	0.137***	0.202***	0.163***	0.184***	0.157***	0.198***	1.000		
Enough money	0.127***	0.188***	0.198***	0.280***	0.258***	-0.103***	-0.166***	0.177***	0.253***	0.280***	-0.038	-0.133***	-0.398***	0.085***	1.000	
Parental support	0.055***	0.177***	0.234***	0.251***	0.239***	-0.109***	-0.149***	0.146***	0.295***	0.275***	0.087***	0.033*	-0.048	0.246***	0.207***	1.000

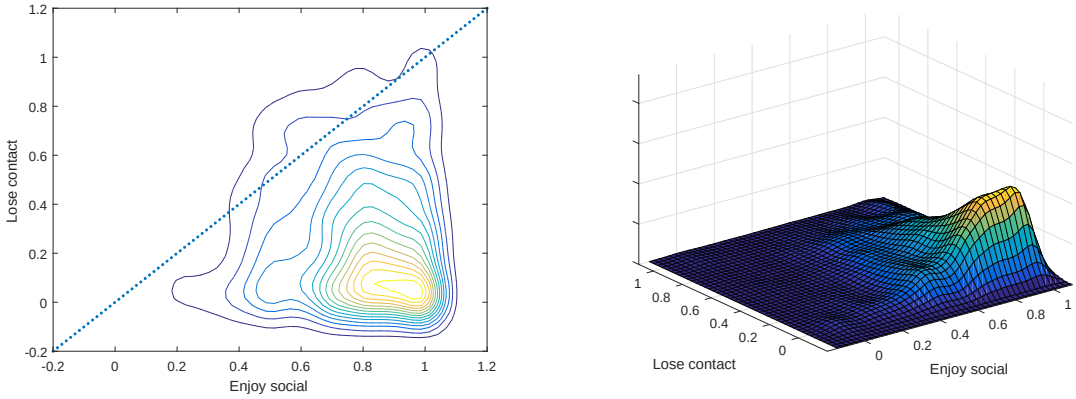
Notes: * p<0.10, ** p<0.05, *** p<0.01. This table shows Spearman rank correlations for beliefs conditional on working. The sample size is 2,540 students. '(Expected) earnings' refers to the expected earnings in pounds at age 30 conditional on having a job. 'Employed' denotes the likelihood of having a job at age 30. '(Enjoy) job' denotes the likelihood of enjoying your job at age 30. In terms of social attributes, '(Enjoy) social', 'Meet (people)' and '(Lose) contact' refer to the likelihood of enjoying the social life, meeting people and losing contact with family/friends respectively. 'Lonely' denotes the perceived belief of feeling alone and not part of a group, and '(Life) partner' reflects the belief of meeting your life partner. In terms of work-specific factors, 'Interest(ing)', 'Enjoy', 'Hard' and 'Stress' refer to the likelihood of finding the tasks at work interesting, enjoying the tasks, finding the tasks hard/workload too high and being stressed respectively. For financial attributes, 'Struggle (financially)', 'Financial (support)' and '(Enough) money' denote the perceived likelihood of struggling financially, having financial support from your parents and having enough money to do what you want to do. 'Parental (support)' is the likelihood of having support from your parents in your decision to work rather than go to university.

Figure D.4. : Distribution of beliefs

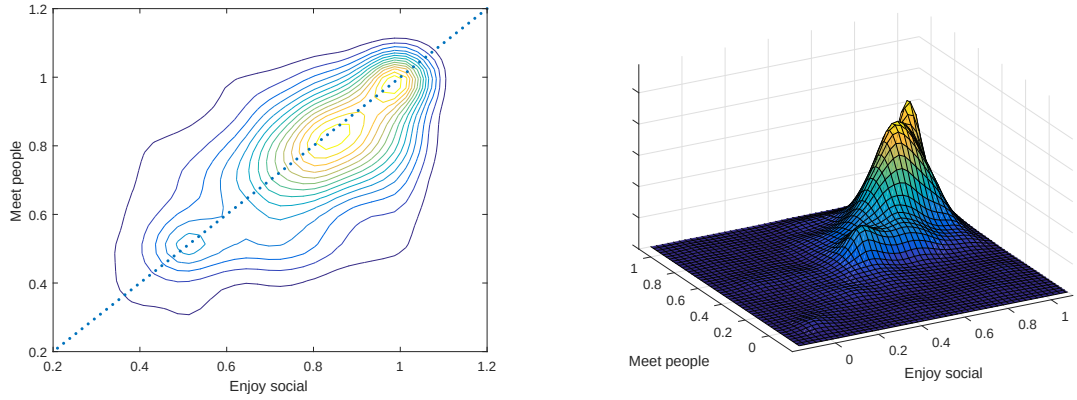
A: Enjoy social life and feel lonely



B: Enjoy social life and lose contact



C: Enjoy social life and meet people



Notes: This figure shows the joint distributions of the perceived likelihood of two social outcomes conditional on going to university. The left panels give the contour plots, while the right panels show density distributions. Panel A looks at the joint distribution of the belief that you will enjoy the social life and feel lonely conditional on going to university, while Panel B looks at the likelihood of enjoying the social life and losing contact with family and friends. Panel C shows the joint distribution for the perceived likelihood of enjoying the social life and meeting people.

Table D.6—: Mean student beliefs (all students)

Belief	University	Work	Difference	N
A: Career attributes				
Expected earnings at age 30 (in £s)	£34,193.728	£23,911.811	£10,281.917***	2,540
(conditional on having a job)	(22,526.09)	(18,379.38)	(16,336.21)	
Have job at age 30	81.20	68.73	12.47***	2,540
	(18.42)	(25.93)	(20.56)	
Enjoy job at age 30	69.17	51.80	17.37***	2,540
	(18.81)	(25.59)	(23.27)	
B: Social attributes				
Enjoy social life and activities	75.66	63.33	12.32***	2,540
	(21.29)	(24.33)	(25.90)	
Meet people	72.95	61.64	11.31***	2,540
	(21.08)	(23.48)	(24.08)	
Lose contact with family/friends	28.32	23.29	5.03***	2,540
	(26.38)	(24.78)	(24.32)	
Feel lonely	31.56	37.41	-5.85***	2,540
	(26.24)	(28.51)	(27.13)	
Meet life partner	44.83	39.98	4.85***	2,540
	(26.06)	(26.69)	(22.81)	
C: Work/university specific factors				
Find material/work tasks interesting	71.01	54.66	16.34***	2,540
	(21.59)	(25.76)	(31.61)	
Enjoy studying/work	70.66	57.19	13.47***	2,540
	(22.08)	(24.79)	(29.16)	
Find material hard/workload high	53.11	42.17	10.94***	2,540
	(23.97)	(24.85)	(27.19)	
Feel stressed	53.10	42.69	10.42***	2,540
	(28.86)	(28.41)	(28.78)	
D: Financial factors				
Struggle financially	52.81	37.97	14.84***	2,540
	(29.72)	(27.60)	(33.14)	
Have financial support from parents	63.84	48.63	15.21***	2,540
	(29.54)	(32.18)	(29.71)	
Have enough money	50.97	56.70	-5.74***	2,540
	(25.40)	(25.33)	(29.26)	
E: Parental support				
Have parental support in your choice	85.07	60.72	24.35***	2,540
	(23.54)	(34.17)	(37.55)	

Notes: standard deviations given in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Columns 1 and 2 give the mean belief for the respective characteristic under the scenarios of going to university and work respectively. Mean beliefs are given on a 0-100 scale other than for expected earnings, which are in pounds. Column 3 gives the mean difference between these two beliefs. A standard t-test is used to test for equality of means.

Table D.7—: Mean beliefs by parental education

	No degree				Has a degree				Diff-in-diff	Diff-in-var
Belief	**University**	**Work**	**Difference**	**N**	**University**	**Work**	**Difference**	**N**		P-value
A: Career attributes										
Expected earnings at age 30 (in £s)	£31,207.51	£22,965.65	£8,241.87***	1,077	£37,090.48	£24,867.32	£12,223.17***	1,334	-£3,981.30***	0.000
(conditional on having a job)	[21,112.58]	[17,258.58]	[14,036.34]		[22,826.81]	[18,609.25]	[17,788.19]		(664.45)	
Enjoy job at age 30	69.09	52.82	16.27***	1,077	69.69	50.96	18.73***	1,334	2.46***	0.103
	[19.46]	[26.00]	[22.63]		[17.86]	[25.02]	[23.72]		(0.95)	
Have job at age 30	81.60	71.47	10.12***	1,077	81.85	67.11	14.74***	1,334	4.62***	0.001
	[18.81]	[25.47]	[19.31]		[17.13]	[25.69]	[21.26]		(0.84)	
B: Social attributes										
Enjoy social life and activities	74.91	63.77	11.14***	1,077	77.19	63.04	14.14***	1,334	3.01***	0.035
	[21.49]	[24.50]	[26.62]		[20.56]	[24.03]	[25.05]		(1.06)	
Meet people	71.69	61.47	10.22***	1,077	74.84	61.68	13.15***	1,334	2.94***	0.107
	[21.67]	[23.88]	[24.71]		[19.75]	[22.88]	[23.58]		(0.99)	
Lose contact with family/friends	27.98	21.26	6.72***	1,077	28.64	24.69	3.95***	1,334	-2.77***	0.208
	[26.97]	[24.35]	[24.71]		[25.85]	[24.74]	[23.83]		(0.99)	
Feel lonely	32.37	35.69	-3.33***	1,077	30.57	38.99	-8.42***	1,334	-5.09***	0.743
	[26.17]	[27.92]	[27.06]		[25.77]	[28.69]	[26.81]		(1.10)	
Meet life partner	42.63	38.48	4.15***	1,077	46.20	40.34	5.85***	1,334	1.71*	0.014
	[26.06]	[26.46]	[21.94]		[25.49]	[25.95]	[23.56]		(0.94)	
C: Work/university specific factors										
Find material/work tasks interesting	69.13	53.99	15.14***	1,077	73.49	55.03	18.46***	1,334	3.31**	0.032
	[21.88]	[25.92]	[32.64]		[20.36]	[25.43]	[30.68]		(1.29)	
Enjoy studying/work	69.61	56.90	12.71***	1,077	72.06	57.23	14.82***	1,334	2.12*	0.292
	[22.15]	[25.33]	[29.72]		[21.55]	[24.29]	[28.83]		(1.20)	
Find material hard/workload high	54.12	42.64	11.48***	1,077	52.31	41.44	10.88***	1,334	-0.60	0.246
	[23.81]	[24.34]	[27.60]		[23.97]	[24.93]	[26.69]		(1.11)	
Feel stressed	54.68	42.59	12.09***	1,077	51.73	42.09	9.64***	1,334	-2.45**	0.149
	[28.18]	[28.09]	[29.52]		[29.17]	[28.09]	[28.32]		(1.18)	
D: Financial factors										
Struggle financially	56.24	37.63	18.61***	1,077	50.44	38.19	12.25***	1,334	-6.37***	0.212
	[29.11]	[28.11]	[33.88]		[29.95]	[27.20]	[32.68]		(1.36)	
Have financial support from parents	60.12	45.94	14.17***	1,077	67.33	50.33	17.00***	1,334	2.82**	0.721
	[31.32]	[33.14]	[29.45]		[27.73]	[31.43]	[29.76]		(1.21)	
Have enough money	46.49	56.94	-10.45***	1,077	54.18	56.08	-1.90**	1,334	8.54***	0.677
	[25.23]	[25.67]	[29.25]		[24.79]	[24.96]	[28.90]		(1.19)	
E: Parental support										
Have parental support in your choice	82.55	62.92	19.63***	1,077	88.45	58.76	29.69***	1,334	10.06***	0.543
	[25.31]	[33.91]	[36.92]		[20.55]	[34.73]	[37.58]		(1.53)	

Notes: standard deviations given in square brackets, and standard errors given in parentheses. * p<0.10, ** p<0.05, *** p<0.01. This table provides mean beliefs by the education level of the respondent's parents. Columns 1-4 are for respondents for whom neither parent went to university, while Columns 5-8 are for respondents for whom at least one parent went to university. Within each group, the first two columns give mean beliefs for the respective characteristic under the scenarios of going to university and work respectively. Mean beliefs are given on a 0-100 scale other than for expected earnings, which are in pounds. The third column gives the mean difference between these two beliefs. T-tests are used to test for equality of means. Column 9 ('Diff-in-diff') gives the average difference for respondents for whom at least one parent has a degree minus the average difference for respondents for whom neither parent has a degree. Column 10 ('Diff-in-var') gives the p-value for a test of equality of variances between the perceived returns.

Table D.8—: Differences in beliefs by SES

	All	Parental Education		
Variable		No degree	Has a degree	P-value
How likely to attend university if achieves grades	0.784	0.741	0.830	0.000
	(0.291)	(0.311)	(0.262)	
Subject preference: Arts or humanities	0.269	0.252	0.278	0.144
	(0.443)	(0.434)	(0.448)	
Subject preference: Life sciences	0.270	0.266	0.277	0.517
	(0.444)	(0.442)	(0.448)	
Subject preference: Physical sciences	0.209	0.190	0.220	0.077
	(0.406)	(0.393)	(0.414)	
Subject preference: Social sciences	0.196	0.212	0.190	0.194
	(0.397)	(0.409)	(0.393)	
Subject preference: Education	0.056	0.081	0.034	0.000
	(0.231)	(0.273)	(0.183)	
How likely to complete degree	0.826	0.815	0.849	0.000
	(0.207)	(0.219)	(0.182)	
How likely to get first or 2.1 if graduate	0.672	0.646	0.705	0.000
	(0.205)	(0.209)	(0.190)	

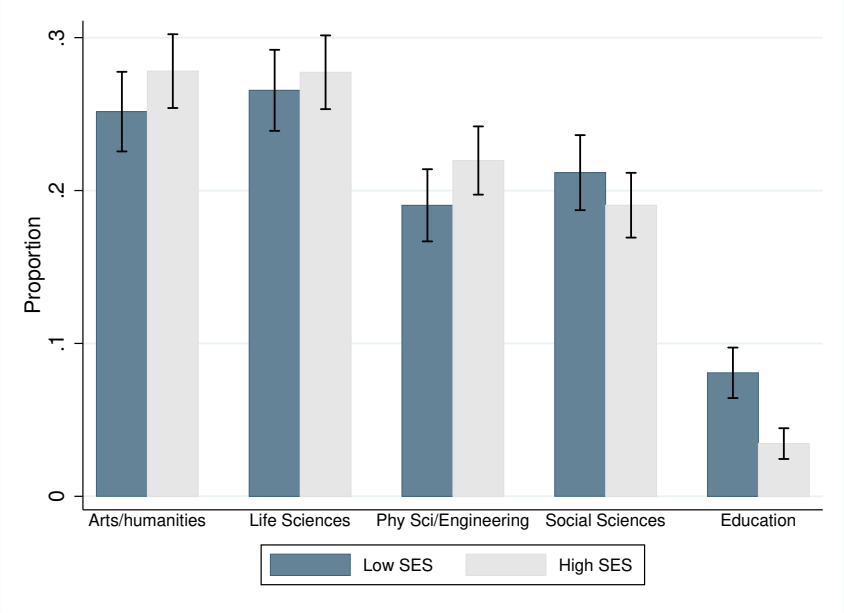
Notes: standard deviations given in parentheses. This table separately provides mean beliefs for whole sample (column 1) and by whether at least one parent has a degree (columns 2 and 3). P-values for a test of difference in means is provided in column 4.

Table D.9—: Differences in beliefs by parental education

	All	Parental Education		
Variable		No degree	Has a degree	P-value
A: If you don't go to university				
Probability of employment at age 30	0.687	0.715	0.671	0.000
	(0.259)	(0.255)	(0.257)	
Expected earnings (cond'l on having a job)	23,911.811	22,965.645	24,867.316	0.010
	(18,379.380)	(17,258.577)	(18,609.251)	
Probability of enjoying job at age 30	0.518	0.528	0.510	0.075
	(0.256)	(0.260)	(0.250)	
B: If you graduate from university with less than a 2:1				
Probability of employment at age 30	0.771	0.780	0.772	0.334
	(0.213)	(0.217)	(0.203)	
Expected earnings (cond'l on having a job)	31,255.906	29,498.607	33,050.975	0.000
	(22,155.806)	(21,369.541)	(22,212.074)	
Probability of enjoying job at age 30	0.666	0.672	0.665	0.411
	(0.215)	(0.222)	(0.207)	
C: If you graduate from university with a first or 2:1				
Probability of employment at age 30	0.856	0.860	0.862	0.792
	(0.175)	(0.177)	(0.163)	
Expected earnings (cond'l on having a job)	37,450.000	34,143.918	40,410.045	0.000
	(24,994.924)	(23,755.780)	(25,028.397)	
Probability of enjoying job at age 30	0.730	0.730	0.734	0.672
	(0.192)	(0.196)	(0.184)	

Notes: standard deviations given in parentheses. This table separately provides mean beliefs for whole sample (column 1) and by whether at least one parent has a degree (columns 2 and 3). P-values for a test of equality of means is shown in column 4.

Figure D.5. : Intended university subject field by socio-economic status (SES)



Note: This graph shows the proportion of respondents intending to study each of the five specified subject fields, separately by socio-economic status. 95% confidence intervals are provided.

Table D.10—: Decomposing actual SES gap and model predictions

	All	Mixed schools	Boys	Girls	Good grades	Final Year
Actual SES gap	8.91	9.05	13.02	6.49	6.15	7.06
Predicted SES gap	6.15	7.08	8.66	4.73	5.36	5.68
(%)	68.54%	78.23%	66.51%	72.88%	87.15%	80.45%
Predicted SES gap (beliefs only)	4.35	5.41	6.83	2.83	2.27	3.46
(%)	48.66%	59.78%	52.46%	43.61%	36.91%	49.01%
Predicted SES gap (pecuniary)	1.08	1.41	1.38	0.98	0.23	0.87
(%)	12.17%	15.58%	10.60%	15.10%	3.74%	12.32%
Predicted SES gap (non-pecuniary)	3.27	4.00	5.45	1.85	2.04	2.59
(%)	36.70%	44.20%	41.86%	28.51%	33.17%	36.69%
Observations	2540	1863	884	1527	1005	759

Notes: Values without '%' indicate percentage points. In the first column the sample includes all individuals, in column 2 it is restricted to coeducation schools, in column 3 we include only boys, in column 4 only girls, in column 5 only those with high grades, i.e. A*-B and in column 6 we only include those in their final year of school.

Table D.11—: Determinants of Perceived Returns (1)

	Earnings	Have Job	Enjoy Job	Social	Meet	Contact	Lonely	Partner
Male student	3852.225***	-0.004	-0.010	-0.010	-0.012	-0.000	-0.002	-0.000
	(713.992)	(0.009)	(0.010)	(0.011)	(0.010)	(0.011)	(0.012)	(0.010)
Parent has degree	2211.277***	0.028***	0.019*	0.025**	0.024**	-0.014	-0.038***	0.015
	(695.631)	(0.009)	(0.010)	(0.011)	(0.010)	(0.011)	(0.012)	(0.010)
# ask about university	447.233***	0.003***	0.003*	0.004**	0.003**	-0.004**	-0.007***	0.002
	(96.139)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Student's school year	-200.653	-0.005*	0.015***	0.026***	0.025***	0.018***	-0.015***	0.009***
	(220.875)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)
Area: unemployment	-258.636	0.011**	0.010**	-0.005	-0.015***	-0.001	0.006	0.001
	(331.856)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Area: degree	89.000***	0.002***	0.001**	0.001	0.000	0.000	0.000	0.000
	(31.272)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Observations	2407	2407	2407	2407	2407	2407	2407	2407
R-squared	0.04	0.03	0.02	0.03	0.04	0.02	0.03	0.01

Notes: Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. All dependent variables are beliefs about returns (university - work), see overview of elicited beliefs about returns in Table 1. Control variables include whether the student is male, self-reported ability (0-100 scale), whether or not at least one of the students' parents has a university degree, the number of people a student can ask about university, the student's school year, the unemployment rate in the local area and the fraction of the population in the local area with a university degree (Source: 2011 Census). OLS estimation is used.

Table D.12—: Determinants of Perceived Returns (2)

	Interest	Enjoy	Hard	Stress	Struggle	Financial	Money	Parental
Male student	0.003	-0.014	-0.014	-0.052***	-0.034**	0.020	-0.001	0.004
	(0.014)	(0.013)	(0.012)	(0.012)	(0.014)	(0.013)	(0.013)	(0.016)
Parent has degree	0.045***	0.029**	0.011	-0.012	-0.028*	0.034***	0.053***	0.068***
	(0.014)	(0.012)	(0.012)	(0.013)	(0.014)	(0.013)	(0.012)	(0.016)
# ask about university	-0.000	0.001	-0.005***	-0.002	-0.005***	0.002	0.005***	0.004*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Student's school year	0.053***	0.035***	0.025***	0.020***	0.030***	0.033***	-0.028***	0.033***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Area: unemployment	-0.020***	-0.006	-0.014**	-0.012**	-0.016**	-0.021***	0.004	0.003
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.008)
Area: degree	0.001	0.000	0.000	-0.001	-0.003***	-0.001*	0.002***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	2407	2407	2407	2407	2407	2407	2407	2407
R-squared	0.07	0.04	0.03	0.02	0.05	0.04	0.05	0.07

Notes: Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. All dependent variables are beliefs about returns (university - work), see overview of elicited beliefs about returns in Table 1. Control variables include whether the student is male, self-reported ability (0-100 scale), whether or not at least one of the students' parents has a university degree, the number of people a student can ask about university, the student's school year, the unemployment rate in the local area and the fraction of the population in the local area with a university degree (Source: 2011 Census). OLS estimation is used.