WORKING PAPER 2022:5

Earnings expectations and educational sorting

An ex-ante perspective on returns to university education

Nikolay Angelov
Per Johansson
Mikael Lindahl
Ariel Pihl

IFAU
INSTITUTE FOR EVALUATION OF LABOUR MARKET AND EDUCATION POLICY

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala.

IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

More information about IFAU and the institute's publications can be found on the website www.ifau.se

ISSN 1651-1166

Earnings expectations and educational sorting: An ex-ante perspective on returns to university education[a]

by

Nikolay Angelov,[b] Per Johansson,[c] Mikael Lindahl,[d] Ariel Pihl[e]

March 3, 2022

Abstract

We estimate means and distributions of ex-ante treatment effects for obtaining university education relative to high school. To achieve this, we conducted a survey which elicited earnings expectations associated with counterfactual educational choices for a sample of high-school students in Stockholm. We find average ex-ante returns to university to be 36%, with higher returns for females, those with high SES backgrounds, and high math scores. The returns vary considerably and are highest for those that choose university, but also positive and sizable for those who do not. Our results imply that students sort into education based on their comparative advantage. Nevertheless, our results suggest that an OLS estimator of the returns to university education should be expected to be quite similar to the average treatment on the treated effect for university education. Additionally, we find evidence that the positive ex-ante earnings returns to high paying fields, among those that do not choose these fields, can (partly) be reconciled by individuals expecting to be compensated through higher non-pecuniary returns to those fields.

Keywords: Ex-ante treatment effects, returns to university, educational sorting, subjective expectations
JEL-codes: I26, J24

---

[a] We are grateful for comments from Joseph Hotz, Martin Nybom, Erwin Ooghe and from participants at seminars at CESifo Area Conference on Economics of Education in Munich, IFAU, Tinbergen Institute, SOFI, JRC/EU commission, Århus University and University of Gothenburg. Mikael Lindahl acknowledges financial support from Jan Wallanders and Tom Hedelius Stiftelse, Tore Browaldh Stiftelse and Torsten Söderberg foundation.
[b] Uppsala Center for Fiscal Studies (UCFS). E-mail: nikolay@angelov.cc
[c] Institute for Evaluation of Labor Market and Education Policy (IFAU), Uppsala University (UCLS and Statistics), Tsinghua University and IZA. E-mail: per.johansson@ifau.uu.se.
[d] Department of Economics, University of Gothenburg, CESifo, IFAU, IZA, and UCLS. E-mail: Mikael.Lindahl@economics.gu.se.
[e] Department of Economics, University of Gothenburg, E-mail: ariel.pihl@economics.gu.se.

Table of Contents

1. Introduction

If returns to education are heterogeneous and educational decisions are made with uncertainty, it is a prospective student's expected return, or the *ex-ante* treatment effect, that is relevant for understanding their educational decision making.[1] However, inferring ex-ante treatment effects from ex-post data requires very strong assumptions such as rational expectations and no unanticipated earnings shocks.[2] By eliciting subjective expected outcomes for counterfactual educational choices through a survey, one can overcome some of these problems and directly estimate ex-ante treatment effects.[3] Although this requires few econometric identification assumptions, it relies on high-quality survey data in which the elicited expectations are informative about individuals' true expected outcomes.

In this paper, we use newly collected subjective expectations data for a sample of high school students from 40 public high schools in Stockholm, Sweden. Following the approach outlined in Arcidiacono et al. (2020), we use this data to expand the focus from college majors and occupational groups. We estimate various ex-ante average and distributional treatment effects of *choosing* university education (relative to stopping at high school), including the average treatment effect (ATE), average treatment on the treated (TT) and average treatment on the untreated (TUT).[4] Our data make this possible since we surveyed high school students about earnings expectations from choosing various university fields-of-study, as well as without continuing to university.

In addition, we expand the ex-ante treatment effect literature using subjective expectations data by explicitly connecting the various ex-ante treatment effects to the descriptive difference-in-means (or OLS) estimator, as well as to various sorting parameters, using the returns to education framework laid out in Heckman, Lochner, and Todd (2006). The fundamental difficulty in estimating ex-post (realized) returns to education is the lack of data on counterfactual outcomes, something which can be overcome by imposing, sometimes very strong, econometric identification assumptions. Estimating mean and distributional treatment

[1] See Carneiro, Hansen and Heckman, 2003; Heckman, Lochner and Todd, 2006; Cunha and Heckman, 2007.
[2] See Arcidiacono, Hotz, Maurel and Romano (2020).
[3] As in Arcidiacono et al. (2020) and Wiswall and Zafar (2020).
[4] Rather than focusing on estimating ex-ante treatment effects from choosing different college majors (as in Wiswall and Zafar, 2020) or from choosing different occupational fields (as in Arcidiacono et al, 2020), conditional on being a college student.

effects is particularly challenging in models that allow for heterogeneous returns, and where students self-select into higher education.[5] We therefore push the usage of subjective expectations data beyond the estimation of ex-ante treatment effects, to also directly estimate the degree and sources of biases in ex-ante OLS estimates of treatment effects. Assuming students have rational expectations, these estimates can be interpreted as showing ex-post relationships net of unanticipated earnings shocks. In addition, this makes it possible to directly estimate ex-ante versions of various sorting parameters and the degree of comparative earnings advantage.

To capture counterfactual outcomes, we designed and implemented a survey among high-school students administered just before their deadline for applying to university. The survey was aimed at eliciting beliefs and expectations about future earnings and other outcomes associated with various educational choices (graduating university in various fields; not attending university). The fact that our sample population is both relatively large, and comes from a varied cross-section of high schools in a large city, is, we believe, an important aspect of this study, since it means that we provide results for a broad population of interest. Although this type of survey is growing in popularity, getting participants to provide valid responses to such hypothetical questions can be challenging (see Dominitz and Manski, 1996; Manski, 2004). To combat this, we took special care in designing the questions and conducting the survey. We have also linked administrative data on students' family background and school performance, as well as follow-up data on application and enrollment at (any Swedish) university which we use to validate the stated educational choices as well as to estimate alternative treatment effects and sorting parameters.

We provide a number of interesting findings regarding the ex-ante returns to university education. First, we estimate the ex-ante average treatment effect to be 36%, indicating that the average prospective student expects about 9% higher earnings for each year of university study. However, these ex-ante returns vary substantially between prospective students, and are positively correlated with being female, coming from a high SES background, and test scores in math, but only weakly related to other measures of school performance. Second, we find evidence of positive sorting effects, meaning that the average ex-ante return to university is

[5] Willis and Rosen, 1979; Card, 1999; Heckman, Lochner and Todd, 2006 Cunha and Heckman, 2007; and Carneiro, Heckman and Vytlacil, 2001 and 2011.

higher for those choosing university than for those choosing high school (so that $TT > TUT$), although there is extensive distributional overlap. This is consistent with students sorting into educations based on their comparative earnings advantage.

Third, despite the apparent heterogeneous returns with respect to schooling choices, resulting in sorting based on comparative advantage, our ex-ante estimates predict that the bias from estimating ATE or TT using an OLS estimator from a regression of log earnings on a university education indicator should be expected to be fairly small. The reason is that the ex-ante estimates of the traditional selection bias and the sorting gain from choosing university (for university choosers) are of opposite sign and of somewhat limited magnitudes. On the other hand, since those who choose to stop education after high school expect to gain much less from university than those who choose university, the ex-ante TUT estimates are typically much smaller than the OLS estimates.

Fourth, in a traditional earnings-education regression model the unobserved ability term is often specified as being one-dimensional, implying hierarchical sorting to educational states. The Roy model instead allows earnings ability to be multi-dimensional in nature, and for sorting based on comparative advantage or absolute advantage at each schooling level to exist. As we estimate the traditional selection bias to be non-positive (meaning that those with higher-than-average expected high school earnings are not more likely to choose university, on average) and positive selection into university (those with higher-than-average expected university earnings choose university), the traditional specification is inaccurate. On the other hand, our data only somewhat support findings of a negative selection bias as found in Willis and Rosen (1986) and Carneiro, Heckman and Vytlacil (2011), using observational data for the US, and Nybom (2017) using observational data for Sweden.

Fifth, we find that the ex-ante TUT estimates are consistently positive.[6] This contradicts the pure Roy model of earnings maximization, and instead suggests that there are other factors affecting these student's choices. Since all universities in Sweden are free and living expenses are covered by generous grants and subsidized loans, financial constraints are unlikely to be

[6] This is in line with results found in Nybom (2017) using Swedish data, but different than what is found in Heckman, Lochner and Todd (2006) and Carneiro, Heckman, and Vytlacil (2011) using data for US, all using instrumental variables and the "MTE-framework" to estimate treatment effects of returns to college using observational data.

important. Using survey measures of non-pecuniary costs, including enjoyment of studies and probability of graduating, we show that these measures are only weakly related to returns, indicating that psychic costs are unlikely to explain this finding. Another possibility is compensating non-pecuniary factors, so that prospective students are prepared to give up earnings gains from choosing university because of higher non-pecuniary benefits from choosing to stop with high school education. We return to this issue below.[7]

To complement the analysis above we also estimate ex-ante treatment effects of choosing high- versus low-paying university fields of study. This makes it possible to compare our ex-ante treatment and sorting earnings effects with those in Wiswall and Zafar (2020), who estimated ex-ante treatment effects for high- versus low paying college majors in the US (science/business vs social sciences/humanities). Interestingly, our findings of sorting based on comparative advantage and positive ex-ante TUT mimic the results in their study. In addition, we are then also able to provide counterfactual subjective ex-ante treatment effects for non-earnings outcomes, including expected social status, enjoyment of study and of work, and work-life balance, many of which are novel in the literature on ex-ante treatment effects.[8] For the non-pecuniary outcomes we find mostly positive and large TT estimates, which are notably higher than the TUT estimates for these outcomes, the latter of which in many cases are negative.

Thus, for the high- and low paying dimension our results show an important compensating role for many non-pecuniary factors, meaning that prospective students are prepared to sacrifice higher expected earnings and instead select a low-paying field which they expect would lead to a job where they would obtain higher non-pecuniary benefits. For instance, individuals choosing low-paying fields expect to enjoy their jobs and sustain a better work-life balance in the occupations resulting from educations in such fields, to a higher degree than they expect to achieve had they instead chosen high paying fields. As the ex-ante treatment and sorting earnings effects for high- versus low paying fields are qualitatively very similar to university versus high school, including a positive and sizable TUT, we believe these results are also informative about why we find positive TUT for university studies relative to high school.

[7] Unfortunately, we do not have access to direct information about non-pecuniary factors in choosing high school.

[8] These non-pecuniary outcomes measures are very similar to those used in Zafar (2013) who investigated determinants of the college major choice across genders. In Pihl, Angelov, Johansson and Lindahl (2019) we use these measures to answer similar questions.

Our paper and findings relate to several strands of literature. Two important recent papers, Arcidiacono et al. (2020) and Wiswall and Zafar (2020), have estimated ex-ante earnings treatment effects (ATE, TT, TUT) for occupational choice and college major, respectively, using subjective expectations data for students enrolled in two US universities.[9] Both studies estimate ex-ante treatment effects between various fields and find evidence of positive ex-ante ATE and positive sorting effects (TT>TUT>0) for earnings.[10] Both these studies find substantial individual heterogeneity in ex-ante earnings returns. They also find that non-pecuniary factors play an important role.[11] In these respects, our results for university choice are quite similar.

We also relate to a small body of literature using subjective expectations data to study university attendance. An early set of papers by Kodde (1986, 1988) find that expecting higher post-secondary earnings in the Netherlands is positively correlated with choosing to attend post-secondary education, while expecting higher earnings with only a high school degree is negatively (but not significantly) associated with further education. More recently, Boneva and Rauh (2020) estimate ex-ante university premiums for the UK, and find that students from low SES backgrounds perceive the earnings returns to university education as being significantly lower compared to students with a high SES background.[12] This is also consistent with our findings. There is also a rapidly growing literature eliciting beliefs and stated preferences about counterfactual choices (see Altonji, Arcidiacono and Maurel, 2016, for a survey), especially those eliciting earnings expectations associated with hypothetical schooling choices (see Dominitz and Manski, 1996; Zafar, 2013, and; Arcidiacono, Hotz and Kang, 2012, for important early contributions).

---

[9] Arcidicaono et al. (2020), use data on 173 (male) Duke University students, from whom they elicited data on earnings beliefs and subjective probabilities of working in five occupation groups, which were matched to actual occupation using data from the social network *LinkedIn*. Wiswall and Zafar (2020) use data on 493 enrolled New York University students and analyze various outcomes, including expected earnings, marital sorting and labor supply, associated with four groups of potential majors.

[10] Wiswall and Zafar (2020) estimate ex-ante treatment effects for Science/Business versus Humanities/Social sciences, and also versus a small sample of drop outs.

[11] Arcidiacono et al (2020) investigate this using an indirect test, whereas Wiswall and Zafar (2020) look at expected spousal earnings and fertility.

[12] Boneva and Rauh (2020) perform an online survey of 2,540 secondary school students' (where 759 students are in their final year) expected earnings and beliefs about some non-pecuniary outcomes. They use their data to estimate the socioeconomic gap in ex-ante earnings premiums and find that students from low SES backgrounds perceives the pecuniary as well as the non-pecuniary returns to university education as being significantly lower.

Our paper is structured as follows. In Section 2, we present the survey and key variables, as well as discuss the validation of these variables. Section 3 presents the conceptual framework for analyzing and interpreting ex-ante treatment effects and sorting parameters. Section 4 presents our main results of earnings returns to university over high school. Section 5 presents the results for high versus low-paying fields and a supplementary analysis using non-pecuniary outcomes. Section 6 concludes.

2. Institutions and the survey

2.1 Education in Sweden

In Sweden, the vast majority of students complete three years of high school (*gymnasium*) after compulsory school (grades 1-9). High school programs are specialized, both for broad academic subjects (e.g., social science, natural science), and vocational tracks. Academic programs contain coursework preparing the students for university, but if students are not qualified for their university program of choice at the end of high school (by completed courses), they can top-up their education with an additional year of high school.

There is a common application for all colleges and universities in Sweden.[13] For degree granting programs and courses beginning in the fall semester, students apply in the spring of the same year. There are no tuition fees for Swedish citizens or permanent residents, and when studying full-time, students are given generous grants and low-interest loans which cover living costs. University programs in Sweden can be divided into the following eight broad categories:

Table 1: Categories of Education and particular programs

These eight categories are adopted from the classification of education that Sweden has used since the 1960s (*SUN*). We use the broadest category (first digit of a possible 3-4). Within each

[13] In Sweden there is a technical difference between a university and a college (högskola), in that the former is legally permitted to award PhDs. Universities are typically more prestigious and selective, but many types of bachelors and masters-level educations can be completed at either type of institution. We use the words college and university interchangeably in this paper.

category there are more specific university programs, some of the most common are listed in the right column.

2.2 Survey implementation

For the purpose of this project, we have collected survey data on a sample of high school students in the municipality of Stockholm. To be part of our population, the students must have attended the final third year of a municipal high school in 2014 and lived in the municipality of Stockholm. Although the fraction of independent (non-municipal) high school students is high in Stockholm, the majority of the students in academic programs attend municipal schools. The municipality of Stockholm includes many suburbs, some well-off and some less so.

A concern with eliciting preferences from survey data is that the result may differ from what would be found in the corresponding real-world choice situation. As our study design is quite similar to a stated preference experiment, advice from the stated preference literature was used when designing the survey. We describe the motivation in designing the survey in more detail in Appendix E, and a translation of the survey in Data Appendix A. The survey timing was chosen carefully to be before the university applications closed, but late enough so that the students had likely put considerable thought into their educational path. Hence, we hope to limit the issue of cognitive dissonance/ex-post rationalization, where students provide biased responses because their field of study selection has already been made.[14] Our study differs from most other studies using subjective expectations data, in that it is drawn from a region rather than a single school, and the students are in high school and not in university.[15] Also, since the sampling is done at the high school level, the students can end up at any university or college, including technical colleges and business schools, also located outside Stockholm.

To maximize saliency and sample size we hired a professional interview company to contact the students and do in-person interviewing. As is typical with voluntary surveys, differences in ease of contact (primarily due to absence of a listed phone number) and willingness to participate mean that the final sample is not entirely representative of the population. However, we do still have a sample composed of students from over 40 different high schools. We

[14] Zafar (2011a) tests for this issue in his sample and shows that those students do not appear to exhibit cognitive dissonance when reporting their beliefs.
[15] Boneva & Rauh (2020) is a notable exception that uses high school students from 37 English schools.

compare the surveyed sample to the population in terms of demographics in Appendix E and Table A1, and their expectations in Section 2.4.

After an introduction, the students were first shown the eight fields of study (listed in Table 1) and asked to think about and choose which of the most common programs in that field they would pursue if they had to pick. As these classifications are standard in Sweden, they should be meaningful to the surveyed students. Then, they were asked to provide their expectations and beliefs about these hypothetical educational choices on 10 different dimensions.[16] They were instructed to imagine their most preferred specific program within each field when answering questions about the broad field-category (e.g., computer science within the natural science category). Two of these ten dimensions were their expected earnings at ages 30 and 40 assuming they had graduated with a degree in the respective field. Then, the students were asked if they expected to go on to pursue postsecondary education, and if so, to which field. The field or no-college option chosen in these two questions is our primary measure of chosen education from the survey. Since the previous questions referred only to college fields, they were finally asked about their expected earnings at age 30 and 40 should they not attend college.[17]

2.3 Data

Our sample includes all 498 individuals who completed the survey and were matched to administrative data from Statistics Sweden. We have matched the students to their parents' incomes and educations (to capture socio-economic status), as well as to demographics such as immigrant background and gender. We observe the high schools the students attended, as well as their coursework and grades, which we use to construct proxies for student ability. Importantly, we have also linked the early 2014 survey to follow up data through early 2019. In these five years we observe all applications to university, as well as enrollment, and eventually graduation and labor market earnings.

[16] These 10 dimensions are, in survey order: probability of passing the degree, probability of enjoying the coursework, expected hours per week of studying, probability that family will approve of choice, probability of finding a job directly after graduation, probability of job satisfaction at age 30, probability of being able to combine work and family life at age 30, expected hours per week of work at age 30, expected earnings at age 30, expected earnings at age 40, and, social status (separate from salary) they associate with the education. See Pihl et al. (2019) for more details, and later sections in this paper were we look at these other outcomes.
[17] At the end of the survey the students were randomized into an offer to see actual average earnings for each of the fields. We found no impact of the treatment on their subsequent application and enrollment behavior. Thus, we feel confident ignoring the experiment for the purposes of this study.

A key variable we use throughout is choice of education. We define this separately for the stated choice as answered in the survey, and revealed choice both through applying as well as through enrolling in a program. When using the administrative follow-up data to define choice of education, we include only degree-granting programs. For field of application, we assign the student to the field of the degree-ranking program that they ranked as their top choice. If they did not apply to university in Sweden between 2014-2019, or they applied by never ranked a program as their top choice (just a course), they were assigned to the no-college choice option. Likewise, for enrollment, if the student was never enrolled in a program in Sweden, we assign them to the no-college choice option.

2.4 Descriptive statistics and validation of the earnings expectation and education measures

Table 2 summarizes the students' responses to the two questions on anticipated earnings after graduating with a degree in each of the eight field of study categories, and the two questions on earnings without going to university. Although we focus on the choice between any college field and no-college, looking at field-specific earnings is useful in order to evaluate how the students responded to the survey. The table uses the mean of each student's expectation at ages 30 and 40, which we think of as a proxy for expected lifetime earnings. At the low end, students expect average earnings of 26,600 SEK per month in the world where they do not attain more education. At the high end, they expect 43,700 SEK per month in the world where they attain a Social Sciences degree.

A common critique of subjective expectations data is that survey respondents may not exert effort in their responses, yielding data that is not a true reflection of their beliefs.[18] If this were the case in our data, we would expect to see expected earnings that did not line up well with reality. In fact, we find that students have reasonable expectations, suggesting that they have thought about this question and acquired relevant information. We can see this in Figure 1, which plots the survey expected earnings on the horizontal axis and population averages (in 2018 for workers around age 40) for each field on the vertical axis. The correlation is 0.862, with the students clearly separating the low-earning categories (no college, teaching,

[18] Cognitive dissonance is also a potential issue with the potential to bias estimates. We discuss this separately in Section 3.4.

humanities, animal/agriculture, services) and the high-earning categories (health, social sciences, engineering and sciences). The levels of earnings the students expect are generally higher than what we observe in the full population in 2018, however when we restrict the population earnings to only workers with jobs in Stockholm, the expected earnings levels are nearly a perfect fit. Additionally, they expect a fairly wide spread of earnings between fields, which tells us that they perceive meaningful differences in remuneration based on education.

Table 2: Mean Expected Earnings by Field

Figure 1: Comparing Expected Earnings to Population Earnings

Another way to show that students took the survey seriously and responded with their true expectations is to use the follow-up data described in the previous section. With this data, we can compare what they said they planned to do in the survey, with what they actually did do over the subsequent 4.5 years. This time horizon covers most of the students' entry into university, but is not far enough to capture their full labor-market potential. Thus, we focus on comparing their stated intention to go to college, to whether they do actually apply or enroll and in which field.

Table 3 summarizes the correspondence between the eight fields of study plus no-college in expectation and the administrative data. We see that roughly 47% of students pursue the same education category that they said they planned to. This is higher, at 55%, if we only look at university fields and those who apply to university. The discrepancy is that while only 23 individuals in the survey said they did not plan to go to college, roughly 100 do not apply to a program (by our definition) by 2019. The correspondence between survey and enrolling in a degree program is similarly high. Of those who enroll in a college program, 51.5% enroll in the same field they said they expected to. There are an additional 58 people who apply to college, but are not enrolled by 2019. This makes the overall match (including no-college) somewhat lower at 38%. These shares are much higher than random allocation to fields, and suggest strong informational content of stated educational choice.[19]

[19] Arcidiacono et al. (2020) and Wiswall and Zafar (2020) are able to use subsamples where they can track actual earnings and compare to expectations and find positive and sizable correlations. This is also something we will be able to do soon for all the survey respondents, through the use of administrative earnings registers and the survey respondents' personal identifiers.

Table 3: Correspondence between stated and revealed educational choices

3 Conceptual framework and estimation issues

With our collected information on each individual's expected earnings in two educational states $S \in \{0,1\}$ we can characterize the ex-ante earnings levels and returns by estimating the means and distributions of the potential outcomes y_{0i}, y_{1i}, and hence the gains $y_{1i} - y_{0i}$. We do this both unconditionally and conditional on treatment status, i.e., choosing S=0 or S=1, where S can represent either the stated choice (in the survey) or the application and enrollment (in administrative data). In our main analysis, and in the framework presented in this section, we frame the choice between university (S=1) and high school (S=0) as the two educational states. However, the reasoning is applicable for any binary categorization, such as high- and low-paying fields of university study.

We will first discuss our parameters of interest and then turn to estimation issues. Note that we always think of the potential outcome variables y_{0i} and y_{1i} as expressed in *log* expected earnings, where y_{1i} is the log of expected earnings in the chosen university field. Although our goal is to estimate means as well as the distributions, the discussion mostly focuses on means.

3.1 Treatment effects

Some ex-ante average treatment effects of interest are:

$$ATE = E[y_{1i} - y_{0i}] \quad (1)$$

$$TT = E[y_{1i} - y_{0i} | S = 1] \quad (2)$$

$$TUT = E[y_{1i} - y_{0i} | S = 0] \quad (3)$$

where ATE is the ex-ante average treatment effect; TT is the ex-ante average treatment effect for those choosing university, and; TUT is the ex-ante average treatment effect for those not choosing university. If $TT \neq TUT$, the expected gain from attending university differs between individuals choosing and not choosing university. Hence, in the case of systematic sorting of individuals based on their expected gains, recovering these separate treatment effects is necessary for education policy.[20] With our data, the ex-ante treatment effects from choosing university can be directly estimated by calculating mean expected earnings for counterfactual outcomes. To fully understand the gains and sorting pattern we also estimate the counterfactual distributions across educational states.[21]

3.2 Characterizing sorting

To connect the various ex-ante treatment effects with parameters for the degree of sorting across educational states, we closely follow the setup in Heckman, Lochner and Todd (2006) (henceforth HLT), who lay out a framework for interpreting treatment effects and sorting parameters when estimating the returns to college using observational data.

The (log) earnings expectations associated with the two educational states $S \in \{0,1\}$ are specified as $y_{0i} = \alpha + u_{0i}$ and $y_{1i} = \alpha + \bar{\beta} + u_{1i}$. The terms u_{0i}, u_{1i} are random variables with $E(u_{0i}) = E(u_{1i}) = 0$ so that the means of the potential outcomes are $E(y_{0i}) = \alpha$, $E(y_{1i}) = \alpha + \bar{\beta}$. Hence:

$$\beta_i = y_{1i} - y_{0i} = \bar{\beta} + u_{1i} - u_{0i} \qquad (7)$$

is the individual expected earnings gain from choosing university over high school, and

$$ATE = E[\beta_i] = E(y_{1i} - y_{0i}) = \bar{\beta} \qquad (8)$$

is the ex-ante average treatment effect (ATE) in the population. Since β_i is heterogeneous in this framework (as long as $u_{1i} \neq u_{0i}$) the average treatment effects can differ in sub-populations, including the ones defined by treatment status S. The ex-ante average treatment

[20] Heckman and Robb (1985) are the first to discuss how to separately estimate ex-post versions of these parameters using observational data.
[21] Ex-ante and ex-post treatment effects are separately estimated using observational data in Cunha, Heckman and Navarro (2005). This literature is further surveyed and discussed in Cunha and Heckman (2007) and Heckman and Vytlacil (2007). Arcidiacono et al. (2020) and Wiswall and Zafar (2020) are the first that use subjective expectations data to estimate ex-ante treatment effects.

on the treated effect $TT = E[\beta_i|S = 1]$ and the ex-ante average treated on the untreated effect $TUT = E[\beta_i|S = 0]$ can be expressed as:

$$TT = E[y_{1i} - y_{0i}|S = 1] = \bar{\beta} + E[u_{1i} - u_{0i}|S = 1] = \bar{\beta} + SE_1 \quad (9)$$

$$TUT = E[y_{1i} - y_{0i}|S = 0] = \bar{\beta} - E[u_{0i} - u_{1i}|S = 0] = \bar{\beta} - SE_0 \quad (10)$$

where the first sorting effect, SE_1, is how much *more* those who actually choose S=1 expect to get in returns to college over an average person (labelled Sorting Gain in HLT). Likewise, SE_0 is how much *less* those who actually choose S=0 expect to get in returns from college, compared to an average person.[22] These sorting effects are of particular interest to us, as they are informative in characterizing ex-ante sorting behavior, and (as we will see below) of the sources of bias in random coefficient models relating schooling to earnings.

If the ex-ante returns vary with choice of educational states, it implies that individuals sort themselves among educational states due to their expected earnings gains (or something correlated with these) associated with a particular educational choice. If individuals sort themselves to the S in which they expect to have an earnings gain advantage compared to the individuals choosing $S' \neq S$, we would have that $y_{1i} - y_{0i}$ is larger among those selecting university than among those selecting high school.[23] Such sorting based on *comparative advantage* in expected earnings gains implies $TT = E[y_{1i} - y_{0i}|S = 1] > E[y_{1i} - y_{0i}|S = 0] = TUT$.[24] The degree to which comparative advantage holds at the mean in the population can therefore be directly estimated by $TT - TUT = SE_0 + SE_1$.[25]

[22] Or, if we think of no-college as the treatment, SE_0 should be viewed as the extra gain that the S=0 types get from no-college relative to the average person.
[23] Which also implies that $y_{0i} - y_{1i}$ is larger for those selecting high school than those selecting university.
[24] The theory of comparative advantage connects earnings and ability formally in the following way (see Sattinger, 1993 and Kirkeboen, Leuven and Mogstad, 2016). Earnings, $Y_i^S = \pi^S q_i^S$, is the product between price (per unit of worker output) π and productivity q, where prices differ only between educational states $\mathrm{S} = \{\mathrm{S}, \mathrm{S}'\}$. In this simple model, an individual *i* is said to have a comparative advantage (i.e., relative productivity advantage) over individual *i'* in state S, and the individual *i'* has a comparative advantage over individual *i* in state S′, if the earnings return from state S for individual *i* is higher than for *i'*: y_i^S-$y_i^{S'}$>$y_{i'}^S$-$y_{i'}^{S'}$, where y equals log (Y). In our ex-ante case, we might think of q as perceived productivity and y as expected earnings, which then depends on perceived productivity and expected wages associated with a certain choice.
[25] Kirkeboen, Leuven and Mogstad, (2016) estimates the sum of these sorting effects across multiple fields of study choice using observational data for Norway.

From this framework we can also infer the following: First, homogeneous gains from treatment or heterogeneous gains which are independent of treatment status both imply identical average treatment effects across educational states, and therefore no sorting based on comparative advantage in the population.[26] Second, if we find evidence in our data of both comparative advantage and $TUT \geq 0$, it indicates a deviation from a pure earnings maximization framework since the gain from university is non-negative also for those who select high school, on average.[27] [28] This finding can be reconciled with earnings maximization in a generalized Roy model including barriers to entry such as costs of schooling, or if the object of maximization is a broader utility measure. We will return to this in Sections 4.2 and 5.2. Third, if individuals choose according to their absolute (expected earnings) advantage at each chosen schooling level we would have $E[y_{1i}|S=1] > E[y_{1i}|S=0]$ and $E[y_{0i}|S=1] < E[y_{0i}|S=0]$.[29] Hence, those that choose university expect to have higher university earnings than what the high school choosers would have had had they chosen university. At the same time, those that choose high school expect to have higher high school earnings than what the university choosers would have had had they chosen high school. Note that absolute advantage implies comparative advantage,[30] but that comparative advantage does not imply absolute advantage.

3.3 Connection to OLS estimates of the returns to education

Here, we use the framework in HLT to illustrate how OLS estimators are related to various treatment effects and sorting parameters which, if data is available on counterfactual outcomes, can be identified directly.[31] With our data on counterfactual expected outcomes, we can directly estimate ex-ante versions of these components and, if certain assumptions are met, use them to infer how sorting affects ex-post returns estimated with observational data.

---

[26] If β_i instead is assumed to be the same for everyone, so that $u_{1i} = u_{0i}$, we have that $SE_0 = SE_1 = 0$, and hence $ATE = TT = TUT$. Alternatively, if $u_{1i} \neq u_{0i}$ but u_{0i} and u_{1i} both are independent of S, we have that $SE_1 = -SE_0$ and hence that $TT = TUT$ as well.

[27] If individuals simply choose the S in which they expect to earn the most, we would have that $y_{1i} > y_{0i}$ for those selecting S=1 and $y_{1i} < y_{0i}$ for those selecting S=0. This implies $TT = E[y_{1i} - y_{0i}|S=1] > 0$ and $TUT = E[y_{1i} - y_{0i}|S=0] < 0$.

[28] Ignoring the earnings that university students forego while studying at university (see Section 3.4 for an additional discussion).

[29] See Willis and Rosen (1979) and Sattinger (1993) for a discussion of the concept of absolute advantage in an earnings-education framework.

[30] Since Absolute advantage implies $E[y_{1i}|S=1] - E[y_{1i}|S=0] > 0 > E[y_{0i}|S=1] - E[y_{0i}|S=0]$ which implies $E[y_{1i} - y_{0i}|S=1] > E[y_{1i} - y_{0i}|S=0]$.

[31] HLT discuss how to identify various treatment effects and sorting parameters in a generalized Roy model allowing for heterogeneous returns to be correlated with educational choice, using observational data and OLS and IV estimation techniques.

If we relate expected earnings to the potential expected earnings outcomes as $y_i = Sy_{1i} + (1-S)y_{0i}$,[32] the probability limit of an OLS estimator (from the regression of expected earnings y_i on a university dummy S) is then:

$$
\begin{aligned}
OLS = E[y_i|S=1] - E[y_i|S=0] &= E[y_{1i}|S=1] - E[y_{0i}|S=0] \\
&= \bar{\beta} + \{E[u_{0i}|S=1] - E[u_{0i}|S=0]\} + E[u_{1i} - u_{0i}|S=1] \\
&= ATE + SB_{u_0} + SE_1
\end{aligned}
\tag{11}
$$

where $SE_1 = E[u_{1i} - u_{0i}|S=1]$ is the sorting gain from university education for those choosing university and $SB_{u_0} = E[u_{0i}|S=1] - E[u_{0i}|S=0]$ is the traditional Selection Bias term, which is the selection bias with respect to high school earnings. $SB_{u_0} > 0$ indicates positive selection into university (or negative selection into high school), and $SB_{u_0} < 0$ indicates positive selection into high school.

By using equations (8)-(10) we therefore also have that:

$$OLS = TT + SB_{u_0} \tag{12}$$

$$OLS = TUT + SB_{u_0} + SE_0 + SE_1 = TUT + SB_{u_1} \tag{13}$$

where $SB_{u_1} = E[u_{1i}|S=1] - E[u_{1i}|S=0]$ denotes the selection bias with respect to university earnings, and where $SB_{u_1} > 0$ if there is positive selection into university.

The Roy model, as discussed and formalized in Willis and Rosen (1979), therefore allows for two types of selection into educational states, which might or might not have the same sign. If $SB_{u_0} > 0$ and $SB_{u_1} > 0$ there is always (regardless of choice) positive selection into university, in that those with higher-than-average expected earnings ability (in either state) end up choosing university, whereas if $SB_{u_0} < 0$ and $SB_{u_1} > 0$ there is positive selection into high

[32] This can be rewritten as $y_i = y_{0i} + (y_{1i} - y_{0i})S = \alpha_i + \beta_i S_i = \alpha + \beta_i S_i + u_{0i}$ which is equivalent to a random coefficient model. Since $y_i = \alpha + \bar{\beta}S + u_{0i} + S(u_{1i} - u_{0i})$, we have in a constant coefficient model implicitly assumed $u_{0i} = u_{1i}$, so that $y_i = \alpha + \bar{\beta}S + u_i$, so that the only source of bias in an estimate from an OLS regression would be the correlation between u_i and S, the traditional selection (or ability) bias. Note that we here also implicitly assume no ex-ante general equilibrium effects, i.e., that a surveyed individuals' expected earnings should not be affected by the educational choices of other individuals.

school (those with higher-than-average expected high school earnings choose high school) and positive selection into university (those with higher-than-average expected university earnings choose university). An earnings-education model with homogenous returns assumes $u_{0i} = u_{1i}$ and, hence, that $SB_{u_0} = SB_{u_1}$. This restricts earnings ability to be one-dimensional and sorting to be hierarchical. The Roy model instead allows earnings ability to be multidimensional in nature and for sorting based on comparative advantage or absolute advantage at each schooling level to exist.

Before we turn to a discussion about estimation issues, including about the assumptions necessary for using our data on counterfactual outcomes to estimate all the parameters in equations (11)-(13), we briefly summarize what we have learned from observational studies of ex-post treatment and sorting effects. In the conventional literature, researchers argued the existence of ability bias, i.e. SB_{u_0}>0, and therefore that OLS estimates of returns to schooling are overestimates of the average returns to schooling in a setting where the individual returns to education are approximately constant in the population. However, at least since Griliches (1977), there has been disagreement among researchers about the existence and magnitude of an "ability bias" in estimations of returns to schooling. Card, 1999, surveys the evidence and conclude that ability bias most likely is positive but small. Papers that have explicitly modelled the choice of education and allowed for heterogeneous returns, in combination with imposing strong econometric identification assumptions, find support for SB_{u_0}<0 (see Willis and Rosen, 1979; HLT; Carneiro, Heckman and Vytlacil, 2011, who all use data for US, and Nybom, 2017, who uses data for Sweden). This would mean that high school only students earn more with that education level than what college goers would, had they chosen high school and would support sorting based on comparative advantage and absolute advantage at each schooling level. [33]

There is less evidence available on the sign of the bias due to heterogeneous returns and of the separate sorting effects (the SEs) from observational studies of returns to college or university. A particular problem arising with heterogeneous treatments effects in studies using instrumental variables, is that IV estimates a weighted average of compliers (Imbens and Angrist, 1994). If the compliers consist of individuals with higher-than-average returns, such

[33] Nybom (2017) finds $TT > OLS$ in the semiparametric model but not in the parametric normal model, and also that observable ability measures leads to a OLS estimate when they are included as control variables, which is in line with a positive ability bias.

studies would overestimate the average return in the population (Card, 1999). Card (2001) models the returns as heterogeneous and showed that (even if $SB_{u_0} = 0$) sorting based on comparative advantage (higher returns to schooling for those selecting more schooling), would be expected to lead to an upward bias of the average marginal return (the ATE).[34] In the studies by HLT and Carneiro, Heckman and Vytlacil (2011), estimates of SE_1 are quite large. The difference between ATE and TUT is positive and large, suggesting a positive SE_0 as well. Hence, both sorting effects are important. In Nybom (2017), TT is significantly larger than ATE, suggesting a positive SE_1. The difference between ATE and TUT is smaller, but still positive, suggesting a positive SE_0.

3.4 Estimation issues

The ex-ante treatment effects and sorting parameters discussed above are summarized in Table 4. They are all constructed from mean expected earnings in the two education states, and hence straightforward to estimate with our data. However, for them to be unbiased estimates, we must make assumptions about the relationship between students' expectation errors and the educational state. These issues are discussed at length in Arcidiacono et al. (2020).

Table 4 about here

If the ex-ante y_{0i}, y_{1i} and S are measured without error, it is straightforward to estimate all the ex-ante treatment effects and sorting parameters from our data. In fact, ex-ante ATE is still identified even in the presence of measurement errors, as long as the measurement errors have the same mean across educational states (Arcidiacono et al., 2020). For the ex-ante TT and TUT we also need that the earnings expectation errors are independent of potential miss-classification of S. Since we have information on S at different stages (survey to enrollment), we have less of an issue with classification error if the treatment of interest is whether an individual graduates or not, although our data do not extend long enough to see the end of most university spells.[35]

[34] In the framework of HLT this follows from the equation (11) since a larger SE_1 will be equivalent to a higher degree of comparative advantage (holding SE_0 constant).

[35] This is similar in spirit to Arcidiacono et al. (2020), who not only defines treatment effects that are ex ante with respect to earnings and choice of occupation, but also that are ex ante with respect to earnings but ex post with respect to choice of occupation.

Estimating the distribution of ex-ante treatment effects correctly in the presence of measurement errors in expected earnings requires stronger assumptions (see Arcidiacono et al., 2020, Appendix A.5). If the measurement errors are classical, the dispersion of the true ex-ante distributions are overestimated. Hence, our estimated ex-ante distributions are likely to be inflated as they partly include measurement errors. For our most important results regarding these distributions, we compare the distribution of ex-ante treatment effects for treated and untreated. Hence, if measurement errors are similar across treated and untreated, the patterns are likely to be similar, even in the presence of classical measurement error.

Previous research has worked to establish that the type of subjective expectations data that we use is not subject to cognitive bias. Specifically, Zafar (2011a; 2011b) collected expectations for the same individuals twice after they chose a major, and does not find that students rationalize their choice by becoming more positive in their expectations of their chosen field (relative to their not chosen fields) over time. This is contrary to a story of cognitive dissonance (i.e., individuals overestimate the benefits of their choice relative to the things they have not chosen). Zafar (2011a) also provides evidence that students exert sufficient mental effort in their response, and that their expectations are well formed and that measurement error in their responses is classical. These other findings are consistent with what we see in our data, although we cannot repeat his test for cognitive dissonance with a single period's observation of expectations.

Inferring what the ex-ante treatment effects say about *ex-post* treatment effects require much stronger assumptions, since this also requires rational expectations and no unanticipated earnings shocks (Arcidiacono et al., 2020). Such assumptions become relevant if we attempt to use our data on earnings expectations to learn about the sources of bias inherent in ex-post OLS estimates of treatment effects when returns are heterogeneous and correlated with educational choice.

The selection bias and sorting parameters discussed in the previous section are expressed in terms of u_{0i} and u_{1i}, which are deviations from the means of ex-ante versions of the potential outcomes y_{0i} and y_{1i}. Using realized earnings data, we can never observe these potential outcomes, but only the realized y_i. However, we can still think of ex-post potential outcomes (or the ex-post potential return) as including two components: one that the individuals cannot forecast at the time when they make their schooling choices (*uncertainty*)[36] and another which they can predict (*heterogeneity*).[37] This distinction is made in the literature attempting to distinguish ex-ante and ex-post returns using observational data (see Carneiro, Hansen and Heckman, 2003; Cunha and Heckman, 2007; and HLT, Chapter 10).[38] Ex-post earnings data capture both the unpredictable and predictable components, whereas ex-ante data only capture the predictable component.[39] Hence, if our estimated biases from ex-ante data differ from those in well-designed observational studies, it could be because schooling choices are correlated with the earnings components that cannot be foreseen at the time when the schooling choices are made.

In Section 2.4, we showed that the expected earnings are highly correlated with the observed average earnings across educational choices (Figure 1). We also note that field of study in Sweden is very highly predictive of ultimate occupation (Björklind et al., 2016).[40] However, we still want to emphasize that our results are only suggestive for those conclusions that require ex-post validity. Additionally, our questions about future earnings are asked conditional on graduating and finding employment, so our estimates are not intended to take risk of drop-out or unemployment into account.

Finally, our main estimates of ex-ante returns are calculated based on the average of individuals' earnings expectations at age 30 and at age 40, rather than lifetime earnings. These could differ because of foregone earnings due to education duration, and because of different earnings growth rates. However, previous research using Swedish administrative data has shown that life-cycle bias is quite low for workers if their earnings are measured at these ages (Böhlmark and Lindquist,2006; Nybom and Stuhler. 2016), at least for men. Discount rates could also differ between individuals, but since our estimates are within individual, they should not impact our results.

[36] Say because of an exogenous shock that changes the returns to university ex-post.
[37] In our case the latter would be what is known to the individual at the time of the survey.
[38] See also Heckman and Vytlacil (2007) who distinguish between the true earnings beliefs and true earnings.
[39] HLT surveys the literature and finds that both are important.
[40] Swedish degree programs are more specialized than American ones (lacking the broad liberal arts foundation).

4 Estimated ex-ante earnings levels and returns distributions: university versus high school

In this section we describe the distributions of ex-ante earnings levels and returns. We begin by describing them overall (Section 4.1), followed by between and within (Section 4.2) the high school and university educational states.

4.1 Ex-ante earnings levels and returns distributions: overall

The overall earnings distributions for high school and university fields of study are shown in Figure 2 for high school (the solid red line) and stated university field of choice[41] (the dashed blue line).[42] In the figure we also show the eight university fields of study alternatives (as solid grey lines). Each distribution consists of all individuals in our sample (regardless of their intent to go to college or not).

Figure 2 about here

The individuals' expectations of high school earnings, i.e., if they were to choose not to continue to university, show that most of responses lie between 20,000 and 35,000 SEK in monthly pre-tax earnings, with a mean of 26,400 SEK (roughly €2,640) and a standard deviation of 6.9. The individuals' expectations of university earnings, i.e., if they were to attend their preferred university field, show a distribution that is located more to the right and is more dispersed. The bulk of the responses are between 25,000 and 50,000 SEK in monthly pre-tax earnings with a mean of 38,600 SEK and a standard deviation of 11.9. The mean is about 46% higher and the standard deviation (SD) is 72% higher for preferred fields of university compared to the high school distribution.

By using the distribution of counterfactual earnings for high school and (preferred) university fields of choice, we can construct the distribution of ex-ante university premiums. Figure 3 shows distributions for preferred university field relative to high school earnings (the dashed

[41] Here we use survey-stated chosen field. If we use chosen field based on application or enrollment, the university distribution (Figure 2) and returns to education distribution (Figure 3) look very similar.
[42] Students who chose the no-college option in the survey don't have a preferred "university field". For them we use a weighted average of all 8 college fields, where the weights are the popularity of these fields among the rest of the sample, as their expected university earnings.

red line) and for the average of all university fields relative to high school earnings (the solid blue line).[43] As the former distribution is located more to the right, individuals' expected earnings for their preferred university field is higher than in non-preferred university fields of study. Hence, prospective students sort into fields where they expect higher earnings. We focus on comparing the preferred university fields with choosing to stop at high school. This captures the ex-ante returns to university education which is most comparable to the ex-post observable university return. The mean return is 36%, but varies quite a lot. We note that 9.5% of the prospective students expect negative returns to university.

Figure 3 about here

Next we investigate which characteristics correlate with an individual's ex-ante returns to university. We want to know if observable characteristics are predictive of expected returns, and to what extent. To do this we perform OLS regressions of the individual return ($\beta_i = y_{1i} - y_{0i}$) on *gender*, a *socio-economic status (SES) index* (having mean zero and standard deviation equal to one) based on parents education and income, if individual is a first- or second-generation immigrant (*foreign*), and on two high-school performance variables determined prior to the survey: *math score*, *English score*. In addition, in some regressions we include the grade point average (*GPA*) at the end of high school, as it is a broader achievement measure and is important for admission to popular university fields of study. The high school achievement variables are all standardized to have mean zero and standard deviation one in the estimations. Table 5 reports the resulting estimates, from OLS regressions on each variable separately (columns 1-6), on all variables combined (columns 7 and 8), with the last column also including field of study indicators.[44] We also checked to see if the sorting patterns differed within high school program specialization, and found no meaningful change relative to column 8.

Table 5 about here

Individual ex-ante returns to university education are correlated with some observable characteristics, including being higher for females, those with higher SES backgrounds and

[43] The detailed field of choice premium distributions (relative to high school) are shown in Appendix Figure A1.
[44] The results using the average of all university fields (rather than the expected earnings in chosen field) relative to high school are qualitatively similar.

those with higher math scores.[45] However, the relationship to previous school performance measures are mixed, as the estimates for English and GPA both are statistically insignificant.[46] Including all these characteristics simultaneously gives an R2 of 0.061, hence a lot of the heterogeneity in the individual return remains unexplained. In the last column, we see that the estimates remain similar using only variation within fields of choice, even though the R2 increases to 0.236. Hence these results are not due to students with different characteristics making different educational choices.

In theory it could also be possible that the differences in expected returns are due to differences in information quality between groups. One piece of evidence against this explanation is an information experiment that we conducted on the same sample at the conclusion of the survey. Students were randomly offered the chance to see true average population earnings within field. If poor information in some groups (e.g., low SES) were the cause of differential expected returns, then we would expect that the information would induce more field switching for these individuals in the follow-up data. We find no evidence of this: no group responded to the information by differentially changing their application and enrollment behavior over the subsequent years.[47]

4.2 Ex-ante earnings levels and returns distributions: by educational states

In this section we estimate the means and distributions of expected earnings levels and returns separately by choice of level of education. This includes estimating various ex-ante treatment effects and sorting parameters (as summarized in Table 4), and inferring what these imply for models of educational choice and for returns to schooling estimations. We estimate these separately by stated choice, application and enrollment (as discussed in Section 2).

4.2.1 Distributions

[45] The result for SES is in line with estimates in Boneva and Rauh (2020) for UK.

[46] Our results are in line with Card's (1999) review in the handbook of labor economics who suggests that return to education is positively related to SES, but only possibly to measured ability, but less in line with results in Nybom (2017) who finds that the returns to education is strongly positively associated with ability, but only with parental earnings unconditional on ability.

[47] The pre-analysis plan for analyzing the information experiment is available at https://www.socialscienceregistry.org/trials/4329/history/49081.

We start by looking at the potential earnings distributions. These distributions are shown separately by the two educational states in Figure 4a-4c. The solid lines are the distributions of earnings for those that stop at high school (blue line) or continue with university (red line). The dashed lines are the counterfactual earnings distributions. Hence, in each figure we illustrate two "actual" (i.e., for the level of education they chose) ex-ante earnings distributions and two counterfactual earnings distributions, for the treated and untreated groups, respectively. The means (in logs) of these four distributions are shown in Panel A of Table 6.

Figure 4 about here

Table 6 about here

We can infer at least two interesting findings from these figures. First, those who choose university have the highest university earnings expectation, higher than what those that stop at high school would have had if they would have chosen university.[48] However, as we see in the figures, there is still a lot of overlap of the distributions. There are a sizable number of high school graduates that, had they chosen university, expect to do better than many university graduates. Second, those who prefer high school have a similar earnings distributions to what those who choose university would have had, had they chosen high school. On average, the expected earnings for high school is slightly higher for those choosing high school, between 2-7%, across the stated, applied and enrolled divisions, although not statistically significantly so.[49]

It is notable that our pattern of results presented in these figures are similar to results in Cunha, Heckman and Navarro (2005) who use factor models applied to representative observational data for the US to create counterfactual earnings distributions.[50] They find that high school graduates are somewhat more successful than college graduates, if the latter would have

[48] This finding holds even if we use a simple mean of all university fields for the high-school choosers.
[49] This can be seen from the "Selection Bias" ($E[y_1|S = 0] - E[y_0|S = 0]$) estimate in the first row in Panel C of Table 6.
[50] See figures 6.1-6.4 in Cunha, Heckman and Navarro (2005). They also find that results for ex-ante and ex-post distributions are similar, which they argue is the in line with heterogeneity (as opposed to uncertainty) explaining most measured variability in earnings. The literature using observational data to analyze counterfactual earnings distributions is surveyed in Cunha and Heckman (2007). This literature is distinguishing between ex-post counterfactual distributions, as well as ex-ante counterfactual distributions, the latter estimated from their ex-post data.

stopped education at high school. However, the overlap of the distributions are substantial. Additionally, they find that college graduates are more successful than high school graduates, if the latter would have gone to college.

4.2.2 Average treatment effects and sorting patterns

In Figures 4d-4f, we use the expected earnings data and calculate the distributions of returns to university separately for those choosing university (S=1, TT) and for those choosing high school (S=0, TUT). Table 6 reports the estimates of the average treatment effects overall and for these two groups (Panel B). Although it is evident from the figures that there is a lot of distributional overlap, the estimates of average treatment effects provide us with some clear findings. The average treatment effects are estimated at 0.36, which implies about 9% per year of college. We find that the ex-ante average treatment on the treated effects are estimated to be only slightly higher, around 0.38, whereas the ex-ante average treatment on the untreated effects are estimated lower but still positive as 0.18 to 0.30, so about 5-7% per year, with larger TUT estimates if choices are revealed instead of stated. That TT>TUT>0 implies sorting with respect to comparative advantage and that those that choose high school have positive expected earnings gains from choosing university, even though they do not choose university. We now turn to a discussion about the implication of these results.

First, our finding of a positive and large TUT goes against the pure Roy model of earnings maximization and instead suggests that there are other factors affecting these students' choices. One such candidate is direct cost of schooling. However, since all universities in Sweden are free and living expenses are covered by generous grants and subsidized loans, this is very unlikely to be the explanation. Another candidate is psychic costs of schooling, which has been shown to be important in rationalizing college education choice behavior in the US (see Heckman, Lochner and Todd, 2006). As part of the survey, we asked the students about their expected enjoyment of study, the probability of graduating on time, the expected number of study hours, and the degree of parental approval associated with a choice of fields of university study, and we use these measures for the preferred field choice as proxies for psychic costs of university education. In Appendix Table A2 we regress the expected university earnings returns on these measures separately for those choosing high school and university. Ignoring the results in the first column (which are very imprecise) there is clear evidence that these measures are unrelated to expected returns for those that do not choose university, whereas they are somewhat related to returns for those that choose university (for which longer expected study hours and (possibly) higher parental approval is associated with higher earnings return). Hence, psychic costs, at least as captured through these measures, are unlikely to explain our finding of a positive TUT.

Another possibility is compensating non-pecuniary factors: that giving up earnings from choosing high school might be expected to benefit individuals later in life in ways other than earnings. Unfortunately, we do not have access to direct counterfactual information about non-pecuniary outcomes in the choice to discontinue education after high school, but we return to this explanation in Section 5.2 using when we investigate ex-ante treatment effect for high versus low paying university fields of choice.

Second, the fact that we find on average $TT > TUT$ means that the expected relative gain from choosing university is higher for those choosing university than for those choosing high school, on average. This is consistent with students sorting based on their comparative advantage in each respective educational state. Since TT is only somewhat larger than ATE, whereas TUT is clearly smaller than ATE, the sorting gain from choosing university, for those who did choose university (compared to the average person) (SE_1), appears to be small. The sorting loss from choosing high school, for those who choose high school (SE_0), is large, although the difference between SE_1 and SE_o decreases when we use revealed choices (see Panel C).[51] As discussed in Section 3.3., we can also express the sum of these two sorting effects as the sum of two selection bias terms: SB_{u_1}, indicating the degree of positive selection into university (if expected university earnings is higher for those choosing university) and $-SB_{u_0}$, indicating the degree of positive selection into high school (if expected high school earnings is higher for those choosing high school). Since TT is somewhat larger than OLS, whereas TUT is clearly smaller than OLS, we estimate a sizable positive selection into university, and a smaller negative (but statistically insignificant) positive selection into high school (see Panel C).

As we discussed in Section 3.3., the magnitudes of these various sorting parameters tell us the bias in using an ex-ante OLS estimate of average return to university education. Since the ex-ante version of the traditional selection bias (SB_{u_0}) is estimated small or negative, the ex-ante

[51] As shown in section 3.3, the difference between the ex-ante TT and TUT estimates consists of the two sorting effects SE_0 and SE_1.

OLS estimate is fairly similar to an ex-ante TT estimate. Since the sorting gain from choosing university, for those who choose university (compared to the average person) (SE_1), is small, the bias in an OLS estimate of ATE due to heterogeneous returns is positive but small (in line with what was argued in Card, 1999, 2008), whereas the selection bias tends to lead to a slight underestimate of ATE. Since they go in opposite directions, the OLS estimate is very similar to the ATE estimate. Inferring TUT from OLS, however, leads to a large downward bias, driven mostly by SE_0.

Note that if we are to compare our results for the estimated ex-ante treatment effects in Panel B with those from observational studies, we find that they are similar to those in Carneiro, Heckman and Vytlacil (2011) and Heckman, Lochner and Todd (2006) in that $TT > ATE > TUT$, where OLS is a downward biased estimate of TT. However, they differ in other dimensions, since they find TUT to be roughly zero; that OLS is an upward biased estimate of ATE and; that the difference between TT and TUT is due to a positive sorting gain (SE1>0). If we instead compare our ex-ante estimates to Nybom (2017), who uses observational administrative data for Sweden, the results are remarkably similar to what we find. Nybom (2017) finds that $TT > ATE > TUT$, where TUT is positive, and that the OLS estimate, without controls, is very similar to the ATE estimate. As OLS<TT, the sign on the selection bias is found to be negative.[52]

Third, a positive selection into university and a positive selection into high school ($SB_{u_1} > 0$ and $SB_{u_0} < 0$, respectively) would suggest positive selection into both educational states and constitute evidence against the one-dimensional earnings-ability model. This pattern was proposed in Willis and Rosen (1986) and found empirical support in papers by, e.g., Carneiro, Heckman and Vytlacil (2011), using observational data for the US. We find somewhat mixed evidence, since the traditional selection bias (SB_{u_0}) is small and negative or zero, whereas SB_{u_1} is quite large and positive.[53] However, it does suggest that we can reject the simple one-dimensional ability model, since such a model would expect that those with high ability do better in both educational states. This is also in line with our finding of a moderate positive

[52] We compare our estimates to the semiparametric estimates reported in Nybom (2017). The only deviation in the results is that the sorting effect in our case is almost entirely driven by SE_0, something which explains why our ex-ante estimates of TT is closer to (the lower) ATE, whereas in Nybom (2017) the TUT estimate is closer to (higher) ATE estimate.

[53] These patterns are notably similar to those in Cunha, Heckman and Navarro (2005), who use factor models applied to observational data for the US.

correlation between individuals' expected earnings with university and high school education:[54] the correlation is about 0.4 and very similar regardless of educational choice (see the last row in Panel C).[55] This is at odds both with results in papers by Carneiro, Heckman and Vytlacil (2011) and Cunha, Heckman and Navarro (2005), who test for this using observational data for US applied to "factor models" and find evidence of a negative correlation.

Fourth, the finding that the ex-ante TUT estimates increase when we use data on application or enrollment, compared to stated choices, suggests that the students who are switched into non-treatment when preferences are revealed, have higher expected returns than those who already during the survey stated that they will not continue to university. This suggests that expected returns vary across prospective students depending on their propensity to enroll at university education.

To see this more clearly, we first divide those who stated that they would choose university into those that later did apply and those that did not, and those that applied to university into those that later did enroll and those that did not. In this way we can estimate average treatment effects for those that are more or less likely to keep being treated, the latter group constituting prospective students that eventually opted out of treatment but likely were closer to actually take up treatment than those that neither stated nor applied to university. These results are shown in Appendix Table A4, and discussed further in Appendix D.[56] We see that the expected return to university among those that stated that they would go to university but did not apply is estimated as 0.263, which is much lower than the 0.400 among those that did apply. A similar pattern is found among those that applied with a higher expected return for those that eventually enroll than for those that did not (0.405 versus 0.342). A division of individuals into only treated and untreated groups therefore gives an incomplete picture of the heterogeneity of returns, and future work using subjective expectations data to estimate ex-ante treatment effects

[54] If we relate expected earnings in high school (y_{0i}) and expected earnings in (preferred) university field of study (y_{1i}) we can directly estimate the association between perceived earnings ability in high-school and university, respectively, since $Cov(y_{0i}, y_{1i}) = Cov(u_{0i}, u_{1i})$ using the framework in Section 3.

[55] See Sattinger, 1993 (Section V.C.), for an insightful discussion about the Roy model, the multidimensionality of ability and the correlation between ability to do well in various chosen states.

[56] In Appendix Figure A2, we elaborate on this by first predicting the propensity to enroll at university, using all the information from the stated, applied and enrollment choices. Relating the expected returns to this predicted propensity to enroll at university, show a positive relationship which further collaborates these results.

should probably continue to explore the distribution of returns with respect to the likelihood of enrolling or choosing treatment.[57]

5 Estimating ex-ante levels and returns distributions across high- versus low paying fields of study choice

In this section we describe similar estimates of means and distributions of treatment effects and sorting parameters as in Section 4, but instead look within university at high- versus low paying fields.[58] The horizontal choice dimension is interesting in its own right, as a comparison with our results for the level dimension in the previous section. In addition, it can be used to make a comparison with the results in Arcidiacono et al., (2020) and Wiswall and Zafar (2020) who estimated ex-ante treatment effects across occupational fields and college majors using data collected from two colleges in the US. As we also have access to counterfactual data on non-pecuniary outcomes associated with different university fields of study, we are also able to provide ex-ante treatment effects for these outcomes, as well as to reconcile them with the results for earnings, which can help rationalize the positive TUT found earlier.

5.1 Ex-ante earnings levels and returns distributions

The estimates of the means and distributions for high- and low paying fields are shown in Table 7 and Figure 5. We see (perhaps unsurprisingly) that expected earnings are higher for the high paying fields, for both those who pursue them and those who do not: the second row of Panel A contains larger mean log expected earnings than those in the first row. The estimated ex-ante ATEs are large and positive (about 0.38) and very similar to the ATEs estimated for university versus high school in the previous section. The estimated ex-ante TTs are larger than the TUTs, but the TUTs are still large and positive. Sorting gain from choosing high paying fields (SE_1) is positive but small, and sorting effects from choosing low-paying fields (SE_0) is positive and

[57] As, for instance, through more explicit estimation of ex-ante versions of treatment effects for those at the margin of participation (Björklund and Moffitt, 1987; Heckman and Vytlacil, 2007; Carneiro, Heckman and Vytlacil, 2011).

[58] We divide the fields of study into high and low paying fields simply by dividing the eight fields of study groups into the four highest and four lowest based on the average expected earnings in Table 2. Hence, the Low-paying fields category consist of fields of study within “Education and teacher training”; “Humanities and Art”; “Agriculture, Forestry and animal health”; and “Services”, and the High-paying fields category consist of fields of study within “Social science, Law, Business, etc.”; “Natural science, Mathematics and Data,”; “Engineering and Manufacturing”; and “Healthcare and social care”.

large. Hence, these results also support sorting based on comparative advantage. The estimated selection bias terms indicate positive selection into both high-paying and low-paying fields (since $SB_{u_1} > 0$ and $SB_{u_0} < 0$) although these terms are only sometimes statistically significant.

Overall, the results here are very similar to those in the previous section where we compare university and high school choice. This is true for the estimates of the average treatment and sorting effects, as well as the distributions. The positive and large TUTs suggests that there are probably other (i.e., non-pecuniary) factors that are very important for this decision. As we collected expectations data on non-pecuniary benefits for field of study, we can investigate this hypothesis directly, something which we do below. We can also compare our results to those in Wiswall and Zafar (2020) who compare TT and TUT for Science/Business versus Humanities/Social Science fields, and find TT>TUT, but that TUT>0, very much in line with what we find here, although their TT estimates are larger than ours.

Figure 5 about here

Table 7 about here

5.2 Ex-ante non-pecuniary returns

In previous sections we showed that the TUTs are positive both for the college choice and high-paying university field choice. This means that individuals systematically leave money on the table when they make their educational choices. One potential reason for this is that there are negative non-pecuniary returns to these same choices which offset the earnings returns in the students' utility function. In this section we investigate if this is the case. We do so by repeating the analysis from Section 5.1. Hence, we estimate ex-ante treatment effects and sorting parameters for these outcomes, and compare the results for those using earnings.

In the survey we asked the students questions about expectations and beliefs about some non-pecuniary outcomes: the probability of finding a job directly after graduation, the probability of job satisfaction at age 30, the probability of being able to combine work and family life at age 30, the social status (separate from salary) they associate with each field of study (and later occupation),[59] and expected hours per week of work at age 30. The work hours question is asked in a scale from 0 to 80 hours, whereas the answers to the other outcomes are provided on a scale from 0-100.[60] Note that we carefully explained what is meant by a probability. For a more detailed description of the general instructions and the specific questions in the survey see Appendix Section A**.** In Appendix Table A3 we show summary statistics in the expected amenities for all fields.[61]

To facilitate comparison between the sizes of the coefficients across variables, we have standardized all expectations amenities to have mean zero and standard deviation one in the full sample. We also reverse hours of work such that a larger value corresponds to a positive outcome (less time spent). These questions were only asked for the college fields because we did not think that imagining the type of job they would get at age 30 or later would be tangible enough for the no-college option.

To test whether these non-pecuniary outcomes explain the positive TUT, we examine returns similarly as we did for the earnings measure in Section 5.1 (and as described in section 3). In Table 8 we condense the presentation of these estimates to the ATEs, TTs and TUTs, as well as for results using data from survey and from enrollment as treatment status.

We find mostly positive and large TT estimates. For instance, those choosing high paying fields of choice expect 1.39 SD higher social status, 0.92 SD higher probability of finding a job, and 1.01 SD higher probability of enjoying the job in a high paying field, compared to what they would have expected to have experienced had they choose a low paying field. Using enrollment in high-paying fields give similar estimates. We also note that the TT estimates for hours worked are positive with those choosing high paying fields expecting to work two-quarter of an hour more.

[59] When we asked the students about perceived social status of the field of study (and resulting occupation) they were specifically instructed to answer independently of the associated earnings level. Because social status is a key concept within sociology we wanted to be able to gauge its importance separately from earnings.
[60] For instance, for each hypothetical choice we asked "How high is the probability that your parents and other family members would approve of your choice of major?" The average response to this question for males was 72.3, meaning that on average they expected that there was a 72.3% chance that their parents would approve the choice.
[61] Expected hours of work per week are high. The mode is 40 hours (28% of the respondents), but over 40% of the respondents provide figures between 41 and 60 hours. This might be due to the survey question which asked about the work hours they need to work, which could be interpreted as full-time work plus overtime etc., and/or that they need to work a lot of hours to keep up in fields where they deem themselves uninterested or untalented.

These TT estimates are notably higher than the TUT estimates for all outcomes. Also, the sign of the TUT estimates are sometimes positive (as for status) and sometime negative (as for enjoying the job and work-life balance). It seems like part of the story for why the TUT estimates for earnings where positive for high versus low paying fields of choice is that prospective students of low-paying fields expect to experience less enjoyment on the job and to be less able to balance work and life, if they would have chosen a high paying field. They give up earnings to instead be compensated in some aspects of non-pecuniary benefits. We also note that given the similarity between the level of education (university/high school) results and the high/low paying fields results, our results here may also be valid for the level results.

6 Conclusions

In this paper we have estimated means and distributions of ex-ante treatment effects for university education relative to high school, as well as to high- and low paying university field-of choice, using elicited earnings expectations associated with counterfactual educational choices. We have shown that average ex-ante returns to university are substantial, with treatment effects for those choosing and enrolling to university being larger than for those who did not choose or enroll at a university. We have also put our results into a framework for estimating the returns to education typically associated with ex-post returns, and found that the traditional (ex-ante) selection bias is small, and possibly even negative, and that although individuals choose in accordance with their expected comparative advantage in earnings, the resulting ex-ante bias due to heterogeneous returns are fairly small.

The use of ex-ante data is not without challenges. It requires high-quality survey data so that elicited expectations are informative about expected outcomes. This is especially true if it is to be compared to future ex-post outcomes. However, it is therefore remarkable how similar our findings are to some of the studies using observational ex-post data and, sometimes, strong econometric identification assumptions, to estimate various means and distributions of treatment effects. For instance, qualitatively, our results from estimating ex-ante treatment effects are much in line with Nybom (2017) who used the MTE framework and observational data in Sweden to estimate ex-post returns to university. Although we find higher average ex-ante returns compared to the estimated ex-post returns in Nybom (2017), we also find agreement with respect to TT to university being larger than the TUT, that the TUT is positive and sizable, and that selection bias is small and negative.

We also estimated ex-ante returns to high- versus low university fields of study (as in Wiswall and Zafar, 2020). We found positive estimates, where TT are larger than TUT and the results in general are qualitatively similar to those for university versus high school. As we also elicited subjective expectations for a set of non-pecuniary outcomes, we find that the choice with respect to traditional high- and low earnings fields, where TUT was found to be mostly positive, can be reconciled with negative TUT returns in some non-earnings factors, such as enjoyment of work and combining family and work. This is in line with an extended Roy model where individuals take into account broader utility when making their educational choices. This is similar to those results in Arcidiacono et al., 2020, and Wiswall and Zafar, 2020, using elicited subjective expectations data for US. Hence, in this way, their results for the US are in line with our results for Sweden, despite the large existing difference in the degree of earnings inequality and system of higher education.

References:

Ajzen, I., Brown, T. C. and Carjaval, F. (2004) “Explaining the discrepancy between intentions and actions: The case of hypothetical bias in contingent valuation”, *Personality and Social Psychology Bulletin*, 30, 1108–1120.

Altonji, Joseph G., Peter Arcidiacono, and Arnaud Maurel. 2015. “The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects.” Working Paper 21655. National Bureau of Economic Research.

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. 2020. “Ex Ante Returns and Occupational Choice.” *Journal of Political Economy* 128 (12): 4475–4522.

Björklind, Karin, Andreas Forsgren, Claes-Håkan Gustafson, Ann-Charlotte Larsson, and Peter Öberg. 2017. “Yrkets Överensstämmelse Med Utbildningen.” Fokus På Arbetsmarknad Och Utbildning. SCB. https://www.scb.se/contentassets/b1ae4493ffd1404987a4d32cbf213ae5/yrkets-overensstammelse-med-utbildningen.pdf.

Björklund, Anders, and Robert Moffitt. 1987. “The Estimation of Wage Gains and Welfare Gains in Self-Selection Models.” *The Review of Economics and Statistics*, 42–49.

Boneva, Teodora, and Christopher Rauh. 2020. “Socio-Economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns.” Mimeo.

Böhlmark, Anders. and Matthew J. Lindquist, 2006. Life-cycle variations in the association between current and lifetime income: Replication and extension for Sweden. *Journal of Labor Economics*, *24*(4), pp.879-896.

Card, David. 1999. “Chapter 30 - The Causal Effect of Education on Earnings.” In *Handbook of Labor Economics*, edited by Orley C. Ashenfelter and David Card, 3, Part A:1801–63. Elsevier.

———. 2001. “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems.” *Econometrica* 69 (5): 1127–60.

———. 2008. “Returns to Schooling.” *The New Palgrave Dictionary of Economics. 2nd Edition, Palgrave Macmillan.*

Carneiro, Pedro, Karsten T. Hansen, and James J. Heckman. 2003. “Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College.” *International Economic Review* 44 (2): 361–422.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil. 2001. “Estimating the Return to Education When It Varies among Individuals.” mimeo.

Carneiro, Pedro, James J. Heckman, and Edward J. Vytlacil. 2011. “Estimating Marginal Returns to Education.” *American Economic Review* 101 (6): 2754–81.

Cunha, Flavio, and James J. Heckman. 2007. “Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling.” *Labour Economics* 14 (6): 870–93.

Cunha, Flavio, James Heckman, and Salvador Navarro. 2005. "Separating Uncertainty from Heterogeneity in Life Cycle Earnings." *Oxford Economic Papers* 57 (2): 191–261.

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

Griliches, Zvi. 1977. "Estimating the Returns to Schooling: Some Econometric Problems." *Econometrica* 45 (1): 1-22.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." *Handbook of the Economics of Education* 1: 307–458.

Heckman, James J., and Richard Robb Jr. 1985. "Alternative Methods for Evaluating the Impact of Interventions: An Overview." *Journal of Econometrics* 30 (1–2): 239–67.

Heckman, James J., and Edward J. Vytlacil. 2007. "Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation." *Handbook of Econometrics* 6: 4779–4874.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2): 467-475.

Jacquemet, N., Joule, R.-V., Luchini, S. and Shogren J. F. (2011) "Do people always pay less than they say? Testbed laboratory experiments with IV and HG values", *Journal of Public Economic Theory*, 13, 857-882.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad. 2016. "Field of Study, Earnings, and Self-Selection." *The Quarterly Journal of Economics* 131 (3): 1057–1111.

Kling, C.L., Phaneuf, D.J. and Zhao, J. (2012) "From Exxon to BP: Has some number become better than no number?", *Journal of Economic Perspectives*, 26, 3-26.

Kodde, David A. 1986. "Uncertainty and the Demand for Education." *The Review of Economics and Statistics*, 460–67.

———. 1988. "Unemployment Expectations and Human Capital Formation." *European Economic Review* 32 (8): 1645–60.

List, J.A. (2001) "Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sportscards", *American Economic Review*, 91, 1498–1507.

List, J.A., Sinha, P. and Taylor M. (2006) "Using choice experiments to value non-market goods and services: Evidence from field experiments", *Advances in Economic Analysis & Policy*, 6.

Manski, Charles F. 2004. "Measuring Expectations." *Econometrica* 72 (5): 1329–76.

Murphy, J., Allen, P., Stevens, T. and Weatherhead, D. (2004) "A meta-analysis of hypothetical bias in stated preference valuation, Department of Resource Economics, University of Massachusetts, Amherst.

Murphy, J., Stevens, T. and Yadav, L (2010) "A comparison of induced value and home-grown value experiments to test for hypothetical bias in contingent valuation", *Environmental and Resource Economics*, 47, 111-123.

Nybom, Martin. 2017. "The Distribution of Lifetime Earnings Returns to College." *Journal of Labor Economics* 35 (4): 903–52.

Nybom, Martin and Jan Stuhler, 2017. Biases in standard measures of intergenerational income dependence. *Journal of Human Resources*, *52*(3), pp.800-825.

Pihl, Ariel, Angelov, Nikolay, Per Johansson and Mikael Lindahl. 2019. "Subjective Expectations, Educational Choice Heterogeneity and Gender: Evidence from a Sample of Swedish High School Students." *University of Gothenburg*, Mimeo.

Sattinger, Michael. 1993. "Assignment Models of the Distribution of Earnings." *Journal of Economic Literature* 31 (2): 831–80.

Taylor, L. Mckee, O.M., Laury, S.K. and Cummings R.G. (2001) "Induced value tests of the referendum voting mechanism", *Economics Letters*, 71, 61–65.

Willis, Robert J., and Sherwin Rosen. 1979. "Education and Self-Selection." *Journal of Political Economy* 87 (5, Part 2): S7–36.

Wiswall, Matthew, and Basit Zafar. 2021. "Human Capital Investments and Expectations about Career and Family." *Journal of Political Economy* 129 (5): 000–000.

Zafar, Basit. 2011a. "Can subjective expectations data be used in choice models? Evidence on cognitive biases" Special Issue Article on 'Measurement and Analysis of Subjective Expectations' *Journal of Applied Econometrics* 26 (3): 520-544.

———. 2011b. "How Do College Students Form Expectations?" *Journal of Labor Economics* 29 (2): 301–48.

———. 2013. "College Major Choice and the Gender Gap." *Journal of Human Resources* 48 (3): 545–95.

1 Tables

Table 1: Categories of education and particular degree programs

Category	Common programs
Education and teacher training	Subject teacher training; Pedagogy and didactics
Humanities and Art	Media production; History and archeology
Agriculture, Forestry and animal health	Veterinary care; Agriculture and forestry
Services	Tourism and travel; Police training
Social science, Law, Business, etc.	Psychology; Business administration; Law
Natural science, Mathematics and Data	Biology; Computer science; Mathematics
Engineering and Manufacturing	Civil engineering; Technical Engineering (mechanical, electrical)
Healthcare and social care	Medical training; Social work and guidance

Table 2: Mean expected earnings by field

Broad Field of Study	E(Earn) 30-40 (1000SEK/mo)
High school	26.6
Teaching	29.5
Humanities	30.6
Animal/Agro	31.7
Services	36.1
Social Sci	43.7
Sciences	39.6
Engineering	38.3
Health	37.4

Table 3: Correspondence between stated and revealed educational choices

	Expected Field
Panel A: Comparison with application data	
Match including everyone (many assigned to “no college”)	46.8% (233/498)
Match for just college fields	55.3% (218/394)
Match for just those with HS in the survey	65% (15/23)
Panel B: Comparison with enrollment data	
Match including everyone (many assigned to “no college”)	38% (189/498)
Match for just college fields	51.5% (173/336)
Match for just those with HS in the survey	70% (16/23)

Table 4: Parameters

Panel A: Returns to schooling estimates		
Individual i's returns	$\beta_i =$	$y_{i1} - y_{i0}$
Observed earnings and education OLS	$OLS =$	$E[y_{i1}\|S_i = 1] - E[y_{i0}\|S_i = 0]$
Treatment on the treated	$TT =$	$E[y_{i1} - y_{i0}\|S_i = 1]$
Treatment on the un-treated	$TUT =$	$E[y_{i1} - y_{i0}\|S_i = 0]$
Average treatment effect	$ATE =$	$E[y_{i1} - y_{i0}] = pTT + (1-p)TUT$
Panel B: Parameters deriving from Panel A		
Selection bias u_0	$SB_{u_0} =$	$OLS - TT = E[u_{0i}\|S = 1] - E[u_{0i}\|S = 0]$
Selection bias u_1	$SB_{u_1} =$	$OLS - TUT = E[u_{1i}\|S = 1] - E[u_{1i}\|S = 0]$
Sorting effect 1	$SE_1 =$	$TT - ATE$
Sorting effect 0	$SE_0 =$	$ATE - TUT$
Comparative advantage	$CA =$	$SE_1 + SE_0 = cov(\beta_i, S_i)/var(S_i)$ $(CA > 0 \implies TT > TUT)$
Earnings ability correlation	$=$	$corr(u_0, u_1) = corr(y_{i0}, y_{i1})$

Note: y_{ic} is individual i's expected earnings in $c = 1$ (college) or $c = 0$ (no college). S_i is their actual expected choice to pursue college or not. p is the share of the population that intends to pursue college.

Table 5: How β_i varies with demographics

	(1) Beta_i	(2) Beta_i	(3) Beta_i	(4) Beta_i	(5) Beta_i	(6) Beta_i	(7) Beta_i	(8) Beta_i
Female	0.0563* (0.0274)						0.0624+ (0.0342)	0.0650* (0.0324)
SES Index		0.0294** (0.00938)					0.0367** (0.0118)	0.0260* (0.0109)
Foreign			0.0476 (0.0365)				0.142** (0.0475)	0.119** (0.0450)
Math Score				0.0320* (0.0157)			0.0475+ (0.0269)	0.0365 (0.0247)
English Score					0.00697 (0.0171)		-0.0351 (0.0289)	-0.0238 (0.0266)
HS GPA						0.00221 (0.0180)	-0.0124 (0.0375)	-0.0216 (0.0330)
Choice FE								Yes
R^2	0.008	0.024	0.004	0.012	0.001	0.000	0.061	0.236
N	498	498	434	397	438	434	343	343

Note: Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Each column is the result of an OLS regression of β_i on the covariates listed on the left. $\beta_i = y_{i1} - y_{i0}$, an individual's expected college premium where y_{i1} is the earnings in stated choice college field for college choosers. For those who do not plan to go to college, we use a weighted average of all the college fields, where the weights are based on popularity among those who stated they intended to go to college in the survey. Both scores and GPA are standardized to mean 0, standard deviation 1.

Table 6: College vs. no-college, college-earnings defined using chosen fields and weighted averages

	Stated		Applied		Enrolled	
	(1) S=0	(2) S=1	(3) S=0	(4) S=1	(5) S=0	(6) S=1
Panel A: Conditional expected log earnings						
No College	3.313	3.239	3.260	3.238	3.258	3.235
College	3.455	3.613	3.533	3.633	3.556	3.615
N	23	475	104	394	162	336
Panel B: Ex-ante treatment effects						
OLS	0.300***		0.373***		0.357***	
	(0.0490)		(0.0290)		(0.0252)	
ATE	0.366***		0.370***		0.353***	
	(0.0137)		(0.0132)		(0.0125)	
TT	0.374***		0.396***		0.380***	
	(0.0138)		(0.0146)		(0.0150)	
TUT	0.182*		0.273***		0.298***	
	(0.0746)		(0.0295)		(0.0219)	
Panel C: Sorting parameters						
SB_{u_0}	-0.0741		-0.0230		-0.0236	
	(0.0484)		(0.0279)		(0.0240)	
SB_{u_1}	0.118*		0.100***		0.0593*	
	(0.0493)		(0.0291)		(0.0252)	
SE_1	0.00887*		0.0257***		0.0269**	
	(0.00385)		(0.00719)		(0.00879)	
SE_0	0.183**		0.0975***		0.0559**	
	(0.0702)		(0.0260)		(0.0179)	
CA	0.192**		0.123***		0.0828**	
	(0.0736)		(0.0328)		(0.0265)	
$corr(u_0, u_1)$	0.370***		0.377***		0.420***	

Note: Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. $S = [0, 1]$ denotes choice ($S = 1$ individuals chose to go to college), while College/No-College denotes hypothetical educational state. $E[y_{i1}|S_i = 1]$ uses on $i's$ chosen field. Because $S = 0$ individuals do not have a chosen field (they 'chose' no college) $E[y_{i1}|S_i = 0]$ is a weighted average of the eight fields, where the weights are the popularity of the fields among the $S = 1$ individuals. These weights are redefined for every time period.

Table 7: High pay vs. low-pay, defined using chosen fields and weighted averages

	Stated		Applied		Enrolled	
	(1) S=0	(2) S=1	(3) S=0	(4) S=1	(5) S=0	(6) S=1
Panel A: Conditional expected log earnings						
Low Pay	3.399	3.293	3.334	3.302	3.350	3.297
High Pay	3.618	3.670	3.590	3.683	3.580	3.663
N	100	375	38	322	52	284
Panel B: Ex-ante treatment effects						
OLS	0.272***		0.349***		0.313***	
	(0.0338)		(0.0393)		(0.0364)	
ATE	0.344***		0.368***		0.345***	
	(0.0113)		(0.0133)		(0.0132)	
TT	0.378***		0.381***		0.366***	
	(0.0125)		(0.0141)		(0.0141)	
TUT	0.220***		0.256***		0.229***	
	(0.0223)		(0.0347)		(0.0317)	
Panel C: Sorting parameters						
SB_{u_0}	-0.106**		-0.0321		-0.0529	
	(0.0327)		(0.0385)		(0.0353)	
SB_{u_1}	0.0522+		0.0930**		0.0837*	
	(0.0282)		(0.0356)		(0.0336)	
SE_0	0.125***		0.112***		0.115***	
	(0.0203)		(0.0330)		(0.0291)	
CA	0.158***		0.125***		0.137***	
	(0.0255)		(0.0369)		(0.0343)	
$corr(u_0, u_1)$	0.526***		0.453***		0.463***	

Note: Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. $S = [0, 1]$ denotes choice ($S = 1$ individuals chose high paying fields), while Low/High Paying denotes hypothetical educational state. $E[y_{i1}|S_i = 1]$ and $E[y_{i0}|S_i = 1]$ use $i's$ chosen field. $S = 0$ individuals do not have a chosen high-paying field, so $E[y_{i1}|S_i = 0]$ is a weighted average of the four high-paying fields, where the weights are the popularity of the fields among the $S = 1$ individuals. We define chosen low-paying field weights similarly for $S = 1$ individuals. These weights are redefined for every time period.

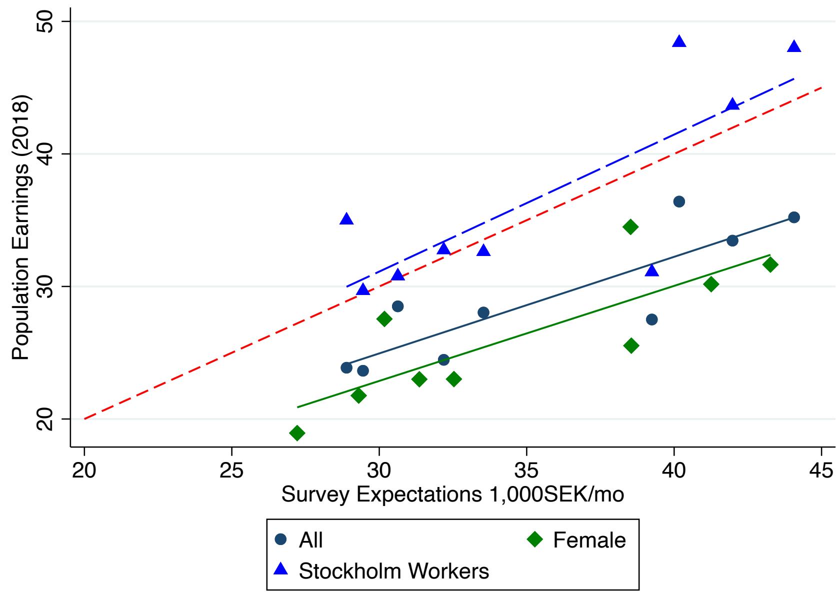
Table 8: Non-earnings returns to high paying versus low paying fields

	(1) Social Status	(2) Find Job	(3) Enjoy Job	(4) Work Hours	(5) Work-Life Bal
Panel A: Sated Choice					
TT	1.391***	0.916***	1.014***	0.664***	0.0610
	(0.0449)	(0.0494)	(0.0490)	(0.0439)	(0.0533)
TUT	0.488***	-0.00540	-1.021***	0.281**	-0.667***
	(0.0826)	(0.108)	(0.0777)	(0.0885)	(0.0984)
Panel B: Application Choice					
TT	1.409***	0.864***	0.928***	0.584***	-0.0903
	(0.0478)	(0.0570)	(0.0553)	(0.0476)	(0.0596)
TUT	0.609***	0.148	-0.737***	0.141	-0.628***
	(0.151)	(0.159)	(0.174)	(0.203)	(0.163)
Panel C: Enrolled Choice					
TT	1.331***	0.750***	0.845***	0.444***	-0.0973
	(0.0534)	(0.0577)	(0.0602)	(0.0513)	(0.0596)
TUT	0.761***	0.146	-0.396*	0.134	-0.315*
	(0.122)	(0.123)	(0.170)	(0.170)	(0.152)

Note: Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. We calculate returns as in Table 7 but replace earnings with the noted variable. The outcomes have been standardized to mean zero, standard deviation 1.

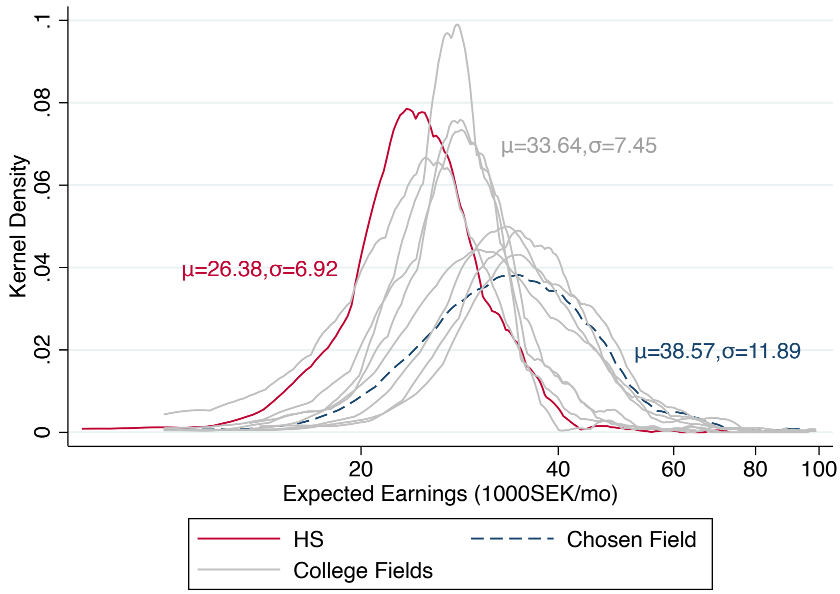
2 Figures

Figure 1: Comparing expected earnings to population earnings



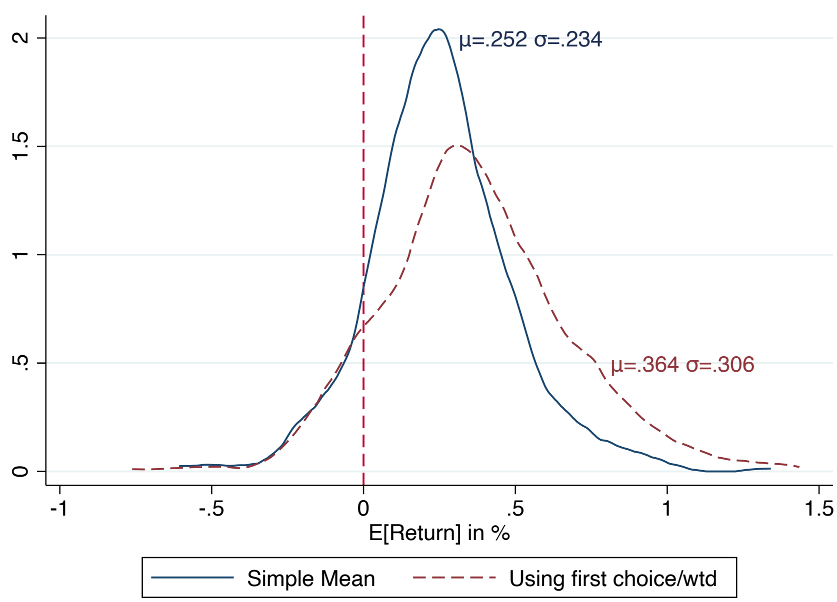
Note: Red dashed line is the 45 degree line. Plots the surveyed mean expected earnings in each field against population mean earnings in administration data (for those aged 40 in 2018), along with linear fit lines. The survey data is the full sample for both "All" and "Stockholm Workers", and the female sample respondents for "Female." The population data for "Stockholm Workers" is all those aged 40 and registered as working in Stockholm municipality in 2018.

Figure 2: Distribution of earnings by university and field



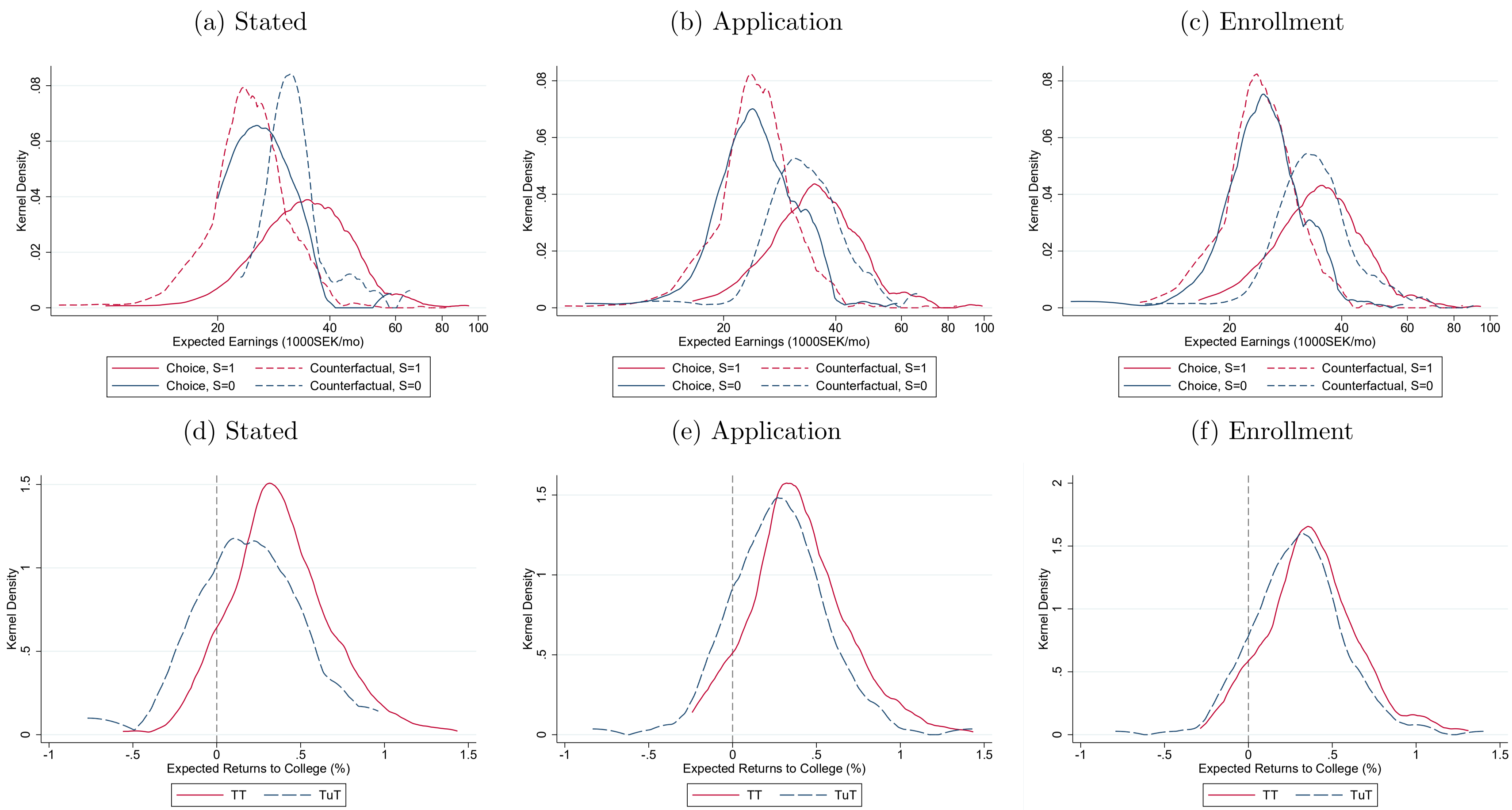
Note: Plots the distributions of expected earnings for no college degree/HS (y_{i0}) in red, and for each individual field (y_{ij}) in grey. The blue dashed line is expected earnings for the stated (survey) chosen field for all individuals.

Figure 3: Distribution of university premiums (β_i)



Note: Shows how the returns to college change when we use the returns to actual chosen college field (rather than the average of all college fields). Since those who don't plan to go to college don't have a "chosen" college field, their returns are the average of the college fields weighted by their popularity in the whole sample.

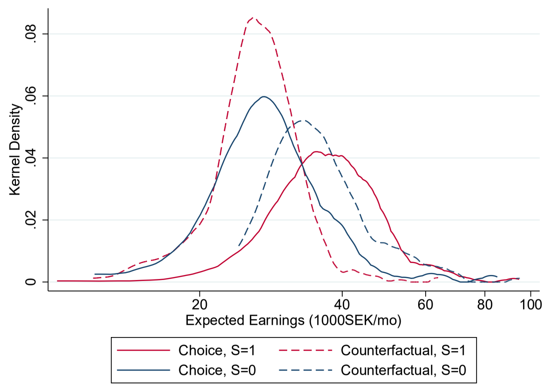
Figure 4: Expected earnings in college vs no-college by treatment status



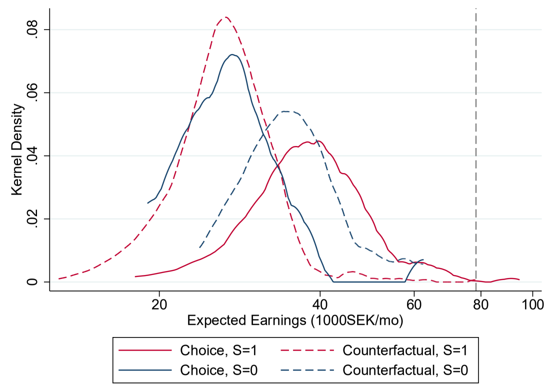
Note (a)-(c): The solid lines are the distributions of earnings for those that stop at high school (blue line) and continue with university (red line). The dashed lines are the counterfactual earnings distributions. Hence, the blue dashed line is the university earnings distribution for those that choose high school, and the red dashed line is the high school earnings distribution for those that choose university. S=1 are people who chose a college field (stated in the survey, or by applying/enrolling), S=0 those did not. Earnings in college uses actual chosen field expected earnings for S=1, and an average of all fields weighted by their sample popularity (amount stated in survey/application/enrollment) for S=0.
Note (d)-(f): Returns to college calculated as log expected college earnings (for either expected field if treated, or a weighted average of all fields if untreated), minus log expected non-college earnings.

Figure 5: Expected earnings in high-paying vs. low-paying fields by treatment status

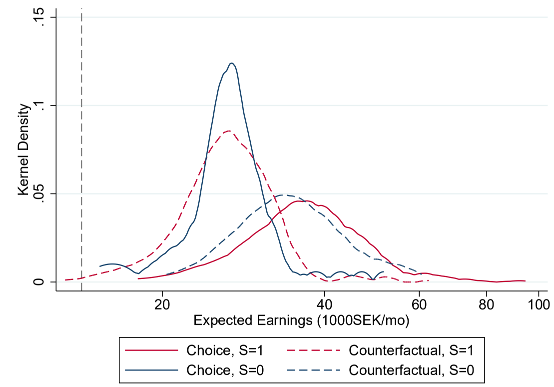
(a) Stated



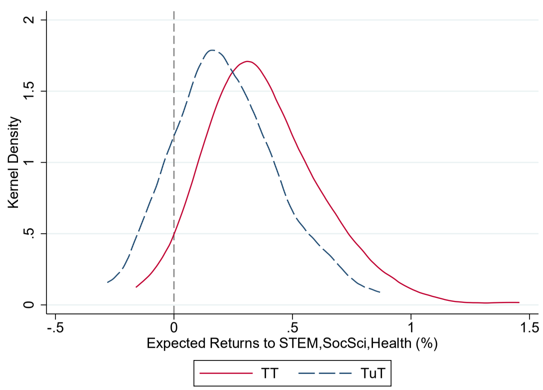
(b) Application



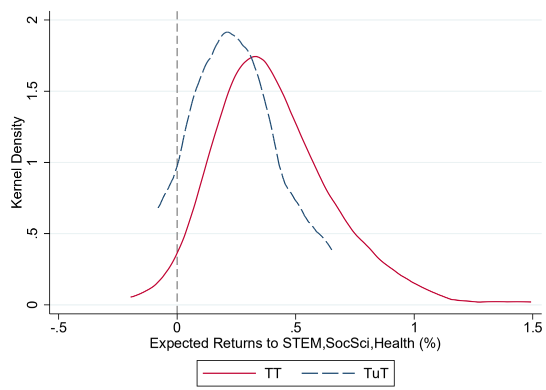
(c) Enrollment



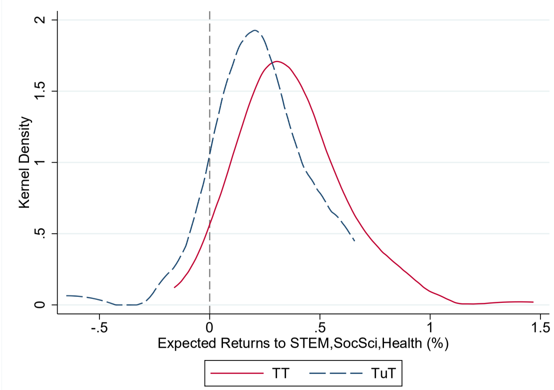
(d) Stated



(e) Application



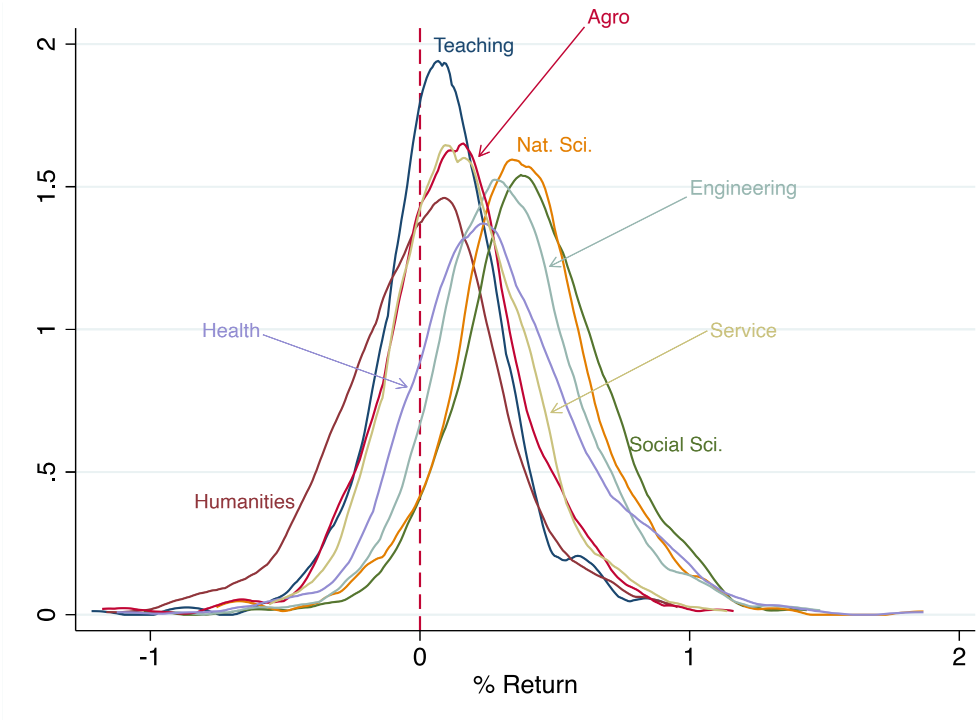
(f) Enrollment



Note (a)-(c): High-paying field categories are: Social Science, Sciences, Engineering and Health care. Low-paying fields are Humanities, Teaching, Agriculture and Services. S=1 is people who chose a high paying field field (stated in the survey, or by applying/enrolling), S=0 those did not. Earnings in high-paying fields uses actual chosen field expected earnings for S=1, and an average of the high paying fields fields weighted by their sample popularity (amount stated in survey/application/enrollment) for S=0. Likewise for low-paying fields.
Note (d)-(f): Returns to high-paying fields calculated as log expected college earnings (for either expected field if treated, or a weighted average of high-paying fields if untreated), minus log expected low-paying field earnings (for either expected field if un-treated, or a weighted average of the low-paying fields if treated.

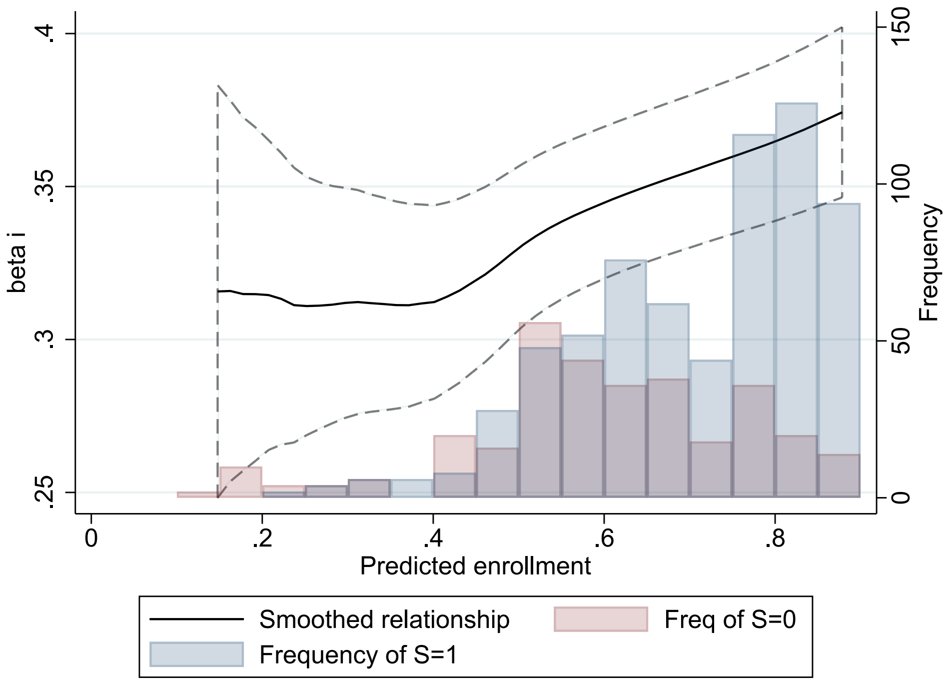
A Appendix Tables and Figures

Figure A1: Unconditional expected returns to field of study relative to no college



Note: X-axis is difference in log earnings between the field and the no college options.

Figure A2: Treatment effect size versus propensity to enroll



Note: X-axis is predicted propensity to enroll in university, where the explanatory variables are from the individual's average ranking of low-paying fields and chosen field in the survey. The line plots a local polynomial for expected returns to college, along with a 95% confidence interval (dashed lines). The histogram uses the right y-axis and shows the number of individual who actually enroll in college (S=1) and don't (S=0) for different predicted enrollment probabilities.

Table A1: Summary statistics on family and high school variables

	Surveyed Sample			Population	
	(1)	(2)	(3)	(4)	(5)
	Male	Female	Total	Stockholm	All
Background Variables:					
Foreign background	0.174	0.208	0.191	0.300	0.174
	(0.380)	(0.407)	(0.394)	(0.458)	(0.379)
Mom went to university	0.510	0.488	0.499	0.390	0.250
	(0.501)	(0.501)	(0.501)	(0.488)	(0.433)
Father went to university	0.500	0.478	0.489	0.378	0.178
	(0.501)	(0.501)	(0.500)	(0.485)	(0.383)
Parent(s) annual income (1000s SEK)	877.6	844.8	865.4	738	661.4
	(789.8)	(537.5)	(680.1)	(677.2)	(403.7)
School Variables:					
Avg. English Test Score (/20)	16.15	15.95	16.06	15.31	13.74
	(3.620)	(3.457)	(3.538)	(3.915)	(4.211)
Avg Math Score (/20)	12.88	12.52	12.70	10.05	8.289
	(5.225)	(5.093)	(5.157)	(5.983)	(6.084)
College Prep Program	0.878	0.910	0.894	0.886	0.625
	(0.328)	(0.287)	(0.309)	(0.317)	(0.484)
STEM Specialized Program	0.504	0.361	0.434	0.361	0.215
	(0.501)	(0.481)	(0.496)	(0.480)	(0.411)
Total Observations	254	244	498	2949	98936
N with all Vars	162	159	321	1600	56635

Table A2: Costs of choosing to go to college and expected returns.

	Survey		Applied		Enrolled	
	S=0	S=1	S=0	S=1	S=0	S=1
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.0967	0.349***	0.264***	0.354***	0.309***	0.331***
	(0.0806)	(0.0265)	(0.0412)	(0.0223)	(0.0272)	(0.0206)
Grad Prob	-0.225	0.00258	-0.0267	-0.0120	0.0322	-0.0258
	(0.182)	(0.0230)	(0.0610)	(0.0281)	(0.0405)	(0.0262)
Enjoy Studies	0.223	0.00841	0.0181	0.0150	-0.0497	0.0361
	(0.157)	(0.0260)	(0.0705)	(0.0262)	(0.0546)	(0.0253)
Study Hours	-0.199*	0.0387**	-0.00504	0.0361*	0.0102	0.0314+
	(0.0783)	(0.0145)	(0.0349)	(0.0157)	(0.0262)	(0.0160)
Fam Approve	0.0259	0.0121	0.00780	0.0496+	-0.0191	0.0603*
	(0.0922)	(0.0248)	(0.0598)	(0.0275)	(0.0384)	(0.0240)
Observations	23	475	104	394	162	336
R^2	0.369	0.019	0.002	0.034	0.010	0.057

Note: Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Outcome is individual expected return (β_i) in all columns. Cost measures are taken from the individual's expectations in the event that they attend college, regardless of whether they did ($S = 1$) or did not ($S = 0$) choose college.

Table A3: Means for all questions by field of study

	Edu	Hum	Soc Sci	Sci/Math	Tech/Eng	Agro	Health	Serv	Total
Prob of passing the degree	76.17	67.40	75.83	64.02	66.53	67.21	69.18	73.40	69.97
	(25.68)	(28.97)	(23.19)	(28.41)	(26.12)	(28.36)	(24.93)	(24.51)	(26.67)
Prob of enjoying coursework	54.88	55.90	67.72	55.72	55.69	44.05	59.95	54.27	56.02
	(25.17)	(29.71)	(24.77)	(29.63)	(27.20)	(28.68)	(26.12)	(25.13)	(27.78)
Expected study hrs/wk	34.79	33.45	44.16	47.47	42.11	33.37	44.90	32.91	39.15
	(17.19)	(18.63)	(19.48)	(20.42)	(19.22)	(16.74)	(20.18)	(17.42)	(19.54)
Parental approval	69.94	61.03	82.64	83.72	75.91	59.35	82.34	62.69	72.20
	(27.43)	(31.18)	(21.55)	(21.34)	(25.70)	(32.37)	(22.60)	(29.48)	(28.43)
Prob find a job	68.87	42.85	64.36	67.11	66.83	55.89	69.74	60.87	62.07
	(25.00)	(25.49)	(21.92)	(23.32)	(22.15)	(25.40)	(23.38)	(22.19)	(25.08)
Prob enjoy job (age 30)	53.13	55.61	67.53	57.23	56.68	47.14	60.33	53.24	56.36
	(25.20)	(29.12)	(22.90)	(27.28)	(25.60)	(28.20)	(25.00)	(23.91)	(26.54)
Expected hrs/wk (age 30)	47.32	39.14	49.24	47.43	45.76	45.06	52.60	44.51	46.38
	(11.53)	(13.33)	(12.18)	(11.74)	(11.09)	(13.26)	(13.40)	(11.45)	(12.80)
Expected earnings at 30	26.29	24.47	37.14	35.71	34.16	27.86	32.85	28.49	30.87
	(5.744)	(8.005)	(11.14)	(10.64)	(10.29)	(8.623)	(11.09)	(8.232)	(10.36)
Expected earnings at 40	30.63	29.45	44.07	41.98	40.17	32.19	39.25	33.54	36.41
	(6.966)	(9.422)	(12.99)	(12.35)	(12.17)	(9.598)	(12.96)	(9.591)	(12.13)
Perceived status for degree	44.94	46.26	78.03	73.27	64.82	40.36	71.09	48.51	58.41
	(20.49)	(21.35)	(14.60)	(17.71)	(18.86)	(20.94)	(23.10)	(20.61)	(24.26)

Table A4: Treatment on the margin

	Stated (S=1)		Applied	
	(1)		(2)	
Panel A: Treatment on the treated and untreated				
TT	0.374***		0.396***	
	(0.0138)		(0.0146)	
TUT	0.182*		0.273***	
	(0.0746)		(0.0295)	
N_1	475		394	
N_0	23		104	
	(3)	(4)	(5)	(6)
	S=1 Apply=0	S=1 Apply=1	Apply=1 Enroll=0	Apply=1 Enroll=1
Panel B: Separating TT into two margins				
TT	0.263***	0.400***	0.342***	0.405***
	(0.0346)	(0.0147)	(0.0438)	(0.0153)
N	89	386	58	336

Standard errors in parentheses. *** p<0.001

Note: Panel A repeats a portion of Table 6. N_1 is the sample size for $S = 1$, e.g. the treated individuals, likewise N_0 is the number of untreated individuals.