# HCEO WORKING PAPER SERIES

## Working Paper

# Expectations in Education:
# Framework, elicitation, and evidence[*]

Pamela Giustinelli[†]

August 19, 2022

## Abstract

This paper reviews the economic literature on subjective expectations in education with a focus on high income countries. It begins with highlighting the motivations that prompted systematic survey elicitation and statistical analysis of youth's expectations of the returns to schooling and with tracing key milestones in the development of this research program. It then proceeds to reviewing the relevant body of research by organizing the discussion around four topics: (i) the analysis of the perceived monetary returns, risks, and costs of schooling; (ii) the analysis of the perceived nonmonetary returns, risks, and costs of schooling; (iii) the analysis of schooling decisions; (iv) the analysis of expectation formation and learning. For each topic, the paper provides: (a) a motivating analytical framework; (b) a methodological discussion of expectations elicitation and a survey of data collections; (c) a review of the empirical evidence. Avenues for future research are discussed in the conclusion.


*JEL Codes*: D84, I26, J24.
*Keywords*: Education; Expectations; Returns to schooling; Schooling decisions; Subjective probabilities; Survey elicitation; Uncertainty.

# 1 Introduction

Subjective expectations have long been important primitives in microeconomic models of decision-making under uncertainty and intertemporal behavior, including but not limited to educational choices and other forms of human capital investment and accumulation.[1] Nowadays, the expression 'subjective expectations' is also commonly used to refer to an increasingly available and popular type of survey data meant to measure those primitives. This paper reviews the economic literature on *subjective expectations as data in* the domain of *education*, with a focus on high income countries.[2]

There was a time when survey data on subjective expectations related to education and its consequences for the individual were scant and when the sole idea of eliciting subjective expectations about the returns to schooling from subjects involved in making educational choices was largely considered radical and unfeasible among economists.[3]

In a seminal paper titled "*Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?*", Manski (1993) examines the implications for contemporary human capital research of the scarcity of interpretable data on youth's perceptions of the returns to schooling. This lack of data had caused two fundamental and interrelated identification problems. "*The first problem is that, not knowing how youth perceive the returns to schooling, one cannot infer their decision processes from their schooling choices. (...) The most one can do is infer the decision rule conditional on maintained assumptions on expectations. (...) The second problem is that, not knowing youth's decision processes, one cannot infer the objective returns to schooling from data on realized outcomes. (...) Hence, one can only infer the objective returns to schooling conditional on the validity of expectations assumptions.*" (Manski (1993), pp. 44-45).[4]

The paper concludes that the question posed in the title could not be answered at the time of writing and that progress would be possible only if economists opened up to the possibility of collecting interpretable subjective data on expectations and/or preferences. While recognizing that the enterprise may be especially challenging in the education context–as it would require elicitation of choice-conditioned forecasts from adolescent respondents–the paper closes unequivocally: "*We shall not know whether this is feasible until we try.*" (Manski (1993), p. 56).

Shortly after, Dominitz and Manski (1996) developed a computer-assisted self-administered interview protocol for survey elicitation of subjective expectations and used it to measure subjective earnings expectations, schooling expectations, and population earnings beliefs from high school students and college undergraduates in Madison, Wisconsin.

The survey featured a number of important innovations over earlier studies of students' expectations. Because a primary goal of the study was to characterize students' perceived uncertainty, the survey measured respondents' *subjective distributions* about own and population earnings by eliciting the probability that earnings exceed multiple

---

[1]For a historical introduction to expectations in microeconomics, see Carter and Maddock (1984). For early models of human capital decisions under uncertainty, see Weiss (1972), Levhari and Weiss (1974), Groot and Oosterbeek (1992)'s extension of the Becker-Mincer model, and Altonji (1993).

[2]See Delavande (2022) for evidence on low and middle income countries.

[3]Scholars in other social sciences have a longer tradition of collecting and analyzing subjective data. Dominitz and Manski (1999) review and discuss the different cultures of research on subjective expectations.

[4]The first problem refers to the difficulty of performing empirical revealed preference analysis when the available data consists of a single distribution of choices within a population of heterogeneous individuals; observing one choice per individual does not reveal the distribution of underlying primitives, since multiple configurations of preferences and expectations may be consistent with a given choice. The identification problem gets harder with intra-family decision-making, peer interactions, or heterogeneous choice sets, all of which are relevant concerns in models of schooling decisions (Giustinelli and Manski, 2018). The second problem is a direct consequence of selection; it "*arises because the youth who choose to enroll in school are those who expect schooling to have favorable outcomes for them. If the expected outcomes are related to the objective ones, then the outcomes experienced by youth who choose to enroll in school differ from those that nonenrollees would experience if they were to enroll. (...) [Any] effort to infer the objective returns to schooling from observations of realized outcomes requires at least some knowledge of the way youth make their schooling decisions.*" (Manski (1993), p. 45).

thresholds, rather than eliciting a single moment (say, expected mean earnings). Because another important goal was to learn how students perceive the returns to schooling, the survey asked respondents their earnings expectations *under hypothetical scenarios* about their school attainment, in addition to their unconditional expectations. Because the study also aimed at collecting expectations that could be compared both within and between students, the survey elicited respondents' subjective probabilities on a *numerical scale of percent chance*, instead of a nominal probability scale.

Finding that respondents' were willing and able to give meaningful answers, Dominitz and Manski concluded that measuring youth's perceptions of the returns to schooling was both feasible and promising. Many more studies of students' expectations followed as a result.

A remarkable example is the Berea Panel Study (BPS), launched by Ralph and Todd Stinebrickner shortly after the Dominitz and Manski (1996)'s study. By following two cohorts of Berea College freshmen for up to fourteen years, and by collecting rich information on the expectations and realizations of participants' education and labor market outcomes during and after college, the BPS has become a landmark among existing surveys of students' expectations.[5]

Over the years, different studies have experimented with different survey modes and question formats, have elicited percent-chance expectations from youth of varying ages and backgrounds in different countries, and have expanded the scope of the outcomes and scenarios posed. Expectations for nonmonetary benefits and costs of schooling have been increasingly collected and analyzed along with those for monetary returns and costs. Most studies have collected expectations at one point in time from a single group of respondents, usually college students and less frequently younger students. Occasionally, expectations have been collected longitudinally, as in the Berea Panel Study (BPS) and in Zafar (2011b, 2013)'s survey of Northwestern University (NU)'s undergraduates, or within field experiments, as in Wiswall and Zafar (2015b,a)'s survey of New York University (NYU)'s undergraduates. Sporadically, they have been elicited from multiple subjects, as in Giustinelli (2016)'s survey of Italian 9th-graders and their parents.

These data have been used to study the perceived monetary and nonmonetary returns to schooling of youth and their families, to estimate random utility models of schooling decisions under uncertainty, and to investigate how students form expectations and update them to arrival of new information.[6]

This paper surveys the above literature, while complementing in scope and content earlier reviews addressing the role of expectations for human capital,[7] as well as multiple chapters in the Handbook of Economic Expectations.[8] Regrettably, some important topics and active research areas could not be covered due to space limitations. These include the literature on parental beliefs about the returns to investments in children,[9] the literature on public opinion and preferences for educational policies,[10] and the literature on students'

---

[5]E.g., see Stinebrickner and Stinebrickner (2003a,b, 2004, 2008a,b, 2012, 2014a,b).

[6]The economic literature on survey expectations related to education and labor earnings (e.g., Dominitz and Manski (1996, 1997), Dominitz (1998, 2001), Fischhoff et al. (2000), Dominitz et al. (2001)) has been the cradle of the broader expectations literature reviewed by Manski (2004, 2018) and by multiple chapters in the Handbook of Economic Expectations. The former has especially contributed to the transition from a first phase of largely exploratory and descriptive studies to a still ongoing second phase where survey expectations have been used to test economic theories, estimate formal models of decision-making and/or learning, and perform counterfactual analyses (e.g., van der Klaauw (2012), Arcidiacono et al. (2012, 2020), Stinebrickner and Stinebrickner (2012, 2014a,b), Zafar (2012, 2013), Wiswall and Zafar (2015b,a), Giustinelli (2016), Kapor et al. (2020)).

[7]These include Heckman et al. (2006), Hartog and Diaz-Serrano (2014), Altonji et al. (2016), Giustinelli and Manski (2018), and Patnaik, Wiswall and Zafar (2020).

[8]For example, Bruine de Bruin et al. (2022), Delavande (2022), Fuster and Zafar (2022), Ilut and Schneider (2022), Koşar and O'Dea (2022), Manski (2022), and Mueller and Spinnewijn (2022).

[9]E.g., see Cunha et al. (2013, 2021), Boneva and Rauh (2018), and Attanasio et al. (2020) on high income countries and Delavande (2022)'s review on low and middle income countries.

[10]E.g., Lergetporer et al. (2018, 2020, 2021), Lergetporer and Woessmann (2019), and Barrows et al. (2021).

experiences, expectations, and outcomes during the COVID-19 pandemic.[11]

The paper is organized as follows. Section 2 deals with expectations for monetary returns, risks, and costs of schooling. Section 3 covers expectations for nonmonetary schooling outcomes and educational expectations. Section 4 reports on the use of survey expectations to model schooling decisions. Section 5 reports on the use of survey expectations to study formation and updating of expectations in education. Section 6 discusses avenues for future research and concludes. A detailed list of 50 surveys of expectations related to education is provided in the Supplementary Appendix.

## 2   Survey Expectations About Monetary Outcomes of Schooling

### 2.1   Motivating Framework

Manski (1993)'s model of information, schooling choices, and outcomes provides a formal framework to motivate the importance of measuring expectations. The framework demonstrates that different assumptions about the nature of youth's expectations and the information on which youth condition their expectations have dramatically different implications for interpretation of the equilibrium distributions of educational choices and realized incomes, even when youth are assumed to form expectations homogeneously.

*Choice environment.* The model considers an overlapping-generations world in which each person lives for two periods. In the first period, each youth chooses between going to school ($j = s$) and going to work ($j = w$). In the second period, everyone goes to work. Each youth is characterized by real-valued levels of ability ($z$) and taste for schooling ($\nu$), which are known to them at the time of the schooling decision. The present discounted log-income that each youth would receive if they were to work immediately is normalized to 0. The present discounted log-income that each youth would receive if they were to enroll in school ($y$) is a random variable whose realization becomes known to them after schooling is completed.

Each youth's value of ($y, \nu, z$) is independently drawn according to the time-stationary process

$$\begin{cases} y = \alpha_1 + \beta_1 z + \epsilon_1 \\ \nu = \alpha_2 + \beta_2 z + \epsilon_2, \end{cases}$$

where $\beta_1 \geq 0$, the vector ($\epsilon_1, \epsilon_2, z$) follows a trivariate normal distribution,

$$\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ z \end{pmatrix} \sim N \left[ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right],$$

and the key restriction is that post-schooling income ($y$) and taste for schooling ($\nu$) are statistically independent conditional on ability ($z$).

Youth's decision rule is

$$j = \begin{cases} s & \text{if} \quad E^*(y|\Omega) + \nu > 0 \\ w & \text{otherwise}, \end{cases}$$

where $E^*(y|\Omega)$ denotes a youth's subjective expected value of $y$ conditional on the common information set ($\Omega$), which is shared by the youth in a given generation and includes the school-work decisions and realized incomes of the prior generation.

*Expectations assumptions.* To complete the model, $E^*(y|\Omega)$ must be specified. Under a benchmark of rational expectations ($R$), all members of the young generation would hold $E^*(y|\Omega^R) = \alpha_1 + \beta_1 z$. Manski (1993) investigates two types of myopic expectations processes, differing in the information youth possess and use to form their expectations.

---

[11]E.g., Aucejo et al. (2020), Aucejo et al. (2021), and Delavande et al. (2021).

Process ($A$) presumes that youth observe the ability levels, school-work choices, and realized incomes of the older generation and condition on the available information when forming expectations for the returns to schooling. Process ($B$) presumes that youth do not observe the ability of the older generation.[12] Formally, the two processes are

$$(A) \quad E^*(y|\Omega^A) = E_0(y|j = s, z)$$
$$(B) \quad E^*(y|\Omega^B) = E_0(y|j = s).$$

*Implications for inference.*    Manski (1993) derives the unconditional and ability-conditioned distributions of schooling enrollment and realized income under expectations processes ($A$) and ($B$) and compares them. Here is a summary of key results.

(i) ($A$) and ($B$) imply different equilibria. Under ($A$), expectations are rational in equilibrium. Under ($B$), and further assuming that taste for schooling does not decrease with ability ($\beta_2 \geq 0$), expectations are fulfilled but incorrect in equilibrium.
(ii) For $\beta_2 \geq 0$, in equilibrium school enrollment is higher under ($A$) than under ($B$) among high-ability youth, and viceversa for low-ability youth.
(iii) Conditional on ability, the mean realized income is the same under ($A$) and ($B$). Instead, the unconditional mean income in the population depends on the expectations process through the distribution of ability among the enrollees. For some configurations of the model parameters mean realized income among the educated youth is higher under ($A$) than under ($B$).
(iv) If youth do not condition their expectations on ability (as in $B$), an econometrician who assumes that they do (as in $A$) will tend to underestimate the importance of the income returns to schooling in driving observed school-work choices.

*Further issues.*    Human capital theory maintains that students take into account the expected lifetime income of different educational paths when making schooling decisions. However, if students are risk averse, the decision will not only be based on the expected value of lifetime income, but also on the risks associated with each educational path.[13] Moreover, if there is unobserved heterogeneity in the effects of schooling decisions, variation in observed schooling outcomes conflates *ex ante* uncertainty with unobserved heterogeneity. These observations carry implications for elicitation and analysis of subjective expectations related to education.[14]

First, it is important to measure individuals' perceptions of the uncertainty they face when making schooling decisions, not just their expected benefits. Starting with Dominitz and Manski's Wisconsin Study and Stinebrickner and Stinebricker's Berea Panel Study (BPS), many surveys of students' earnings expectations have elicited students' subjective earnings distributions under alternative schooling scenarios, not just the first moment of these distributions. This, in turn, has enabled researchers to analyze higher moments of students' subjective earnings distributions, to investigate their role in education and occupation choices, and to perform tests of rational expectations that take into account the information contained in higher moments.

Second, uncertainty may be multidimensional, corresponding to different sources and types of risks and returns. For instance, some of the risks individuals bear when investing in education are monetary (e.g., financial risks associated to student debt and income

---

[12]Manski (1993) discusses alternative specifications. E.g., youth may condition their expectations on the ability levels, choices, and incomes of their peers or neighbors.

[13]Hartog and Diaz-Serrano (2014) review existing models of human capital investment under risk, including extensions of the Becker-Mincer model with uncertain payoffs, mean-variance models, option value models, and models incorporating consumption decisions. Later I will discuss the possibility that some of the uncertainty faced by students and families confronting educational choices may be more germane to deep uncertainty or ambiguity than measurable risk (see also Ilut and Schneider (2022)'s chapter).

[14]Heckman et al. (2006) review methods for separately inferring the components of the returns to schooling that are predictable by individuals at the time of schooling decisions (heterogeneity) from those that are not predictable (uncertainty). These methods do not require or pursue direct elicitation of individuals' expectations.

volatility), whereas others are nonmonetary (e.g., school dropout risk). Similarly, the benefits and costs of schooling can be both monetary and nonmonetary. Perceived non-monetary benefits, risks, and costs may be as important drivers of educational choices as monetary ones. Indeed, they have received increasing attention in recent analyses without and with expectations data.

## 2.2 Survey Elicitation

### 2.2.1 The Dominitz-Manski Study

The overarching goal of Dominitz and Manski (1996)'s study was to provide a proof of concept for survey elicitation of unconditional and conditional subjective expectations about earnings and schooling from school-going youths, using a numerical scale of percent chance. In practice, the study has become a template for survey measurement of subjective expectations well beyond the original outcomes and population.[15] Within this overarching goal, the study had three specific aims. First, to characterize students' uncertainty about their future earnings. Second, to learn how students participating in the survey perceive the returns to schooling. Third, to investigate whether and, if so, how, students' expectations about their future earnings are related to students' perceptions about the population distribution of current earnings.

*Survey design.* The survey consisted of an interactive computer-assisted self-administered interview (CASI). The CASI software enabled the authors to design a survey that appeared straightforward to respondents, but that actually incorporated an extensive question-branching algorithm as well as tools such as training screens, help screens, and review-and-revise screens, intended to aid respondents in expressing meaningful expectations.

The survey sequentially elicited:

1. *Unconditional earnings expectations.* Each respondent was queried about his or her subjective probability distribution of real earnings at ages 30 and 40.

2. *Expected earnings under alternative schooling scenarios.* Each respondent was asked about his or her expected earnings at ages 30 and 40 under the hypothetical scenario that the respondent continues in school at least through receipt of a bachelor's degree, and under an alternative scenario assuming that less schooling is completed (a high school diploma without further schooling for high school students, the current semester of college without further schooling for college students).

3. *Schooling expectations.* Respondents were asked to state the probability that they will (a) attend college before age 21 (for high school students only); (b) receive a bachelor's degree before age 30; and (c) still be in school at age 30.

4. *Beliefs about current earnings distributions.* Male (female) respondents were asked about the current distributions of earnings among American men (women) of age 30 who have attained at least a bachelor's degree and among those who have attained a high school diploma but no further schooling.

For each outcome, the survey first elicited the median of the respondent's subjective distribution in order to characterize the distribution's central tendency. Then, it elicited the probability that earnings would exceed a sequence of thresholds, posed in increasing order, to characterize additional features of distribution.

---

[15]Some of the subsequent studies of students' expectations have employed partially different designs in terms of interview type and survey mode (e.g., in-person paper-&-pencil, mailed paper-&-pencil, phone interview, online survey), elicitation format (e.g., type-in box, bins-and-balls, slider), response form (e.g., expected value vs. distribution, point vs. interval probability), use and type of incentives (e.g., no reward vs. reward for survey completion vs. reward for belief accuracy). See Bruine de Bruin et al. (2022), Delavande (2022), Fuster and Zafar (2022), Manski (2022), and other chapters in the Handbook of Economic Expectations for methodological discussions and examples.

1. *Eliciting median earnings.* "*Look ahead to when you will be [X] years old. Think about the kinds of jobs that will be available for you and that you will accept. What is the median amount of money that you think you will earn per year by the time you are [X] years old?*"

2. *Eliciting probabilities that earnings exceed thresholds.* "*When you are [X] years old, what do you think is the percent chance that you will earn more than [Y] thousand dollars per year? That is, what are the chances out of 100 that you will earn more than [Y] thousand dollars per year?*"

Each respondent answered these questions with $X$ taking the values 30 and 40, and $Y$ taking three values that depend on the respondent's elicited median, so the thresholds were person specific.

The choice of starting with elicitation of median earnings was motivated by a desire to avoid response anchoring to specific thresholds chosen by the researcher. The respondent's own median provides a natural self-anchor for later selection of thresholds. The choice of asking questions about a range of thresholds spanning the support of the respondent's subjective distribution was driven by the observation that this approach should yield more information about the shape of the distribution than would the same number of questions asked about a narrower or wider range of thresholds. Finally, the choice of using three thresholds aimed at balancing the tension between a natural desire to collect as refined data as possible and that of avoiding excessive response burden and survey length.[16]

This procedure was devised to perform probabilistic elicitation of expectations and beliefs about earnings, a continuous outcome. Probabilistic elicitation of expectations or beliefs about discrete events, such as school enrollment or graduation, can be similarly performed by asking the respondent to report the percent chance between 0 and 100 that the event in question will occur. Using this approach, Dominitz and Manski (1996) elicited from high school participants the likelihood of attending college before age 21, and from all participants the likelihood of receiving a bachelor's degree before age 30.[17]

When survey length and response burden are not a major concern, a researcher interested in eliciting subjective belief distributions about discrete-valued variables may elicit subjective probabilities for all values that the variable can take and empirically test whether respondents' probability reports are consistent with the sum-to-one restriction implied by probability theory. When a survey asks subjective probabilities for a series of nested events, additional logical or consistency checks become possible. For example, Dominitz and Manski (1996) investigate violations of monotonicity.

To derive measures of the perceived returns to schooling and subjective uncertainty about future earnings, Dominitz and Manski (1996) additionally elicited students' earnings expectations under alternative schooling scenarios. This is the version administered to high school respondents.

1. *Introducing hypothetical scenarios.* "*The next sets of questions ask you to put yourself in one of two hypothetical situations. In the first situation, you assume that you continue in school until you finish your senior year of high school and obtain your diploma, and you do not continue in school after that. In the second hypothetical situation, you assume that you continue in school at least until you finish your senior year of college and obtain your college diploma (a bachelor's degree). When responding to these questions, please attempt to fully place yourself in the hypothetical situation as it is described.*"

---

[16]Dominitz (2001), Bellemare et al. (2012), and others demonstrate how to use this type of data to estimate subjective cumulative distribution functions.

[17]Similarly, Fischhoff et al. (2000) elicited from a subset of teen respondents in the National Longitudinal Study of Youth 1997 (NLSY97) the likelihood of working for pay more than 20 hours per week at different horizons, both unconditionally and conditional on being in school vs. not being in school. Naturally, if the discrete variable of interest can take more than two values, as in the case of letter grades or GPA intervals, more questions are needed. For instance, the Berea Panel Study (BPS) repeatedly asked college-going respondents the percent chance that their GPA will be in each of five intervals (Stinebrickner and Stinebrickner, 2014*a*,*b*).

2. *Specifying schooling scenarios.* "*In the first hypothetical situation, assume that you continue in school until you complete your senior year of high school and obtain your high school diploma. Please respond under the assumption that you do not return to school at any time after high school. Remember, this is a hypothetical situation. Just think about the kinds of jobs that would be available for you and that you would accept. Think about the amount of money you would make on these jobs. Again, you should ignore the effects of price inflation on earnings.*"

3. *Eliciting median earnings under specified schooling scenarios.* Following the introduction of the schooling scenarios, respondents were asked to state their median earnings at age [X] under each of the specified scenarios.

4. *Eliciting probabilities that earnings exceed thresholds under specified schooling scenarios.* After the median, respondents where asked to state the likelihood that their earnings at age [X] would exceed three thresholds under each of the specified scenarios. As before, the three thresholds were determined on the basis of the respondent's subjective median.

*Interpretation and analysis.* Dominitz and Manski (1996) use the elicited median earnings at age 30 and age 40 under alternative schooling scenarios to derive measures of respondents' perceptions of the returns to schooling at those ages. These are simply obtained as the respondent-specific difference between the elicited median earnings at a given age, assuming that the person attains a bachelor's degree, and the subjective median earnings at the same age, assuming that the person does not do so. Formally, let $M_i(S, X)$ denote the median earnings at age $X$ elicited from respondent $i$ under schooling scenario $S$. Respondent $i$'s perceived return at age $X$ of attaining a bachelor's degree ($S = 1$) versus not ($S = 0$), is computed as $M_i(1, X) - M_i(0, X)$, where $X \in \{30, 40\}$.

To characterize the spread of respondent $i$'s expectations, Dominitz and Manski (1996) use the interquartile range (IQR) of the respondent-specific parametric probability distribution fitted to the elicited probabilities. When applied to the probabilities that earnings at age $X$ exceed specified thresholds under schooling scenario $S$, this procedure yields a schooling and age-specific measure of IQR at the individual level, $IQR_i(S, X)$.

To investigate how students' perceived uncertainty about future earnings varies with education, Dominitz and Manski (1996) contrast the sample distributions of $IQR_i(S, X)$ across schooling scenarios within $X$. The authors do not define or derive any specific measure of earnings risk. They instead maintain a prudent view of the interpretability of $IQR_i(S, X)$ as an individual measure of subjective uncertainty, as it is derived indirectly rather than being directly elicited from respondents.

The logic of Dominitz and Manski (1996)'s elicitation procedure may be better understood through the lenses of potential outcomes.[18] Let $S$ denote levels of completed schooling. Let $Y(S, X)$ denote the potential income that a survey respondent would earn at age $X$, *if* he or she were to complete schooling level $S$. And let $\Omega$ denote the information available to a respondent at the time of the survey. Ideally, one would like to learn the distribution of life-cycle earnings that the respondent associates with various schooling options, i.e., the respondent's subjective distribution $P[Y(S, X) \; \forall X | \Omega]$. In practice, because elicitation of entire subjective distributions of life-cycle earnings is impractical, Dominitz and Manski (1996) focus on earnings at two ages under two schooling options. Thus, their elicitation procedure seeks to learn $P[Y(S, X) | \Omega]$ for $X \in \{30, 40\}$ and $S \in \{0, 1\}$.

The wording used by the survey to describe the two schooling scenarios offered no reason why one or the other scenario might be realized. This was intentionally done in an attempt to avoid having respondents draw from the scenario descriptions information that might influence their expectations. To make this point formally, let $L$ denote the schooling level that a respondent $i$ will eventually choose, and observe that in general $L$ may be unknown to $i$ at the time of the survey. When so, $L$ is not part of respondent

---

[18]Originating in statistics from the work of Rubin (1974), see Imbens (2020) for a recent essay on its relevance for empirical economics.

$i$'s information set, $\Omega_i$, at the time of the survey. Now suppose that in the description of a scenario, the respondent is asked to assume that he or she chooses schooling level $S$. Then, respondent $i$ should report earnings expectations $P[Y(S,X)|\Omega_i, L = S]$ rather than $P[Y(S,X)|\Omega_i]$. Dominitz and Manski (1996)'s procedure aims at eliciting the latter distribution rather than the former. The difference between the two is related to the selection problem discussed by Manski (1993). A typical aim of studies of the returns to schooling is to compare $P^*[Y(S,X)]$ across alternative schooling levels $S$, where now $P^*$ stands for a frequentist probability distribution instead of a subjective one. Because the realizations of $Y(S,X)$ can be observed only among those persons who choose schooling level $S$, the data reveal $P^*[Y(S,X)|L = S]$ but not $P^*[Y(S,X)]$.

### 2.2.2   Subsequent Expectations Surveys

*Earnings and monetary returns.*   Measuring the perceived returns to schooling by first eliciting students' subjective earnings distributions under alternative schooling scenarios and then contrasting the medians or means of these distributions between pairs of scenarios at the individual level has become common practice in the survey expectations literature following Dominitz and Manski (1996). Studies that have collected and analyzed survey data on students' subjective earnings expectations under alternative scenarios include Menon Eliophotou (1997*b*,*a*), Carvajal et al. (2000), Wolter (2000), Botelho and Pinto (2004), Brunello et al. (2004), Webbink and Hartog (2004), Wolter and Zbinden (2005), Menon Eliophotou (2008*a*,*b*), Mazza and Hartog (2011), Schweri et al. (2011), Zafar (2011*b*, 2012, 2013), Arcidiacono et al. (2012), Hartog et al. (2014), Hastings et al. (2015, 2016), Wiswall and Zafar (2015*b*,*a*, 2020), Reuben et al. (2017), Ruder and van Noy (2017*b*), Osikominu and Pfeifer (2018), Gong et al. (2019, 2020*b*), Kiessling et al. (2020), Belfield et al. (2020), Ehrmantraut et al. (2020), Madeira (2020*b*), Patnaik, Venator, Wiswall and Zafar (2020), Boneva and Rauh (2021), Boneva, Golin and Rauh (2022), Crossley et al. (2021), Boneva, Buser, Falk and Kosse (2022), Briel et al. (2022), and Diaz-Serrano and Nilsson (2022). In these studies, elicitation scenarios span alternative attainment levels (e.g., high school diploma, college dropout, college degree, and so on), achievement levels (e.g., GPA intervals), study concentrations (e.g., college majors), and work experience.

A smaller set of studies has investigated the spread of individuals' subjective earnings distributions, viewed as a measure of *ex ante* risk. These studies include Mazza and Hartog (2011), Schweri et al. (2011), Hartog et al. (2014), Hastings et al. (2015, 2016), Wiswall and Zafar (2015*b*,*a*, 2020), Zambre (2018), Gong et al. (2019, 2020*b*), Patnaik, Venator, Wiswall and Zafar (2020), Crossley et al. (2021), and Diaz-Serrano and Nilsson (2022).

Some studies have alternatively, or sometimes additionally, elicited respondents' beliefs about current population earnings or respondents' expectations about future population earnings. Examples of these studies are Betts (1996), Botelho and Pinto (2004), Huntington-Klein (2015, 2016*b*), Wiswall and Zafar (2015*b*,*a*), Hastings et al. (2015, 2016), Ruder and van Noy (2017*a*), Baker et al. (2018), Bleemer and Zafar (2018), Osikominu and Pfeifer (2018), Klossner and Pfeifer (2019), Kerr et al. (2020), Conlon (2021), and Briel et al. (2022).

Expectations about own earnings and expectations (or beliefs) about population earnings are different concepts. Hence, respondents' reports about them may differ. Firstly, respondents may think that future earnings distributions will differ from current ones. Secondly, respondents should form expectations about own outcomes conditional on the information available to them; such information may include personal attributes and other factors that distinguish the respondent from the relevant population. At the same time, it is possible that when forming expectations for own future earnings respondents take into account what they know about current earnings in a relevant population, as in Manski (1993)'s framework. This provides a rationale for collecting both sets of data and investigating their relationship, as done by Dominitz and Manski (1996), Wiswall and Zafar (2015*a*), and others.

*Other monetary outcomes.* Multiple studies have elicited expectations about monetary outcomes related to education other than earnings, especially costs. For instance, Menon Elio-photou (1997b,a) considers students' expectations of the direct costs of higher education. In the BPS, Stinebrickner and Stinebrickner (2008b) elicit expectations for expenses during college that exclude college-related costs (i.e., consumption expectations), and expected amount of borrowing under specified borrowing scenarios. Hastings et al. (2015, 2016) elicit and analyze students' own expectations and population beliefs about college program costs. Huntington-Klein (2016b) considers students' tuition and aid expectations. Bleemer and Zafar (2018) measure respondents' beliefs about college costs in the population. Belfield et al. (2020) analyze students' expectations about tuition costs. Boneva and Rauh (2021) and Boneva, Golin and Rauh (2022) look at students' expectations of struggling financially, having enough money, and receiving financial support from family, under alternative schooling/work scenarios. Conlon (2021) elicits expected student debt.

Some studies have also collected data on students' expectations or beliefs about labor market outcomes other than earnings. The most common ones are employment status, unemployment risk, and labor supply (e.g., Carvajal et al. (2000), Fischhoff et al. (2000), Brunello et al. (2004), Zafar (2011b, 2012, 2013), Huntington-Klein (2015, 2016b), Wiswall and Zafar (2015b,a, 2020), Giustinelli (2016), Baker et al. (2018), Gong et al. (2019), Kiessling et al. (2020), Gong et al. (2020a), Belfield et al. (2020), Delavande et al. (2020), Ehrmantraut et al. (2020), Boneva and Rauh (2021), and Boneva, Golin and Rauh (2022)). While these outcomes are not directly expressed in monetary terms, having a job or losing it have direct financial consequences for individuals and their households. Additional labor market outcomes with financial implications about which subjective expectations have been collected from students, include the time of job arrival (Gong et al., 2019), working hours (Carvajal et al. (2000), Fischhoff et al. (2000), Reynolds and Pemberton (2001), Zafar (2011b, 2012, 2013), Kiessling et al. (2020), Ehrmantraut et al. (2020), Wiswall and Zafar (2020)), occupation (Carvajal et al. (2000), Schweri et al. (2011), Schweri and Hartog (2017), Arcidiacono et al. (2020)), type of employer/sector (Carvajal et al., 2000), and location of first job (Carvajal et al., 2000).

## 2.3 Empirical Evidence

*Are elicited earnings expectations meaningful?* Expectations studies have evaluated the informativeness and quality of survey expectations by investigating the prevalence and predictors of item non-response, by checking the coherence of expectations reports with respect to probability rules, by comparing expectations with realizations for the same or comparable individuals based on future or historical realizations data, and by investigating whether expectations covary with respondents' characteristics and other observables in similar ways as realizations do (Manski, 2004).

Early quality assessments are sometimes performed in the survey pretest phase. For example, Dominitz and Manski (1996) held multiple debriefing sessions with pretest respondents. These sessions revealed that students generally understood the questions well, could articulate meaningful answers, and greatly enjoyed the interactive features of the software. At the same time, these preliminary discussions also revealed some critical issues with the interpretation of specific questions. For example, when reporting expectations of their own future earnings, pretest respondents tended to condition on being employed full-time even though the scenario did not specify future employment.[19] Dominitz and Manski addressed these issues during the design phase by appropriately revising the text and wording of the relevant questions.

The consistency checks and error screens embedded in Dominitz and Manski (1996)'s CASI protocol did ensure that students' final responses were logically consistent. Neverthe-less, the authors took advantage of the fact that the software captured the entire sequence of responses and revisions given by each student to investigate the internal coherence of

---

[19]For discussions and evidence on the role of conditioning information in expectations reports, see Dominitz (1997), Giustinelli and Shapiro (2019, 2021), Hudomiet et al. (2021), and references therein.

the students' initial responses. They focused on three types of errors: (1) reporting a probability outside the unit interval; (2) reporting a sequence of probabilities violating monotonicity of the cumulative distribution; (3) reporting a probability inconsistent with the previously reported median. None of the respondents made the first type of error, but some made errors of the second or third type. The analysis reveals three main patterns. First, (3) errors were more common than (2) errors. Second, the error rates in the early sections of the survey were higher than the error rates in later sections, suggesting that respondents tend to learn. Third, investigation of person-specific patterns of responses across questions revealed that some respondents are more error-prone than others.[20]

Dominitz and Manski (1996) additionally investigate bunching of responses. They find very little bunching among reports of expectations and beliefs about median earnings and a little more bunching among probability reports on the 0-100 scale. They do not find excessive bunching at so-called "focal" responses of 0, 50, and 100 percent.[21]

Subsequent studies perform additional analyses aimed at assessing the quality of expectations elicited from students. For example, Zafar (2012) shows that the expectations of earnings at age 30 reported by NU students imply relative earnings rankings across majors similar to those implied by statistics constructed with data from the Baccalaureate & Beyond Study, indicating that students are aware of the existing earnings differences across majors. Wiswall and Zafar (2020) compare the distributions of expected and realized own earnings of NYU undergraduates and find that they are remarkably similar. Mean expected earnings are very close to the mean of the actual earnings six years later: $73.5K versus $75K.

Wiswall and Zafar (2015a) quantify the extent of measurement error in the earnings expectations of NYU students and assess the impact of measurement error on inference. To do so, the authors administer a separate survey to a second group of students, acting as a control group and not used in the main analysis. In the first stage of the survey, control students were asked the same questions as the treatment students. In the second stage, control students were reasked their expectations without being treated with information about the population earnings. Hence, in the control group any observed changes in own earnings expectations between stages should reflect measurement error only. Under the maintained assumption of classical measurement error, Wiswall and Zafar (2015a) estimate a reliability ratio of 0.984 (equal to 1 minus the noise-to-signal ratio), implying a minimal attenuation bias in their estimates of about 1.6 percent.

The majority of surveys eliciting earnings expectations from students has focused on college students. A smaller number of studies has elicited earnings expectations from high school students.[22] A natural question is whether elicitation from younger students presents special challenges or their reports have been found to differ in quality or informativeness relative to those of older students. The evidence on this question is scarce at best.

Wolter (2000) appears to be the only study, after Dominitz and Manski (1996), to have elicited earnings expectations from both high school and university students. Using a survey whose design followed closely that of Dominitz and Manski (1996), the author elicited expectations for own and population earnings in a sample of Swiss students. The sample included students attending their final year in a college-preparatory high school or a vocational high school and students attending university. Wolter (2000) investigates the

---

[20]Schweri et al. (2011) report that elicitation of earnings expectations in their sample of Swiss university students worked well. Two thirds of the respondents gave responses consistent with the laws of probability and one fifth of respondents needed only a single reminder.

[21]Bunching of probability reports at 0 percent and at multiples of 5 or 10 percent may be an indication of rounding, a nonclassical form of measurement error. Manski and Molinari (2010) hypothesize that rounding of probability reports may reflect the preferred granularity with which some respondents communicate the chances of uncertain events or a lack of precise probabilistic expectations. For evidence and methods on rounded probabilities, see Manski and Molinari (2010), de Bresser and van Soest (2013), Kleinjans and van Soest (2014), and Giustinelli et al. (2020, 2022). See also Manski (2022) in this handbook.

[22]E.g., Menon Eliophotou (1997b,a), Wolter (2000), Menon Eliophotou (2008a,b), Mazza and Hartog (2011), Hastings et al. (2015, 2016), Huntington-Klein (2015, 2016b), Belfield et al. (2020), and Boneva and Rauh (2021).

predictors of measures of deviations between students' expectations and current population earnings and finds no systematic differences by age or type of school attended.

Using a broadly similar survey instrument, Mazza and Hartog (2011) elicit labor market information and expectations in a sample of Dutch students attending their last year of high school in one of three high school tracks of vocational, general, or college-preparatory type. A novelty of their survey protocol is that the questionnaire was fielded online. However, no feedback was provided for logically incorrect answers. When investigating non-response, the authors document a high prevalence of early quitters, which they attribute to the online mode. The heterogeneity analysis shows that male respondents and respondents from a low-education background are significantly more likely to be early quitters. Conditional on not quitting in the early part of the survey, the authors do not find differential non-response rates to the questions eliciting earnings expectations vis-à-vis other questions. When investigating response mistakes, the authors find that, *ceteris paribus*, male students, immigrant students, and students with a low-education background make significantly more mistakes.

Giustinelli (2016) elicits subjective probabilities and expectations over multiple outcomes following hypothetical choice of alternative high school curricula in a sample of Italian 9th-graders, whose modal age was 14, and their parents.[23] The author reports difficulties with elicitation of earnings expectations but not with elicitation of expectations or probabilities about other track-specific outcomes. The majority of survey participants–interestingly, among both students and parents–were unable or unwilling to report their expectations about the student's future earnings under alternative schooling scenarios.[24]

Belfield et al. (2020) and Boneva and Rauh (2021), who elicit earnings expectations and subjective probabilities over multiple outcomes of sixth-form and university education from UK students aged 13-14 and 13-18 respectively, do not report of any particular issue with elicitation from younger respondents.

*Patterns and heterogeneity of earnings expectations.* Dominitz and Manski (1996) analyze the empirical distributions of the median and interquartile range (IQR) of respondents' subjective belief distributions about their future earnings at ages 30 and 40 under three attainment scenarios: (1) unconditional on schooling, (2) conditional on obtaining a high school diploma (for high school respondents) or terminating school after the current semester (for college respondents), and (3) conditional on obtaining at least a bachelor's degree. The authors define four groups according to the respondent's gender and current schooling level (high school or college). They find limited variation in the median and IQR by gender and schooling, possibly due to their sample consisting of respondents from a high and relatively homogeneous socioeconomic background. But they document the existence of substantial heterogeneity within groups. When comparing the empirical distributions of the subjective median and IQR across scenarios, the authors find that those corresponding to the unconditional scenario are similar to those corresponding to the bachelor-degree scenario, consistently with the high probability of obtaining a bachelor's degree reported by most students. When they further compare these measures across age scenarios, the authors find that students expect their earnings to increase with age and to increase more should they obtain a bachelor's degree than a lower attainment level.

Carvajal et al. (2000) elicit earnings expectations among US college seniors enrolled in a randomly selected set of courses at the College of Business of Florida International University in Miami. They find significant heterogeneity by gender, with female students expecting lower mean earnings than their male peers.

---

[23]Elicitation from students was done via a paper-&-pencil survey administered in the schools by trained enumerators. Elicitation from parents was done via an at-home self-administered paper-&-pencil survey.

[24]Among students, item non-response rates to the earnings expectations questions were as high as 65%. Respondents provided nevertheless informative feedback. The majority of students reported being clueless about the order of magnitude of monthly salaries, while a minority gave estimates based on their parents' earnings. Some parents left written notes on the questionnaire, expressing their perceived difficulty of providing meaningful earnings forecasts and stating that they did not regard future earnings as an important factor for the choice of high school track.

Brunello et al. (2004) analyze the earnings expectations of students specializing in Economics or Business from twenty-six universities across Europe, spanning Austria, Finland, Germany, Greece, Italy, Portugal, Sweden, Switzerland, and UK. Like Carvajal et al. (2000), they focus on expected mean earnings. But, like Dominitz and Manski (1996), they elicit expectations about students' own future earnings at two points in time (at labor market entry and after 10 years on the job) and under two attainment scenarios (the student obtains a college degree and the counterfactual in which the student had not continued onto college after high school). The analysis documents substantial variation in subjective earnings expectations across students. A part of this variation is explained by students' personal characteristics or circumstances. For example, female students expect significantly lower college and high-school earnings than male students, *ceteris paribus*.[25] The implied gender difference widens when expectations about earnings after 10 years on the job are considered, consistent with females expecting to enter jobs with lower relative earnings growth. Students in their second and third year of college expect lower college wages at labor market entry than freshman students, but these differences shrink when considering expectations for earnings after 10 years on the job.[26] Students who expect to take longer than required to complete college have lower expected college wages than their counterparts, but similar expected high school wages, for both horizons. Students with a more educated mother have higher expected earnings after college and high school. Students who rank themselves above the average in their class and those who report having acquired information about earnings from personal sources or from the daily or weekly press have higher expected college earnings. On the other hand, students who report having chosen their university based on costs or by an external assignment have lower earnings expectations. Finally, students from countries that have higher wages and those enrolled in more selective universities have higher expectations.

Moving to younger students, Mazza and Hartog (2011) uncover five main patterns in their data on expected median earnings of Dutch high school students.[27] First, earnings expectations display higher variation within groups than between groups, as in Dominitz and Manski (1996)'s data. Second, female students expect significantly lower median earnings than male students, consistent with findings of most expectations studies but Dominitz and Manski (1996)'s. Third, for the scenario in which the student does not continue to university, the expected median varies by attended high school tracks. Fourth, expected median earnings following university are generally higher than those from going to work right after high school. Fifth, the subjective earnings distributions shift upward with hypothesized work experience, suggesting that respondents are aware of the increasing wage profile over the life cycle.

Other studies have elicited earnings expectations under hypothetical scenarios specifying alternative fields of study (e.g., college majors), rather than attainment levels. For example, Arcidiacono et al. (2012) elicit expectations about earnings 10 years after graduation under alternative major-career combinations in a sample of male undergraduates attending Duke University. The authors additionally elicit students' subjective probabilities of alternative careers. A number of patterns emerge from the analysis. First, the way in which students' subjective probabilities of entering each career vary across majors suggests that some careers are perceived to be tied to certain majors, whereas other careers less so. For example, careers in Science or Health are perceived to be associated with majors in Science or Engineering. Likewise, a career in Education is perceived to be asso-

---

[25]These findings are based on multivariate regressions.

[26]This latter pattern concurs with that documented by Betts (1996) (reviewed below), using expectations about population earnings (instead of own earnings) collected in a sample of University of California, San Diego (UCSD) undergraduates.

[27]Like Brunello et al. (2004), they elicit students' earnings expectations in four scenarios: entry wage with secondary education, entry wage with tertiary education, wage after 10 years of work with secondary education, and wage after 10 years of work with tertiary education. Like Dominitz and Manski (1996), they elicit for each respondent the expected median (instead of mean) earnings, plus two additional points on the respondent's expected earnings distribution. Using the latter, the authors construct a measure of variance and one of skewness of students' subjective earnings distributions similar to that used by Schweri et al. (2011).

ciated with a major in the Humanities. In contrast, students' subjective probabilities of entering a Business career are relatively high for all majors, even though it is particularly high following a major in Economics. Second, also students' earnings expectations vary systematically across majors. For example, majoring in Science or Engineering is perceived to lead to higher earnings in Science and Health careers, while majoring in Economics is perceived to lead to higher earnings in Business. Third, students in later years tend to report lower earnings expectations than students in earlier years, irrespective of their major. This difference is statistically significant for selected majors and is consistent with findings from other studies. The authors attribute it to differential information between groups due to learning over time. Fifth, the observed patterns in students' earnings expectations is consistent with income sorting in choice of majors. Specifically, with the exception of students majoring in Humanities, students associate the highest or second-to-highest expected earnings to their own major. However, most students also expect equally high or higher earnings from a major in Economics, regardless of their actual major.

Reuben et al. (2017) investigate the gender gap in earnings expectations among NYU undergraduates. The authors document a large, and increasing with age, gender gap in expected future earnings, amounting to 31% lower expected earnings at age 30 and 39% lower expected earnings at age 45. These figures are partly attributable to gender differences in major/career choices, since male students are more likely to select high-earning majors such as Business and less likely to select low-earning such as Humanities than female students. However, the authors show that a large gap remains when analyzing students' earnings expectations for all majors, including the non-chosen ones. On average across all majors, female students expect to earn 19% less than men at age 30 and 23% less at age 45.[28]

*Perceived monetary returns to schooling.* Despite the substantial within-group variation in earnings expectations and the associated heterogeneity in the perceived returns across students in their sample, Dominitz and Manski (1996) find that students share a common belief that the monetary returns to a college education are positive and increasing with age.[29] These two patterns of students' perceived monetary returns to schooling–positive in sign but heterogenous in magnitude–have been repeatedly confirmed by subsequent studies (e.g., Menon Eliophotou (1997*b*,*a*), Wolter (2000), Botelho and Pinto (2004), Brunello et al. (2004), Webbink and Hartog (2004), Mazza and Hartog (2011), Schweri et al. (2011), Belfield et al. (2020), Boneva and Rauh (2021), Boneva, Golin and Rauh (2022), and others).

Some studies have found that the variation in perceived monetary returns across students–and not just that in the underlying school-contingent earnings expectations–is partly associated with variation in students' characteristics, circumstances, and information. For example, Botelho and Pinto (2004) find that, *ceteris paribus*, female students expect significantly lower monetary returns to a college education than male students and that, irrespective of gender, senior students expect lower returns than freshman students. These patterns concur with those documented above with respect to school-contingent earnings expectations. Moreover, the authors find that students expect for themselves returns that are above the average ones, with this tendency being stronger among male and freshman students.

Brunello et al. (2004), too, find that female students expect lower college wage premia at both labor market entry and after 10 years, but the difference is statistically significant only for the latter case. The authors additionally find that older students have lower expected college wage premia, conditional on year of enrollment. Students from countries

---

[28]Using experimental measures of competitiveness, overconfidence, and economic preferences, Reuben et al. (2017) show that, conditional on the chosen major, gender differences in competitiveness and overconfidence explain nearly 20% of the gender gap in earnings expectations.

[29]Recall that a person's perceived return to a college education is defined as the person-specific difference in subjective earnings expectations between the scenario in which the person obtains a college degree and that in which the person does not.

that have higher college wage premia have higher perceived college wage premia, but only based on their subjective expectations about earnings after 10 years on the job.

Among younger UK students in 9th grade, Belfield et al. (2020) find that female students and low socioeconomic (low-SES) students report, on average, lower expected earnings across all schooling scenarios (high school, sixth form, and university).[30] However, in multivariate regressions conditioning on a rich set of covariates, both female students and low-SES students are shown to hold significantly higher perceived monetary returns to university and insignificantly higher perceived monetary returns to sixth form. In a separate analysis of UK students aged 13-18, Boneva and Rauh (2021) find that low-SES students expect a lower earnings premium from going to university than high-SES students.

*Perceived earnings risk.* When analyzing the interquartile ranges (IQRs) of the fitted subjective earnings distributions, Dominitz and Manski (1996) find that most students perceive substantial uncertainty with respect to their future earnings. Specifically, the IQRs are of the same order of magnitude as the reported subjective medians. The authors additionally show that students tend to be more uncertain about their future earnings with a bachelor's degree than about their earnings with a lower attainment level. On the other hand, students tend to be equally uncertain about their future earnings at ages 30 and 40, conditional on attainment.

Similarly, Mazza and Hartog (2011) find that most respondents in their sample of Dutch high school students have subjective belief distributions about future earnings characterized by a large variance, which the authors interpret as an indication that students have a high perceived wage risk. Perceived wage risk does not vary systematically across schooling levels (scenarios), but does vary by gender, with female students perceiving higher wage risk than male students. Irrespective of gender, students' expectations imply a perceived decreasing wage variance over time, that is, between entry wage and wage after 10 years.

Schweri et al. (2011) study the subjective distributions of future own earnings held by a sample of Swiss university students. The authors are especially interested in the relationship between expected wage (mean) and wage risk (variance) and in whether students expect wage risk to be compensated on the labor the market. Like the other two studies, the authors find that students' belief distributions about post-schooling earnings are highly dispersed. However, differently from Mazza and Hartog (2011), they find that the variance constructed from students' responses varies systematically across scenarios. In particular, the variance tend to be higher in high-age and high-education scenarios (age 40 vs. 30 and tertiary vs. secondary education). The authors regress the (log of) expected median wage on personal characteristics, year and scenario dummies, and the expectations-based measures of wage variance and skewness. As predicted by theory, higher perceived risk (variance) significantly increases expected median wage and higher skewness reduces it. Specifically, an increase in the wage variance from 0 to 1 is associated with a more than 40% higher expected median wage. The authors show that the risk elasticities implied by their expectations-based estimates are comparable to those estimated using risk-augmented Mincer earnings equations with realizations data. The authors further show that these effects are not mechanically due to the fact that students' subjective earnings distributions are lognormal, as they are not.

Schweri et al. (2011)'s analysis additionally shows that students expect median earnings to increase with both age and education and that male students expect significantly higher median earnings than female students, but finds no differences by family background or by ability, as measured by math grades. The authors view these patterns in students' subjective belief distributions as being consistent with observable labor market patterns, but inconsistent with students having/using private information when reporting their earnings

---

[30]After completing 11 years of compulsory schooling (ages 5–16) and obtaining the General Certificate of Secondary Education (GCSE), UK students face a first consequential choice between remaining in school for two more years ('sixth form') and leaving school. The two extra years lead to A level qualifications and to a second important choice between going or not to university. About 82% of UK students aged 16–18 years are enrolled in full-time education, while 48% of any given cohort continues to university.

expectations. However, as noted by Gong et al. (2019) within the context of their analysis of the earnings expectations of Berea College students, the uncertainty-reducing effect of the greater or better information that high-ability or high-SES students might have may well be counterbalanced by the uncertainty-enhancing effect of their having greater access to job opportunities. Gong et al. (2019), too, find no difference by American College Testing (ACT) score in the extent of uncertainty held by Berea students at college entrance about their future labor market earnings. They do find differences by race, though. Specifically, black students have significantly larger earnings uncertainty when they enter college, but resolve more uncertainty than the other students during college. So, the race gap in earnings uncertainty is closed by the end of college.

*Beliefs about population earnings.* Manski (1993)'s question of what information students condition on when forming expectations about future earnings and the monetary returns to schooling clearly underlies most of the analyses of survey-elicited earnings expectations reviewed thus far. Some studies have tried to address this question more directly by eliciting students' beliefs about current population earnings, by comparing these beliefs with actual population earnings, and by investigating whether/how students' expectations about own future earnings are related to their beliefs about population earnings.

Betts (1996) elicits beliefs about population earnings from undergraduate students at the University of California, San Diego (UCSD). His elicitation focuses on mean earnings but spans a wide range of subpopulations, covering workers with different attainment levels, from different fields of specializations, and of different age/work experience. The author finds that while students are aware that wage profiles are positively sloped, their estimates of population earnings vary greatly across personal characteristics and wage subpopulations. For example, students in later study years provide systematically lower earnings estimates for multiple subpopulations than students in earlier years. Students from lower income families also provide lower estimates of the earnings of college graduates than students from more affluent families.[31] Students provide higher estimates of the average starting salary of graduates in their own field than for graduates in other fields. Betts (1996) shows that this is due to students underestimating population earnings in fields other than their own; indeed, students hold more accurate beliefs about the population earnings of graduates from their own field than for other fields. Another interesting finding is that students' accurate knowledge of starting salaries in their own field does not carry over to salaries of more experienced workers; for example, students are found to greatly underestimated the slope of the wage profile for engineers with a bachelor's or a master's degree. Betts (1996) further documents that belief accuracy varies by year of study, family socioeconomic background, and GPA. Specifically, students in later years of study make smaller mistakes, consistently with learning over time. Students from poorer families make larger errors when estimating salaries of college graduates, consistently with the possibility that students form their beliefs based on the reality of their everyday experience. Higher GPA students make smaller errors when estimating starting salaries.

In a sample of community college students in California, Baker et al. (2018) find that less than 15% of students can rank broad categories of majors accurately in terms of labor market outcomes and that students believe that salaries are, on average, 13% higher than they actually are. In contrast, Conlon (2021) finds that Ohio State University (OSU)'s freshmen underestimate the average salaries of workers from most fields by about $15K (19% percent of the average true value) or more, on average. Additionally, the author shows that students are significantly misinformed about the differences in average salary between fields. Across all pairs of fields, the mean error in the difference in average salary is close to $20K (84% percent of the average true difference).

Bleemer and Zafar (2018) and Ruder and van Noy (2017a) provide evidence from non-student samples. Bleemer and Zafar (2018) find that nearly three-quarters of their respondents, a sample of household heads from the New York Fed's Survey of Consumer

---

[31] Betts (1996) hypothesizes that this may be because young people form beliefs about the returns to education by observing workers in their neighborhood.

Expectations (SCE), underestimate average returns to a college degree. Moreover, the authors find that both college-educated and higher-income respondents make significantly lower absolute errors than their counterparts. Ruder and van Noy (2017a) find that US respondents recruited through Amazon's Mechanical Turk (MTurk) hold incorrect estimates of the population distributions of earnings for graduates from different fields. The authors' analysis focuses on respondents' perceptions about the dispersion of the population earnings distributions, or earnings risk. The analysis finds that respondents believe that graduates in economics and mathematics face similar earnings risk, implying that relative to the actual population distributions respondents overestimate the earnings risk for mathematics and underestimate it for economics.

Dominitz and Manski (1996) and Wiswall and Zafar (2015a) are, to my knowledge, the only two studies to have elicited both population earnings beliefs and own earnings expectations from students within the same survey. Dominitz and Manski (1996) investigate students' beliefs about the current distributions of earnings among 30-year-old Americans of the same gender as the respondent and with either a college degree or a high school diploma and no further educations. They find that these beliefs follow a pattern similar to that of students' expectations: limited variation across genders and schooling levels, but substantial variation within groups. The authors compare respondents' beliefs with estimates obtained using data from the Current Population Survey (CPS) on actual earnings and schooling. They find that the median male respondent has accurate perceptions of the current median earnings of males, whereas the median female respondent overestimates the median earnings in the female population. They speculate that this may be due to a tendency of respondents to report their beliefs about the earnings of full-time, year-round workers rather than of all persons. When examining the IQRs, the authors find that students are aware that the spread of the earnings distribution among college graduates is larger than that among high school graduates, but tend to overestimate the degree of earnings inequality.[32] Finally, the authors investigate the relationship between population beliefs and own expectations in the subsample of high school students. They find the subjective medians of the two distributions to be highly correlated with each other, suggesting that students take their knowledge of the distributions of current earnings into account when forming expectations about own future earnings.

Wiswall and Zafar (2015a) find that their respondents have, on average, significantly biased beliefs about the population distribution of earnings, despite their being high-ability undergraduates at New York University (NYU), a selective private university. For example, on average, sample respondents underpredict annual average earnings of male workers with no college degree by nearly $10K and overpredict average earnings of male graduates in Economics/Business by nearly $35K. The authors find that errors in students' beliefs about population earnings vary widely across respondents, with some students overestimating and others underestimating population average earnings. However, these errors are found to be only weakly correlated with students' observable characteristics. Using an experimental design that randomly gives students information about population earnings, the authors show that students' expectations about own earnings are causally linked to students' beliefs about the population earnings. However, the latter do not fully explain the former, indicating that students form their expectations about own earnings also based on other factors.

Moving to evidence from countries other than the US, Botelho and Pinto (2004) compare the average market returns to college as perceived in their sample of Portuguese university students with estimates of the marginal rate of return to college obtained from data on schooling and earnings of the Portuguese Ministry of Labor. They find that on average students tend to overestimate the actual returns, but female and senior students have significantly more accurate beliefs than male students and students in earlier years of university.[33]

---

[32] Wolter (2000), instead, finds that Swiss students tend to underestimate the extent of the spread in wages.

[33] A distinctive feature of Botelho and Pinto (2004)'s elicitation procedure is that it used a scoring rule to provide financial incentives for accurate reporting to a random subsample of respondents. The authors find no

Hastings et al. (2015) find that the accuracy of the beliefs of Chilean college applicants about the earnings of past graduates varies by selected student's characteristics. In particular, while high-achieving students hold nearly correct beliefs on average, low-income and low-achieving students who apply to low-earning programs (institution-major combinations) overestimate earnings for past graduates by over 100%.

Kerr et al. (2020) evaluate the impact of a randomized intervention providing a large sample of high school students in Finland with information about the labor market prospects associated with alternative post-secondary programs, including program-specific information about employment rates, average monthly earnings, a graph on the distribution of monthly earnings (first and ninth deciles, quartiles, and median) for current 30–34-year-olds, the two most common occupations, and the share of graduates in those occupations. The authors did not elicit students' prior knowledge or beliefs about population earnings and the other statistics provided by the intervention. Thus, they are unable to investigate whether students hold accurate beliefs. However, they find that treated students do update their expectations about own labor market prospects in sensible directions.

*Are elicited earnings expectations rational?*  Having collected a substantial amount of data from students on their subjective earnings expectations, a natural question to ask is whether the expectations students report are consistent with the economic theory benchmark of rational expectations. A few studies have tried to address this question by comparing students' earnings expectations with earnings realizations. Early analyses have focused on the means of the expectations and realizations distributions. More recent analyses have also looked at higher moments of the two distributions. Sometimes, elicited expectations have been compared with subsequent realizations in a relevant population; whenever possible, they have been compared within the same sample.

For example, Webbink and Hartog (2004) compare earnings expectations and subsequent realizations within the same sample of Dutch students. Expectations were elicited in the first survey wave, which was fielded in the year in which students started post-secondary education. Actual earnings were elicited after four years. The authors cannot reject the null hypothesis of equality of means between expected and realized earnings. They further show that this finding does not vary by respondent's characteristics.

More recently, D'Haultfoeuille et al. (2021) and Crossley et al. (2021) have developed a series of new tests of rational expectations that take into account the information on higher moments (beyond the first), contained in survey expectations for continuous variables whenever these expectations are elicited as multiple points on the respondent's subjective belief distribution.[34] Crossley et al. (2021) use their tests to assess the rationality of earnings expectations in the Berea Panel Study (BPS). They find that earnings expectations become more accurate as BPS students progress through college, and especially after they leave college. The analysis almost always rejects the null hypothesis of rational expectations based on the data collected in the in-school period, but the evidence against rationality of expectations is much weaker on basis of the data collected in the post-college period. In related work, Crossley et al. (2022) propose a novel measure of the average ex post accuracy of expectations of economic agents and provide a decomposition into different sources of inaccuracy, including failure of rational expectations, subjective uncertainty, and aggregate shocks. They illustrate using BPS data.[35]

---

significant effects of incentives on the accuracy of respondents' beliefs.

[34]Some of these tests can be applied using microdata on expectations and realizations observed in different datasets that cannot be matched, while others require observing expectations and realizations for the same individuals. Under some restrictions, some versions of the tests are made robust to measurement errors and can be extended to account for aggregate shocks.

[35]D'Haultfoeuille et al. (2021) apply their tests to earnings expectations elicited in the Labor Market module of the New York Fed's Survey of Consumer Expectations (SCE). While the SCE surveys a different population from the BPS (as SCE respondents are household heads rather than students), D'Haultfoeuille et al. (2021) also find evidence of substantial heterogeneity in the rationality (or lack thereof) of respondents' earnings expectations. In their setting, the null of rational expectations is systematically rejected for SCE respondents without a college degree, but not for college-educated workers. While more extensive testing is needed before

*Other labor market outcomes.*  Beyond eliciting students' expectations about earnings, many studies have elicited students' expectations for additional labor market outcomes such as employment status (e.g., Carvajal et al. (2000), Fischhoff et al. (2000), Brunello et al. (2004), Zafar (2011*b*, 2012, 2013), Huntington-Klein (2015, 2016*b*), Wiswall and Zafar (2015*b*,*a*, 2020), Giustinelli (2016), Baker et al. (2018), Gong et al. (2019), Gong et al. (2020*a*), Belfield et al. (2020), Delavande et al. (2020), Boneva and Rauh (2021), and Boneva, Golin and Rauh (2022)). Many of these studies find that these expectations tend to follow empirical patterns that are similar to or consistent with those documented for earnings expectations, for example, in terms of how expectations vary by students' characteristics.

*Monetary costs.*  Measurement of expectations about the monetary costs of education, financing opportunities and conditions, student debt, and related outcomes is highly topical, as these data have the potential to shed light on the unsettled debate regarding the importance and role of credit constraints in human capital investment and educational choices.[36] The available evidence to date is still relatively scarce.

Hastings et al. (2015, 2016) find that Chilean college applicants have correctly centered but noisy expectations about the costs of college. Hastings et al. (2016) investigate the heterogeneity in students' beliefs about annual costs of college attendance and compare them with administrative values. The analysis shows that 39.6% of sample students claims not to know about costs at their stated first-choice degree programs in college and that this proportion vary by students' socioeconomic status (SES for short) and performance in the college admission exam (*Prueba Seleccion Universitaria* or PSU). Low-SES students are 10.6 percentage points (46.7% versus 36.1%) more likely than high-SES students to claim not to know about degree costs. Low-PSU students are 19.2 percentage points (48.2% versus 29.0%) more likely than high-PSU students to claim no knowledge of costs. The analysis further shows that 50% of students have cost expectations that are approximately within 11% of the cost value observed in the administrative data. Low-SES and low-PSU students have significantly noisier cost expectations.

Huntington-Klein (2016*b*) conducts a broadly similar analysis to that of Hastings et al. (2016) in a sample of high school junior and seniors from thirteen schools in King County, Washington. The author compares students' expectations of tuition, fees, and book costs at four local colleges or universities with the actual in-state listed prices. Differently from Hastings et al. (2016) who show that Chilean students have correctly centered cost expectations, Huntington-Klein (2016*b*) finds that students overestimate costs at their first-choice institution by 40.6% at the median. However, the heterogeneity pattern in students' expectations and errors is similar between the two analyses. For example, low-SES students are more likely to skip the tuition question and exhibit larger IQR of the error distribution.[37]

Bleemer and Zafar (2018) find that about 60% of their SCE respondents overestimate average college net costs in the US, with no significant differences by education or income.

Belfield et al. (2020), Boneva, Golin and Rauh (2022), and Boneva and Rauh (2021) provide evidence on the UK. Belfield et al. (2020) investigate high school students' beliefs about university tuition costs. The authors conducted their survey the year after the government allowed UK universities to increase their annual tuition fees to between £6K and £9K.[38] In their analysis they find that, on average, students believe that they will pay

---

reaching firm conclusions, these initial findings suggest that the hypothesis of rationality of earnings expectations may be more credible for certain subpopulations than others.

[36]See Lochner and Monge-Naranjo (2012, 2016) for recent reviews on credit constraints in education.

[37]Obviously, there are key differences between Huntington-Klein (2016*b*)'s and Hastings et al. (2016)'s studies. The institutional settings are very different. Moreover, Huntington-Klein (2016*b*) analyzes partially younger students and a different school-choice margin (college enrollment/attaiment instead of degree program).

[38]At the same time, Boneva, Golin and Rauh (2022) and Boneva and Rauh (2021) note that UK students can easily fund their education through loans provided by the government. These loans are usually not contingent on the student or family income, can be used to pay for course fees and living costs, and don't need to be repaid until students start earning an income above a certain threshold. Moreover, students from low-income

£7.3K per year, but the model reported value is actually £9K, consistently with the fact that most universities transitioned to the maximum fee. They also find that the response distribution is similar across gender and socioeconomic groups, with responses bunching at £5–6K and £9–10K and a few outliers above £10K. However, high-SES students are found to expect somewhat higher costs than low-SES students.

Boneva, Golin and Rauh (2022) analyze university students' expectations of struggling financially, having enough money, and receiving financial support from family if the student continues to postgraduate education vis-à-vis enters the labor market.[39] Boneva and Rauh (2021) analyze a similar set of expectations elicited from students attending secondary school with respect to their university enrollment decision. Boneva, Golin and Rauh (2022) find that, on average, female university students report a higher probability of struggling financially than their male peers. Moreover, students with better-educated parents (high SES) who attended a state school perceive a significantly larger difference in the probability of struggling financially if they pursue postgraduate education versus not than students who attended a private school, despite giving lower subjective estimates of the immediate costs of postgraduate education. Boneva and Rauh (2021) find that, on average, students report a higher probability of struggling financially, a lower probability of having enough money, and a higher probability of receiving financial support from their parents, if they go to university than otherwise. Beliefs, however, vary significantly across SES groups. On average, low-SES students report significantly higher probabilities of struggling financially and lower probabilities of having enough money and receiving financial support from their parents than their high-SES peers.

## 3  Survey Expectations About Nonmonetary Outcomes of Schooling

### 3.1  Motivation

Economists have long recognized that the outcomes of schooling and other human capital investments are not limited to earnings, but include also many outcomes of nonmonetary nature, also known as nonpecuniary, nonmarket, or nonproduction outcomes. Reviews of theory and evidence on the latter include Grossman (2006), Lange and Topel (2006), Lochner (2011), and Oreopoulos and Salvanes (2011).

Nonmonetary outcomes of schooling span benefits (or returns), risks, and costs. For example, progressing through the stages of formal education gives students the option value of acquiring further education,[40] but it also exposes students to the risks of non-passing grades, grade retention, and school dropout.[41]

Attending school, doing homework, studying, and taking tests require effort, which may be psychologically costly and generate disutility for students. Heckman et al. (2006) refer to these as the "psychic costs" of schooling. Yet, some of these very activities may be enjoyable and may give direct utility to some students. After all, attending school means learning new things, exchanging ideas, hanging out with one's schoolmates, and enjoying the amenities that certain schools offer. Thus, the process of acquiring education can have direct consumption value.[42] In practice, acquiring education may be differentially costly and/or enjoyable depending on students' ability and tastes (Arcidiacono, 2004).[43]

---

households are eligible for maintenance grants, which do not need to be repaid at all (Boneva and Rauh, 2021).

[39]Boneva, Golin and Rauh (2022)'s survey also asks students whether they hold a student loan and about their current university experience. The authors find that 82% of high-SES students and 95% of low-SES students hold a loan. Among the students without a loan, 85% of high-SES students and 50% of low-SES students report being helped financially by their family. Additionally, the authors find that low-SES students are more likely to report that they struggle financially and that they work while studying.

[40]This has been recognized at least since Weisbrod (1962) and Comay et al. (1973, 1976). For empirical analyses, see Stange (2012), Lee et al. (2015), Belfield et al. (2020), and Gong et al. (2020b).

[41]E.g., see Manski (1989), Altonji (1993), Eckstein and Wolpin (1999), and Stinebrickner and Stinebrickner (2012, 2014b).

[42]This, too, has been long recognized by economists; e.g., Schultz (1963) and Lazear (1977). For empirical evidence, see Jacob and Lefgren (2007), Jacob et al. (2018), Huntington-Klein (2016a), and Gong et al. (2021).

[43]Heckman et al. (2006) wonder why "(...) *if high school graduates have such positive returns to attending*

Education can also yield nonmonetary benefits after its completion. For example, in addition to facilitating labor market entry and yielding higher earnings, obtaining a more advanced degree or specializing in a certain field of study may increase access to jobs with desirable nonmonetary attributes such as flexibility or status.[44]

Dominitz and Manski (1996) focused on elicitation and analysis of students' expectations about earnings, paving the way for the large number of studies reviewed in the previous section. The initial focus on the monetary outcomes of schooling was a natural starting point, but mounting evidence on the importance of nonmonetary outcomes, and the inability of pecuniary factors such as monetary returns and financial constraints to satisfactorily account for observed patterns in human capital investment and educational choices, eventually led expectation researchers to expand the scope of elicitation and analysis of expectations to a wider range of outcomes, particularly nonmonetary ones.

## 3.2 Survey Elicitation

### 3.2.1 Nonmonetary Returns, Risks, and Costs of Schooling

Survey elicitation of expectations of nonmonetary schooling outcomes and related perceived returns and risks has followed closely the template described in Subsection 2.2. Thus, instead of discussing methodological aspects of elicitation, this subsection provides an encompassing list of nonmonetary outcomes for which subjective expectations data have been collected, along with references to the studies that have elicited and analyzed these expectations. The list distinguishes between expectations of nonmonetary outcomes that are realized during school and those that are realized after individuals leave school.

*Nonmonetary outcomes that are realized in school.*

1. *School performance, study effort, and own ability.* Studies that have collected percent-chance expectations about students' future grades, study effort, and related aspects of academic performance and difficulty, under alternative schooling scenarios, include Zafar (2011b, 2012, 2013), Stinebrickner and Stinebrickner (2012, 2014a,b), Wiswall and Zafar (2015b), Giustinelli (2016), Giustinelli and Pavoni (2017, 2019), Baker et al. (2018), Belfield et al. (2020), Delavande et al. (2020), Boneva and Rauh (2021), Boneva, Golin and Rauh (2022).[45] Fewer studies have elicited students' perceptions about own ability or relative ranking, within their actual schooling setting and in counterfactual ones (see Arcidiacono et al. (2012) and Wiswall and Zafar (2015b)). Some studies have used elicited expectations about own grades and study effort across multiple schooling scenarios to derive measures of perceived ability (see Stinebrickner and Stinebrickner (2012)).

2. *Subject interest and schoolwork enjoyment.* Studies that have elicited percent-chance expectations about students' interest in school subjects, enjoyment of schoolwork, enjoyment of the school environment, or related outcomes, include Zafar (2011b, 2012, 2013), Stinebrickner and Stinebrickner (2012, 2014b), Gong et al. (2021), Wiswall and Zafar (2015b), Giustinelli (2016), Giustinelli and Pavoni (2017, 2019), Baker et al. (2018), Belfield et al. (2020), Boneva and Rauh (2021), Boneva, Golin and Rauh (2022).

3. *Social life and mating.* In two recent surveys of secondary school and university students respectively, Boneva and Rauh (2021) and Boneva, Golin and Rauh (2022) elicit students' subjective probabilities about events characterizing students' social

---

*college, don't more attend?"* They argue that *"People do not only (or even mainly) make their schooling decisions by looking at their monetary returns in terms of earnings. Psychic costs play a very important role. More able people have lower psychic costs of attending college."* (Heckman et al. (2006), pp. 434-436).

[44]E.g., see Maestas et al. (2018) and Wiswall and Zafar (2018) for recent US evidence.

[45]Nonprobabilistic and/or nonnumerical measures of expectations about school grades and study effort have been collected for a long time in education surveys. For instance, Madeira (2018, 2020a) analyze grade expectations elicited from students, parents, and teachers participating in the Beginning School Study (BSS).

life in alternative schooling settings. The outcomes are: enjoying social life and activities, meeting people the student easily gets along with, and losing contact with family and friends. Boneva and Rauh (2021) additionally ask university students their subjective probability of finding a life partner.

4. *Mental health.* Boneva and Rauh (2021) and Boneva, Golin and Rauh (2022) also ask responding students their subjective probability that they would feel lonely and their subjective probability that they would feel stressed in alternative schooling settings.

5. *Support of parents or peers.* A small set of studies has elicited students' probabilistic expectations about parental approval and (non-financial) support of alternative schooling options (e.g., Zafar (2012), Giustinelli (2016), Giustinelli and Pavoni (2017, 2019), Boneva and Rauh (2021), Boneva, Golin and Rauh (2022)). Fewer studies have elicited students' probabilistic expectations about peers' choices and inter-peer coordination (e.g., Giustinelli (2016) and Garbin (2021)).

*Nonmonetary outcomes that are realized after school.*

1. *Work flexibility and work-life balance.* Zafar (2011b, 2012, 2013) elicit students' subjective probability of being able to reconcile work and family at the available jobs after graduating from alternative majors. Boneva, Golin and Rauh (2022) elicit students' subjective probability of achieving a good work-life balance.

2. *Work stability.* Baker et al. (2018) elicit students' beliefs about employment stability following alternative majors.

3. *Work enjoyment.* Multiple studies have elicited students' expectations of enjoying working at the available jobs after graduation or of closely related outcomes (e.g., Zafar (2011b, 2012, 2013), Giustinelli (2016), Delavande et al. (2020), and Boneva and Rauh (2021)).

4. *Social status and career/job satisfaction.* Multiple studies have elicited students' expectations about social status on the job or in society and/or their expectations about career/job satisfaction (e.g., Zafar (2011b, 2012, 2013), Delavande et al. (2020), and Boneva, Golin and Rauh (2022)).

5. *Contribute to society.* Delavande et al. (2020) and Boneva, Golin and Rauh (2022) additionally ask students' their subjective probability of contributing to society.

6. *Family-related outcomes.* A number of studies have elicited students' subjective expectations over family-related outcomes; for example, being held back in career by family-related reasons (Kunz and Staub, 2020), being married and having children in specific age brackets (e.g., Gong et al. (2020a) and Wiswall and Zafar (2020)), and having children (Boneva, Golin and Rauh, 2022).

### 3.2.2 Educational Attainment, Opportunities, and Choices

Elicitation of attainment expectations and schooling plans has a long tradition in education surveys such as those conducted by the National Center for Education Statistics (NCES).[46] These include the National Longitudinal Study of the H.S. Class of 1972 (NLS72), the National Education Longitudinal Study of 1988 (NELS88), the High School and Beyond (HS&B), the Education Longitudinal Study of 2002 (ELS2002), and the High School Longitudinal Study of 2009 (HSLS09), reviewed in the Supplementary Appendix. These studies have typically not allowed respondents to express uncertainty about their future schooling and work behavior, or those of their children. In the few occasions in which

---

[46]The NCES is the primary federal entity for collecting and analyzing data related to education in the US and other nations.

respondents were given the opportunity to respond probabilistically, the response scale was typically based on broad nominal categories of chance.[47]

To the best of my knowledge, the only large-scale US survey that has collected data on youth's expectations using a numerical scale of percent chance is the National Longitudinal Study of Youth 1997 (NLSY97). The base year questionnaire (Round 1) asked respondents aged 15–16 a battery of questions eliciting their unconditional probabilities of various schooling, work, parenthood, crime, and mortality outcomes.[48] For example, schooling expectations include the 1-year-ahead probability of being enrolled in a regular school, the probability of receiving a high school diploma by one's 20th birthday, and the probability of earning a four-year college degree by one's 30th birthday. Similar questions were administered in subsequent waves (Rounds 4, 5, and 6). Additional information on the expectations modules and other aspects of the NLSY97 are given in the Online Appendix. Fischhoff et al. (2000), Dominitz et al. (2001), Reynolds and Pemberton (2001), and Cowan (2017) provide analyses of the responses to these questions. Notably, a subset of the expectations questions asked to children in Round 1 were also asked in the parent questionnaire. Dominitz et al. (2001) analyze child and parent responses to the common set of expectations questions jointly.[49]

Outside the US, since 2000 the German Socio-Economic Panel (SOEP) has fielded a Youth Questionnaire for household members aged 16-17, featuring a battery of percent-chance expectations about education and labor market outcomes.[50] The outcomes are: to be accepted for a desired apprenticeship/place at university, to complete training/university successfully, to find a job in the desired career, to have job-related success, to be held back in career by family-related reasons, to become self-employed, to work in a foreign country, to get married, to cohabit with a partner (without being married), and to have one child and to have two or more children.

Following Dominitz and Manski (1996)'s small exploratory study and the larger-scale collections on the NLSY97 and SOEP, an increasing number of studies has elicited expectations about educational attainment, school dropout, school enrollment, and field choice. Fewer studies have collected data on beliefs about school or program admission, and about the preparedness and choice opportunities early educational stages provide for later ones. Here is a list of studies that have elicited expectations in each of these categories.

1. *Attainment and dropout.* Multiple studies have collected survey data on respondents' subjective expectations of graduating from specific degrees or programs (e.g., high school, college, high school track, college major), dropping out of school, and closely

---

[47]E.g., the student questionnaire of the NELS88 (second follow-up) asks respondents their expectations that they will graduate from high school, go to college, have a job that pays well, and have a job that the person enjoys doing. However, these questions elicit respondents' expectations using either a verbal 5-point scale of chance (from very low to very high chances), or in terms of expected age bands within which each event might occur. Other NCES surveys have asked responding parents to indicate "how far in school" they expect their child to go (i.e., the highest educational degree he or she will attain). Again, these questions do not allow parents to express uncertainty about their children's future education. For an extensive analysis of the educational expectations collected by the NCES, see Jacob and Wilder (2010), who combine data from HS&B, NELS88, and ELS2002 to document and analyze time trends and socioeconomic differences in educational expectations of US youths. More recently, Gershenson et al. (2016) and Papageorge et al. (2020) have investigated teachers' expectations for students' attainments in the ELS2002.

[48]After the initial development, the questions were refined on the basis of feedback obtained through a set of one-on-one focused interviews with a group of teens in the Pittsburgh area. The final questions were administered to NLSY97 respondents using computers brought into their homes and included two warm-up questions, on the chances of eating pizza and contracting flu during the year ahead, which were designed to ease respondents into the task and see how they used the response mode and scale on simple topics. Subsequent expectations surveys have used similar warm-up questions.

[49]A small set of expectations questions have also been asked in the The National Longitudinal Study of Adolescent to Adult Health (Add Health); for example, adolescents' perceived chance of living to age 35 and of attending college. However, response choices were constrained to a limited number of categories, spanning "almost no chance", "some chance, but probably not", "50/50 chance", "a good chance", and "almost certain".

[50]The SOEP is a longitudinal study of private households in the Federal Republic of Germany from 1984 to 2019 and in the eastern German länder from 1990 to 2019.

related outcomes. These studies include Jacob and Wilder (2010), Mazza and Hartog (2011), Zafar (2011*b*, 2012, 2013), Stinebrickner and Stinebrickner (2014*a*,*b*), Gong et al. (2020*b*), Giustinelli (2016), Giustinelli and Pavoni (2017, 2019), Baker et al. (2018), Belfield et al. (2020), Kunz and Staub (2020), Boneva and Rauh (2021), Boneva, Golin and Rauh (2022), and Boneva, Buser, Falk and Kosse (2022).

2. *Enrollment and field choices.* Multiple studies have additionally collected survey data on respondents' subjective expectations of enrolling in or selecting specific schools or programs (e.g., high school, college, high school track, college major). These studies include Mazza and Hartog (2011), Zafar (2011*b*, 2012, 2013), Stinebrickner and Stinebrickner (2014*a*,*b*), Wiswall and Zafar (2015*b*), Giustinelli (2016), Giustinelli and Pavoni (2017, 2019), Baker et al. (2018), Belfield et al. (2020), Kunz and Staub (2020), Boneva and Rauh (2021), Boneva, Golin and Rauh (2022), and Boneva, Buser, Falk and Kosse (2022).

3. *Application and admission in centralized choice.* Fewer studies have elicited applicants' beliefs about school admission chances in centralized school choice settings, where admission is not guaranteed to all applicants. Kapor et al. (2020) elicit parents' subjective probability that their child will be admitted to the high schools indicated in the application form in New Haven, Connecticut. Similarly, Arteaga et al. (2022) elicit beliefs about admission chances from school applicants within the centralized school choice system in Chile, which comprises both primary and secondary schools, and Larroucau et al. (2021) from high school students applying to college through the Chilean centralized college admissions system.

4. *Choice flexibility and opportunities.* A few studies have elicited expectations about choice flexibility or opportunities following alternative schooling paths. For example, Giustinelli (2016) investigates the likelihood of facing a flexible choice between university and work options following graduation from alternative high school tracks in a sample of 9th-graders and their parents in northern Italy, a country whose educational system features early curricular specialization or tracking. A similar question is included in a related study of expectations and high school track choice among Italian 8th-graders and their parents, described by Giustinelli and Pavoni (2017). The same study additionally elicits students' and parents' percent-chance expectations that alternative high school tracks will prepare the student best for work vis-à-vis university, will prepare the student best for specific university fields, and will enable the student to make a flexible choice among university fields. Belfield et al. (2020) elicit students' expectations about whether sixth form is instrumental for college in the UK.[51]

## 3.3 Empirical Evidence

*Are elicited probabilities meaningful? Rational?* Fischhoff et al. (2000) evaluate the probability reports of NLSY97 teen respondents aged 16-17 to 18 questions eliciting percent-chance expectations for significant life events, including "Be a student in a regular school one year from now", "Have received a high school diploma by the time you turn 20", and "Have a 4-year college degree by the time you turn 30". For each question, Fischhoff et al. (2000)'s analysis addresses: (1) the overall picture of teens that probability responses provide, (2) the accuracy of responses against existing statistics for similar individuals, (3) the correlations across probability judgements, and (4) the coherence of responses with those to related questions in other modules of the NLSY97. The authors' overall assessment is that the relative accuracy of most aggregate estimates, the lack of stable individual-difference tendency to give high or low probabilities, the sensible correlations between probability responses and responses to related NLSY97 questions, the respondents' use of

---

[51]Gong et al. (2020*b*) obtain estimates of the perceived option value of college from combining a college dropout model with respondents' subjective dropout probabilities and perceived monetary and nonmonetary returns to college education in the BPS.

the entire response scale, and between-group differences reflecting social realities provide confidence in the ability of the instrument to capture respondents' beliefs.[52]

With respect to the schooling questions, the analysis shows that the majority of respondents expects to be in school almost surely in a year, with a mean probability of 92 percent and a median and mode of 100 percent. These numbers match favorably with the statistics that 95% of the teens in the sample report being currently in school. Teens not in school give lower probabilities (46 vs. 95 percent), yielding a rank order correlation of 0.87 between school expectations and current status. The analysis further shows that these probabilities are correlated positively with self-reports of the "percent of peers who plan to go to college" and negatively with the "percent of peers who have cut classes or school". Respondents have similarly high subjective probabilities of achieving a high school diploma by age 20: 93 percent, on average. Students currently in school report higher probabilities than those who are not (95 vs. 61 percent). These responses, too, are correlated with peers' college plans but not with peers' tendency to cut classes. Compared to the 84% completion rate of 20-year-olds, these judgments look somewhat optimistic. Even more optimistic are the expectations of earning a 4-year college degree by age 30, where the mean subjective probability is more than twice the current rate: 73 vs. 30 percent. Also in this case, teens currently in school report higher college expectations than those no longer in school (75 vs. 37 percent). These probabilities are found to correlate positively with the highest-level math class taken, the number of math classes taken, and the percent of peers seen as planning to go to college.

Taking advantage of the longitudinal nature of the NLSY97, Cowan (2017) compares teens' college expectations with their subsequent college outcomes and investigates the predictors of the observed discrepancies between the two. The author finds that teens' expectations are highly predictive of their future college completion. For example, a 1 percentage point increase in a youth's subjective likelihood of college graduation is associated with a 0.6 percentage point increase in the probability of attaining a college degree by age 30 and expectations alone explain about 14% of the variation in college attainment.[53] At the same time, Cowan (2017)'s analysis confirms Fischhoff et al. (2000)'s earlier finding that teens' college expectations are overly optimistic.[54] For example, only 24% of respondents who give a subjective probability of attaining college by age 30 between 41 and 80 percent are subsequently observed to attain a college degree and only 52% percent of those who give a probability between 81 and 100 percent are. The analysis further shows that the expectation bias is highly negatively correlated with youths' scholastic aptitude and that, once AFQT test scores are accounted for, other variables such as gender, race/ethnicity, parental education, and family income have little or no predictive power. Interestingly, the empirical relationship between youths' expectations and family income tracks closely that between actual college completion and family income. The author interprets these results as suggesting that while teenagers seem to properly anticipate the role of family income in college outcomes, they tend to have a poor understanding of the importance of academic preparation for college success.[55]

_____

[52]Focusing on 3 of the 18 NLSY97 expectations questions (the probability of receiving a high school diploma by age 20, of serving time in jail or prison by age 20, and of becoming the mom/dad of a baby by age 20), Dominitz et al. (2001) show that item response rates are very high (exceeding 99%) and response bunching is low-to-limited (of the order of 3-4% percent for the diploma and jail questions and 12% percent for the parenthood question).

[53]Thus, the author concludes that youths take the expectations question seriously. Domina et al. (2011) further show that college expectations are an important predictor of student effort in high school.

[54]This is also consistent with earlier evidence based on nonprobabilistic expectations (e.g., Reynolds and Pemberton (2001) and Jacob and Wilder (2010)) and subsequent evidence based on probabilistic expectations in other samples (e.g., Zafar (2011b) and Stinebrickner and Stinebrickner (2012)).

[55]Reynolds and Pemberton (2001) find that 71% of the NLSY97 15- and 16-years-olds give themselves a greater than 50 percent chance of completing a college degree by age 30 and 57% of them a greater than 75 percent chance. The authors note that while these expectations are excessively optimistic when compared with college graduation rates (about 28% in 1998), they compare favorably with the two-thirds of high school graduates aged 16-to-24 enrolled in college in 1997. They hypothesize that youth may not understand the challenges of completing a college degree conditional on starting college. Later evidence confirms that indeed

Cowan (2017) investigates the relationship between expectations and outcomes more formally by performing a "regression test" of rational expectations. Specifically, he estimates a mean linear regression of the binary indicator for obtaining a college degree by age 30 on teenage college expectations and a constant and tests the joint hypothesis that the expectation coefficient is equal to 1 and the constant is equal to 0, corresponding to a null hypothesis of rational expectations.[56] The regression estimates confirm that youths' college expectations are overly optimistic, although the degree of bias shrinks as the AFQT percentile increases. The author rejects the null hypothesis for the whole sample as well as for each subsample corresponding to different AFQT terciles.

Subsequent studies perform additional analyses assessing the quality of students' probability judgements about nonmonetary education outcomes. For example, Zafar (2011a) examines cognitive biases in subjective expectations of NU undergraduates, with a focus on whether cognitive dissonance affects reporting[57] and whether respondents exert sufficient mental effort when answering. Taking advantage of a two-wave panel of percent-chance expectations over multiple outcomes, the analysis rules out cognitive dissonance being a serious concern and finds no systematic patterns in mental recall of previous responses, in the extent of rounding, or in other nonclassical measurement error.

Exploiting the availability of multiple expectations reports per respondent for a subset of events and respondents within a narrow time window, Boneva and Rauh (2021) perform test-retest checks. The authors find that the test-retest correlations are high and do not vary by socioeconomic background. They further validate their questions by comparing respondents' subjective probabilities of enrolling in university with actual application decisions among students in their final year of high school and by assessing the reliability of responses to questions that are similar but reverse-coded.

Stinebrickner and Stinebrickner (2012) show that subjective expectations data collected in the BPS outperform standard assumptions on expectations (e.g., rational expectations) in satisfying a basic theoretical implication of their model; that is, at the end of the first year of college, both past grades (a stock variable) and expectations about future grades (a flow variable) should be related to college drop out. Analyses of this type can be highly valuable for assessing the credibility of subjective expectations data vis-à-vis traditional economic assumptions.[58]

*Academic performance, study effort, and ability.* Zafar (2011b) documents that NU undergraduates are on average overoptimistic about their academic performance in college, measured as their subjective probability of graduating with a GPA of at least 3.5. For example, the mean probability varies from a high 83 percent for Engineering and Math & Computer Studies to a high 0.95 for Social Sciences I.[59] Using the same data, Zafar (2012, 2013) show that, even though students are largely aware of the relative difficulty of the various majors,[60] their expectations about academic performance vary widely by respondent characteristics and major. For example, Zafar (2013) documents systematic differences across genders. He finds that about 60% percent of male respondents reports a probability of graduating with a GPA of at least 3.5 in Engineering greater than 50 percent, whereas only 30% female respondents do so. On the other hand, the fractions of male and female students reporting a probability higher than 50 percent for Literature and Fine Arts are much closer to each other: 95% and 90% respectively. Zafar (2012) documents differences between students pursuing a single versus double major. He finds that

_____

students are too optimistic at college entrance about their college prospects but revise downward their beliefs as they learn about their ability (Stinebrickner and Stinebrickner, 2012, 2014a,b).

[56] In the case of binary outcomes the regression test is equivalent to some of the more sophisticated tests recently developed in the literature (e.g., D'Haultfoeuille et al. (2021)).

[57] This is the concern that respondents may report beliefs that rationalize their choices.

[58] I thank Todd Stinebrickner for pointing this out.

[59] Including Anthropology, Gender Studies, History, Linguistics, Political Science, Psychology, and Sociology.

[60] For example, less than 40% of respondents believes that there is a greater than 60 percent chance of graduating with a GPA of at least 3.5 in Engineering, whereas more than 80% of the respondents believes that to be the case for Literature and Fine Arts, consistently with available NU-wide statistics.

double-major students report a higher probability of graduating with a GPA of at least 3.5 and lower expected hours per week spent on coursework than single-major students for most majors. This is consistent with a positive selection of double-major students, who indeed tend to have a higher GPA than single-major students.

Consistently with Fischhoff et al. (2000)'s and Cowan (2017)'s evidence from the NLSY97 and with Zafar (2011b, 2012, 2013)'s evidence on NU undergraduates, Stinebrickner and Stinebrickner (2012) find that also Berea students are too optimistic about grade performance, especially at college entrance. The authors further show that overoptimism is concentrated among students with a high school GPA in the bottom third of the sample (students with a high school GPA in the top third have instead accurate beliefs) and among male students (female students are also overoptimistic but less so). While Berea students are also slightly overoptimistic about study effort, the authors show that optimism about academic performance is mostly driven by overoptimism about ability (about 70%). In terms of gender differences, the authors find that while male and female students take courses of the same difficulty, on average male students expect to study half an hour less per day than female students (3.44 vs. 3.93) and actually do so (3.13 vs. 3.67). Because male and female students are similarly overoptimistic about how much they will study, the gender difference in overoptimism about ability is similar to the gender difference in overoptimism about grade performance.[61]

Similarly, Boneva, Golin and Rauh (2022) find that university students in the UK are also quite optimistic about their academic performance. On average, students report a perceived likelihood of obtaining first-class honors conditional on graduating of 56 percent, whereas only 28% of students who graduate from a full-time university program obtain first-class honors. Also consistently with earlier US evidence, the authors find that students in their final year of university and female students are significantly less overoptimistic about the chances of obtaining first-class honors than their first/second year and male counterparts.

Belfield et al. (2020) examine the subjective probability of obtaining in 11th grade the necessary grades to go to sixth form and of obtaining in sixth form the necessary grades to go to university, among 9th-graders aged 13-14 in the UK. Mean responses are 78 and 75 percent, respectively. The authors document substantial response variation across groups, especially by student academic performance, gender, and socioeconomic background. For example, they find a strong and positive association between students' subjective probabilities about their future academic performance and students' standardized test scores in grade 6 and household's income quartile. They also find that female students and students with better-educated parents report significantly higher probabilities of obtaining the necessary grades for sixth form and for university. Interestingly, controlling for students' subjective probability of enjoying going to sixth form/university, which the authors interpret as a measure of students' perceived "consumption value" of education,[62] significantly reduces the gender and socioeconomic differences in students' perceptions about the probability of receiving the necessary grades to pursue further education. The authors directly investigate the associations between students' expectations about their future academic performance and those about other outcomes. They find that students' beliefs about the consumption value of sixth form and of university are significant predictors of students' beliefs about the likelihood of obtaining the required grades for sixth form. Similarly, students' beliefs about the consumption value of university is a significant predictor of students' beliefs about the likelihood of obtaining the required grades for university. On the other hand, students' beliefs about the monetary returns to sixth form or university do not have predictive power on the perceived likelihood of obtaining the required grades for

---

[61]The authors show that gender differences in study amounts are related to gender differences in the disutility of studying: male students require, on average, 15% more compensation to study an additional hour keeping grades fixed ($8.47 vs. $7.40), with this difference being statistically significant.

[62]The authors define the consumption value of education as consisting of "*different non-pecuniary benefits and costs associated with being in full-time education, such as the (dis)utility from acquiring new skills, experiencing new things and places, socializing with new people, or participating in social events and student activities*" (Belfield et al. (2020), p. 490).

either sixth form or college.[63] The authors interpret this evidence as potentially suggesting that students exert more effort in school to achieve higher grades if they believe that by getting the required grades they can pursue further education from which they derive a high consumption value.

In a sample of UK high school students spanning a wider age range (13-18), Boneva and Rauh (2021) confirm that relative to high-SES students low-SES students have, on average, a significantly lower subjective probability of obtaining the required grades to go to university (73 vs. 67 percent), of graduating from university if they enroll (85 vs. 82 percent), and of obtaining high grades if they graduate (71 vs. 65 percent).

*"Enjoying" education and other nonmonetary outcomes.* Belfield et al. (2020), Boneva and Rauh (2021), and Boneva, Golin and Rauh (2022) provide extensive evidence about the consumption value of education perceived by high school and university students in the UK. Gong et al. (2021) provide evidence in the BPS.

Belfield et al. (2020) measure students' perceived consumption value of sixth form (university) by asking high school students aged 13-14 to report their subjective probability that they would enjoy sixth form (university) if they were to enroll. Boneva and Rauh (2021) ask the same question with reference to university in a sample of high school students aged 13-18. And Boneva, Golin and Rauh (2022) ask the same question with reference to post-graduate education in a sample of university students. The latter two studies expand the battery of questions capturing aspects of the consumption value of education, for example, by additionally asking students their subjective probabilities of enjoying studying, finding the study material interesting, finding the material/work tasks hard, feeling stressed, enjoying social life, meeting people with whom to easily get along, feeling lonely, losing contact with family/friends.[64]

The average response in Belfield et al. (2020)'s sample is 77 percent for the first question (sixth form) and 73 percent for the second (university), but with substantial response heterogeneity across respondents as captured by a cross-sectional standard deviation of 20 and 22 percent, respectively. The authors show that students' beliefs about the consumption value of education vary systematically by gender and socioeconomic background, with female students reporting higher enjoyment probabilities by 0.25 standard deviations for both sixth form and university, and high-SES students reporting higher enjoyment probabilities by 0.15 standard deviations for sixth form and 0.29 standard deviations for university. The authors further show that the correlations between the perceived consumption values of education and the perceived monetary returns to education are low (0.07 for sixth form and 0.15 for university).[65]

Boneva and Rauh (2021) find that students from different SES backgrounds share the beliefs that they would be more likely to enjoy the content, find the material interesting, find the material hard, enjoy social life, meet people they easily get along with, lose contact with their family/friends, and feel stressed, and that they would be less likely to feel lonely, if they were to go to university instead of work. However, the magnitudes of these differences vary significantly between SES groups. In particular, low-SES students have smaller perceived (positive) returns to university in terms of their subjective likelihoods of enjoying the content, finding the material interesting, enjoying social life, and meeting people they easily get along with, and larger perceived (negative) returns to university in terms of their likelihoods of feeling lonely and losing contact with their family/friends.

---

[63]These findings are consistent with those of Baker et al. (2018), who find that beliefs about salary are uncorrelated with beliefs about ability.

[64]These questions were partially inspired by Belfield et al. (2020)'s finding that students' primary stated reasons for going to university are "to experience new things and places" (41% among high-SES and 48% among low-SES students), "to enjoy the social life" (25% among high-SES and 21% among low-SES students), and "to enjoy education" (15% among high-SES and 13% among low-SES students).

[65]These findings are consistent with those of Baker et al. (2018), who find that salary beliefs are uncorrelated with enjoyment beliefs, but in contrast with those of Wiswall and Zafar (2015*b*), who find that salary expectations are correlated with taste for majors.

Similarly, Boneva, Golin and Rauh (2022) report broadly similar findings among university students with respect to the possibility of continuing to post-graduate education vis-à-vis starting to work.

Gong et al. (2021) use the Euler equation, expectations data, and other measures collected in the BPS to assign an overall monetary value to the consumption value of college attendance. They estimate an average annual consumption value of college as high as $15K with considerable heterogeneity across students.

*Nonmonetary outcomes in the labor and marriage markets.* Zafar (2012, 2013) analyze NU students' beliefs about the prospects of being able to reconcile family and work at the available jobs following graduation from different majors and find that these beliefs vary widely across both majors and respondent characteristics. For example, Zafar (2012) shows that the belief distribution for Natural Sciences is first order stochastically dominated by that for Literature and Fine Arts, consistent with the general perception that the typical jobs in the pure science and medical profession entered by the NU graduates with a major in Natural Sciences have hectic work schedules.

Using the same data, Zafar (2013) finds that the belief distributions of male and female students are similar for Social Sciences II,[66] but dramatically different for Engineering. In the latter case, the belief distribution of male students first order stochastically dominates the belief distribution of female students. For example, about 50% of male respondents believes that the probability of being able to reconcile family and work in jobs following a major in Engineering is greater than 75 percent, while less than 20% percent of female respondents expect that to be the case.

Boneva, Golin and Rauh (2022) analyze students' beliefs about the likelihoods of being satisfied with their career, having a high status in society, being able to contribute to society, and having a good work-life balance, if they obtain a post-graduate qualification versus not. The authors find that on average students report higher likelihoods for these outcomes in the post-graduate scenario and find no statistically significant differences in average perceived returns across SES groups.

Wiswall and Zafar (2020) investigate the perceived marriage returns to human capital among NYU undergraduates and document that students perceived large returns to human capital in terms of marriage prospects, potential spouse's earnings, and fertility. They also find differences in perceived returns and penalties between genders. Specifically, male and female students share the belief that the marriage return to completing a college degree is positive, but the average perceived returns differ in magnitude between the two groups: with female students believing that their likelihood of being married at age 30 would be nearly 13 percent higher on average if they completed a college degree and men believing that their likelihood would be over 35 percent higher. The authors document other significant differences across genders. For example, female students believe that there is a marriage market penalty of about 15 points on the percent chance scale to completing a degree in Science or Business relative to one in the Humanities or Social Sciences. In contrast, male students' expectations imply no effect of college major on their probability of getting married.

The authors additionally elicit students' beliefs about the earnings of their potential spouse, if the latter were to attain different human capital levels, and about fertility prospects. They find that students perceive a large spousal return to completing high-earning degrees, indicating that they believe that investing in human capital will not only lead to higher earnings for themselves but also to a match with a higher-earning spouse. With respect to fertility, students expect that by age 30 they would have less children following a major in Science or Business than one in the Humanities. However, these differences in expected fertility by major are halved when the question is asked with reference to age 45, suggesting that students anticipate that human capital investments will have an effect on the timing of fertility, in addition to having an effect on its level.

---

[66]Including Economics and Mathematical Methods in the Social Sciences.

Boneva and Rauh (2021) document that students from all SES groups share the belief that they are more likely to meet their life partner if they go to university than if they start working, but the perceived marriage premium from attending university is significantly larger among high-SES students on average.

*Attainment and dropout.*  Reynolds and Pemberton (2001) study the rise in US youth's college expectations between the late 1970s and the late 1990s by comparing the expectations reports about the prospects of receiving a college degree given by 15- and 16-year-old respondents in the NLSY97 with the earlier ones given by similarly-aged respondents in the NLSY79.[67]  The authors show that in 1979 less than 40% of respondents expected to achieve 16 or more years of education. By 1997, over 70% of respondents reported they had more than a 50 percent chance of completing college by age 30 and over 50% believed their chances were greater than 75 percent. The authors examine and compare the relative influence of gender, race/ethnicity, family characteristics, and local labor market conditions on college expectations in 1979 and 1997. They find that in 1979 college expectations vary by race/ethnicity, household type, family size, parental education and income, and the county economic and educational context, but not by gender. Specifically, racial/ethnic minorities have significantly higher college attainment expectations than non-Hispanic Whites, *ceteris paribus*. College attainment expectations are also significantly positively associated to the parental education and income, to the percent of the county population with a college degree, and the county's unemployment rate.[68] On the other hand, youth in two-parent families where only one parent is related to the youth through birth or adoption have significantly lower college attainment expectations than youth living with both biological parents. Similarly, youth with more siblings have lower college expectations. By 1997, the associations with race/ethnicity, family resources, and county context become weaker and those with family size remain unchanged relative to 1979. The most noticeable differences are the growing gaps between male and female youths and among family arrangements. In 1997, high school girls have significantly higher college attainment expectations than their male peers and, according to enrollment data, are significantly more likely to attend.  Similarly, high school youth living with both biological parents have significantly higher college expectations than youth with any other type of family arrangement.  For 1997, the authors further investigate the role of academic performance, peer group characteristics, and students' reports of the school environment.[69] They find that grades received in eighth grade have a strong impact on expectations: a one-letter grade increase is associated with a 10-point increase in the subjective probability of obtaining a college degree. Perceptions of teachers matter as well: youth that indicate having good teachers who are interested in their students hold a 4-to-5-point higher college attainment probability than other students, on average. Some peer group behaviors matter, too: students reporting that more than 50% of their peers plan to go to college have a 9-point higher mean subjective probability of completing college than other students.

Stinebrickner and Stinebrickner (2012) show that Berea students' overoptimism about GPA and ability translates into overoptimism about completion.  When asked directly, BPS respondents believe on average that they have an 86 percent chance of graduating, whereas approximately 60% of entering students eventually graduate from Berea.  Using the same data, Stinebrickner and Stinebrickner (2014a) provide additional evidence that students are substantially overoptimistic about graduation at entrance. BPS respondents assign a probability of about 13 percent, on average, to the possibility that they will drop out of college and only 5% of them believe that dropping out is their most likely outcome. In reality, over 35% of Berea students drops out.

Wiswall and Zafar (2015b) document similar optimism among NYU undergraduates.

---

[67]A caveat with the comparison is that the NLSY97 elicits college expectations probabilistically, whereas the NLSY79 asks for a point prediction.

[68]The authors interpret this as being due to a "warehouse effect", whereby youth postpone entry into the labor market until job conditions improve.

[69]This information is not available for 1979.

While NYU dropout rates are significantly lower than those of most US universities,[70] Wiswall and Zafar (2015b)'s respondents give an average subjective probability of not graduating lower than 3 percent.

In addition to eliciting university students' subjective probability of enrolling in a post-graduate degree if they obtain the required grades, Boneva, Golin and Rauh (2022) elicit students' subjective probabilities of completing their undergraduate degree and of obtaining a post-graduate degree if they enroll. On average, students believe that they have a 89 percent chance of completing university. This estimate does not vary significantly by students socioeconomic background and compares favorably with UK statistics showing that the percentage of full-time first-degree students who do not obtain a degree ranges between 10.1 and 10.7. On the other hand, low-SES students report significantly lower probability of graduating from a post-graduate degree than high-SES students, with the difference being about 6 percentage points on the 0-100 scale.

Kunz and Staub (2020) provide evidence on Germany using SOEP data. Focusing on respondents aged 16-17, the authors show that most students with a high subjective probability of completing a post-graduate degree are overly optimistic about their graduation prospects in the medium run. Nevertheless, the information content of students' subjective probabilities is comparable to that of predictions obtained using standard econometric models with perfect foresight. The authors investigate the predictors of students' expectations and compare their relative importance for expectations and actual post-graduate enrollment and completion. They find that the most important predictors of students' completion expectations are their academic ability and personality traits, whereas the most important predictors of actual enrollment and completion are measures of family background (parental education and household income) and local labor market conditions (youth unemployment and education demand and supply). The authors show that this difference holds within as well as between groups characterized by different observables.

*Education plans.*  Stinebrickner and Stinebrickner (2014a) document that at college entrance BPS students are highly uncertain about their college major. For example, on average, students assign a probability of only 31 percent to choosing the major that they are actually observed to choose in the administrative data. The authors show that this partly occurs because initially students are open to the possibility of majoring in Science, with about 20% of them indicating Science as their most likely major, but relatively few students (about 7%) end up majoring in Science. The three main reasons why the actual number of Berea college graduates majoring in Science is much smaller than the number that would be expected given students' initial beliefs are directly tied to some of the evidence previously discussed. First, students are overoptimistic about their probability of graduating. Second, students who initially indicate science as their most likely major are too optimistic about their probability of remaining in Science. Third, students who initially indicate a major other than Science as their most likely major are too optimistic about the probability that they will change it to Science.

Wiswall and Zafar (2015b), too, investigate students' expectations about their major. In the case of NYU students, Humanities/Arts is the major with the highest average probability of being selected (42.6 percent), followed by Economics/Business (30 percent). The authors show that for most university fields the cross-sectional distribution of students' choice probabilities is bimodal, due to a considerable mass of students reporting a zero-to-small probability of majoring in each field and another mass reporting a large-to-one probability of doing so. Using the same data, Patnaik, Wiswall and Zafar (2020) find that, on average, female students report a significantly higher probability of majoring in Humanities than their male peers (49.9 vs. 31.4 percent) and a significantly lower probability of majoring in Economics/Business (23.8 vs. 36.7 percent) or Engineering (5.6 vs. 8.7 percent).[71]

---

[70]For example, full-time first-time students in the NYU class of 2015 had a four-year graduation rate of 75% and a six-year graduation rate of 83%.

[71]These figures refer to students' baseline beliefs. The effect of Wiswall and Zafar (2015b,a)'s information

In the UK context, Belfield et al. (2020) document that high school students aged 13-14 have an average probability of continuing to sixth form (conditional on obtaining the required grades) of 85 percent and an average probability of going to university (again conditional on obtaining the required grades) of 73 percent. Once combined with students' subjective probabilities of obtaining the necessary grades, these figures imply an average unconditional probability of leaving education after year 11 without going to sixth grade of 30 percent, an average unconditional probability of going to sixth form but not to university of 24 percent, and an average unconditional probability of going to university of 46 percent. These figures compare favorably to UK enrollment statistics, whereby about 82% of students aged 16–18 are enrolled in full-time education while 48% of students in any given cohort continues to university. Using data from the British Household Panel Study (BHPS) and the UK Longitudinal Household Study (UKLHS), the authors show that students' enrollment probabilities correlate positively and significantly with actual enrollment. Additionally, they find substantial heterogeneity in beliefs by gender and socioeconomic status. Female students report a 5.9 percentage points higher probability of going to sixth form and a 7.4 percentage points higher probability of going to university than their male peers. Students who have at least one parent with a university degree report a 3.9 percentage points higher probability of going to sixth form and a 9.6 percentage points higher probability of going to university than other students. And students in the top income quartile report a 8.7 percentage points higher probability of going to sixth form and a 10.8 percentage points higher probability of going to university than students in the bottom income quartile.

Among older UK students, Boneva, Golin and Rauh (2022) show that university students give an average probability of enrolling in post-graduate education of 49 percent, but with substantial response heterogeneity both within and between SES groups. The mean subjective probability of enrollment is 47 percent among first-generation students whose parents never went to university, and 52 percent among continuing-generation students with at least one parent who went to university. Among the latter, students who have at least one parent with a postgraduate degree report an 8 percentage points higher enrollment probability than the other students.

*Parental approval and parental beliefs.* The evidence on parental beliefs and roles in educational choices, especially those taken during students' adolescence years, is scant. However, some of the most recent expectations studies have elicited students' beliefs about the chances that parents would approve or support their choice. For example, Zafar (2012) analyzes students' responses within the context of college major choice and finds that the mean subjective probability of gaining parental approval varies from a low of 59 percent for Literature and Fine Arts to a high of 87 percent for Natural Sciences. Noticing that students associate Natural Sciences with a higher social status (mean value of 0.72 on a 0–0.9 scale vs. 0.38 for Literature and Fine Arts) and higher returns in the labor market (mean expected earnings at age 30 of $90K vs. $50K for Literature and Fine Arts), the author interprets the evidence as suggestive that students believe parents to be more likely to approve of majors associated with higher status and returns in the labor market.

In the context of post-graduate enrollment choice in the UK, Boneva, Golin and Rauh (2022) find that on average students' perceived likelihood of receiving nonmonetary parental support in their choice is similar across the two choice alternatives: 79 percent for stopping with an undergraduate degree and 78.9 percent for pursuing a post-graduate degree. However, the authors find significant differences across SES groups. In particular, while high-SES students report higher subjective probabilities of having parental support if they continue to post-graduate education than otherwise (81.5 vs. 79.5 percent), the opposite pattern is observed among low-SES students (76.4 vs. 78.5 percent). The authors show that the greater subjective likelihood of parental approval attributed to the post-graduate option by the high-SES students is driven by the subgroup of students whose parents have a postgraduate qualification.

treatment on students' beliefs is discussed in later sections.

Only a handful of studies to date has separately elicited percent-chance expectations from students and their parents and have compared the two sets of expectations. For example, Fischhoff et al. (2000) and Dominitz et al. (2001) do so using NLSY97 data. Fischhoff et al. (2000) document that parent and teen response distributions are quite similar in terms of means, medians, standard deviations, and use of 50 percent. Hence, NLSY97 parents share their teens' generally high expectations for educational attainment, with no significant child-parent difference in the mean probabilities of being in school in a year and of having a diploma by age 20. At the family level, the authors document high correlations between the judgments of paired teens and parents, declining slightly as the time period increases: 0.71 for being in school in one year, 0.65 for having a diploma by 20, and 0.51 for having a college degree by 30.

Dominitz et al. (2001) further document that most youths and parents see no chance at all of an adverse schooling outcome, with the parents being more optimistic than the students on average: 84.2% of parents and 78.7% of youths report a 100 percent chance that the student will receive a high school diploma by age 20. They also find that, while the cross-sectional median of the within-family differences between parent and child reports is uniformly 0, the mean difference varies across groups. For example, the latter is significantly lower for female youth than for male youth. In families where the youth report unfavorable past experiences such as school suspension, grade repetition, sexual intercourse, and cigarette or marijuana smoking, both youth and parents report lower subjective probabilities that the youth will receive a diploma by age 20, irrespective of the youth's gender.

In the context of high school track choice in Italy, Giustinelli (2016) similarly finds that students and parents hold broadly similar beliefs about the chances of a variety of choice consequences or outcomes at the time of choice. However, while the cross-sectional mean and median of the within-family differences in beliefs are generally very small, especially for outcomes such as school enjoyment, the standard deviation and extreme quantiles indicate the existence of a fair amount of variation in child-parent differences across families. Based on a Wilcoxon matched-pairs signed-ranks test, the null hypothesis of expectations equality between the child-parent matched pairs is rejected for the majority of choice alternatives and outcomes. Moreover, some outcome-specific patterns are found. For example, consistently with Dominitz et al. (2001)'s evidence from the NLSY97, also Italian parents appear to be more optimistic about the students's performance than students are. Moreover, parents expect on average smaller differences in the student's study effort and academic performance across high school tracks than students do.

## 4 Analysis of Schooling Decisions With Survey Expectations

### 4.1 Choice Framework

#### 4.1.1 Basics

The choice problem of Manski (1993)'s adolescent econometricians reviewed in Subsection 2.1 can be reframed as a problem of decision-making with an uncertain state.[72] Let $\mathcal{J}$ denote the choice set, $\Theta$ the space of feasible states of nature, $\theta^* \in \Theta$ the true state, and $f(\cdot, \cdot) : \mathcal{J} \to \mathbb{R}$ the decision maker's objective function. The decision maker wants to select the alternative that solves

$$\max_{j \in \mathcal{J}} f(j, \theta^*). \tag{1}$$

To solve problem (1) the decision maker must find a way to deal with not knowing $\theta^*$. A standard behavioral assumption in economics is that the decision maker places a subjective probability distribution over the feasible values of the state and chooses the

---

[72]This part follows Manski (2004), but with the notation modified to maintain consistency throughout.

alternative that maximizes his or her subjective expected utility (SEU),

$$\max_{j \in \mathcal{J}} \int u\left[f(j, \theta^*)\right] dP, \tag{2}$$

where $P$ denotes the decision maker's subjective distribution on $\Theta$ and $u(\cdot) : \mathbb{R} \to \mathbb{R}$ is an increasing transformation of $f$. The assumption appeals to Savage (1954)'s axiomatic derivation of subjective expected utility. It implies that the decision maker is subjectively rational, in the sense that he or she makes SEU-maximizing decisions; but it places no restrictions on the decision maker's expectations, $P$. In particular, SEU maximization does not require or imply rationality of expectations. A decision maker holding rational expectations about the uncertain state, $\theta^*$, would actually know the objective probability distribution, denoted by $\Pi$, that generates $\theta^*$, where the latter is now viewed as the realization of a random variable.[73]

In the motivating framework of Subsection 2.1, the unknown state of nature relevant to the youth's school-work decisions is represented by the returns to schooling. Because the youth do not necessarily know $\Pi$, they need to learn its features from the observed choices and outcomes of the prior generation. To do so, the adolescent econometricians must combine available information (data) with assumptions on the problem's unobservables. In particular, when forming expectations, they must make assumptions about the outcomes that members of the prior generations would have experienced had they made alternative (counterfactual) choices.

Instead of making unverifiable assumptions on how youths and their families form expectations relevant to schooling decisions, in the past fifteen years an increasing number of studies has employed the elicitation procedures described earlier to directly measure decision makers' subjective expectations over choice-relevant uncertainties and to estimate random utility models of educational choices using elicited expectations. Most studies have maintained the assumption that decision makers choose among schooling alternatives by maximizing their SEU. The SEU has been typically specified as a function of the decision maker's subjective expectations over the uncertain consequences of choice as elicited in the survey,[74] and of preference parameters to be estimated. Direct observation of decision makers' expectations along with their choices has opened the possibility of using elicited expectations on the right-hand side of the model to aid identification and estimation of the structural parameters characterizing decision makers' preferences, thus dispensing with the strong assumptions on expectations traditionally maintained for identification.

The uncertainty faced by the adolescent econometricians populating the simple framework of Subsection 2.1 is limited to the returns to schooling. But when making schooling decisions in real life, individuals may be uncertain about multiple aspects of choice. Hence, $\theta^*$ may be better viewed as a vector. Studies that have specified random utility models of educational choices with uncertain consequences and have estimated them using elicited expectations have recognized this feature of the problem and have addressed it by allowing the decision maker's SEU to depend on multiple monetary and nonmonetary outcomes. To do so, they have collected data on respondents' subjective expectations over rich sets of monetary and nonmonetary outcomes of schooling within the *same* survey.

To choose among feasible alternatives a decision maker must form subjective expectations over the consequences of *all* alternatives under consideration for choice. These include the selected alternative as well as those that were *ex ante* feasible but were not chosen *ex post*. A practical implication is that studies employing subjective expectations data for estimation of choice models must elicit from survey respondents their choice-conditioned expectations over consequences of *all* choice alternatives in the person's choice set.

---

[73]SEU maximization does require, however, that decision makers can assign precise (point) probabilities to the consequences of each action.

[74]Corresponding to the distribution $P$ over $\Theta$ in problem (2).

*Model specification.* Decision maker $i$ faces a choice among a finite set of schooling alternatives, $\mathcal{J}$.[75] These may be college majors as in Arcidiacono et al. (2012) and Zafar (2013), high school tracks as in Giustinelli (2016), or other educational options.[76]

The choice is static and once and for all.[77] Decision makers are forward-looking, so their choice depends on what they expect to result from it in the future. Individual $i$ derives utility $U_i(\vec{\theta})$, where $\vec{\theta} = \{\theta_n\}_{n=1}^N$ is a finite vector of outcomes, some of which are realized during school and others after $i$ graduates from school (or drops out). Because the realizations of the components of $\vec{\theta}$ are uncertain at the time of choice, in order to make a decision individual $i$ forms subjective probabilities, $\{P_{ij}(\vec{\theta})\}_{j \in \mathcal{J}}$, over the consequences of choosing each alternative $j \in \mathcal{J}$. Then, $i$ chooses his or her SEU-maximizing alternative,[78]

$$j_i^* = \arg \max_{j \in \mathcal{J}} \int U_i(\{\theta_n\}_{n=1}^N) dP_{ij}(\{\theta_n\}_{n=1}^N). \tag{3}$$

Both Zafar (2013) and Giustinelli (2016) specify (3) as additively separable with respect to the elements of $\vec{\theta}$ and, for each element of $\vec{\theta}$, multiplicatively separable in probabilities and utilities. Under these assumptions, and denoting binary outcomes with $\{b_n\}_{n=1}^{N_B}$ and continuous outcomes with $\{c_n\}_{n=1}^{N_C}$, the decision problem (3) can be written as

$$j_i^* = \arg \max_{j \in \mathcal{J}} \sum_{n=1}^{N_B} \{P_{ij}(b_n = 1) \cdot u_i(b_n = 1) + [1 - P_{ij}(b_n = 1)] \cdot u_i(b_n = 0)\} +$$

$$\sum_{n=1}^{N_C} \gamma_{in} \cdot E_{ij}(c_n) \equiv \arg \max_{j \in \mathcal{J}} \sum_{n=1}^{N_B} P_{ijn} \cdot \Delta u_{in} + \sum_{n=1}^{N_B} u_i(b_n = 0) + \sum_{n=1}^{N_C} \gamma_{in} \cdot E_{ijn}, \quad (4)$$

where $P_{ijn} \equiv P_{ij}(b_n = 1)$ denotes decision maker $i$'s subjective probability that outcome $b_n = 1$ will result if alternative $j$ is chosen; $\Delta u_{in} = u_i(b_n = 1) - u_i(b_n = 0)$ represents the difference in utility that $i$ derives from $b_n = 1$ relative to $b_n = 0$ following any choice; $E_{ijn} \equiv E_{ij}(c_n)$ denotes $i$'s subjective expectation for continuous outcome $c_n$ following choice of $j$; $\gamma_{in}$ represents the associated utility. Being constant across choice alternatives, $\sum_{n=1}^N u_i(b_n = 0)$ drops out of the choice.[79]

The choice problem in (4) is formulated directly in terms of those utilities and subjective marginal probabilities the decision maker uses to choose among the available choice alternatives, thus dispensing with the whole Savage (1954)-type state space. In addition to being cognitively and descriptively plausible (e.g., see Gilboa and Schmeidler (2004)), this assumption simplifies elicitation.[80] As a downside, it rules out any perceived interdependencies among subjective probabilities and/or utilities of different outcomes.[81]

---

[75]In principle, the number and composition of choice alternatives may vary across decision makers due to heterogeneity in constraints and/or consideration. In practice, the vast majority of studies of schooling decisions has assumed that decision makers choose from a common set of available alternatives, i.e., $\mathcal{J}_i = \mathcal{J} \; \forall i$. Giustinelli and Manski (2018) consider extensions with heterogeneous choice sets.

[76]This part follows closely Zafar (2013)'s model and the unitary-family version of Giustinelli (2016)'s model. It also borrows from Arcidiacono et al. (2012). See Koşar and O'Dea (2022) for a more extensive treatment of microstructural models with subjective expectations data.

[77]In Zafar (2013)'s analysis the assumption is justified by the focus on sophomore students, in Giustinelli (2016)'s setting by the fact that track switching is rare.

[78]Indifference between alternatives is assumed to occur with probability zero.

[79]An implication of the linear SEU specification is that decision makers are assumed to be risk neutral with respect to continuous outcomes, notably expected earnings. Zafar (2013) and Giustinelli (2016) maintain this assumption for survey tractability, since relaxing it would require elicitation of multiple points on the respondent's belief distribution of own future earnings for each choice alternative. Arcidiacono et al. (2012) let expected earnings enter the student's SEU in logarithms, implying a specific coefficient of risk aversion. Patnaik, Wiswall and Zafar (2020)'s model allows for both heterogeneous risk and time preferences.

[80]Additive separability rules out interactions between outcomes assumed separable, although in principle individual elements of $\vec{\theta}$ could be defined as joint events. Use of additive separability is partly motivated by tractability in data collection, since measuring subjective probabilities of separable binary outcomes only requires elicitation of one marginal probability per outcome and choice alternative, instead of joint probabilities.

[81]Although a standard feature of canonical SEU (Savage, 1954), multiplicative separability rules out the

[Zafar](2013) and [Giustinelli](2016) specify the SEU as a function of 11 outcomes, encompassing monetary and nonmonetary consequences of college major choice and high school track choice, respectively.[82]

*Econometric implementation.*    A common econometric formulation for (4) is

$$d_{ij^*} = \mathbf{1} \left\{ \sum_{n=1}^{N_B} P_{ij^*n} \cdot \Delta u_n(X_i) + \sum_{n=1}^{N_C} \gamma_n(X_i) \cdot E_{ij^*n} + \varepsilon_{ij^*} > \right.$$
$$\left. \sum_{n=1}^{N_B} P_{ijn} \cdot \Delta u_n(X_i) + \sum_{n=1}^{N_C} \gamma_n(X_i) \cdot E_{ijn} + \varepsilon_{ij} \quad \forall j \neq j^* \right\}, \quad (5)$$

where $\mathbf{1}\{\cdot\}$ denotes the indicator function and where preference heterogeneity is now expressed by letting the utility parameters depend on a vector of observed characteristics of $i$, $X_i$, and by introducing an additive term, $\varepsilon_{ij}$, capturing the components of $i$'s SEU that are unobserved to the econometrician.[83]

Under the standard revealed preference argument that each person's choice coincides with his or her SEU-maximizing alternative, observing a cross-section of decision makers' choices, characteristics, and subjective expectations enables identification of the unknown utility parameters, given assumptions on the distribution of unobservables. On the right-hand side of (5), subjective expectations and individual characteristics are directly elicited in [Zafar](2013)'s and [Giustinelli](2016)'s surveys. On the left-hand side of (5), [Giustinelli](2016) observes actual enrollment choices, whereas [Zafar](2013) elicits students' intended major as well as their preference ranking over feasible majors.[84] Both [Zafar](2013) and [Giustinelli](2016) assume i.i.d. type-I extreme value unobservables, leading to closed-form conditional choice probabilities of the logit type.[85]

*Modelling with subjective choice probabilities.*    In addition to using survey-elicited expectations over the consequences of alternative choices on the right-hand side of the model, an increasing number of papers has employed data on subjective choice probabilities, in place of or combined with actual choices, on the left-hand side of the model. Examples

---

possibility that a person's subjective probability of an event depends on his/her (dis)utility for the event, as in models of utility-based or motivated beliefs (see [Benabou and Tirole](2016)'s review). A simple way to partially relax this assumption – and also that of outcome separability – would be to allow the utility of an outcome, say $\hat{n}$, to depend on the decision maker's subjective probability of another outcome, say $\check{n}$, hypothesized to be related to the first. This would be equivalent to introduce heterogeneity in the utility of outcome $\hat{n}$ with respect to the expectation for outcome $\check{n}$.

[82] E.g., the outcomes entering [Zafar](2013)'s SEU specification are: 1. successfully completing a field of study in 4 years; 2. graduating with a GPA of at least 3.5 in the field of study; 3. enjoying coursework; 4. hours per week spent on coursework; 5. parents approve of the major; 6. obtain an acceptable job immediately upon graduation; 7. enjoy working at the jobs available after graduation; 8. be able to reconcile work and family at the available jobs; 9. hours per week spent working at the available jobs; 10. social status of the available jobs; 11. income at the available jobs. Outcomes 1-3, 5, and 6-8 are treated as binary, 4 and 9-11 as continuous.

[83] This has been standard practice in econometric analysis of revealed preferences since [McFadden](1973). To accommodate the fact that individuals with identical observed characteristics and choice environments are sometimes observed to make different choices, an individual-and-choice-specific random term is added to the utility, from which the expression *random utility model*.

[84] [Arcidiacono et al.](2012) observe actual majors for students who already declared their majors and elicit intended majors from the other students. Choice intentions, stated choices, and choice rankings are all forms of stated preference (SP) data, usually collected in experiments or surveys. The term *stated preferences* refers to the fact that the respondents state what choices they would make in the hypothetical situations, or predict what choices they will make in the future. It is used in contrast to the term *revealed preferences* (RP), which relate to people's actual choices in real-world situations. For introductory treatments of discrete choice modelling with SP data, alone or combined with RP data, see [Ben-Akiva et al.](1994), [Train](2009), and references therein.

[85] Other distributional assumptions are of course possible. Different distributional assumptions on the unobservables imply more or less restrictive substitution patterns among choice alternatives and can have implications for identification and estimation. See [Train](2009) for a textbook treatment.

include Stinebrickner and Stinebrickner (2014a), Wiswall and Zafar (2015b), and Wiswall and Zafar (2020) on college major choice in the US, Schweri and Hartog (2017) on nursing college enrollment in Switzerland, and Belfield et al. (2020), Boneva and Rauh (2021), Boneva, Golin and Rauh (2022) on sixth form, college, and postgraduate enrollment in the UK.

Survey-elicited choice probabilities may be thought of as a probabilistic form of stated preference (SP) data. So, in general, they feature similar strengths and weaknesses as SP data vis-à-vis RP data.[86] However, subjective choice probabilities can have distinctive advantages over more traditional forms of SP data.[87]

Manski (1999) introduces a formal framework for interpretation and analysis of choice expectations. Formally, person $i$'s subjective probability at the time of survey, of choosing option $j^*$ at some specified future time or in a specified hypothetical scenario, can be expressed as

$$q_{ij^*} = Q_i \left[ \sum_{n=1}^{N_B} P_{ij^*n} \cdot \Delta u_{in} + \sum_{n=1}^{N_C} \gamma_{in} \cdot E_{ij^*n} + \epsilon_{ij^*} > \right.$$
$$\left. \sum_{n=1}^{N_B} P_{ijn} \cdot \Delta u_{in} + \sum_{n=1}^{N_C} \gamma_{in} \cdot E_{ijn} + \epsilon_{ij} \quad \forall j \neq j^* \right], \quad (6)$$

where for simplicity I dispense from introducing notation for time.

The right-hand side of equation (6) provides a subjective random utility interpretation of elicited choice probabilities. It says that person $i$ holds a subjective probability, $q_{ij^*}$, that choosing option $j^*$ will (would) be optimal at the time of choice (in the posed scenario). The term $\epsilon_{ij}$ may be thought of as a composite term, $\epsilon_{ij} = \vartheta_{ij} + \xi_{ij}$, where $\vartheta_{ij}$ captures factors that are unknown to the econometrician but known to the decision maker (like $\varepsilon_{ij}$ in (5)), while $\xi_{ij}$ captures factors that are unknown to *both* the decision maker and the econometrician. According to Manski (1999)'s taxonomy, $\xi_{ij}$ represents so-called *resolvable uncertainty*, including all those factors that are unknown to $i$ when he or she is asked to make predictions $\{q_{ij}\}_{j \in \mathcal{J}}$, but would be known to him or her in the actual choice situation. Thus, the subjective distribution $Q_i$, defined over $\{\xi_{ij}\}_{j \in \mathcal{J}}$, expresses person $i$'s resolvable uncertainty on his or her future (or hypothetical) optimal action. If at the time of prediction a respondent perceives no resolvable uncertainty about his or her optimal action in the future (or hypothetical) scenario, he or she can express this lack of uncertainty by assigning a corner probability of 100 percent to the optimal action and corner probabilities of 0 percent to the remaining actions.

A decision maker may also face *unresolvable uncertainty*. This type of uncertainty does not lead to interior choice probabilities, since in this case the decision maker would form expectations over the uncertain consequences of choice and condition on these expectations when making or planning a decision.

---

[86]As put by Train (2009) (pp. 152-153), "*Revealed-preference data have the advantage that they reflect actual choices. (...) However, such data are limited to the choice situations and attributes of alternatives that currently exist or have existed historically. (...) Even for choice situations that currently exist, there may be insufficient variation in relevant factors to allow estimation with revealed-preference data. (...) The advantage of stated-preference data is that the experiments can be designed to contain as much variation in each attribute as the researcher thinks is appropriate. (...) The limitations of stated-preference data are obvious: what people say they will do is often not the same as what they actually do.*"

[87]Traditional elicitation of stated choices has asked respondents to select from a given list of options the alternative that they prefer, or expect to choose, in the future or in some hypothetical scenario. Manski (1990) shows that responses to this kind of questions have limited information content. Intuitively, observing that a respondent selects a certain option only reveals that he or she assigns a subjective probability higher than a certain threshold to the event of choosing that option, where the threshold depends on the loss function used by the respondent in forming his or her prediction. Unfortunately, the respondent's subjective probability, threshold, and loss function are typically unobserved. Thus, Manski (1990) recommends directly eliciting respondents' subjective probabilities over choice alternatives. See Juster (1966), Stinebrickner and Stinebrickner (2014a), and Manski (2016) for additional theoretical arguments, empirical evidence, and in-depth discussions about the greater informativeness of subjective choice probabilities over stated intention data.

To implement (6) econometrically and obtain estimates of the model parameters, one needs assumptions on the components of $\{\xi_{ij}\}_{j \in \mathcal{J}}$. For instance, if one assumes that these components are i.i.d. Type 1 Extreme Value according to *both* the econometrician and the decision maker, the choice probabilities, $\{q_{ij}\}_{j \in \mathcal{J}}$, have the familiar exponential form,

$$q_{ij} = \frac{e^{\sum_{n=1}^{N_B} P_{ijn} \cdot \Delta u_n(X_i) + \sum_{n=1}^{N_C} \gamma_n(X_i) \cdot E_{ijn} + \vartheta_{ij}}}{\sum_{h \in \mathcal{J}} e^{\sum_{n=1}^{N_B} P_{ihn} \cdot \Delta u_n(X_i) + \sum_{n=1}^{N_C} \gamma_n(X_i) \cdot E_{ihn} + \vartheta_{ih}}}, \tag{7}$$

and are therefore invertible by applying the natural logarithm to each side of (7).[88]

Application of the log-odds transformation to (7) yields a linear specification,

$$ln[q_{ij}/q_{i1}] \equiv ln[q_{ij}] - ln[q_{i1}] = \alpha_j + \sum_{n=1}^{N} \beta_n(X_i) \cdot (p_{ijn} - p_{i1n}) + v_{ij}, \tag{8}$$

where $j = 1$ is the reference alternative; the alternative-specific constant for the reference alternative has been normalized to 0 (i.e., $\alpha_1 = 0$); $\beta_n(X_i)$ denotes a generic element of the vector of utility parameters, $\vec{\beta}(X_i) \equiv \{\{\Delta u_n(X_i)\}_{n=1}^{N_B}, \{\gamma_n(X_i)\}_{n=1}^{N_C}\}$, to be estimated; $N \equiv N_B + N_C$; and $v_{ij} = (\vartheta_{ij} - \vartheta_{i1})$.

The parameters in (8) can be estimated via least squares, using data on respondents' subjective probabilities over choices on the left-hand side of the model and on respondents' subjective probabilities (expectations) over choice consequences on the right-hand side of the model.[89]

Once the model has been estimated, it can be used to shed light on the determinants of observed choices (or choice probabilities) and to address policy-relevant questions. For example, model estimates off survey expectations have been often used to assess the relative importance for choice of different outcomes, to quantify the relative contribution of expectations vis-à-vis utilities, and to perform post-estimation analyses such as group decompositions, policy counterfactuals, and welfare analyses. Concrete examples are provided in later subsections reporting on the empirical evidence.

### 4.1.2 Extensions

The basic framework described in Subsection 4.1.1 relies on a number of simplifying assumptions and abstracts from a number of issues.

First, unless one insists on risk neutrality of preferences, a comparison of expected returns would generally involve measuring not just the means but the whole outcome distributions conditional on choosing each alternative. More generally, applicability of the simple framework is restricted to choice models in which relevant expected future returns can be fully captured by a finite set of measurable summary statistics. As observed by van der Klaauw (2012) (p. 526), "(...) *if realizations of outcomes are correlated across choice alternatives, one would generally need to measure the entire joint subjective distribution of future choice-specific outcomes.*"

Second, elicited right-hand side expectations may often reflect unobserved heterogeneity in preferences or effort, raising concerns about their endogeneity. For example, reported expectations about the likelihood of obtaining high grades in a certain program or field of study may reflect expected study effort as well as preferences over grades. If these aspects are not measured explicitly and incorporated in the model, grade expectations alone will be endogenous. A similar argument applies to major-specific expectations for earnings and other outcomes that may depend on unobserved taste for majors.

---

[88]When the elicited choice probabilities in $\{q_{ij}\}_{j \in \mathcal{J}}$ have implied corner values of 0 or 1, the most common practice has been to recode them to values just above 0/below 1. See Blass et al. (2010) for additional details and for alternative distributional assumptions on the unobservables capturing resolvable uncertainty.

[89]If robustness of parameter estimates to recoding of corner choice probabilities is a concern, the model may be estimated via least absolute deviations (LAD). Median regression is invariant to transformations that do not alter the ordering of values relative to the median. However, it requires that the unobservables are symmetrically distributed around zero conditional on the observables.

Third, the basic framework is static. Yet, educational choices and other investments in human capital are sequential in nature, calling for analysis of dynamics. As observed by van der Klaauw (2012) (p. 526) "(...) *For example, in the case of college major choice or occupational choice models, in making current choice decisions individuals may consider the option to switch in the future. Generally, in making current choices, individuals may consider future benefits associated with sequences of future choice decisions. In that case, a comparison of returns or utility levels associated with choice alternatives would involve consideration of the expected outcomes conditional on any possible sequence of choices up to that future period, as well as the probabilities of making these sequential choices.*".

Fourth, because some of the most consequential schooling decisions are made during adolescence, they should be viewed as the outcome of a process that may involve multiple actors; first and foremost the student and the student's parents, but potentially also the student's teachers, peers, older friends and siblings, etc. This observation raises the question of whose expectations and preferences are most relevant for secondary and tertiary schooling decisions. More generally, it calls for modelling and measurement extensions that address the locus of decision-making within the family (Giustinelli and Manski, 2018), as well as the roles of potentially important actors outside the family such as the student's teachers and peers.[90]

Fifth, schooling decisions can be quite complex, due to a variety of institutional and/or market factors. These may include the presence of a large number of options of possibly heterogeneous quality, as in large US cities like NYC; the existence of centralized choice systems with complex application-and-admission procedures, as in Boston, Barcelona, Santiago, and many other cities worldwide; the presence of early curricular specialization or tracking, shaping the connections between early and later choices and outcomes, as in Germany, Italy, and other OECD countries.[91] These aspects raise important questions. For example, whether students and parents are aware of all choice alternatives; whether they know the workings of the application-and-admission mechanism; whether they understand the connections between early education stages and later ones, and can form sufficiently precise beliefs about the consequences of early choices on later ones. More generally, they call for modelling and measurement extensions that address these questions.

Recent analyses of schooling decisions with elicited expectations have sought to address some of these questions by improving and innovating on both modelling and measurement. For example, Wiswall and Zafar (2015b) extend Zafar (2013)'s model of college major choice by allowing students' earnings expectations to be correlated with unobserved tastes for majors. To address the endogeneity of earnings expectations, the authors experimentally generate a panel of expectations via a randomized information treatment providing students with information about population earnings. Patnaik, Venator, Wiswall and Zafar (2020) further extend Wiswall and Zafar (2015b)'s analysis by incorporating heterogeneity in risk and time preferences across students.

Hastings et al. (2015) develop a model of college degree program choice in Chile, where the student's utility depends on the degree program's characteristics and the interaction of the student's tastes and the degree program's attributes, including future earnings and program costs. The model allows for uncertainty and learning in students' beliefs about population earnings and program costs. Identification and estimation take advantage of survey data on students' expectations and beliefs about earnings and costs, as well as of a randomized intervention providing college applicants with information about population earnings and costs which creates exogenous variation in learning.

Delavande et al. (2020) extend previous analyses of college behavior by developing and estimating a discrete choice model of time allocation decisions where students consider investments in academic and non-academic activities such as job placements or volunteering.

---

[90]E.g., see Giustinelli (2016), Carlana (2019), Carlana et al. (Forthcoming), Papageorge et al. (2020) on the roles and importance of teachers. See Manski (2000), Epple and Romano (2011), Sacerdote (2011), Bramoullé et al. (2020), Garbin (2021), and references therein, on the roles and importance of social and peer interactions.

[91]E.g., see Abdulkadiroglu and Sonmetz (2003), Agarwal and Somaini (2018), and Son (2020) on the complexities of centralized school choice systems. See Ariga et al. (2005), Betts (2011), and Altonji et al. (2016) on the prevalence and implications of curricular specialization.

Notably, the framework allows for unobserved choice set heterogeneity across students.

Stinebrickner and Stinebrickner (2014a) extend the basic static framework of college major choice by considering a model where a student's final major is viewed as the end result of a learning process. The analysis relies on the longitudinal information on subjective expectations and outcome realizations collected in the Berea Panel Study (BPS). Importantly, because in this case beliefs are elicited before choices are made, BPS students are allowed to express uncertainty about the choice alternative they will eventually choose. In this case, estimating students' preferences and other relevant parameters requires specification of the process by which the decision maker updates his or her beliefs before the final decision.[92]

Congdon-Hohman et al. (2017) propose that students making college major decisions may face deep uncertainty (ambiguity) about future major-specific wage distributions. They develop a model where uncertainty-averse students respond to the uncertainty about future major-specific wage distributions by choosing the major that performs best under the worst-case subjective expectation. The model implies that greater uncertainty about a particular major reduces the likelihood that students choose that major. Lacking expectations data, the authors are unable to estimate the model or test the empirical validity of its premises or predictions.

Giustinelli and Pavoni (2017), on the other hand, take advantage of novel survey questions eliciting students' and parents' perceived awareness of the choice alternatives and allowing respondents to quantify their degree of confidence in their expectations for program-specific consequences by means of probability intervals, to study a framework of high school track choice with limited choice awareness and belief ambiguity about the student's likelihood of timely graduation from high school.[93]

Giustinelli (2016) extends the basic unitary model of high school track choice by opening the family "black box". The paper considers Bayesian models of high school track choice with subjective risk and nonstrategic modes of child-parent decision-making. Model specification is informed by survey questions eliciting the decision roles of the student and the student's parents, covering as main cases unilateral choice by the student, choice by the student with parental input, and student-parent joint decision. The models are estimated by combining actual choices with survey information about students' and parents' subjective probabilities over choice consequences, individually preferred choices, and decision roles. The framework and data enable separate identification and estimation of the parameters characterizing the student's utility, the parent's utility, and the mode of interaction between student and parents. The latter set of parameters capture how the student incorporates parental expectations and preferences in his or her own (choice by the student with parental input), or how student and parent resolve misalignments in their preferences over tracks (student-parent joint decision).

Garbin (2021) extends the basic model of high school track choice in a different direction, by allowing the SEU to depend on the student's expectations about the choices of his or her friends. The model is estimated using multiple waves of a survey that elicits students' expectations about the choices of the peers in their network in addition to their expectations over own outcomes.

Kapor et al. (2020) deal with modelling of school choice within a centralized admission mechanism. They enrich the standard framework by incorporating parental subjective beliefs about the student's admission chances.

## 4.2 Survey Elicitation

The observations on survey elicitation presented in Subsections 2.2 and 3.2 continue to apply when collection of expectations data is purposely conducted with the goal of studying

---

[92]Taking a similar perspective and modelling approach, Stinebrickner and Stinebrickner (2014b) study college dropout. Because of their focus on the role of learning, these studies are reviewed in greater detail in Section 5.

[93]Because of the paper's focus on the evolution of awareness and ambiguity over the decision process, also this paper is reviewed in greater detail in Section 5.

the determinants of schooling decisions. Hence, the discussion of this subsection is limited to selected aspects of the collection of expectations data when its goal is choice modelling.

*Data needs for the basic model.* As noted in Subsection 4.1.1, the typical data on expectations that have been collected and used for estimation of random utility models of field-of-study choice are right-hand side expectations over the consequences of choosing alternative schooling options. These data may include a student's subjective probability of enjoying the coursework in each study program, his or her expected earnings following graduation from each program, and so on. The key observation here is that in order to estimate the model one needs to elicit subjective expectations over choice consequences *for each and every choice alternative* in the decision maker's choice set, including the chosen option as well as the unchosen ones. For example, Zafar (2013)'s survey elicited the subjective probability of each outcome following hypothetical choice of each of the 8 majors in the students' choice set, Giustinelli (2016)'s survey for each of 10 curricula.

A second observation concerns a common tension between the desire of collecting as rich as possible expectations data and the need of keeping response burden under control. For example, Zafar (2013)'s SEU specification includes 11 expectations per major, Giustinelli (2016)'s 11 per curriculum. These are relatively large numbers of expectations to elicit. Both studies sought to maximize the variety of outcomes entering the SEU by specifying most outcomes as binary, so that only one subjective probability per outcome needed to be elicited. Alternatively, given the same number of questions, one may focus on a smaller set of outcomes and elicit richer expectations on the latter. For example, elicitation of subjective belief distributions for continuous outcomes such as future earnings, or for joint discrete outcomes, will generally require more questions.

To estimate the model one also needs information on schooling decisions, that is, left-hand side data. These data may consist of actual choices, stated choices, or both (e.g., Arcidiacono et al. (2012) and Giustinelli (2016)). Stated choices may be elicited in the form of choice intentions (e.g., Arcidiacono et al. (2012)), ratings or rankings of choice alternatives (e.g., Zafar (2013)), or choice probabilities (e.g., Stinebrickner and Stinebrickner (2014a), Wiswall and Zafar (2015b), and Boneva, Golin and Rauh (2022)). Thus, additional survey questions may be required, especially if the analysis is based on stated choices.

Choice measures can vary substantially in their implied response burden as well as in their degree of informativeness. Survey questions eliciting subjective choice probabilities may be more burdensome than forced-choice questions, especially when the number of choice alternative is large or when elicitation is repeated over multiple choice scenarios. Yet, responses to the former type of questions tend to be more informative than responses to the latter type of questions. Moreover, allowing a decision maker to express uncertainty by providing an interior choice probability is conceptually the correct thing to do whenever a survey precedes the actual decision.

Actual choices, on the other hand, may be measured via a single survey question or may be observed in administrative data, implying limited or no burden for respondents. Moreover, actual choices have the advantage that they capture actual behavior. Yet, they may not always be available or may not display sufficient variation across decision makers.

A third and related consideration regards the timing of the survey relative to the time of choice. Ideally, one would like to measure the expectations over choice consequences that decision makers use to make their choice along with the choice made. In practice, this can be very challenging (if not impossible) to achieve, especially within a single survey. In many real-life decisions, the time of choice may vary across decision makers, making it difficult to time the survey so as to ensure observability of the choice-relevant expectations for all respondents. At any given point in time, different respondents may be observed at different stages of their decision process. Respondents who are observed early in the decision process may still be in the process of forming their expectations over choice consequences. Respondents who are observed after they made their choice may have already updated their expectations over consequences in response to post-choice new information, or to

41

ex-post rationalize their choice. Even if they have not updated their expectations since the time of choice, they might not recall them perfectly.[94]

A design that collects expectations over choice consequences before the choice is made, but reasonably close to it, and then observes actual choices in a follow-up survey, or in an administrative data set that can be linked to the initial survey, may provide a reasonable solution to that trade-off. A longer and high-frequency longitudinal design such as that of the Berea Panel Study (BPS), which observes students frequently, at key times of information arrival (e.g., through grades), and for a prolonged period of time, provides a more effective but also more costly solution.[95]

A final observation concerns the alignment of survey recipients and decision makers. Ideally, one would like to measure the expectations of the individuals who actually make the decision being modelled. In practice, the identity of the decision makers may not be obvious or known ex ante, especially for decisions related to schooling and human capital acquisition. Schooling laws set the ages at which youth can enter and leave school, but they do not rule on whether or when they can autonomously make schooling decisions. Schooling decisions are better viewed as family decisions, whose locus of decision-making may depend on factors such as the child's age and development, the parents' parenting style, and other characteristics of family members or the choice environment. Thus, when looking within families and over time, one would expect a trajectory whereby parents are completely in charge of early schooling decisions (e.g., over elementary schools), followed by a phase in which the child acquires decision-making agency and increasingly participates in schooling decisions that are taken during adolescence (e.g., high school choice or college enrollment), until the child becomes an autonomous decision maker (e.g., over college majors, post-graduate education, and beyond). However, when looking within decisions and across families, the locus of family decision-making may actually vary across families, especially for decisions that occur during adolescence.

Expectations surveys conducted with the goal of modelling college major or dropout decisions and post-graduate decisions have universally targeted the students (e.g., Arcidiacono et al. (2012), Zafar (2013), Stinebrickner and Stinebrickner (2014a,a), Wiswall and Zafar (2015b), and Boneva, Golin and Rauh (2022)). Whereas expectations surveys conducted with the goal of modelling earlier decisions, such as those of the high school track or college enrollment, have taken different approaches, with some surveys targeting the students (e.g., Belfield et al. (2020) and Hastings et al. (2015)), others the parents (e.g., Kapor et al. (2020)), and yet other ones targeting both (e.g., Giustinelli (2016) and Giustinelli and Pavoni (2017)). The still scant evidence from studies that have been able to investigate and compare the predictive power of students' and parents' expectations on secondary and post-secondary schooling decisions (e.g., Attanasio and Kaufmann (2014) and Giustinelli (2016)) points to the importance of measuring the expectations of both the students and their parents, along with information about their decision-making roles.

*Data needs for richer models.* The studies reviewed in Subsection 4.1.2 develop extensions of the basic choice framework that make the model more realistic and use additional data that make identification more credible. These improvements are not free of costs, since the need for additional data can increase the complexity of survey design for researchers and the response burden for respondents.

For example, Wiswall and Zafar (2015b)'s identification strategy relies on the collection of a panel of right- and left-hand side expectations via a randomized information treatment embedded in the survey. Hence, the authors had to elicit pre- and post-treatment expectations from their respondents within the same survey. Patnaik, Venator, Wiswall and Zafar (2020) further relies on the collection of incentivized measures of students' risk and time preferences.

Stinebrickner and Stinebrickner (2014a,b)'s analyses take advantage of the longitudinal

---

[94]For further discussions and evidence see, e.g., Zafar (2011a), Giustinelli (2016), and Kapor et al. (2020).

[95]However, see Kapor et al. (2020)'s discussion about the possibility that surveying individuals during the decision process may affect the process and the decision.

and high-frequency features of the expectations collected in the Berea Panel Study (BPS), of the richness of the elicited expectations over grades, effort, graduation, major selection, and of the ability to observe realizations within the sample.

Giustinelli (2016)'s identification strategy relies on the joint availability of data on actual choices (viewed as a family outcome), on right-hand side expectations over choice consequences and left-hand side choice preferences (separately elicited from the student and the student's parents), and on students' and parents' decision-making roles within the family. Garbin (2021) further relies on information about each student's network of peers and on the student's expectations over the choices of his or her peers.

Giustinelli and Pavoni (2017, 2019) use information on respondents' perceived choice awareness and perceived degree of confidence in their expectations for track-specific outcomes, elicited separately from the students and their parents and repeatedly over the decision process.

## 4.3 Empirical Evidence

*Monetary returns and risks.* The existence of large earnings differentials across schooling levels and fields of study is well documented.[96] Yet, observational studies have typically found modest effects of expected earnings on schooling decisions.[97] Turning to the evidence from the literature employing survey expectations, Zafar (2013) finds a positive but small and insignificant effect of subjective expectations about income at age 30 on college major choices of NU undergraduates, irrespective of gender and other stratifications. Arcidiacono et al. (2012) find somewhat larger and statistically significant effects among male Duke undergraduates, implying for example that a one standard deviation increase in expected earnings in a business career increases the fraction of students majoring in economics from 19.7% to 22.9%. Wiswall and Zafar (2015b), too, find that expectations about future earnings are a significant determinant of major choice among NYU undergraduates, although their estimated average elasticities are modest (between 0.03 and 0.07). They provide evidence that major choice elasticities vary across groups, with younger students displaying higher responsiveness to changes in future earnings than older students, presumably because older students have a higher cost of switching major.[98]

Importantly, Wiswall and Zafar (2015b) show that the estimated elasticities shrink once the correlation between unobserved tastes for majors and expected earnings has been taken into account. They further show that accounting for students' subjective uncertainty about future earnings is also very important. They estimate a large degree of risk and find that unaccounted-for risk aversion tends to inflate the estimated responsiveness of major choice to changes in expected earnings. In follow-up work, Patnaik, Venator, Wiswall and Zafar (2020) additionally show that ignoring heterogeneity in risk and time preferences also overstates the importance of expected earnings for major choice.

Some of these studies have employed survey expectations to investigate whether well-documented differences in schooling decisions across demographic or socioeconomic groups can be explained by differences in students' subjective expectations or are perhaps due to other factors such as differences in preferences. For example, Zafar (2013) studies the determinants of the gender gaps in college majors, whereby female students are significantly less likely than male students to pursue majors such as Engineering or Computer Science and more likely to pursue majors such as Education or Health. The author finds that the gender gaps in college majors observed in his sample of NU undergraduates cannot be explained by differences across genders in either expectations or preferences for

---

[96]E.g., Card (1999), Heckman et al. (2006), Altonji et al. (2016), and Psacharopoulos and Patrinos (2018).

[97]E.g., see Arcidiacono (2004) and Beffy et al. (2012) for evidence on the US and France with respect to the postsecondary field of study. See also Altonji et al. (2016)'s review.

[98]Using nonprobabilistic expectations in the NLSY, Montmarquette et al. (2002) find that earnings expectations play an important role in college major choices, but their impact vary widely by gender and race. Baker et al. (2018) also find that expected labor market outcomes matter for college major choice of community college students in California. Their experimental estimates imply that a 10% increase in salary is associated with a 14-to-18% increase in the probability of choosing a specific category of majors.

future earnings. Patnaik, Venator, Wiswall and Zafar (2020) further show that while male students are less risk averse and patient than female students, gender differences in expectations about future earnings, risk aversion, and patience cannot explain gender gaps in major choice in their sample of NYU undergraduates.

These findings refer to college major choices of US undergraduates. The evidence for other countries and other schooling decisions, especially those involving younger students, is scarce.[99] Hastings et al. (2015, 2016) provide evidence on college enrollment and major choice in Chile. Belfield et al. (2020) and Boneva and Rauh (2021) provide evidence on sixth form and university enrollment among high school students in the UK and Boneva, Golin and Rauh (2022) about postgraduate enrollment among UK university students.[100]

Hastings et al. (2015) study how expectations about earnings and costs held by college applicants in Chile affect their choices about where to enroll and what to study (institution-major pairs). The study combines a large-scale field experiment giving applicants information about college- and major-specific earnings and costs (a disclosure policy), with a survey collecting students' choice preferences, expectations about own earnings and costs, and beliefs about population earnings and costs. The analysis shows that the intervention decreases the demand for degree programs in the bottom tercile of the returns distribution by 3.3% overall, 4.6% for low-income students, and 3.9% for low-achieving students, while leaving college enrollment unaffected. To understand why disclosure reduces but does not eliminate the demand for low-return programs, the authors estimate a structural model of college demand that allows for belief uncertainty and learning about earnings and costs. They show that the intervention affects students' college decisions by reducing the uncertainty in their earnings beliefs, but the effect of learning is dampened by students' preferences for nonmonetary attributes.

Belfield et al. (2020) study the determinants of sixth form and university enrollment among secondary school students aged 13-14 in the UK. The authors find that students' perceptions of the monetary returns to further education are statistically associated with their plans to continue in full-time education, but the perceived monetary returns can explain only a small share of the variation in education plans relative to other perceived benefits of schooling.

Boneva, Golin and Rauh (2022) investigate the socioeconomic gap in postgraduate education in the UK. They compare students who are the first generation in their family to go to university (low SES) with students whose family members attended university (high SES), as the former are significantly less likely to pursue postsecondary education than the latter. This is reflected in the educational expectations of respondents: low-SES students report a 5 percentage points lower likelihood of continuing to postgraduate education than high-SES students, on average.[101] The analysis shows that the perceived monetary returns to postgraduate education are higher among high-SES students and that the significant estimated impact of expected full-time earnings at age 35 on postgraduate enrollment plans is driven by high-SES students. These differences, however, are not a major source of the observed enrollment gap, which is better explained by differences in other factors.

Boneva and Rauh (2021) consider younger UK students (aged 13-18) and investigate the gap in university enrollment between students who have at least one parent with a university degree (high SES) and students whose parents have no university degree (low SES). Using survey expectations over a rich battery of monetary and nonmonetary outcomes similar to those collected and analyzed by Boneva, Golin and Rauh (2022), the authors document similar patterns among younger students. Low-SES students expect a

---

[99]As seen in Section 2, many studies of earnings expectations have been conducted outside the US, especially among university students and to a lesser extent among high school students. However, only a handful of them estimate expectations-based models of schooling decisions.

[100]For additional evidence from studies conducted in low and middle income countries, see Delavande (2022)'s review in this handbook.

[101]The authors find substantial differences also within high-SES students: students with at least one parent holding a postgraduate degree report a 8 percentage points higher probability of enrolling in a postgraduate degree than students who have at least one parent with a first degree, but no parent with a postgraduate degree.

lower earnings premium from going to university than high-SES students. Nevertheless, these SES differences in perceived monetary returns account for only a trivial fraction of the observed SES enrollment gap.

*Monetary costs.*    Recent decades have been characterized by increasing returns to schooling and labor market risk in both the US and other high income countries, especially for postsecondary education. The costs of postsecondary education have also risen along with the returns, especially in the US. These combined trends have increased the demand for credit, simultaneously raising the question of whether college students in the US can borrow enough (or are credit constrained) and the concern that some students may be actually taking on too much college debt (Lochner and Monge-Naranjo, 2016). The recent evidence on the high degree of complementarity between early and late human capital investments has further raised the concern that postsecondary aid policies may simply come too late to be effective in helping disadvantaged youth (Lochner and Monge-Naranjo, 2012).

Given the ongoing debate, it would seem natural to ask whether students' expectations about tuition and other monetary costs of postsecondary education and their beliefs about borrowing opportunities and terms play a role in college decisions and outcomes such as enrollment, dropout, and major. In practice, the evidence on the US is scarce.[102] The analyses of college major choice reviewed above largely abstract from these considerations, as they focus on students who are already in college, either approaching or after major selection. Moreover, they consider selective institutions, whose students prevalently come from non-disadvantaged backgrounds and where dropout rates are modest. The Berea Panel Study (BPS), on the other hand, focuses on disadvantaged students. However, since all Berea students receive full tuition and large room-and-board subsidies, the above concerns are likely to be of limited relevance for them. Indeed, Stinebrickner and Stinebrickner (2008b) show that difficulties with borrowing money to pay for consumption during college do not have an important effect on dropout decisions.

To my knowledge, the only two studies providing some evidence on these issues in the US context are Huntington-Klein (2016b) and Bleemer and Zafar (2018).[103] Bleemer and Zafar (2018) examine the role of information gaps about college monetary costs and returns on college enrollment expectations by providing respondents to the New York Fed's Survey of Consumer Expectations (SCE) randomized information about the monetary costs of college or the monetary returns to college. The authors find that the cost information treatment has no significant impact on college attendance expectations in either the full sample or any of the demographic subgroups.[104] The return information treatment, on the other hand, is found to have a significant impact on college attendance expectations. Respondents' subjective probability of sending their own child to college increases by 4.9 percentage points on average and that of recommending college for a friend's child by 2.3 points, corresponding to a 0.2 standard deviation increase in college attendance expectations. The effects are larger for disadvantaged respondents. In the short run, the return information treatment closes the education and income gaps in parents' college attendance expectations by around 30% and the recommendation gaps by about 15%. These effects are shown to be persistent two months after the initial survey.

Additional evidence is provided by Hastings et al. (2015, 2016) on Chile and by Belfield et al. (2020), Boneva, Golin and Rauh (2022), and Boneva and Rauh (2021) on the UK. Hastings et al. (2015) find that the disclosure intervention does not cause students to choose lower-cost programs, consistent with the fact that they hold accurate cost beliefs on average. Thus, Hastings et al. (2015)'s and Bleemer and Zafar (2018)'s findings point to a similar conclusion: information provision about the monetary returns to college signif-

---

[102]This includes studies that do not employ survey expectations. Altonji et al. (2016) review the studies examining the impact of financial aid and tuition fees on college attendance, but note that very little is known about their impact on college major choice.

[103]Huntington-Klein (2016b) replicates Hastings et al. (2016)'s earlier analysis for Chile, reviewed below.

[104]These findings are in line with findings from earlier studies showing modest effects of information interventions about costs; e.g., Bettinger et al. (2012) and Hoxby and Turner (2013).

icantly affects college decisions and plans, whereas information provision about monetary costs does not.

Hastings et al. (2016) use the same data as Hastings et al. (2015) to evaluate the predictions of an informal model of college enrollment, where college applicants face uncertainty about costs and earnings for different degree programs, decide whether to engage in costly degree-specific search, and further learn about the chosen degree after enrollment.[105] The last prediction–stating that students with less accurate expectations about costs or earnings are more likely to enroll in degrees where past students have performed poorly in those dimensions and more likely to drop out when the uncertainty is resolved–is especially relevant here, as it connects students' cost and earnings expectations with students' college decisions. Empirical findings are partly consistent with the prediction: students who overestimate costs are less likely to matriculate in any degree program and more likely to drop out; students who overestimate earnings matriculate at similar rates as other students, but choose degree programs where past students have performed poorly (less likely to graduate, earned less in starting jobs, and more likely to default on student loans). Notably, students with a stated preference for labor market-related degree characteristics are shown to be less likely to overestimate earnings outcomes and to choose degrees where past students have gone on to earn more, while the opposite is true for students with a stated preference for enjoyment of the curriculum.

Moving to the UK, Belfield et al. (2020) find no evidence that students' expectations about tuition costs play a role in their sixth-form and college enrollment plans. The authors interpret these results as being consistent with the fact that participation rates in higher education have not dropped after the introduction of higher university tuition fees.

Boneva, Golin and Rauh (2022) find that financial considerations play a role in students' postgraduate plans. Students' subjective expectations of struggling financially and their expectations about the immediate costs of college (tuition fees plus forgone earnings) are negatively related to students' subjective probability of enrolling in a postgraduate program. They estimate that students would be willing to trade-off £0.38 in immediate costs for a £1 increase in earnings at age 35. However, nonmonetary factors are found to play a larger role. Boneva and Rauh (2021) find similar results with respect to younger students' college enrollment plans. Financial considerations related to the possibility of struggling financially during college and receiving financial support from parents play a greater role than earnings expectations, but a smaller role than other nonmonetary factors.

*Nonmonetary factors: ability, taste, and beyond.* Given the limited ability of expected monetary returns and costs to explain observed schooling decisions and to rationalize observed differences in educational choices across demographic and socioeconomic groups, the literature has turned to investigate the explanatory power of a variety of nonmonetary factors or outcomes. These include outcomes that students experience in school and are typically related to students' academic abilities and "tastes", for example, students' academic performance, study effort, and coursework enjoyment. They also include post-schooling outcomes in the labor market, such as job satisfaction and flexibility, and in other markets such as the marriage market.

A robust finding across expectations-based analyses of schooling decisions is that non-monetary aspects of education and work are predominant drivers of students' schooling decisions and plans, although the importance of specific nonmonetary factors over others can vary across settings (e.g., types of decision or countries) and decision makers (e.g., sociodemographic groups or students versus parents).

Two factors whose importance has been long emphasized in the human capital literature

---

[105]The predictions are: (1) students facing higher search costs have less accurate beliefs about program attributes such as monetary costs and returns; (2) students who place low value on costs or earnings when making the enrollment decision have less accurate beliefs about those outcomes; (3) students know more about programs in their consideration set; and (4) students with less accurate expectations about costs or earnings are more likely to enroll in degrees where past students have performed poorly in those dimensions and more likely to drop out when the uncertainty is resolved.

are students' abilities and tastes, as they capture key dimensions of the quality of the match between a student and each of the educational alternatives available to him or her (Arcidiacono, 2004). Indeed, Arcidiacono et al. (2012) find that major-specific abilities are important determinants of major choice among Duke undergraduates, above and beyond expected earnings, and that students prefer majors at which they are good. Concretely, the model estimates show that equalizing student abilities across majors would drop the fraction of humanities majors from 9.3% to 5.9%, while increasing the fraction of economics majors to 23.8%.

Wiswall and Zafar (2015b) emphasize the importance of the residual unobserved taste-for-major component, which the authors identify as the dominant factor driving major choice among NYU undergraduates. The importance of students' tastes for majors, however, is found to vary across students. Specifically, it plays a much larger role for older students (juniors) than for younger ones (freshman and sophomores), consistent with a large and increasing cost of switching majors as students progress through college.[106]

The majority of education studies that have collected survey expectations with the goal of estimating choice models has elicited respondents' expectations over many nonmonetary schooling outcomes in addition to those over monetary ones. These studies have thus been able to directly assess the relative importance of different nonmonetary factors and to quantify their overall importance relative to that of monetary factors. Moreover, they have shed light on the ability of expectations and preferences for nonmonetary factors to explain educational gaps across gender and socioeconomic groups.

For example, Zafar (2013)'s estimates show that the most important factors in the choice of major of NU undergraduates are enjoying the coursework, enjoying the work at potential jobs, and gaining parental approval. Moreover, his estimates imply that nonmonetary factors explain over 50% of the variation in college major choices of male students and over 80% of the choice variation among female students. This finding is driven by male and female students having significantly different preferences over workplace attributes, with female students valuing nonmonetary aspects such as enjoying work on the job and reconciling work and family much more than male students. Monetary aspects of the workplace explain as much as four times of the choice for males than do nonmonetary aspects of the workplace; for females, the two are equally important. Female and male students, on the other hand, have similar preferences over in-school outcomes, but hold different expectations about specific nonmonetary outcomes such as enjoying the major-specific coursework. Zafar (2013) uses his model to simulate a number of policy-relevant counterfactuals. He shows that replacing female students' subjective expectations about enjoying the major-specific coursework with those of male students would reduce the choice gap in Engineering by 50 percent, whereas replacing female students' subjective expectations about ability and future earnings with those of male students would reduce the gap by a much lower 15 percent. Zafar (2013) concludes that policies aiming at sensitizing females' preferences for STEM fields may be more promising than policies providing information about the monetary returns to STEM fields.

Baker et al. (2018) find that the main determinants of major choice among community college students are beliefs about course enjoyment and grades. Whereas Wiswall and Zafar (2020) provide evidence that students sort into college majors partly based on their perceived returns in terms of potential spouse's earnings and fertility. The authors find that family expectations are particularly important for the major choices of NYU female students.

In the context of high school track choice in Italy, Giustinelli (2016) finds that the student's enjoyment of the track's core subjects is by far the most valued attribute by both students and parents, irrespective of the family's decision process (solo choice by the

---

[106]The authors discuss the implications of their results for the modelling of tastes in choice models. While prevalent approaches have assumed that tastes are orthogonal to everything else in the model, the substantial correlation between tastes and observables documented by Wiswall and Zafar (2015b) suggests that this assumption may be problematic. Moreover, the fact that observables are able to explain only a limited share of the variation in tastes (about 20%) suggests that allowing for an unrestricted distribution of tastes, as Wiswall and Zafar (2015b) do and as opposed to imposing a particular parametric distribution, may be important.

student, student choice with parental input, child-parent joint decision). The importance of other short-term outcomes (e.g., academic performance or study effort) relative to longer-term ones (e.g., facing flexible college and work choices after graduation) is instead highly heterogeneous across family members and decision processes. Using her model to perform a number of policy counterfactuals, the author shows that a "math sensitization" campaign increasing students' perceived likelihood of enjoying math in the general track by 10 points on the 0-100 percent chance scale, a relatively small movement in expectations, can have a large impact on the distribution of choices due to the large utility weight that students and parents attach to coursework enjoyment. Subjecting university access to graduation from one of the general curricula–an example of an institutional policy tightening curricular specialization–would also generate a significant impact on choices. Instead, providing information about population graduation rates or college enrollment by graduation track would have little-to-no impact on choices. This is partly due to the modest utility weights estimated for the latter outcomes and partly due to the fact that families' expectations about these outcomes are already on target.

The findings of Belfield et al. (2020), Boneva, Golin and Rauh (2022), and Boneva and Rauh (2021) for the UK also point to a predominant importance of nonmonetary factors. For example, Belfield et al. (2020)'s analysis of sixth-form and university enrollment plans shows that students' perceived consumption value of schooling plays a highly prominent role in students' plans to continue in full-time education: the perceived consumption value alone explains up to 43% (51%) of the variation in students' subjective probability of going to sixth form (university). Moreover, the existing heterogeneity in perceived consumption value across socioeconomic groups–whereby high-SES students perceive the consumption value of sixth form (university) to be 0.15 (0.29) standard deviations higher than perceived by low-SES students–helps explain a high share of the observed SES gaps in sixth-form and university enrollment probabilities. Similarly, the heterogeneity in perceived consumption value of university across genders (a quarter of a standard deviation higher for females) explains a substantial share of the gender gap in university enrollment probabilities. Belfield et al. (2020)'s analysis also shows that students' plans take the dynamic nature of the decision problem into account. Specifically, students' perceived benefits of both sixth form and university play an important role in students' sixth-form decisions: students who expect larger monetary returns or consumption value of university are significantly more likely to plan to go to sixth form.

The estimates of Boneva, Golin and Rauh (2022)'s choice model of postgraduate enrollment and the results of their post-estimation decomposition show that of the 3.9 percentage point SES gap predicted by the model, 91% can be explained by SES differences in expectations and 9% by SES differences in preferences. Expectations about the returns to postgraduate education are the most important in explaining the enrollment gap, accounting for approximately 70% of it. While both perceived monetary and nonmonetary returns play a role, the latter appear more relevant empirically.

Boneva and Rauh (2021) perform a similar exercise with respect to younger students' intentions of enrolling in university, this time following a model-free approach. The authors employ a statistical learning algorithm to determine the predictive importance of students' expectations about different monetary and nonmonetary factors with respect to their subjective probability of going to college. They find that differences in perceived returns across SES groups explain 35% of the SES gap in enrollment probability. Among the nonmonetary factors, students' expectations about having parental approval, enjoying studying, and job satisfaction play the most important role. Monetary factors and parental wealth play minor roles relative nonmonetary factors.

*Parents and family decision-making.* Economic models of school choice have traditionally assumed that parents make schooling decisions for their children while children are minors and that students become solo decision makers when they come of age.[107] Accordingly,

---

[107]Only limited attention has been devoted to the roles of agency acquisition by children, child-parent interactions, and family decision-making in schooling decisions. Lundberg et al. (2009) and Dauphin et al. (2011)

demand models of postsecondary schooling have typically taken the student as the main decision maker or have implicitly viewed the family as a unitary entity.[108] However, Zafar (2013)'s and Boneva and Rauh (2021)'s findings that gaining parental approval is among the most important factors explaining college major choices of NU undergraduates and university enrollment plans of UK high school students suggest that parents may play a more active role in postsecondary schooling decisions than previously assumed.[109] Zafar (2012) provides further evidence related to students' selection of their majors' composition among students who pursue a double major, about one quarter of US undergraduates. The analysis shows that students believe that their parents are more likely to approve majors associated with high social status and high returns in the labor market. Estimates of a choice model of double majors allowing for specialization concur with Zafar (2013)'s finding: enjoying the coursework and gaining parental approval are the most important determinants also for double majors choice. The model estimates further show that gaining parental approval and enjoying the coursework are important factors for both majors in a student's major pair. Thus, students who pursue two majors do not seem to select one to satisfy their own interests and the other to please their parents. On the other hand, double-major students do tend to choose majors that differ in their chances of completion and difficulty and in finding a job upon graduation.

Estimation of a school choice model that explicitly incorporates the subjective expectations and preferences of both the student and the student's parents requires elicitation of subjective expectations and choice preferences of multiple members within the same family. To the best of my knowledge, to date this has never been done for the purpose of studying postsecondary schooling decisions, but it has been done by Giustinelli (2016) with the goal of analyzing high school track choices among Italian adolescents.

Using direct survey evidence from questions eliciting aspects of the family decision process and the roles played by the student and the student's parents, Giustinelli (2016) identifies three main processes of decision-making: (PR1) solo choice by the student, occurring in the 19.82% of sample families;[110] (PR2) choice by the student with parental input, in 38.05% of families;[111] and (PR3) child-parent joint decision, in 42.13% of families.[112] Combining actual enrollment choices with survey information on individual choice

---

investigate children's decision agency in domains other than school choice. Doepke and Zilibotti (2017) and Doepke et al. (2019) address the mechanisms through which parenting styles affect children's human capital and related outcomes.

[108] Two notable exceptions are Kalenkoski (2008) and Attanasio and Kaufmann (2014). Kalenkoski (2008) develops a cooperative bargaining model in which child and parents have conflicting preferences over the child's level of post-secondary schooling and parental transfers. The author derives testable implications about parental altruism and income pooling based on a comparison between the bargaining model and the unitary preference model. An empirical test using data from the High School and Beyond Surveys rejects the unitary model in favor of the bargaining one. Attanasio and Kaufmann (2014) investigate the predictive power of perceived monetary returns and risks on high school and college enrollment decisions in Mexico. Taking advantage of the ability to observe youths' and mothers' expectations within the same data set (albeit not within the same families), the authors analyze whose expectations predict each of the two choices and how this varies by youth's gender and age. They find that expected returns and risks as perceived by mothers are predictive of all schooling decisions. Expectations of male students in their last year of high school predict the decision to enter college, whereas expectations of male students in their last year of junior high school does not predict the decision to enter high school. Girls' expectations do not predict either of the two decisions, whereas mothers' expectations are particularly strong predictors of their daughters' decisions.

[109] See also Hotz et al. (Forthcoming), who provide evidence that American parents finance students' college enrollment off family income and college completion off housing wealth.

[110] This is intended as a process where parents do not participate explicitly in the decision process. However, in the model the student's SEU is allowed to depend on his or her perception of parental choice preferences over tracks. The model also allows for the student's choice to be affected by the student's teachers (through their choice recommendations) and peers (through the student's expectations of his or her best friends' choices).

[111] In this process, the student makes the final decision based on his or her own preferences and expectations, but accounts for parental expectations over track-specific outcomes via outcome-specific weights.

[112] In this process, the student's and parent's SEUs are aggregated via family member-specific weights. In families where both parents are present, the analysis treats parents as a unit and uses the expectations and preferences of the responding parent.

preferences independently reported by students and parents, Giustinelli (2016) documents that in nearly 50% of families the actual choice does not coincide with the stated-preferred choice of at least one between the student and the parent. In the majority of these cases, the actual choice coincides with the stated-preferred choice of the student but not of the parent (33% of the whole sample). Further interacting this information with information about the family decision process, Giustinelli (2016) shows that the latter figure is substantially higher among solo-choice children (40%) and lower among families where child and parent make a joint decision (29%). Giustinelli (2016) estimates a model that explicitly allows for the documented heterogeneity in family decision processes and in the ways in which student and parent expectations and preferences over track-specific outcomes affect high school track choice. The analysis uncovers significant differences in estimated preferences over track-specific outcomes between students and parents and, within them, across decision processes. These differences involve, for instance, the disutility of study effort and the importance of pleasing one's parent (more important for PR1 students), the importance of a regular path in high school and of being in school with one's friends (more valued by students than parents), and the importance of being able to make a flexible college field choice (more valued by PR2-PR3 students and PR3 parents). The estimates further show that among PR2 families parental expectations have differential influence on the student's choice through different outcomes. For example, parental expectations about the child's academic performance in high school matters more than the child's own expectations, whereas the opposite is true about post-graduation choice flexibility. Among PR3 families, parents have a greater influence on the final choice than their children, with an estimated decision weight of about two-thirds. Counterfactual simulations further confirm the importance of explicitly accounting for students' and parents' expectations and decision roles. For example, mistakenly taking the parent to be the main (or representative) decision maker in all families is found to overestimate the magnitude of enrollment response to the simulated math sensitization campaign. On the other hand, taking process heterogeneity into account shows that a disclosure policy publishing population statistics on graduation rates would have the largest impact on the choices of solo-choice students. It also shows that if parents alone were made aware of an institutional policy making curricular tracking more rigid, the policy's impact would be smaller than if children too were made aware.

*Peer effects.*    Giustinelli (2016)'s estimates suggest that students prefer high school tracks in which they expect their best friends to enroll. In her setting, the underlying mechanism for this preference is that students value the possibility of attending the same high school as their best friends and, thus, of sharing with them the following five years of their school lives.[113] An alternative but not necessarily mutually exclusive possibility is that students value choice alternatives that they expect to be chosen by many of their peers because of the implied "popularity" of these alternatives. In both cases, the peer effects operate through the students' expectations or perceptions about their peers' choices. Yet, there is very limited research to date using survey-elicited expectations to study peer effects in schooling decisions.[114] Garbin (2021) is a recent exception. The paper considers a model of high school track choice, where the students' subjective expected utility depends on their expectations about the high school track choices made by the students in their network. In this case, the peer effects operate through the second mechanism just described, that is, students valuing popular tracks. The paper finds that the expectations about the choices of one's peers matter more for own choice than the expectations over important academic

---

[113]Most high schools in Italy offer a single curriculum or at most a small number of curricula within the same track. Moreover, in small- and medium-sized municipalities each curriculum or track is often offered by one school. Hence, in many cities including that where Giustinelli (2016) collected her data, choosing the same high school track implies attending the same school, and viceversa.

[114]See Brock and Durlauf (2001a,b) and Lee et al. (2014) for discrete choice models with social interactions where decision makers' expectations about others' choices play an explicit role. These works, however, do not employ data on expectations. See also Epple and Romano (2011), Sacerdote (2011), and Bramoullé et al. (2020) for reviews of the vast literature on peer effects in education.

outcomes such as the student's enjoyment of the track's core subjects and expected study effort.

*Centralized school choice.*   In many places around the world, school choices are determined within centralized school choice systems. In such systems students are matched to schools on the basis of students' or parents' self-reported preferences over schools and of schools' capacity and other constraints, via algorithms developed from mechanism design and matching theory. Different matching mechanisms can have different properties, for example, in terms of matching stability, allocation efficiency, or strategy proofness.[115] However, these theoretical properties hold under the strong and often unrealistic assumptions that the canonical version of the mechanism is actually implemented and that applicants are fully informed, hold rational expectations, and make rational choices. In practice, students and families may not be fully familiar with the application-and-admission rules, may not hold accurate expectations about the student's chances of being admitted into different schools, and may consequently make application mistakes. These problems may disproportionally affect disadvantaged youth and families and may lead to individually suboptimal outcomes and undesirable distributional properties. Mounting evidence supports these concerns.[116]

In light of these concerns, it would seem important to develop and field survey questions that measure students' and parents' beliefs about the application-and-admission rules, their familiarity with the choice alternatives, and their expectations about the student's chances of being admitted into different schools. These data could then be used to investigate the importance of those aspects in application behavior and incorporate them in models of centralized school choice. In practice, the task has been challenging due to the very complexities of centralized school choice (e.g., complex rules and many choice alternatives) and by the need to collect survey data that can be linked to administrative information on application behavior and allocation outcomes. Despite these challenges, a small number of papers has begun to systematically address these issues.

Kapor et al. (2020) study the application behavior of families of rising 9th-graders in New Haven, Connecticut, whose school district has used a centralized matching algorithm similar to the Boston mechanism to assign graduating middle schoolers to high schools for over 20 years. The authors conducted in-person interviews with the families and then linked the survey responses to administrative records of the school placement process, including families' applications and assignment outcomes. The survey included questions eliciting respondents' expectations about the child's admission chances for schools ranked first and second in two hypothetical applications containing a mix of nearby schools, high-performing schools, and popular schools at the district level. It also included questions eliciting respondents' first- and second-most preferred schools if they could choose to attend any school with certainty, a measure of ordinal preferences, along with a measure of cardinal preferences based on a lottery choice task. The descriptive analysis of the survey data combined with the administrative data shows that many families engage in strategic reporting behavior, but do so on the basis of an inaccurate understanding on the rules and of inaccurate beliefs about the admission chances. This, in turn, leads to application mistakes. Indeed, the fraction of respondents that correctly describe key features of the assignment mechanism is found to be lower than that implied by random guessing. Moreover, respondents are shown to report subjective beliefs about admission probabilities that differ from a rational expectations benchmark constructed by the authors by a large mean (absolute) value of 37 percentage points. As a consequence, families end up underestimating the extent to which ranking a school higher or lower on their application affects the admission chances. The authors develop a model of centralized school choice that incorporates these patterns and use its estimates to study the welfare properties of two counterfactuals relative to the status quo. The first counterfactual simulates a switch to a strategy-proof deferred acceptance (DA) mechanism, where applicants do not need to

---

[115]See Hakimov and Kübler (2021) for a survey of mechanisms and their properties in the education context.
[116]E.g., see Sattin-Bajaj (2014), Calsamiglia and Güell (2018), Son (2020), and references therein.

understand assignment probabilities to play an optimal strategy. The second counterfactual simulates an informational intervention enabling households to play the Bayes Nash equilibrium in the game induced by the status quo mechanism. The authors show that switching to truthful reporting in the DA mechanism would yield welfare improvements over the baseline given the belief errors observed in the data, but the opposite conclusion would be reached if families were assumed to have accurate beliefs.

In follow-up work, Larroucau et al. (2021) investigate the role of beliefs about admissions chances in centralized choice settings with strategy-proof mechanisms and show that beliefs still matter, as they influence applicants' search behavior over schools. Combining a model where applicants engage in costly school search with data from a large-scale survey of school applicants in Chile, the paper provides evidence that learning about schools is costly, beliefs about admissions chances affect searching, and applicants underestimate the risk of not being placed by the system. Using RCT and RD research designs, the authors evaluate policies that give live feedback about admission chances to applicants within the Chilean and New Haven choice systems. They find that 22% of applicants submitting applications where non-placement risk is high respond to the feedback by adding schools to their lists and that this reduces non-placement risk by 58%.

In companion follow-up work, Arteaga et al. (2022) analyze the prevalence and relevance of application mistakes in the Chilean strategy-proof centralized college admissions system, focusing on admissibility mistakes that consist in applying to programs without meeting all requirements. The authors field a nationwide survey, collecting information about students' true preferences, their subjective beliefs about admission probabilities, and their level of knowledge about admission requirements and admissibility mistakes. The data show that 2-4% of students do not list their true most preferred program, only a fraction of this behavior can be rationalized by belief biases, and students tend to underestimate the probability of extreme events and to underpredict the risk of not being placed. The authors design and evaluate an information intervention aimed at reducing application mistakes. They find that showing personalized information about admission probabilities significantly reduces the risk of not being placed and the incidence of admissibility mistakes.

Tincani et al. (2021) further document the importance of students' biased beliefs for college admission and enrollment distortions in Chile and how those interact with preferential admission policies. While both high- and low-ability students are overconfident about their admission prospects, this induces the former to under-provide effort and the latter to over-provide it. This, in turn, generates under-admission and under-enrollment distortions among high-ability students and over-admission and over-enrollment distortions among low-ability students. The authors show that preferential admissions exacerbate these distortions, leading to more frequent over-enrollments of low-ability students (who still over-provide effort) and to more frequent under-enrollments of high-ability students (who under-provide effort even more than under no preferential admissions). The authors conclude that belief biases lead overconfident but underprepared students to enter college in response to preferential admissions.

# 5   Analysis of Expectation Formation and Learning With Panel Data on Expectations

## 5.1   Learning Framework

### 5.1.1   Basics

The basic choice framework presented in Subsection 4.1.1 and many of its extensions reviewed in Subsection 4.1.2 take elicited expectations as given and use them as data inputs in estimation of econometric models of schooling decisions. That approach assumes that decision makers have well-formed expectations over the consequences of schooling alternatives, act on their expectations when making schooling decisions, and can accurately report their expectations in surveys; it does not require that the econometrician specifies

the process that decision makers use to form their expectations. Indeed, if the main goal of the analysis is to recover the distribution of preferences across decision makers in the context of a static choice, specification of the process of expectation formation is not necessary. However, understanding how individuals facing school choice form the expectations they use in decision-making and update them as new information unfolds becomes necessary when the goal is more ambitious.

First, recall from Manski (1993)'s analysis reviewed in Subsection 2.1 how different assumptions about the nature of decision makers' expectations and the information on which they condition those expectations can have very different implications for interpretation of the equilibrium distributions of choices and outcomes, even when expectations are formed homogeneously across the population. In real life, decision makers may condition on different information and may form expectations heterogeneously, making the inferential problem even more challenging.

Second, understanding the nature of decision makers' expectations becomes essential when the goal is to evaluate the quality of observed choices. If decision makers have limited or incorrect information and, as a result, they form partial or biased expectations over the consequences of alternative options, their choices may be suboptimal.

Third, understanding formation and updating of expectations is also very important when the goal is to predict schooling decisions in novel policy scenarios that could affect expectations in ambiguous directions. Last but not least, dynamic models of sequential schooling decisions over time do require specification of the expectations process.[117]

This subsection introduces a formal framework where students update their expectations for own future earnings upon receiving information about the population distribution of earnings. The framework, which follows closely Wiswall and Zafar (2015a)'s model and analysis of earnings expectations among NYU undergraduates, connects naturally to Manski (1993)'s motivating framework where youth form expectations about own future earnings based on the schooling decisions and income realizations observed in the previous generation, and to Dominitz and Manski (1996)'s analysis of the relationship between expectations about own earnings and perceptions about population earnings among Wisconsin students.

*Framework setup.* Let $E_{it} = E(y|\Omega_{it})$ be student $i$'s expectation at time $t$ about own earnings at some future date, where $y$ denotes future earnings and $\Omega_{it}$ denotes $i$'s information set at $t$. Student $i$ reports this expectation as

$$E_{it} = E(y|\Omega_{it}) = \int y dG_i(y|\Omega_{it}) = h_i(\Omega_{it}), \tag{9}$$

where $G_i(y|\Omega_{it})$ is $i$'s belief about the distribution of future earnings conditional on the information set, $h_i(\cdot)$ is a scalar-valued function providing the mapping between $i$'s information set and $i$'s expectation, and the major's subscript $j$ is omitted for simplicity.

The student's information set may contain own information (e.g., perceived own ability in different fields) and population information (e.g., the perceived population distribution of earnings). This distinction is key, since Wiswall and Zafar (2015a) provide randomized information about the latter and study its effect on students' revision of own earnings expectations, on which the choice of college major is based. It is therefore useful to partition the information set as $\Omega_{it} = \{\Phi_{it}, \Lambda_{it}\}$, where $\Phi_{it}$ is the student's belief about the population statistic about which the experiment provides information,[118] and $\Lambda_{it}$ captures the remaining components of the information set. Let $\Omega_{i,t+1} = \{\Phi_{i,t+1}, \Lambda_{it}\}$ denote $i$'s new information set after the information treatment and $E_{i,t+1}$ denotes $i$'s new expectation

---

[117]Microeconometric models of dynamic discrete choice have traditionally maintained rational expectations. However, recent contributions have shown that identification without rational expectations is feasible; e.g., see An et al. (2021). Bridging the emerging literature on dynamic discrete choice modelling without rational expectations with the literature on survey expectations appears a promising avenue for future research.

[118]Here assumed to be a scalar for simplicity of exposition. In Wiswall and Zafar (2015a)'s experiment, students learn the distribution of earnings of various subgroups in the US population.

about own future earnings.

Wiswall and Zafar (2015a) consider updating under two conditions.

1. *Condition 1.* The information is new and credible: $\Phi_{it} \neq \Phi_{i,t+}$, so $\Omega_{it} \neq \Omega_{i,t+1}$.

2. *Condition 2.* The information is relevant: $h_i(\Omega'_{it}) \neq h_i(\Omega_{it})$, where $\Omega'_{it} = \{\Phi'_{it}, \Lambda_{it}\}$ and $\Phi'_{it} \neq \Phi_{it}$.

The first condition requires that at least some students have incorrect beliefs about the relevant population statistic and, hence, the information they receive is new to them. It also requires that these students view the received information as credible. The second condition requires that the information is relevant for forming expectations, causing students for whom the information is new to update their expectations about own future earnings. Students for whom conditions 1 and 2 hold update their own earnings expectations: $E_{it} \neq E_{i,t+1}$.

*Forms of updating.* The framework is quite general and can accommodate multiple forms of updating. As noted by Wiswall and Zafar (2015a), how individual students respond to population information will generally depend on $h_i(\cdot)$ and $E_{it}$. The paper considers two forms of updating, each assuming a different form for $h_i(\cdot)$.[119]

1. *Skill price updating.* The student believes that earnings are the product between the individual's level of skill and the unit price skill. The students know their level of skill, but is uncertain about the price of skill. Discovering that mean earnings in the population are higher than initially believed leads the student to revise upward his/her belief about the price of skill in the economy and, hence, also his/her own earnings expectation. The magnitude of the revision is given by the product between the belief error about the population earnings and the ratio of the student's belief about his/her own skill level and the population average.

2. *Relative skill updating.* The student believes that earnings are based on the individual's level of skill relative to the population average skill. The student knows his/her level of skill, but is uncertain about the population average skill. Discovering that mean earnings in the population are higher than initially believed leads the student to revise his/her belief about the average level of skill in the economy upward, and his/her relative position downward. The direction of the revision in own earnings expectations depends on the relative rate at which earnings are related to the individual's skills vis-à-vis the population skill level. The magnitude of the revision depends on the magnitude of the belief error.

The paper does not take a stand on the form of updating followed by students. It investigates the heterogeneity in students' updating empirically with a minimal structure.

*Econometric analysis of expectations updating.* To study the relationship between errors in population beliefs and revisions in own earnings expectations, Wiswall and Zafar (2015a) estimate a series of regressions of the form

$$E_{ij,t+1} - E_{ijt} = D_i^{Control} + D_i^{SpecT} + D_i^{GenT} + \beta_1(Error_{ij} \times D_i^{SpecT})$$
$$+ \beta_2(Error_i \times D_i^{GenT}) + \varepsilon_{ij}. \quad (10)$$

The dependent variable is the revision between before and after treatment in student $i$'s expectation about own earnings at age 30 if $i$ were to graduate from major $j$. The dummy $D_i^{Group}$ equals 1 if $i$ is assigned to treatment group $Group$, and equals 0 otherwise, where $Group \in \{Control, SpecT, GenT\}$ and where $SpecT$ denotes specific treatments and $GenT$ denotes general treatments. Students assigned to a specific treatment were given information about population earnings in a specific subpopulation of the US (e.g., 30-year old workers of a given gender who graduated from a specific major), whereas students assigned to a general treatment were given information about population earnings in the

---

[119]For a formal treatment of these cases, see Wiswall and Zafar (2015a)'s Appendix A.

general population of 30-year old workers (all or college graduates). $Error_{ij}$ denotes $i$'s belief error about population earnings in major $j$ (defined as true minus perceived population earnings), so it refers to specific treatments. $Error_i$ refers to the belief error about population earnings in general treatments. The parameters of interest are $\beta_1$ and $\beta_2$, expected to be different from zero under learning.

To explore heterogeneity in updating, Wiswall and Zafar (2015a) further estimate variants of regression (10) investigating whether updating varies as a function of the type of information provided, the respondent's personal characteristics, or the respondent's initial degree of uncertainty.

*Characterizing the heterogeneity in updating.* To characterize the observed heterogeneity in updating of students' expectations about own earnings, Wiswall and Zafar (2015a) construct a quasi-Bayesian benchmark of the following form:

$$E_{i,t+1}^B = \omega_i E_{it} + (1 - \omega_i)\Phi_{i,t+1}. \tag{11}$$

$E_{i,t+1}^B$ denotes the quasi-Bayesian posterior.[120] $E_{it}$ denotes the prior. $\Phi_{i,t+1}$ denotes the new information provided by the experiment. $\omega_i \in [0,1]$ denotes the weight student $i$'s places on the prior belief relative to the new information, formally:

$$\omega_i = \frac{V(E_{it})^{-1}}{V(E_{it})^{-1} + V(\Phi_{i,t+1})^{-1}}, \tag{12}$$

where $V(E_{it})$ is the variance of the prior and $V(\Phi_{i,t+1})$ is the variance of the information. The relative weight placed on the information is $V(E_{it})/V(\Phi_{i,t+1})$, so the responsiveness to the information is directly proportional to the uncertainty in the prior beliefs. A value $\omega_i \neq 1$ or $V(\Phi_{i,t+1}) < \infty$ corresponds to the case in which the respondent finds the information new and credible (condition 1).

Quasi-Bayesian updating places restrictions on the $h_i(\cdot)$ function in (9). It implies that $h_i(\cdot)$ is linear and separable in the population information, $\Phi_{i,t+1}$, and assumes a particular weight for it. At the same time, the updating model in (11) is standard for any variable that has a distribution belonging to the generalized beta family (including normal and lognormal distributions). Controlling for the precision of the respondent's prior and of his/her signal (information), the quasi-Bayesian benchmark can be thought of as a "one-for-one updating model", which provides a testable implication.

Wiswall and Zafar (2015a) investigate how the heterogeneity in updating observed in the data compares to the quasi-Bayesian benchmark. To do so, they compare the actual revised expectation about own future earnings, $E_{i,t+1}$, with the quasi-Bayesian posterior, $E_{i,t+1}^B$, constructed using the prior, $E_{it}$, and the information treatment, $\Phi_{i,t+1}$. They use the data on individual beliefs about the distribution of own future earnings to estimate the person-specific weights, $\omega_i$, and the underlying variances, $V(E_{it})$ and $V(\Phi_{i,t+1})$. The empirical distribution of population earnings is used to estimate $V(\Phi_{i,t+1})$.

The paper characterizes the updating heuristics used by the students by classifying each student into one of five updating types, depending on how the student's observed posterior compares with his/her benchmark posterior. A student is classified as (1) a unit updater, if his/her posterior is within a certain band around the benchmark posterior; (2) an alarmist, if his/her response is more exaggerated than implied by the benchmark; (3) a conservative, if the respondent updates in the right direction but less than the benchmark; (4) a contrary, if the updating is in the opposite direction than prescribed by the benchmark model; and (5) a nonupdater, if there is no response to the information.

While in a textbook Bayesian framework any form of non-Bayesian updating would be considered irrational, this is not the case here. A student who already knows the information provided by the treatment, or does not view it as relevant for forming own

---

[120] Wiswall and Zafar (2015a) use the term *quasi-Bayesian benchmark* because their study design does not match that of typical experimental studies where subjects are given noisy signals about the same quantity over which belief updating is analyzed.

expectations, would not revise his/her expectations. Moreover, as pointed out by Wiswall and Zafar (2015a), students may rationally revise their expectations in the opposite direction than their population belief errors, if the information lead them to primarily update their beliefs about own relative ability.

### 5.1.2  Extensions

Wiswall and Zafar (2015a) establish that college students' expectations about own future earnings are causally linked to the students' beliefs about the distribution of earnings in relevant populations, so the former are formed at least partially based on the latter. The experimental design enables the authors to credibly study the process by which students update their earnings expectations in response to known shocks to each student's information set. Similarly, Hastings et al. (2015) investigate the effects of learning about program costs and population earnings on the cost and earnings expectations and college choices of Chilean college applicants within the context of a randomized disclosure policy, which provides students with information about population costs and earnings.

In real life, students form earnings expectations over time, as they acquire information about population earnings and other relevant inputs (e.g., own ability, tastes, labor market conditions other than earnings). The amount of uncertainty about future earnings that students perceive at the time of important decisions such as college enrollment, and how quickly students expect the uncertainty to resolve, will affect their decisions. Similarly, subsequent college decisions such as dropping out, selecting a major, or switching major, will depend on how perceived uncertainty evolves over time and the reasons behind its resolution (or lack thereof). These are important issues whose investigation requires high-frequency measurement of students' belief distributions about own future earning and other outcomes over time, as carried out in the Berea Panel Study (BPS).

Exploiting these and other features of the BPS, Gong et al. (2019) characterize how much earnings uncertainty is present for Berea students at college entrance, how quickly this uncertainty is resolved and, when not completely resolved, how much uncertainty remains at the end of college. Additionally, they investigate how uncertainty about college GPA, uncertainty about college major, and their resolution during college influence the resolution (or lack thereof) of income uncertainty.

Gong et al. (2019) builds on and extends earlier analyses of expectations formation and learning about monetary and nonmonetary outcomes among college students, including Stinebrickner and Stinebrickner (2012, 2014a,b). The latter are noteworthy, as they formally connect students' learning about own grade performance/academic ability to specific college choices; namely, college dropout (Stinebrickner and Stinebrickner, 2012, 2014b) and major selection (Stinebrickner and Stinebrickner, 2014a).

In a different choice context, that of high school track choice among Italian families, Giustinelli and Pavoni (2017) investigate the extent and evolution of limited awareness about choice alternatives and belief ambiguity about specific choice-related outcomes. They interpret their data through the lens of a formal framework of learning with limited choice awareness and belief ambiguity applied to high school track choice.

### 5.2  Survey Elicitation

Econometric analysis of expectations formation and learning requires panel data on expectations. Panel data on expectations may be obtained by collecting expectations at regular time intervals from the same subjects over time, or by embedding a randomized information treatment in a survey of expectations so as to generate an experimental panel of expectations. Stinebrickner and Stinebrickner's BPS provides a template for the first approach. Wiswall and Zafar's NYU Study provides a template for the second approach.[121] Building on and extending the design of Dominitz and Manski (1996)'s proof-of-concept

---

[121]An alternative but not necessarily mutually exclusive possibility is to elicit expectations under alternative hypothetical scenarios that provide hypothetical information (e.g., Giustinelli et al. (2022) outside education).

study in multiple directions, the BPS and the NYU Study represent the state of the art in theory-motivated survey elicitation of students' expectations. This subsection summarizes the main distinctive features of the two studies.

*The Berea Panel Study (BPS).*    The BPS is a longitudinal study that surveyed the 2001 and 2002 freshman cohorts of Berea College ten-to-twelve times a year throughout college and yearly, for up to ten years, since leaving college. These surveys collected participants' expectations longitudinally on multiple outcomes, including consumption expenses, academic performance, study effort, and college major while in college, and a variety of post-college outcomes related to labor income, marriage, and children. Another key feature of the study is that its design enabled observation of outcome realizations along expectations for the same individuals through the linking of survey responses to detailed administrative data from Berea College.

To understand the motivation, scope, and design of the study, it is important to consider its setting. Berea College is a four-year institution located in central Kentucky, whose mission is to provide an education to students of limited economic resources. By providing a full tuition subsidy and large room and board subsidies to all entering students, Berea attracts especially students from low-income families. However, despite being a no-tuition institution, dropout rates at Berea have been historically high and comparable to those of low-income students at other schools. The PBS was purposely designed to study the determinants of dropout decisions during college, including credit constraints (Stinebrickner and Stinebrickner, 2008*b*), learning about own academic ability (Stinebrickner and Stinebrickner, 2012, 2014*a*,*b*), and learning about future income (Gong et al., 2019).

Clearly, the BPS' ability to observe both expectations and realizations for multiple related outcomes throughout college is a major feature of the study that has enabled descriptive and structural analyses of how students form and update expectations about academic performance and earnings and how such learning and reduction of uncertainty translate into college decisions such as whether or not to drop out, which major to pursue, and whether to change major.

A perhaps equally important feature of the BPS is that it was designed to survey students at specific times, especially those that are most relevant for modelling students' learning process and related choices. Thus, it was crucial to elicit students' beliefs at college entrance before students actually started to take classes and to learn about their tastes and abilities. It was equally crucial to survey students at those times when they received important pieces of information, especially grades.

Because BPS participants were followed for up to a decade after leaving college, the most recent analyses of BPS data have turned attention to post-college outcomes; for example, career and skill mismatch among college graduates (Agopsowicz et al., 2020); marriage, children, and labor supply expectations and outcomes (Gong et al., 2020*a*); income expectations and realizations (Crossley et al., 2021).

The BPS' ability to follow respondents and measure their expectations and realized outcomes repeatedly at regular time intervals for up to ten years after leaving college is, indeed, another major feature of the BPS that will soon enable unprecedented analyses of expectations and outcomes.

*The NYU Study.*    Wiswall and Zafar's NYU Study consists of a set of related online surveys of New York University (NYU) undergraduates, whose design builds on Zafar's earlier study of Northwestern University (NU) students (Zafar, 2011*a*,*b*, 2012, 2013). The first and main survey of the NYU Study was administered to a sample of full-time freshman, sophomore, and junior students, 18 or older, and US citizens, between May and June 2010. Two follow-up surveys were administered to a subsample of the original sample in early 2012 and 2016. The surveys were designed to study the determinants of students' expectations about future earnings (Wiswall and Zafar, 2015*a*) and of students' college major choices (Wiswall and Zafar (2015*b*, 2020) and Patnaik, Venator, Wiswall and Zafar

([2020]).[122] Since NYU is a large, selective, private university located in New York City, the study population is dramatically different from that of the BPS. As discussed by the principal investigators and documented by the students' Scholastic Aptitude Test (SAT) Score, the participants in the NYU study are high-ability students. Hence, the NYU Study complements in multiple ways the BPS.

Upon recruitment through the NYU's Center for Experimental Social Sciences (CESS), the students were sent an online link to the first main survey constructed using the Survey Monkey software. The students were given a few days to respond and were asked to do so in one sitting. The survey took approximately 90 minutes and had multiple sections. Students were not allowed to skip questions or to go back and change their previous answers in the survey. Many of the questions had built-in logical checks. Respondents were paid a fixed compensation for completing the survey, but reporting of beliefs and expectations were not incentivized on the ground that scoring rules tend to induce biased responses when respondents are not risk neutral (e.g., Winkler and Murphy (1970)).

The survey instrument consisted of three stages. In the initial stage, students were asked their prior expectations about own earnings and other outcomes, conditional on completing various majors. Earnings expectations were elicited for the student's first job after college and for later periods at ages 30 and 45, assuming full-time work. Respondents were instructed to consider the possibility they might receive an advanced/graduate degree by age 30, to ignore the effects of price inflation, and to answer on the basis of their own beliefs without using outside information. In addition to expected (average) earnings, students were asked the percent chance that their future earnings will exceed two thresholds ($35K and $85K), whose choice was informed by the population distribution of earnings in the various fields.

In the intermediate stage, students were randomly assigned to one of four information treatments, each revealing statistics about the earnings and labor supply of one of the following US subpopulations: (1) all 30-year-old US workers, (2) all 30-year-old college graduates, (3) 30-year-old female college graduates, and (4) 30-year-old female college graduates. These statistics were calculated by the authors using data from the Current Population Survey and the National Survey of College Graduates. Before receiving the information, students were asked their beliefs about the earnings distribution in the assigned subpopulation. After receiving the information, students were reasked their expectations for own earnings and their population beliefs for majors for which they were not provided information.

In the final stage, students were given all information of the four treatments and were reasked their expectations.

### 5.3 Empirical Evidence

*Earnings.* Taking advantage of the experimental design of their NYU Study, Wiswall and Zafar (2015a) document a positive relationship between students' revisions in own earnings expectations and population errors. Treated students who discover that they underestimated (overestimated) population earnings revise their earnings expectations upward (downward). The estimated mean response is inelastic, with an error of $1K in population earnings resulting in an average revision of $0.18K in expectations. This suggests that expectations about own earnings are only partially linked to publicly available information about population earnings. At the same time, the response is heterogeneous across treatments, with larger revisions in response to the more relevant major-specific information. Students are highly heterogenous in their information-processing heuristics. In the overall sample, 20% of students do not update their expectations. Among those who do, the most common heuristic is unit updater (28% of the sample), followed by alarmist (19%),

---

[122]Related surveys and experiments with NYU undergraduates by the same authors have focused on gender differences in expectations and preferences over majors (Reuben et al., 2017) and over job characteristics (Wiswall and Zafar, 2018).

contrary and conservative (12% each), and undefined (9%).[123] This distribution further varies with students' characteristics and type of information provided. For example, students who have been in college longer are less likely to revise their expectations and less likely to react excessively to information relative to the quasi-Bayesian benchmark (unit updater). Students whose prior population beliefs exceed actual population earnings are more likely to update conservatively. No difference across genders is found. The information treatment also induces revisions in choice probabilities over majors, implying a switch in the model major for about 12% of the students.

In an earlier study of NU undergraduates, Zafar (2011b) examines how students' predictions of starting salaries of recent graduates evolve over time and how they compare to actual earnings of recent graduates, using two waves of data collected in consecutive years. The author finds that prediction errors in the student's own major become smaller over time, indicating learning; whereas prediction errors in non-pursued majors tend to remain constant over time, with students' beliefs about population salaries remaining below actual salaries on average.

As clear from Wiswall and Zafar (2015a)'s study, a thorough analysis of expectations updating for continuous outcomes, such as earnings, requires information on the entire distribution describing students' subjective expectations. While this information is not available in Zafar (2011b)'s data, it is available longitudinally in the BPS. Gong et al. (2019) take advantage of this feature of the BPS data to investigate the extent and evolution of perceived uncertainty about future earnings among Berea students. The authors find that at college entrance, the cross-sectional average of the standard deviation of students' subjective belief distributions about future own earnings at age 28 ranges between $9.6K and $14.1K a year, depending on the approach used. While the importance of this initial earnings uncertainty relative to that of heterogeneity is well below 50%, the uncertainty itself is quite persistent, as about 65% of it is still present at the end of college. Exploiting the richness of information collected in the BPS, the authors show that 15%-18% of the variance representing the uncertainty about earnings at age 28 at college entrance can be attributed to uncertainty about grade performance at college entrance and 11%-17% to uncertainty about college major. The uncertainty about these two factors combined accounts for about 18%-24% of the initial overall uncertainty about future income. This shows that the portion of uncertainty that is resolved during college is largely attributable to students' learning about their academic ability and college major. With respect to the remaining uncertainty, the authors provide evidence that transitory factors, such as search frictions, are likely to play an important role in generating the initial uncertainty about future earnings.

*Academic performance.* Zafar (2011b)'s analysis of expectations updating considers students' expectations about a variety of major-specific binary outcomes; for example, graduating in 4 years, enjoying the coursework, having parental approval about the major, and finding a job upon graduation. The focus of the analysis, however, is on students' expectations about their academic performance, elicited as the probability of graduating with a GPA of at least 3.5. This focus is motivated by the availability of students' GPA realizations, which Zafar (2011b) uses to construct an "information metric" that identifies new information about own academic ability the students receive between survey waves. The author shows that, on average, students are initially overconfident about their academic performance in all majors, but they adjust their short-term predictions in response to the new information sensibly: upward upon receiving very positive news, downward upon receiving very negative news, no adjustment for news in the intermediate positive/negative range. No significant adjustments are observed in long-term GPA expectations. The

---

[123]Recall that a student is classified as (1) a unit updater, if his/her posterior is within a certain band around the benchmark posterior; (2) an alarmist, if his/her response is more exaggerated than implied by the benchmark; (3) a conservative, if the respondent updates in the right direction but less than the benchmark; (4) a contrary, if the updating is in the opposite direction than prescribed by the benchmark model; and (5) a nonupdater, if there is no response to the information.

information metric and the revisions in the expectations about weekly hours spent on coursework are found to be negatively related, suggesting that students view ability and effort as substitutes. A comparison of the updating patterns between GPA expectations in the student's own major and GPA expectations in non-pursued majors reveals that the two are very similar (statistically equal), suggesting that learning has a general component. That is, by learning about match quality in one's own major, students learn also about their match quality in other majors. While the analysis reveals the presence of substantial heterogeneity in the students' information-processing rules, a number of patterns documented in the analysis suggests that students revise their expectations in ways that are broadly consistent with Bayesian learning.

Stinebrickner and Stinebrickner (2014a), too, study the role of learning about one's academic ability and other major-specific factors within the context of college major choice. Motivated by the view that a student's final major is the outcome of a learning process in which the student's learn about the quality of his or her match with each possible major, Stinebrickner and Stinebrickner (2014a)'s analysis takes advantage of the BPS design to provide a characterization of what students entering Berea expect about their final major, to relate these expectations to actual major outcomes, and to investigate the reasons behind the gap between students' initial expectations and their final outcomes. The authors show that at college entrance there is substantial uncertainty in students' expectations about their final major. On average, students' subjective probability of graduating with the major with which they eventually graduate is about 30 percent. The authors additionally show that this is partly due to the fact that students are initially open to majoring in science, but only few students do so. Specifically, about 20% of entering students views science as their most likely outcome, a higher proportion than for any other major. Moreover, students' average perceived likelihood of completing a science degree is 16 percent, similar to that of other majors. Yet, only 7% of students complete a degree in science. To understand the reasons behind the observed gap between students' overoptimistic initial expectations about completing a degree in science and students' final outcomes, the authors develop a model that explicitly accounts for the fact that before making a final decision about one's future major, students will have the opportunity to resolve some of the uncertainty about major-specific factors entering their major-specific utilities, while some uncertainty might remain. The model focuses on two major-specific factors, average semester grade performance and average annual future income. The model is used to examine how expectations over final majors would change if a student's initial expectations about grade performance and future income were replaced (1) by the expectations about grade performance and future income that students hold at later stages in college, or (2) by the expectations about grade performance and future income that correspond to actual distributions of these variables. The authors find that students' initial overoptimism about completing a degree in science can be largely attributed to their beginning school with substantial misperceptions about their likelihood to perform well in science. By examining the process by which students update their expectations about grade performance in each major, the authors show that science is special relative to other majors, in that students learn about their ability in science primarily by taking science classes, whereas substantial learning about one's ability in other majors occurs by taking classes outside those majors.

Stinebrickner and Stinebrickner (2012) and Stinebrickner and Stinebrickner (2014b) investigate the role of learning about academic ability within the context of college dropout, once again viewed as the end result of a process in which students learn about a variety of factors affecting their utility of attending and graduating from college. Stinebrickner and Stinebrickner (2012)'s analysis is motivated by the inability of financial factors and credit constraints to satisfactorily explain the educational attainment of low-income students and by existing differences in dropout rates within low-income students, with female students having a significant completion advantage over male students. Taking advantage of BPS data, the authors examine an alternative explanation based on students' learning about their academic ability through grade performance. They find that the substantial gender difference in dropout in the BPS sample is predicted almost entirely by academic differences

by gender, as measured by both first-year grades and students' expectations about future grades, with male students holding significantly more overoptimistic beliefs about future grades than female students. Learning about grade performance/ability is found to play a prominent role in students' dropout decision: the authors' estimates imply that dropout would be reduced by 41% if no learning occurred about these factors. Entering students, especially males, are too optimistic about their future grade performance due to their being overoptimistic about their ability. Despite the fact that low-performing students update their expectations substantially over time, they remain overoptimistic about their future performance due to their not understating the importance of permanent factors in determining grades. This evidence mitigates the widespread concern among policy makers that students may be leaving college prematurely.

Building on Stinebrickner and Stinebrickner (2012)'s analysis, Stinebrickner and Stinebrickner (2014b) develop a dynamic model of college dropout with learning and estimate it using BPS data. They use the model to quantify the overall importance of students' learning about grade performance for college dropout outcomes and to assess the empirical relevance of alternative mechanisms through which this kind of learning can affect dropout. The model considers three main mechanisms why low-performing students may dropout: (1) they are forced out of college by grade requirements even though they would prefer to stay; (2) they leave because getting low grades lowers their financial returns to staying in college; (3) they leave because getting low grades reduces enjoyability of college. Using the model to simulate a counterfactual in which students hold initially correct expectations about future grade performance, the authors find that about 45% (36%) of dropout in the first two (three) years would disappear and show that the dropouts observed in early years of college is primarily due to students' learning about their academic performance. With regard to the mechanisms through which learning operates, the analysis provides empirical support for (2) and (3), but not for (1). That is, students who perform poorly in the first years leave mainly because staying is not sufficiently valuable, rather than because they are at risk of failing. Poor performance substantially lowers enjoyability of college and affects expectations about post-college earnings. Finally, the authors investigate heterogeneity by gender and show that males are more likely to drop out relative to females because of academic issues rather than because of inherent taste factors. These findings have important policy implications discussed extensively in the paper.

Giustinelli and Pavoni (2017) study the evolution of expectations for future academic performance, along with that of choice awareness, among Italian 8th-graders. After establishing that students' choice awareness/familiarity and expectations ambiguity measured at the time of pre-enrollment are predictive of pre-enrollment decisions, the authors examine the extent of choice awareness (or lack thereof) and expectation ambiguity at the time of choice and their heterogeneity across students' characteristics. Subsequently, they characterize the evolution of students' choice awareness and expectations ambiguity between the beginning of 8th grade and the time of pre-enrollment across different tracks and across respondents' characteristics.[124] The authors find that at the beginning of 8th grade students' awareness about the choice alternatives is incomplete and their expectations about experiencing a regular path in high school are largely ambiguous. These forms of limited knowledge are not uniform across choice alternatives and students' characteristics. They tend to be especially concentrated among students' low-ranked alternatives and among newly introduced alternatives. By the time of pre-enrollment, students' learning is still incomplete. Moreover, learning itself tends to be concentrated on specific alternatives, this time those most highly ranked by students. Both initial knowledge and learning patterns display substantial heterogeneity across demographic and socioeconomic groups. The authors identify five main patterns as a function of students' gender, 8th-grade GPA, mother's education (high SES), father's occupation (low SES), and country of birth. For example, students with a high-educated mother display intensive learning, which is quite differentiated across high school tracks, conditional on a set of controls including the student's GPA. Conversely, students with a father working in a blue-collar occupation display

___

[124]Giustinelli and Pavoni (2019) carry out a parallel analysis using parents' responses.

a more focused learning pattern concentrated on curricula of the technical and vocational (non-general) tracks. Foreign-born students start 8th grade with lower choice awareness and higher expectations ambiguity relative to native students and follow a biased learning pattern whereby their level of expectations ambiguity for curricula of the general track tends to increase over time. The analysis also shows that a sizeable fraction of students changes their preference ranking over choice alternatives over the survey period. For example, close to one fourth of the alternatives that were not ranked first at the beginning of 8th grade (wave 1) becomes top-ranked by the time of pre-enrollment (wave 3). This suggests that the learning process might be crucial for the choice of a sizable fraction of students, as these do not yet know their most preferred alternative at the beginning of 8th grade. The 'selective learner' might become stuck with a coarse view of the world that associates a high degree of ambiguity to initially unattractive options, thus reinforcing undesirability of the latter in later periods.

## 6 Conclusion

The economic literature on survey-elicited probabilistic expectations in education was put in motion nearly 30 years ago by a question lacking empirical answer, "*How do youth infer the returns to schooling?*" (Manski, 1993), and by a small set of pioneering studies such as Dominitz-Manski's study and Stinebrickner-Stinebrickner's BPS, which paved the way for subsequent surveys ad analyses of subjective expectations, both within and outside education. The amounts and varieties of survey measurements, analytic tools, and empirical evidence accumulated since have been extraordinary. Having now taken stock of them, it is natural to ask what they have taught us about Manski (1993)'s initial question. Upon careful review, it appears that they have generated a number of important answers. And yet, in the process of doing so, they have also expanded the scope of the initial question and raised important new ones.

We have learned that students' expectations about own future earnings are causally linked to students' beliefs about population earnings, but only partially determined by them. We have further learned that students form expectations about own future earnings and about the monetary returns, risks, and costs of schooling heterogeneously, but only a part of this heterogeneity can be accounted for by observable factors such as students' characteristics and institutional or market conditions.

In fact, the latter have received much less attention than the former. For example, Arcidiacono et al. (2020) find that a larger share of their study participants ended up in a business occupation than respondents themselves had predicted when they were still Duke undergraduates, while smaller shares ended up in government and law occupations. The authors hypothesize that these differences may be the result of an aggregate shock to the labor market, especially as the Great Recession began in December 2007 and their study surveyed Duke undergraduates in early 2009.[125] This suggests an interesting avenue for future research investigating how expectations of the college premium and of college attendance respond to the business cycle. This may be addressed by collecting expectations longitudinally over the business cycle or by eliciting expectations under alternative business cycle scenarios.

Systematic investigation of earnings expectations across alternative schooling scenarios has shown that students do believe that higher levels of education, or specific fields of study, are more remunerative and, in some cases, more risky than others. Yet, these perceived monetary returns and risks of schooling do not seem to be the major determinants of students' schooling plans and decisions, as nonmonetary returns, risks, and costs are often more or equally important.

An underinvestigated issue concerns the perceived sources of individuals' subjective returns to schooling; namely, the main causal models underlying students' and families'

---

[125]Existing evidence points to a post-Great Recession negative shock, which affected entry into the legal profession and was likely not fully anticipated by prospective graduates.

perceived monetary returns to schooling.[126]

We have learned that students form expectations about the nonmonetary consequences of schooling partly on the basis of their personal experiences and socioeconomic circumstances. Students and their families tend to be overly optimistic about the student's educational attainment, with this tendency being more prominent among male students and low-ability students. Among the latter, substantial learning and uncertainty resolution occurs during school, especially about own ability and monetary labor market returns. However, the learning process remains often incomplete, partly due to students' unobserved tastes for programs and their preferences for nonmonetary aspects of schooling. Similar learning patterns have started to emerge also in settings where the information is exogenously provided to students rather than endogenously acquired by them.

Last but not least, we have just began to learn about the dangers of ignoring family interactions, peer interactions, and the role of teachers in expectations formation and related schooling decisions and paths.

The experience and evidence accumulated so far suggest that measurement-wise we need more longitudinal surveys capable of tracking the evolution of expectations and realizations for multiple related outcomes over a long period of time, like Stinebrickner-Stinebrickner's BPS. We need more surveys embedding randomized field-experimental components, like Wiswall-Zafar NYU study and the Chilean studies by Hastings, Neilson, and collaborators. We need more surveys linkable to administrative schooling data, like the BPS and the Chilean studies. We need more multi-actor surveys of probabilistic expectations enabling the linking of responses across students and parents within families and across students within networks of peers, like the NLSY97 and Giustinelli's studies in Northern Italy. We also need probabilistic measures of expectations from teachers that are linkable to students' and parents' expectations. We need more studies willing to experiment with elicitation formats that enable respondents to express deep uncertainty, like Giustinelli-Pavoni's study. Eventually, we need studies that can combine a number of these features on a larger scale.[127]

We should also not discount the already large and continuously increasing number of surveys that have collected the same or largely overlapping sets of expectations related to education in different populations or using different survey mediums. Systematizing the survey instruments and expectations data already collected and making them more visible and accessible could lead to novel analyses and new findings.

On the modelling and estimation fronts, we need extensions of the current analytical tools that will enable us to employ more systematically survey-elicited expectations to investigate settings with dynamics, interacting decision makers, and deep uncertainty.

On substantive grounds, there are still many fruitful avenues to explore. We have very limited evidence about when, where (from whom), and how individuals involved in making or counseling schooling decisions acquire information; what information they acquire; and how they incorporate the information in their expectations and decision process. We have very limited evidence about who are the relevant actors for different schooling decisions,

---

[126]The two main explanations put forward by economic theory are the human capital theory, positing that education increases individuals' labor market productivity and earnings through skill acquisition (e.g., Becker (1962), Schultz (1963), Mincer (1974)), and the signalling theory, positing that education serves as a mere signalling device enabling employers to identify high-ability workers (e.g., Spence (1973), Stiglitz and Weiss (1990)). A third hypothesis emphasizes the screening or selection role of (higher) education, whereby premia arise because university attendance induces students to resolve the uncertainty about their returns (e.g., Chiswick (1973), Lange and Topel (2006)). To my knowledge, the only analysis investigating this issue is Ehrmantraut et al. (2020), which finds evidence consistent with the signalling hypothesis and inconsistent with the screening or selection hypothesis using German data.

[127]A potential challenge comes from the tension between the opposing strengths and weaknesses of small researcher-led studies and larger-scale multi-purpose studies. The former can afford greater richness of data to study a specific topic in depth, but must pay the price of smaller samples, lower representativeness, and more limited scope and focus, and viceversa. A potential solution may be the pursuit of larger and more representative studies, but with embedded innovation tools such as the experimental modules included in the U.S. Health and Retirement Study (HRS) or the rotating-panel structure of the NYFed's Survey of Consumer Expectations (SCE).

who are the main decision makers, whether these are the same as those who acquire choice-relevant information or shape students' choice set, and whether/how these aspects vary across groups.[128] We have even less evidence about teachers' information, expectations, and objectives when counseling schooling decisions. And we have close to no evidence on the precise or ambiguous nature of the expectations held by the relevant actors and on how all these aspects interact with one another in determining observed schooling decisions and subsequent outcomes.

Similarly to 30 years ago, the path ahead looks both exciting and challenging. But perhaps differently from 30 years ago, the accumulated experience and progress provide some confidence that the enterprise is again feasible and rewarding.

### SUPPLEMENTARY APPENDIX: A List of Surveys That Have Collected Data on Subjective Expectations Related to Education

## A    Introduction

This Supplementary Appendix contains a detailed list of studies that have collected expectations data related to education. For each study, it provides information about the study's institution(s) and principal investigator(s); the target population; the study type, scope, and relevant content; a link to the study's webpage and/or support documentation (when available); and references to relevant research output. Section B deals with studies that have collected survey data on subjective expectations and/or beliefs using a numerical scale of percent chance. Section C covers studies where subjective expectations and/or beliefs were elicited using non-probabilistic or non-numerical formats. In both sections, the studies are listed in chronological order with respect to the year of fielding.

## B    Surveys of Expectations Elicited Probabilistically on a Numerical Scale of Percent Chance

### B.1    Dominitz-Manski's Wisconsin Study

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Jeff Dominitz and Charles F. Manski.
- **Study population:** High school students from Memorial high school in Madison, Wisconsin (U.S.A.) and college undergraduates (including freshmen, sophomores, juniors, and first-term seniors) from the University of Wisconsin-Madison in 1992.
- **Study type:** Cross-sectional study, fielded in the spring of 1992 via a computer-assisted self-administered interview (CASI).
- **Scope:** To elicit students' subjective earnings distributions under alternative scenarios about future schooling, to measure the perceived returns to schooling, and to investigate the relationship between students' expectations about own future earnings and their beliefs about population earnings.
- **Relevant content:** Expectations about future own earnings distribution (median plus probability of earnings falling below/above specified thresholds determined from the subjective median) and beliefs about population earnings distribution (same elicitation), under alternative schooling scenarios; expectations about own schooling.

---

[128]The mismatch between the "normative model" of family decision-making assumed by the NYC Department of Education and school administrators and the quite distinct and heterogeneous decision-making processes followed by actual NYC families documented by Sattin-Bajaj (2014)'s ethnographic work is eye-opening.

- **Documentation:** Dominitz and Manski (1996) provide a detailed description of the study and of the survey instrument.

- **Relevant research:** Dominitz and Manski (1996).


## B.2 National Longitudinal Study of Youth 1997 (NLSY97): Expectations Modules

- **Link to study:** NLSY97 Data Overview and NLSY97 Expectations Overview.

- **Institution(s):** U.S. Bureau of Labor Statistics.

- **Principal investigator(s):** (Of the initial expectations module) Baruch Fischhoff, Andrew M. Parker, Wandi Bruine de Bruin, Julie Downs, Claire Palmgren, Robyn Dawes, and Charles F. Manski.

- **Study population:** For the NLSY97, U.S. adolescents aged 12-17 in 1997. For the first expectations module, a subsample of 15- and 16-year-old respondents.

- **Study type:** Longitudinal study that follows the lives of a representative sample of American youth born between 1980 and 1984, beginning in 1997. Expectations modules appear in Youth Round 1 (1997-1998) for respondents aged 15-16; Youth Rounds 4 (2000-2001) and 5 (2001-2002) for respondents of all ages; Youth Round 6 (2002-2003) for 16-year-old respondents; Responding Parent Round 1 (1997-1998).

- **Scope:** To assess adolescents' risk perceptions in various domains and to study how adolescents manage risks.

- **Relevant content:** The initial module (Youth, Round 1) elicits youth's expectations for the following 18 events. (i) One-year-ahead questions: arrested (whether rightly or wrongly); dead from any cause (e.g., crime, illness, accident); enrolled in a regular school; pregnant or have gotten someone pregnant; seriously drunk at least once; victim of a violent crime; working for pay for $> 20$ hours per week and enrolled; working for pay for $> 20$ hours per week and not enrolled. (ii) By the youth's 20th birthday questions: mothered or fathered a baby; dead from any cause (e.g., crime, illness, accident); received a high school diploma; served time in jail or prison. (iii) By the youth's 30th birthday: earned a four-year college degree; working for pay for $> 20$ hours per week. A subset of these questions are also elicited from the youth's parents in Round 1 and from the youth in later rounds. Additional questions asked in later rounds include: (i) One-year-ahead questions: get married; moved from current residence. (ii) Five-year-ahead questions: arrested (whether rightly or wrongly); enrolled in a regular school; pregnant or have gotten someone pregnant; Seriously drunk at least once; dead from any cause; victim of a violent crime; working for pay for $> 20$ hours per week and enrolled; working for pay for $> 20$ hours per week and not enrolled; get married. (iii) Across entire life: Have any/have more children, how many.

- **Documentation:** Survey instruments and data are available at the following links: NLSY97 Questionnaires and Accessing NLS Data.

- **Relevant research:** Fischhoff et al. (2000), Dominitz et al. (2001), Reynolds and Pemberton (2001), Walker (2001), Domina et al. (2011), Cowan (2017).


## B.3 Wolter's Swiss Study

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Stefan C. Wolter.

- **Study population:** High school and college students in German-speaking Switzerland in 1998.
- **Study type:** Cross-sectional study, fielded via a CASI survey between April and November 1998 among final-year students in two high schools (A-levels) with concentration in economic studies, students of one business college, and three classes of students in the Economics Department of the University of Applied Sciences of Bern.

- **Scope:** To study the role of students' expectations about wages in influencing students' academic and professional decisions and to allow comparisons with findings from existing studies of students' expectations in the U.S. (e.g., Dominitz and Manski (1996) and Betts (1996)).
- **Relevant content:** Expectations about future own earnings distribution (median plus probability of earnings falling below/above specified thresholds determined from the subjective median) and beliefs about population earnings distribution (same elicitation), under alternative schooling scenarios and age scenarios.
- **Documentation:** -
- **Relevant research:** Wolter (2000).

## B.4  Schweri-Hartog-Wolter's Bern-UAS Study

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Juerg Schweri, Joop Hartog, and Stefan C. Wolter.
- **Study population:** Four starting cohorts of university students (1998-2001) in the Economics Department of the University of Applied Sciences in Bern, Switzerland.
- **Study type:** Repeated cross-sections. A CASI survey was administered to four successive student cohorts (1998-2001). Respondents from the 1998 cohort were recruited within Wolter's Swizerland Study (described below).
- **Scope:** To measure the extent of uncertainty in students' wage expectations associated with alternative schooling paths and to test whether individuals expect compensation for wage risk associated with alternative schooling choices.
- **Relevant content:** Expectations about own future earnings and beliefs about population earnings (median earnings, subjective probability of earning at most 80% of the median, and subjective probability of earning at least 120% of the median), under 10 scenarios varying age and education.
- **Documentation:** Excerpt of survey instrument available in Schweri et al. (2011)'s appendix.
- **Relevant research:** Schweri et al. (2011).

## B.5  Berea Panel Study (BPS)

- **Link to study:** Berea Panel Study.
- **Institution(s):** Centre for Human Capital and Productivity (CHCP), University of Western Ontario (Canada).
- **Principal investigator(s):** Ralph Stinebrickner and Todd Stinebrickner.
- **Study population:** Two freshman cohorts entering Berea College, Kentucky (U.S.A.) in 2000 and 2001.

- **Study type:** Multi-cohort longitudinal study. Students were surveyed 10-12 times per year throughout college and annually for up to 10 years after college. The survey data is also linked to administrative data from Berea College.

- **Scope:** To elicit information about students' subjective expectations and outcomes during and after college, especially factors that may influence college dropout decisions and subsequent labor market outcomes.

- **Relevant content:** Expectations and realizations of outcomes regarding students' academic performance, college major, college dropout, nonmonetary benefits of college, labor market, marriage, and children.

- **Documentation:** Excerpts of survey instruments are provided in the appendix of individual papers.

- **Relevant research:** Stinebrickner and Stinebrickner (2003*a*,*b*, 2004, 2008*a*,*b*, 2012, 2014*a*,*b*), Conley et al. (2017), Gong et al. (2019, 2020*a*,*b*, 2021), Agopsowicz et al. (2020), Crossley et al. (2021), Crossley et al. (2022).

## B.6  German Socio-Economic Panel (SOEP): Youth Questionnaire's Expectations Battery

- **Link to study:** SOEP webpage.

- **Institution(s):** German Institute for Economic Relevant Research (DIW Berlin).

- **Principal investigator(s):** The list of principal investigators of the SOEP Core 1984-2019 is available at the SOEP webpage.

- **Study population:** Private households in the Federal Republic of Germany from 1984 to 2019 and the eastern German länder from 1990 to 2019. Since 2000, the study has included a Youth Questionnaire for household members aged 16-17, featuring a battery on expectations about education and labor market outcomes.

- **Study type:** Longitudinal study. Respondents have been surveyed yearly since 1984 via face-to-face interviews with all household members aged 12 years and over.

- **Scope:** To track long-term psychological, economic, societal, and social developments.

- **Relevant content:** The youth questionnaire includes a section eliciting respondents' probabilities for the following future events: (i) being accepted for a desired apprenticeship/place at university; (ii) completing training/university successfully; (ii) finding a job in a desired career; (iii) achieving job-related success; (iv) experiencing long(er) unemployment; (v) being held back in career by family-related reasons; (vi) becoming self-employed; (vii) working in a foreign country; (viii) getting married; (ix) cohabiting with a partner (without being married); (x) having one child; and (xi) having two or more children.

- **Documentation:** Survey instruments are available at the following link: SOEP Survey Instruments. Instructions about data access are provided at the following link: SOEP Data Access.

- **Relevant research:** Kunz and Staub (2020).

## B.7  Madeira MBA Study

- **Link to study:** Survey of Lifetime Income Expectations (Kellogg MBA students, 2004)

- **Institution(s):** -

- **Principal investigator(s):** Carlos Madeira.

- **Study population:** MBA students of the Northwstern University's Kellogg Graduate School of Business.

- **Study type:** Cross-sectional study. Student data was collected by means of an in-school, paper&pencil questionnaire administered by the principal investigator. The sample of 64 MBA students of the Kellogg Graduate School of Business was interviewed in November of 2004 (academic year 2004/2005).

- **Scope:** To study the subjective returns to an MBA degree and the uncertainty around the earnings of MBA students.

- **Relevant content:** Students' probabilistic expectations about their earnings after completion of their degree and in the counterfactual scenario in which they had chosen not to pursue any postgraduate education. The survey elicits students' subjective median earnings and four additional subjective percentiles of students' subjective earnings distributions (70%, 90%, 110%, and 130% of the median). Students' subjective earnings distributions (median and four subjective percentiles around the median) are elicited for three horizons after graduation (1, 5, and 15 years) to measure the returns to the MBA degree for both the short and long term. The survey also elicits students' beliefs about earnings of two past cohorts of Kellogg's MBA students.

- **Documentation:** The questionnaire and data are available at the Mendeley Data repository.

- **Relevant research:** Madeira (2020*b*).


## B.8   Mazza-Hartog-NIBUD Study

- **Link to study:** -

- **Institution(s):** National Institute for Family Finance Information (NIBUD).

- **Principal investigator(s):** Jacopo Mazza and Joop Hartog.

- **Study population:** Final-year high school students in the Netherlands from three types of secondary schools: preparatory middle-level vocational education (VMBO), higher general continued education (HAVO), and preparatory scientific education (VWO).

- **Study type:** Cross-sectional study, fielded as an open Internet survey whose link was made available on the web between March and July 2006. A battery of questions on expectations was added to the regular 2006 NIBUD's questionnaire (only for students in their last high school years).

- **Scope:** To provide empirical evidence on the information and expectations that Dutch high school students possess before their decision on tertiary education participation.

- **Relevant content:** Students' subjective earnings distribution (median earnings, probability that wage exceeds 125% of the subjective median, and probability that wage falls short of 75% of the subjective median), right after graduating from high school and 10 years after graduation; students' intentions about whether/where to continue their studies after high school and the probability of doing so.

- **Documentation:** -

- **Relevant research:** Mazza and Hartog (2011).

**B.9 Hartog-Ding-Liao's Beijing Study**

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Joop Hartog, Xiaohao Ding, and Juan Liao.
- **Study population:** Third- and fourth-year university students of six universities in Beijing in the 2006/2007 academic year.
- **Study type:** Cross-sectional study. In June of 2007, the survey instruments were transmitted to the Students Affairs Department of the six universities, which took care of the survey distribution among third- and fourth-year students with consideration to representativeness in terms of major/gender/family background.
- **Scope:** To investigate whether the intention to acquire further education via a master degree is related to expected earnings and/or anticipated earnings uncertainty.

- **Relevant content:** Expectations about future events and outcomes, including students' subjective earnings distribution (expected earnings, probability of earning less than 75% of stated expected earning, and probability of earning more than 125% of stated expected earning), right after obtaining a bachelor degree and 10 years after graduation (repeated under the scenario of a master degree); students' schooling/work expectations after obtaining a bachelor.
- **Documentation:** -
- **Relevant research:** Hartog et al. (2014).

**B.10 Zafar's NU Study**

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Basit Zafar.
- **Study population:** Sophomores at Northwestern University (U.S.A.), with an intended major in the Weinberg College of Arts & Sciences (WCAS), in the academic year 2006/2007.
- **Study type:** Panel study with two waves of data collection. The baseline survey was conducted between November 2006 and February 2007 in the Kellogg Experimental Lab and was computer-based. The follow-up survey was fielded between November 2007 and February 2008.
- **Scope:** To study (i) the determinants of students' college major choices (among fields and between single and double majors), with a focus on the roles of expectations and utilities; (ii) the gender gap in college major choice; and (iii) how students form expectations about various major-specific outcomes.
- **Relevant content:** Subjective expectations for a variety of major-specific outcomes, spanning during-college and post-graduation outcomes as well as monetary and non-monetary outcomes (e.g., college enjoyment, academic performance, parental recognition, labor market earnings, etc.).
- **Documentation:** Survey excerpts available in the main text or online appendices of individual papers (Zafar, 2011a,b, 2012, 2013). For papers published in journals with replication policies requiring publication of the data and codes used for the analysis, these can be found at the journal site under supplementary materials.
- **Relevant research:** Zafar (2011a,b, 2012, 2013).

### B.11    Giustinelli's Verona Study (VRS)

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Pamela Giustinelli, Paola Dongili, and Diego Lubian.
- **Study population:** 9th-grade students attending any public high school of the municipality of Verona (Italy) in the fall of the 2007/2008 school year and the students' parents.
- **Study type:** Cross-sectional study. Student data was collected by means of an in-school, paper&pencil questionnaire administered by trained enumerators. Parent data was collected by means of an at-home, self-administered, paper& pencil questionnaire.
- **Scope:** To study high school track choice by Italian families, with a focus on the role of uncertainty and child-parent interactions in decision-making.
- **Relevant content:** Actual choices; student and parent probabilistic expectations for main consequences of high school track choice, including outcomes during school (e.g., enjoyment, study effort, school performance, being in school with friends, pleasing parents) and post-graduation outcomes (e.g., flexibility of choice between college and work choice, college enrollment, flexibility of college among college fields, employment); student and parent perceptions of their roles in the decision-making process.

- **Documentation:** Survey instruments available at the author's research page.
- **Relevant research:** Dongili and Giustinelli (2009), Giustinelli (2016), Giustinelli and Manski (2018).


### B.12    Duke College Major and Expectations Survey (DuCMES)

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Peter Arcidiacono (Phase 1&2), V. Joseph Hotz (Phase 1&2), Arnaud Maurel (Phase 2), Teresa Romano (Phase 2), and Songman Kang (Phase 1).
- **Study population:** Male undergraduates of Duke University (U.S.A.) in the academic year 2008/2009.
- **Study type:** Panel study encompassing two survey waves and two administrative-data components from Duke University and LinkedIn. The first wave consists of a computer-based survey, which was conducted on campus between February and April 2009. The second wave consists of an online survey conducted between February and April of 2016 among a subset of wave 1 participants whose contact information was recovered by linking the information on students' name, major, graduation year, and occupation between the Duke Alumni Database and LinkedIn.
- **Scope:** The purpose of the initial survey is the collection of data on students' earnings expectations and subjective assessments of their abilities in chosen and counter-factual majors to estimate a model of college major choice and quantify the importance of these factors in driving students' sorting into majors. The purpose of the follow-up survey and the data linkage with LinkedIn is (i) to assess the predictive validity of students' expectations with respect to their education and labor market outcomes; (ii) to recover ex ante treatment effects of major-occupation combinations on earnings; and (iii) to investigate the relationship between ex ante treatment effects and individual choices within a model of occupation choice with subjective expectations.

- **Relevant content:** The first survey collects information on current or intended major, students' probabilistic expectations about future careers/occupations under alternative scenarios about graduation major, and students' expectations about future earnings under alternative scenarios about major-career combinations. The second survey collects relevant outcomes and updated expectations.

- **Documentation:** The survey instruments are available at the following links: DuCMES Survey 1 and DuCMES Survey 2. For papers published in journals with replication policies requiring publication of the data and codes used for the analysis, these can be found at the journal site under supplementary materials.

- **Relevant research:** Arcidiacono et al. (2012), Arcidiacono et al. (2020).

## B.13 Wiswall-Zafar's NYU Study

- **Link to study:** -

- **Institution(s):** NYU Center for Experimental Social Sciences (CESS).

- **Principal investigator(s):** Matthew J. Wiswall and Basit Zafar.

- **Study population:** Full-time undergraduates at New York University (U.S.A.), who were adult U.S. citizens enrolled in their freshman, sophomore, or junior years in the 2009/2010 academic year and whose email was included in the recruitment list used by the Center for Experimental Social Sciences (CESS) at NYU.

- **Study type:** Study encompassing cross-sectional, longitudinal, and panel-experimental components. The first survey, including an experimental component in the form of a randomized information intervention, was fielded online between May and June 2010. It was followed by a follow-up survey, conducted in early 2012 on a sub-sample of the initial participants. A second separate survey was fielded in May 2012 in the CESS Computer Lab located on the main NYU campus, accompanied by an experimental component.

- **Scope:** The first survey and its follow-up were designed to study (i) the determinants of college major choice, with a focus on separating the roles of expectations and preferences; (ii) the manner in which college students' expectations about own future earnings are related to students' beliefs about population earnings; (iii) the manner in which college students revise their expectations about own future earnings in response to randomly-provided information about population earnings; and (iv) students' perceptions of a range of monetary and nonmonetary returns to human capital, including own earnings, potential spouse's earnings, marriage, and fertility. The second survey was designed to study (i) preferences for workplace attributes, including earnings and earnings growth, work flexibility, and job stability, and (ii) gender differences in college major choices, with a focus on overconfidence, preferences for competitiveness and risk, and earnings expectations.

- **Relevant content:** The first survey and its follow-up elicit students' expectations related to college major choice and its consequences, including future earnings, employment, marriage prospects, potential spousal characteristics, and fertility. They additionally elicit students' beliefs about population earnings and include a randomized information treatment providing students with statistics on earnings and labor supply in relevant groups of the U.S. population. Earnings expectations are elicited multiple times, before and after an information treatment providing information about population earnings. The second survey elicits probabilistic measures of stated preferences over monetary and nonmonetary job attributes. In the experimental component used to study the gender gap in college major choice, experimental measures of overconfidence, risk preference, and competitiveness are also elicited.

- **Documentation:** Survey excerpts are available in the main text or online appendices of individual papers. Complete survey instruments are available from the authors upon request. For papers published in journals with replication policies requiring publication of the data and codes used for the analysis, these can be found at the journal site under supplementary materials (e.g., on the ReStud's webpage for Wiswall and Zafar (2015*b*), on the QJE's webpage for Wiswall and Zafar (2018), etc.).

- **Relevant research:** Wiswall and Zafar (2015*b*,*a*, 2018, 2020), Reuben et al. (2017), Patnaik, Venator, Wiswall and Zafar (2020).

## B.14 Schweri-Hartog Nursing Education Study

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Juerg Schweri and Joop Hartog.
- **Study population:** Healthcare employees in their third and last year of apprenticeship training in fall 2010/winter 2011 in any canton of Switzerland (24 out of 26 cantons covered in practice).
- **Study type:** Two-wave panel study. The baseline survey, adapted from Trede and Schweri (2014)'s earlier one, was paper-and-pencil and fielded during class, with the assistance of the students' teachers. The follow-up survey was administered online in the summer of 2012, one year after the trainees' graduation.
- **Scope:** To study the effect of ex ante expected wages the decision (planned and actual) to pursue a nursing college education in Switzerland.
- **Relevant content:** First survey: wage expectations for the two main career options (working as a healthcare employee and working as a nurse) in five scenarios ((i) work as a fully trained healthcare employee directly after training; (ii) work as a fully trained healthcare employee at age 25; (iii) work as a fully trained healthcare employee at age 35; (iv) work as a registered nurse at age 25; and (v) work as a registered nurse at age 35); probability to each career option looking ahead to the year 2014; background and personal characteristics, including non-cognitive skills, time preference, and risk aversion.
- **Documentation:** Excerpt of survey instrument available in Schweri and Hartog (2017)'s appendix.
- **Relevant research:** Schweri and Hartog (2017).

## B.15 Assessing Perceived Costs and Benefits of Post-High School Opportunities Survey (APCAB)

- **Link to study:** -
- **Institution(s):** University of Washington.
- **Principal investigator(s):** Nick Huntington-Klein.
- **Study population:** High school junior and senior students in King County, Washington in the school year 2011/2012.
- **Study type:** Cross-sectional study, fielded in homerooms, assembly settings, and open environments such as cafeterias, between April and June 2012.

- **Scope:** To evaluate high school students' estimates of the wage return to and costs of different levels of college attainment and the change in the employment rate associated with degree attainment; to compare subjective returns and costs to projections calculated using data on realizations; and to investigate the heterogeneity in subjective returns, costs, and their deviations from realization-based estimates across students' demographic characteristics.

- **Relevant content:** Students' beliefs about wages and employment rate of local population and of education costs over six different levels of education (high school dropout, high school graduate, some college but no degree, 2-year degree, 4-year degree, and advanced degree).

- **Documentation:** Excerpts of the survey instrument are reported in appendix to individual papers. Data used for specific analyses may be available in form of supplementary material on the publication journal's website (e.g., JEBO's website for Huntington-Klein (2015)).

- **Relevant research:** Huntington-Klein (2015, 2016*b*,*a*).


## B.16   briq Family Panel (bfp)

- **Link to study:** briq.

- **Institution(s):** briq (in cooperation with the SOEP of the DIW Berlin for data collection).

- **Principal investigator(s):** Armin Falk and others.

- **Study population:** Families living in the Bonn and Cologne (Germany) in fall 2011 and having children born between September 2002 and August 2004.

- **Study type:** The briq Family Panel (BFP) is an ongoing panel study. Since its start in 2011, 9 waves have been completed and the 10th wave is in preparation, with about 500 families still participating. The study tracks the impact of a mentoring program on children's personality and skill development, and thus have both survey and experimental components. After the baseline survey in 2011, a first follow-up was conducted shortly after the end of the 1-year mentoring intervention, in early 2013. Since fall 2014, yearly follow-ups have followed.

- **Scope:** To track the impacts of social environment and mentoring (the latter provided via a randomized program) on children's personality and skill development, and to assess their role in socioeconomic differences in outcomes (e.g., gender differences in competitiveness and earnings expectations).

- **Relevant content:** Competitiveness of children and their parents; labor market expectations (students' likely net monthly earnings at age 30 if they were to work full-tim in three attainment scenarios: (i) 10th grade, (ii) Abitur, or (iii) a university degree); perceived likelihood of: (i) continuing with formal schooling after 10th grade, (ii) succeeding in obtaining the Abitur if they continue with formal schooling, and (iii) obtaining a university education if they complete the Abitur.

- **Documentation:** -

- **Relevant research:** Boneva, Buser, Falk and Kosse (2022). See also Falk and Kosse (2021) for information on the sampling design, data collection procedures, and details on the mentoring intervention.


## B.17   Giustinelli's Vicenza Study (VIS)

- **Link to study:** -

- **Institution(s):** -

- **Principal investigator(s):** Pamela Giustinelli, Paola Dongili, and Diego Lubian.

- **Study population:** 8th-grade students attending any public junior high school of the municipality of Vicenza (Italy) in the fall of the 2011/2012 school year and the students' parents.

- **Study type:** Short longitudinal study, encompassing four waves of data collected between October 2011 and April 2012. The first two waves were fielded in October of 2011 and December of 2011, before students' pre-enrollment into high school. The third wave was fielded in mid February 2012, close to the deadline for submitting the high school pre-enrollment form. The fourth and last wave was fielded in April 2012. The first three waves include separate surveys for students and parents. The fourth survey was administered to students only. Participating students come from 10 of the 11 public middle schools located inside the municipality of Vicenza at the time of the study. The survey was paper&pencil, self-administered at home, and returned in a sealed envelope.

- **Scope:** To study high school track choice by Italian families, with a focus on the role of uncertainty and child-parent interactions in decision-making.

- **Relevant content:** Actual pre-enrollment choices; student and parent probabilistic expectations for main consequences of high school track choice, including outcomes during school (e.g., enjoyment, study effort, school performance, being in school with friends, pleasing parents) and post-graduation outcomes (e.g., flexibility of choice between college and work choice, college enrollment, flexibility of college among college fields, employment); student and parent perceptions of their roles in the decision-making process. Relative to the VRS, the VIS features a number of innovations, including repeated questions about student and parent activities during the months leading to pre-enrollment choice; repeated questions about student and parent perceptions of choice alternatives; elicitation of subjective probabilities over choice consequences as probability intervals; student and parent perceptions of each other's choice preferences; an extensive battery of questions on the decision-making roles of children and parents, expectation- and realization-based measures of choice vetoing behavior by parents; elicitation of the student's network of friends, the student's subjective expectations about his/her friends' choices, and choice coordination attempts; student and parent perceptions of teacher orientation criteria.

- **Documentation:** Survey instruments are available at the author's research page. For papers published in journals with replication policies requiring publication of the data and codes used for the analysis, these can be found at the journal's website under supplementary materials (e.g., see RED's website for Giustinelli and Pavoni (2017)'s analysis).

- **Relevant research:** Giustinelli and Pavoni (2017, 2019), Giustinelli and Manski (2018), Garbin (2021).

## B.18 Belfield-Boneva-Rauh-Shaw-Kantar Public's UK Study

- **Link to study:** -

- **Institution(s):** Professional survey company Kantar Public UK.

- **Principal investigator(s):** (Of the expectations module) Chris Belfield, Teodora Boneva, Christopher Rauh, and Jonathan Shaw.

- **Study population:** 9th-grade students participating in the Kids Omnibus Panel.

- **Study type:** Cross-sectional study, with a survey and an administrative-data component. The survey was fielded online by the Kantar Public UK survey company in

July/August 2013. Main survey respondents were 9th-grade students participating in the company's Kids Omnibus, but household information was provided by the students' parents. A subsample of survey participants was matched with the National Pupil Database, containing information on students' performance on standardized tests.

- **Scope:** To study (i) the role of beliefs about monetary and nonmonetary benefits and costs of education in students' decisions to pursue further education ("sixth form" and university); (ii) whether students' beliefs about the benefits and costs of education predict students' beliefs about own school performance; and (iii) whether differences in beliefs about the benefits of further education can explain gender and socioeconomic differences in students' beliefs about own school performance.

- **Relevant content:** Students' perceptions about the likelihood of obtaining the grades necessary to pursue further education; students' beliefs about the likelihood of continuing to further education, if they get the necessary grades; students' beliefs about monetary and nonmonetary benefits and costs of education; and time and risk preferences. Beliefs are elicited separately for sixth form and university education using hypothetical investment scenarios.

- **Documentation:** Survey questions are reported in Belfield et al. (2020)'s appendix.

- **Relevant research:** Belfield et al. (2020).

### B.19 DiazSerrano-Hartog-Nilsson-van Ophem-Yang's Multi-Country Study

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Luis Diaz-Serrano, Joop Hartog, William Nilsson, Hans van Ophem, and Po Yang.
- **Study population:** Undergraduate students from the Faculty of Economics and Business at the University of the Balearic Islands (UBI) and the Faculty of Economics and Business at the Rovira i Virgili University (URV) in Catalonia (Spain); from the Faculty of Economics and Business at the University of Amsterdam (Netherlands); and from various majors at the University of Peking (China).
- **Study type:** Multi-country cross-sectional study, consisting of four surveys (one per location) fielded between 2013 and 2015. Some of the surveys were administered online, others in class.
- **Scope:** To estimate the relationships of students' expectations of university graduation and of labor market earnings with their personal abilities and attitudes and to investigate whether (and if so how) they vary across students in different countries.
- **Relevant content:** Students' expectations about own future earnings (median or mean, probability of earning more than 25% above the median/mean, and probability of earning less than 25% below the median/mean), in the current major and in counterfactual ones; students' beliefs about earnings of an average graduate in the respondent's field of study at age 45; probability of graduating under alternative schooling scenarios.
- **Documentation:** Survey questions are provided in Diaz-Serrano et al. (2016)'s appendix.
- **Relevant research:** Diaz-Serrano et al. (2016), Diaz-Serrano and Nilsson (2022) (the latter uses UBI data only).

### B.20 Baker-Bettinger-Jacob-Marinescu's Bay Area Study

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Rachel Baker, Eric Bettinger, Brian A. Jacob, and Ioana Marinescu.
- **Study population:** A sample of undergraduates at two community colleges of the California's Bay Area (U.S.A.) in the academic year 2013/2014, with priority given to students taking classes that had no pre-requisites, fulfilled general education requirements, were not strictly pre-vocational, and enrolled a high percentage of first year students.
- **Study type:** Cross-sectional study, with a survey and an administrative-data component. Classes were recruited through a request sent to course instructors via the departmental representatives to the faculty senate. The survey was administered in person, using a paper&pencil format. The final sample consists of 12 classes whose instructors were willing to use 30-40 minutes of class time for the survey. The classes were in a number of disciplines, including psychology, environmental science, language arts and biology. The survey data were linked to administrative data on the labor market outcomes of community college students.
- **Scope:** To study the accuracy of community college students' beliefs about labor market outcomes in different majors and to examine to what extent community college students factor in labor market outcomes when choosing a degree.
- **Relevant content:** Students' beliefs about the labor market outcomes of graduates from their own college (likelihood of finding a job, expected salaries, and likelihood of choose majors in a range of categories); students' own curricular plans; perceptions of their ability; expected enjoyment; and expected grades.
- **Documentation:** Survey instrument is provided in Online Appendix A of Baker et al. (2017)'s NBER working paper.
- **Relevant research:** Baker et al. (2017, 2018).

### B.21 Conlon's OSU Study

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** John J. Conlon.
- **Study population:** Freshman cohort of 2014/2015 at the Ohio State University (U.S.A.).
- **Study type:** Cross-sectional study with experimental and administrative-data components. The survey, which includes a randomized information intervention, was fielded online in November 2014. The data collected through the survey was linked to students' transcript data from their freshman through their senior year.
- **Scope:** To investigate whether correcting undergraduates' beliefs about salaries by providing them with salary information can affect students' choices of major and classes observed in the transcript data.
- **Relevant content:** Students' beliefs about the average salary of workers who graduated with degrees in various fields collected in the survey, plus transcript information on students' classes and major.
- **Documentation:** Survey instrument is provided in Conlon (2021)'s appendix.
- **Relevant research:** Conlon (2021).

### B.22 Expectations Studies on the Fachkraft 2030

- **Link to study:** -
- **Institution(s):** For Fachkraft 2030, Constata and the Department of Labor Economics at Maastricht University.
- **Principal investigator(s):** -
- **Study population:** University students sand recent graduates from all universities, study fields, and regions in Germany in 2014 and 2015.
- **Study type:** Survey of over 15,000 students and recent graduates, recruited as part of the Fachkraft 2030 (formerly Fachkraft 2020), and fielded online in September 2014 and March 2015. (The Fachkraft 2030 comprises a series of studies on the economic and general living situation of students and young graduates in Germany. The first study, which was carried out by Constata in cooperation with the Department of Labor Economics at Maastricht University, was fielded at the beginning of the 2012/2013 winter semester. Since then, 16 more studies have followed, surveying over 300,000 students across Germany on recurring and sometimes varying topics.)
- **Scope:** To provide large-scale evidence on gender wage expectations and investigate their determinants; to study students' perceived returns to education and whether/how they map in economic theories (e.g., human capital, signalling, etc.).
- **Relevant content:** Wage expectations for counterfactual study trajectories and over the entire prospective working life; expectations about future labor force participation, working hours, child-rearing plans, and wage negotiations; perceived and actual ability, personality, IQ, beliefs and preferences.
- **Documentation:** Selected questions are reported in the Ehrmantraut et al. (2020)'s appendix. Additional information and documentation can be found at the following links: Fachkraft 2020, Fachkraft 2030 (1) , Fachkraft 2030 (2).
- **Relevant research:** Kiessling et al. (2020), Ehrmantraut et al. (2020).

### B.23 Kapor-Neilson-Zimmerman's New Haven Study

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Adam J. Kapor, Christopher A. Neilson, and Seth D. Zimmerman.
- **Study population:** Parents (or guardians) of 8th-grade students enrolled in New Haven (U.S.A.) public schools in the school years 2014/2015 and 2016/2017.
- **Study type:** Two cross-sections collected via computer-assisted in-person interviews (CAPI) in the summers of 2015 and 2017.
- **Scope:** To elicit parental beliefs about their children's admission chances; to investigate accuracy (rationality) of beliefs and whether these affect school application behavior; and to study how welfare outcomes in centralized school choice depend on the assignment mechanism when participants are not fully informed.
- **Relevant content:** Preferences over schools; consideration sets; beliefs about admissions probabilities; sources of information.
- **Documentation:** Survey instruments and data used by Kapor et al. (2020) are available at the following links: Survey (Appendices E and F) and Data.
- **Relevant research:** Kapor et al. (2020).

### B.24  Bleemer-Zafar Study on the Survey of Consumer Expectations (SCE)

- **Link to study:** SCE Main.

- **Institution(s):** Federal Reserve Bank of New York (NYFed).

- **Principal investigator(s):** Of the specific survey modules for the Bleemer-Zafar Study, Zachary Bleemer and Basit Zafar. A complete list of researchers and staff involved in running the SCE can be found at the SCE website, SCE Staff.

- **Study population:** The SCE is a study of household heads in the U.S. The Bleemer-Zafar modules surveyed SCE panelists who were not first-time participant in the January 2015 administration of the SCE.

- **Study type:** The SCE is a nationally representative, internet-based survey of a rotating panel of household heads. Respondents participate in the panel for up to twelve months, with an approximately equal number rotating in and out of the panel each month. The monthly Core survey is supplemented by survey modules on special topics. The Bleemer-Zafar study is composed of two SCE modules with an embedded randomized information intervention. The baseline survey was fielded in January 2015, while the follow-up survey was fielded in March 2015.

- **Scope:** The SCE has been collecting data on respondents' expectations to advance research investigating how consumers form, update, and act on their expectations for a range of economic variables and outcomes. Bleemer-Zafar surveys and embedded field experiment were designed to evaluate the immediate and medium-term effects on respondents' beliefs about college costs and returns of an information experiment providing respondents with objective information about average college returns or costs.

- **Relevant content:** The SCE has been collecting data on respondents' expectations about inflation, job prospects, earnings growth, and other topics. Bleemer and Zafar's SCE modules (baseline and follow-up) elicit respondents' population beliefs (average earnings of current 40-year-olds working full-time conditional on their having or not having a college degree, or average annual total net and sticker cost including room, board, and tuition of four-year public as well as non-profit private universities); self child's beliefs (only for respondents with children under age 18: likelihood that the respondent's oldest child will attend college, beliefs about the child's earnings at age 30 conditional on having or not having a college degree, and beliefs about the annual total cost of the child's college attendance); self friend's child's beliefs (all respondents: likelihood of recommending a four-year college education for a hypothesized 15-year-old child of a friend who is currently performing well in high school; only for respondents without a child under age 18: beliefs about the friend's child's earnings at age 30 conditional on having or not having a college degree).

- **Documentation:** The text of the questionnaires is provided in Bleemer and Zafar (2018)'s appendix. All SCE questionnaires and microdata can be downloaded from the SCE Databank webpage.

- **Relevant research:** Bleemer and Zafar (2018).


### B.25  BOOST2018 Study

- **Link to study:** BOOST2018.

- **Institution(s):** University of Essex's Institute for Social and Economic Research (ISER).

- **Principal investigator(s):** -

- **Study population:** First-year non-returning undergraduates, starting Bachelor's courses at the Colchester campus of the University of Essex (UK) in October 2015.

- **Study type:** Twelve-wave longitudinal study, with a survey and an administrative component. Eligible students took three online surveys per year and participated in one lab session per year over three consecutive years (2015-2018).

- **Scope:** To understand students' life at university, including the issues they face in their daily life, their aspirations, expectations, attitudes, and their outcomes.

- **Relevant content:** The online surveys collect information on students' academic investments (e.g., hours of study), nonacademic investments (e.g., working for pay, participation in volunteering groups, etc.), and their expectations about future academic achievement and labor market outcomes (e.g., earnings, probability of employment and of nonmonetary job attributes). The survey data was linked to the university's administrative records, including information on the student demographics, socioeconomic status, participation rate in their neighbourhood of domicile, grades, timetable of scheduled lectures and classes, and weekly records of attendance.

- **Documentation:** Selected survey questions used for Delavande et al. (2020)'s analysis are described in the appendix of the paper.

- **Relevant research:** Delavande et al. (2020).

## B.26 Ruder-Van Noy's MTurk Study

- **Link to study:** -

- **Institution(s):** -

- **Principal investigator(s):** Alexander I. Ruder and Michelle Van Noy.

- **Study population:** U.S. adults recruited through Amazon's Mechanical Turk (MTurk), with MTurk approval rates above 98%.

- **Study type:** Cross-sectional study with an experimental component. The survey was conducted online through Qualtrics.

- **Scope:** To study the impact on perceived earnings risk and college major preferences of an information intervention that provides individuals with information on the earnings risk of college majors.

- **Relevant content:** Respondents' expectations about the distribution of future earnings of a fictional sophomore student and college major's recommendations for the fictional sophomore. The survey includes four experimental conditions, corresponding to different manipulations of a vignette about a fictional sophomore college student.

- **Documentation:** The survey instrument is provided in Ruder and van Noy (2017a)'s appendix.

- **Relevant research:** Ruder and van Noy (2017a).

## B.27 Ruder-Van Noy's U.S. University Study

- **Link to study:** -

- **Institution(s):** -

- **Principal investigator(s):** Alexander I. Ruder and Michelle Van Noy.

- **Study population:** Undergraduate students from a large U.S. public university with three socioeconomically diverse campuses.

- **Study type:** Cross-sectional study with an experimental component. The survey was conducted online through Qualtrics in 2015.

- **Scope:** To assess (i) how labor market information influences student expectations about postgraduate earnings in different college majors; (ii) how this information relates to student expectations about completing a degree in these different majors; and (iii) how these aspects vary across students' family backgrounds (e.g., low vs high income, first college generation vs not).

- **Relevant content:** Earnings expectations following graduation from each major and the probability of completing a degree in each major. The experiment provides treatment students with labor market information about the median earnings of the U.S. population of college graduates from four-year universities.

- **Documentation:** -

- **Relevant research:** Ruder and van Noy (2017*b*).


## B.28    Boneva-Rauh's UK University Plans Study

- **Link to study:** -

- **Institution(s):** -

- **Principal investigator(s):** Teodora Boneva and Christopher Rauh.

- **Study population:** UK secondary school students aged 13-18 in the school year 2016/2017.

- **Study type:** Longitudinal study with two survey waves and an administrative-data component. The first survey was fielded online in November 2016, followed by a short follow-up in January 2017. The survey data was merged with administrative school data.

- **Scope:** To shed light on the determinants of the observed socioeconomic gaps in university enrollment between first- and continuing-generation students in the UK, with a focus on the role of a variety of perceived monetary and nonmonetary returns and costs.

- **Relevant content:** Students' beliefs about different immediate and later-life returns to university education (e.g., social life, subject interest, stress, parental support, financial standing, and labor market outcomes); students' perceptions about the likelihood of obtaining the required grades to go to university, graduating conditional on enrolling, and obtaining high grades conditional on graduating; students' beliefs about the likelihood that they would work while studying, if they decided to go to university; students' university application choices (from the follow-up survey and only for students in their last year of high school).

- **Documentation:** Survey questions are provided in Boneva and Rauh (2021)'s appendix.

- **Relevant research:** Boneva and Rauh (2021).


## B.29    Boneva-Golin-Rauh's UK Postgraduate Plans Study

- **Link to study:** -

- **Institution(s):** -

- **Principal investigator(s):** Teodora Boneva, Marta Golin, and Christopher Rauh.

- **Study population:** UK university students aged 18-27 in the 2018/2019 academic year, drawn from a survey company's online panel.

- **Study type:** Cross-sectional study, conducted by a professional survey company in the fall semester 2018. Respondents were selected to be representative of full-time university students in England.
- **Scope:** To shed light on the determinants of the observed socioeconomic gaps in post-graduate enrollment between first- and continuing-generation university students in the UK, with a focus on the role of a variety of perceived monetary and nonmonetary returns and costs.
- **Relevant content:** Students' beliefs about different immediate and later-life returns to postgraduate education (e.g., social life, subject interest, stress, parental support, financial standing, and labor market outcomes); students' perceptions about the likelihood of enrolling in postgraduate education and the likelihood that they will obtain the necessary qualifications to enroll in a postgraduate degree (i.e., completing an undergraduate degree and obtaining First-class honors, conditional on completing an undergraduate degree); students' beliefs about the likelihood of graduating, conditional on enrolling in a postgraduate degree; information on how students allocate their time among different activities as well as on students' finances.
- **Documentation:** Survey questions are provided in Boneva, Golin and Rauh (2022)'s appendix.
- **Relevant research:** Boneva, Golin and Rauh (2022).

## B.30 Arteaga-Kapor-Neilson-Zimmerman's Chile&New Haven Smart Platforms Study

- **Link to study:** -
- **Institution(s):** Chilean Ministry of Education (Mineduc) and ConsiliumBots.
- **Principal investigator(s):** Felipe Arteaga, Adam J. Kapor, Christopher A. Neilson, and Seth D. Zimmerman.
- **Study population:** The study encompasses administrative data on centralized school choice in Chile and in New Haven, Connecticut (U.S.A.), and a survey component. The Chilean administrative data refer to all applicants who used the centralized platform between 2018 and 2020. Survey respondents were drawn from school choice applicants in 2020. The New Haven administrative and survey data refer to 2020 applicants.
- **Study type:** Two-country study encompassing survey, administrative, and experimental components. The experimental component consists of two randomized interventions. The first is an interactive pop-up embedded in the application platform, which computes a predicted risk value for each application submitted through the application platform and returns a pop-up warning for all applications with a non-placement risk greater than 30%, encouraging applicants to add more schools to their applications. The second intervention consists of supplemental reminders to applicants who submitted risky applications. These reminders were delivered via text or WhatsApp messages in Chile and via email in New Haven.
- **Scope:** To study the importance of costly search in centralized school choice and to evaluate the effect on no-placement risk of a "smart matching platforms" intervention that provides live feedback on admissions chances.
- **Relevant content:** The survey deals with school-choice topics, including modules about choice preferences, beliefs (e.g., admission chances), and search behavior.
- **Documentation:** The survey instruments are provided in Online Appendices G and H of Arteaga et al. (2021)'s NBER working paper.
- **Relevant research:** Arteaga et al. (2021, 2022).

## C Surveys of Expectations Elicited in Forms Other Than Probabilistic and Numerical

### C.1 Monitoring the Future (MTF)

- **Link to study:** Monitoring the Future.
- **Institution(s):** Survey Research Center (SRC) inside the Institute for Social Research (ISR) at the University of Michigan (UM).
- **Principal investigator(s):** The profiles of principal investigators are provided at the MTF's webpage.
- **Study population:** U.S. students in senior classes in 1975 and U.S. 8th- and 10th-graders in 1991.
- **Study type:** A repeated series of surveys in which the same segments of the population (8th-, 10th-, and 12th-graders; college students; and young adults) are presented with the same set of questions over a period of years to see how answers change over time. The first survey was fielded among senior students from public and private high schools nation-wide in 1975. After that, the survey was repeated yearly. Beginning in 1991, similar surveys of nationally representative samples of 8th- and 10th-graders have been conducted annually. The questionnaires are group administered in classrooms during a normal class period whenever possible, but are occasionally administered in larger groups. Beginning with the class of 1976, a randomly selected sample from each senior class has been followed up biannually after high school on a continuing basis. These respondents receive a mail questionnaire at their home, which they complete and return to MTF.
- **Scope:** Also known as National High School Senior Survey, the MTF has multiple purposes. Among them is to study changes in the beliefs, attitudes, and behavior of young people in the United States with regard to issues such as government and politics, alcohol and other drug use, gender roles, and protection of the environment. The results of the study are meant to aid policymakers' monitoring of society's progress toward national health goals.
- **Relevant content:** MTF collects a wide array of data, including information on students' risk perceptions and risky behaviors. Beginning in 1976, MTF has collected data annually on a national representative sample of 12th-graders, which has consistently asked them about their educational expectations. MTF began surveying 8th- and 10th-graders in 1991, asking them similar questions.
- **Documentation:** Questionnaires and data can be accessed through the University of Michigan's ICPSR Repository at the following link.
- **Relevant research:** Jacob and Wilder (2010).

### C.2 High School & Beyond (HS&B)

- **Link to study:** HS&B Home.
- **Institution(s):** National Center for Education Statistics (NCES).
- **Principal investigator(s):** -
- **Study population:** U.S. 10th- and 12th-graders in 1980, their parents, and teachers.

- **Study type:** Nationally representative longitudinal study. The first wave was administered in 1980 (Base Year) via a Computer Assisted Telephone Interview (CATI) survey. The following waves were fielded in 1982 (First Follow-up), 1984 (Second Follow-up), and 1986 (Third Follow-up). The sophomore cohort was further interviewed in 1992 (Fourth Follow-up) and 1993 (Fifth follow-up). The study contains

information on high school transcripts in 1983 and postsecondary transcripts in 1993.

- **Scope:** To track students' academic growth over time and to study the relationship between early high school experiences and students' subsequent educational experiences in high school and thereafter, in order to aid studies of how family, community, school and classroom factors promote or inhibit student learning.

- **Relevant content:** Among others, expectations about educational attainments.

- **Documentation:** Access to HS&B survey instruments and data requires obtaining (or amending) an NCES restricted data license. For information see Restricted-use Data License Link 1 and Restricted-use Data License Link 2.

- **Relevant research:** Jacob and Wilder (2010).

## C.3 Beginning School Study (BSS)

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Doris R. Entwisle and Karl L. Alexander.
- **Study population:** First-grade children with a disadvantaged background from 20 public elementary schools of Baltimore (U.S.A.) in 1982, their parents, and their teachers.
- **Study type:** Longitudinal study, collecting multiple observations per year from 1982 to 2002. The data were collected by means of face-to-face interviews with the students and self-administered questionnaires for parents and teachers. School records were examined for data on school grades, test scores, and other variables. In later years, most of the parent questionnaires were answered by mail or by phone.
- **Scope:** To observe how urban school children negotiate the beginning of school and subsequent school transitions and, by monitoring the same individuals over time, to explore links to later trajectories of personal and academic development. The study had multiple foci, including retention and dropout; how patterns in the early grades play out over the longer term; and how seasonal learning patterns(school year vs summer), family structure, and work outside school might affect progress in elementary and middle school. An overarching goal of the study was to identify sources of resilience in the participants' lives that enabled them to overcome the effects of poverty and other conditions in order to finish school or training programs, to find steady jobs, to enter stable relationships, to avoid substance abuse, to stay out of trouble with the law, and to attain a sense of personal satisfaction or fulfillment with their lives. A history of the study is provided by Entwisle et al. (2002).
- **Relevant content:** The BSS measures a large set of outcomes, including school achievement (based on standardized test scores and grades given by teachers), students' expectations about own achievement (grades), self-esteem, locus of control, and other non-cognitive characteristics. As participants left middle childhood, measures of dating behavior, vocational planning, substance use and many other adolescence topics were added to the collection. A Young Adult Survey (1998-2002) gathers work and education histories and data on personal relationships, marital status, and parenthood. Contextual data from families and schools were also collected, including school and neighborhood racial mix, classroom-grouping practices, socioeconomic level, and other related areas. Most years, teachers provide information on grading standards, feelings about their jobs and schools, their expectations about the performance of individual children in their classes, ratings of children's personal traits, grouping practices, and so on. Responding parents report about family structure and family changes, their occupations and education, their expectations for their

child's schooling and school performance, their feelings about the school, and their understanding of its evaluation procedures.

- **Documentation:** Questionnaires and data for the first 8 waves of the BSS can be accessed through the Harvard Dataverse Repository's Murray Relevant research Archive Dataverse at the following link.

- **Relevant research:** Madeira (2018, 2020a).


## C.4 Longitudinal Study of East Germany Students

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** -
- **Study population:** Students of two parallel cohorts, one born in 1973 and one born in 1976, observed in East Germany over the period 1985 to 1995.
- **Study type:** Longitudinal study. Students in the younger cohort were surveyed between grade 3 in the academic year 1985/1986 and grade 12 at the end of the academic year 1994/1995 (i.e., between ages 9 and 18), while students in the older cohort were surveyed between grade 6 (in the academic year 1985/1986) and up to three years after grade 12 in 1995 (i.e., between ages 12 and 21).
- **Scope:** To understand the determinants of the development of cognitive abilities of children and teenagers, as well as their economic, social, and political preferences. The study was continued after the German Reunification. A history of the study is provided by Kuhnke (1997).
- **Relevant content:** The study's surveys cover a wide range of topics, including educational achievement and attainment, preferences, family development, social relations and psychological wellbeing measures. The survey includes questions eliciting students' educational (Abitur) plans at several points in time. Outcomes are also observed within the data.
- **Documentation:** The data can be obtained from the Central Archive for Empirical Social Relevant research (University of Cologne).
- **Relevant research:** Azmat and Kaufmann (2021).


## C.5 National Educational Longitudinal Survey 1988 (NELS88)

- **Link to study:** NELS88 Home.
- **Institution(s):** National Center for Education Statistics (NCES).
- **Principal investigator(s):** -
- **Study population:** U.S. 8th-graders in 1988, their parents, teachers, and school principals.
- **Study type:** Nationally representative longitudinal study. The first wave was administered in the 1987-1988 school year (Base Year) via a paper&pencil survey. The following waves were fielded in 1990 (First Follow-up), 1992 (Second Follow-up), 1994 (Third Follow-up), and 2000 (Fourth Follow-up). The study contains information on high school transcripts in 1988 and postsecondary transcripts in 2000.
- **Scope:** To track students at critical transitions as they leave middle or junior high school and progress through high school and into postsecondary institutions or the work force. The data was collected to ensure comparability with NLS72 and HS&B data.

- **Relevant content:** Among others, expectations about educational attainments and related outcomes.

- **Documentation:** Survey instruments and data are available at the following links: NELS88 Questionnaires and NELS88 Data.

- **Relevant research:** Jacob and Wilder (2010), Jerrim (2015).


## C.6  Dutch Continued Education Study (Verder Studeren)

- **Link to study:** Verder Studeren.

- **Institution(s):** -

- **Principal investigator(s):** -

- **Study population:** Dutch students whose higher education started in 1991, representing all years, levels, and types of higher education in the Netherlands (including higher vocational education).

- **Study type:** Panel study, spanning the years 1991-1997 and following students during their educational career (including students' choice among different types of secondary schools and HBO/university enrollment).

- **Scope:** To study issues related to school choice processes, including school careers (drop out, transfer to other courses, educational output); the roles of parents, teachers, and events; student financing; background variables (basic characteristics, place of birth, residence, household characteristics, parental education, parental occupation and employment, household income and capital assets).

- **Relevant content:** The subsample of students starting postsecondary education in 1991 was asked about their position and history in education and their motives for choosing this type of education. These students were also asked an open question about their expected starting salary after graduation. In the following waves, these respondents were asked every year about their position in or outside higher education, the motivation for the decisions made, and their future plans. In the last survey wave in 1995, students were asked about their actual earnings in 1995.

- **Documentation:** A description of the study and of the expectations questions is provided by Webbink and Hartog (2004) and references therein.

- **Relevant research:** Webbink and Hartog (2004).


## C.7  Betts' UCSD Study

- **Link to study:** -

- **Institution(s):** -

- **Principal investigator(s):** Julian R. Betts.

- **Study population:** Undergraduate students across all undergraduate faculties at the University of California, San Diego (U.S.A.) in the academic year 1992/1993.

- **Study type:** Cross-sectional study, fielded over a period of four months beginning in November 1992 using a paper&pencil survey.

- **Scope:** To examine undergraduates' knowledge of population salaries by type of education.

- **Relevant content:** Beliefs about population salaries. These include three sets of questions. First, starting salaries at the time of the survey for workers with a bachelor's degree in chemical, electrical, mechanical, or civil engineering (four questions); a master's degree or Ph.D. in chemical and electrical engineering (four questions);

a bachelor's degrees in chemistry and psychology; an MBA preceded by a technical degree (science and engineering); an MBA preceded by a nontechnical degree. Second, average salaries for engineers at the time of the survey by their highest degree and years of experience. Third, average earnings in 1990 of workers aged 25-34 years, working full-time.

- **Documentation:** The survey instrument is described in Betts (1996)'s appendix.
- **Relevant research:** Betts (1996).

## C.8  Menon's Cyprus Study I

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Maria E. Menon.
- **Study population:** Final-year secondary school students in Cyprus in the school year 1993/1994.
- **Study type:** Cross-sectional study of a representative sample of students (mostly aged 17-18) from seven secondary public school specializations in Cyprus in the academic year 1993/1994.
- **Scope:** To study perceived rates of return to higher education in Cyprus.
- **Relevant content:** Expectations about education and labor market outcomes, including intention to pursue higher education vis-a-vis employment; earnings expectations at three points in time (point of starting work, after four years of work, age 46), with and without a higher education diploma; and expectations of the direct costs students or their family would have to incur if they entered higher education.
- **Documentation:** -
- **Relevant research:** Menon Eliophotou (1997*b*,*a*).

## C.9  Menon's Cyprus Study II

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Maria E. Menon.
- **Study population:** Final-year secondary school students in Cyprus in the school year 2003/2004.
- **Study type:** Cross-sectional study of a representative sample of students (mostly aged 17-18) from secondary school specializations in Cyprus in the academic year 2003/2004.
- **Scope:** To study perceived rates of return to higher education in Cyprus after 10 years from the first study.
- **Relevant content:** Expectations about education and labor market outcomes, including intention to pursue higher education vis-a-vis employment; earnings expectations at three points in time (point of starting work, after four years of work, age 46), with and without a higher education diploma; and expectations of the direct costs students or their family would have to incur if they entered higher education.
- **Documentation:** -
- **Relevant research:** Menon Eliophotou (2008*a*,*b*).

### C.10 Menon's Cyprus Study III

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Maria E. Menon.
- **Study population:** Third- and fourth-year students attending the University of Cyprus (except those studying Medicine).
- **Study type:** Cross-sectional study, with a retrospective component tracing students' academic progress before and during university.
- **Scope:** To study perceived rates of return to higher education in Cyprus and if/how they vary by students' socioeconomic status and educational experiences.
- **Relevant content:** The survey covers three main areas: students' educational experience prior to university entrance (e.g., academic attainment and taking private lessons); students' academic attainment after university entrance (i.e., GPA in the first, second, third, and–when applicable–fourth year); and students' expectations about (un)employment and earnings after graduation. Expected earnings are elicited at three horizons: first job after graduation, after four years of employment, and after 20 years of employment, so as to enable estimation of expected lifetime earnings.
- **Documentation:** -
- **Relevant research:** Menon Eliophotou and Pashourtidou (2017).

### C.11 Carvajal-Bendana-Bozorgmanesh-Castillo-Pourmasiha-Rao-Torres's Florida Study

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Manuel J. Carvajal, David Bendana, Alireza Bozorgmanesh, Miguel A. Castillo, Katayoun Pourmasiha, Priya Rao, and Juan A. Torres.
- **Study population:** 1994 graduates and 1996 senior undergraduates from the College of Business at Florida International University, Miami (U.S.A.).
- **Study type:** Study with multiple cohorts, fielded via a paper&pencil survey between December 1995 and February 1996 among alumni who obtained a baccalaureate degree in 1994, and in the spring of 1996 among senior fourth-year students enrolled in thirty-six senior-level courses.
- **Scope:** To compare the labor market expectations of current students with the labor market experiences of recent graduates and explain potential differences.
- **Relevant content:** Expectations (collected from senior undergraduates) and realizations (collected from alumni) about employment status, type of employer, type of occupation, hours worked per week, location of first job, and annual salary in the first job after graduation.
- **Documentation:** The survey instrument is provided in Carvajal et al. (2000)'s appendix.
- **Relevant research:** Carvajal et al. (2000).

### C.12 Brunello-Lucifora-Winter-Ebmer's European Study

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Giorgio Brunello, Claudio Lucifora, and Rudolf Winter-Ebmer.
- **Study population:** Undergraduate students in economics and business from universities of 10 European countries.
- **Study type:** Multi-country cross-sectional study, implemented via an in-class paper&pencil survey targeting undergraduate students in twenty-six economics and business faculties across 10 European countries, including Austria, Finland, Germany, Greece, Italy, Portugal, Sweden, Switzerland, and UK, during the academic year 1999/2000.

- **Scope:** To study the determinants of wage expectations and expected employment probabilities, the variability of expectations and their variation across countries, and the tradeoff between expected starting wages and expected wage growth.
- **Relevant content:** Expectations about future earnings and employment outcomes under alternative schooling scenarios.
- **Documentation:** Survey instruments and data are available as supplementary materials at the Journal of Human Resources' webpage.
- **Relevant research:** Brunello et al. (2004), Wolter and Zbinden (2005) (the latter analyzes the Swiss data only).

### C.13 Botelho-Costa Pinto's Minho Study

- **Link to study:** -
- **Institution(s):** -
- **Principal investigator(s):** Anabela Botelho and Ligia Costa Pinto.
- **Study population:** Undergraduate students (freshmen and seniors) from the College of Business and Economics at the University of Minho (Portugal) in the academic year 2000/2001.
- **Study type:** Cross-sectional study with an experimental component, conducted on campus in April 2001.
- **Scope:** To compare students' subjective beliefs about own and population economic returns to college education, based on elicitation of expectations about own and population earnings under multiple work experience scenarios. The experiment was designed to evaluate the effect of financial incentives for accurate reporting of beliefs about current population earnings.
- **Relevant content:** Beliefs about current population earnings and expectations about future own and population earnings under alternative scenarios about work experience.
- **Documentation:** -
- **Relevant research:** Botelho and Pinto (2004).

## C.14 Educational Longitudinal Survey 2002 (ELS2002)

- **Link to study:** ELS2002 Home.

- **Institution(s):** National Center for Education Statistics (NCES).

- **Principal investigator(s):** -

- **Study population:** U.S. 10th-graders in 2002 and 12th-graders in 2004, their parents, and teachers.

- **Study type:** Nationally representative longitudinal study. The first wave was administered in 2002 (Base Year) via a paper&pencil survey. The following waves were fielded in 2004 (First Follow-up), 2006 (Second Follow-up), and 2012 (Third Follow-up). The study contains information on high school transcripts in 2005 and postsecondary transcripts in 2013.

- **Scope:** To evaluate students' trajectories from the beginning of high school into postsecondary education, the workforce, and beyond. The ELS2002 is the fourth in a series of school-based longitudinal studies, after the NLS72, HS&B, and NELS88. All of these studies deal with the transition of American youth from secondary schooling to subsequent education and work roles. The ELS2002 looks back to the previous three studies and builds on them to guarantee basic comparability. At the same time, it updates the content of the survey and extends the time line by a decade.

- **Relevant content:** Among others, expectations about educational attainments (from students, parents, and teachers) and job occupation right after high school and at 30 (from students only).

- **Documentation:** Survey instruments and data are available at the following links: ELS2002 Questionnaires and ELS2002 Data.

- **Relevant research:** Jacob and Wilder (2010), Gershenson et al. (2016), Papageorge et al. (2020).

## C.15 High School Longitudinal Study of 2009 (HSLS09)

- **Link to study:** HSLS09 Home.

- **Institution(s):** National Center for Education Statistics (NCES).

- **Principal investigator(s):** -

- **Study population:** U.S. 9th-graders in 2009, their parents, their math and science teachers, school administrators, and school counselors.

- **Study type:** Nationally representative longitudinal study. The first wave was administered in 2009 (Base Year). The following waves were fielded in 2012 (First Follow-up), and 2016 (Second Follow-up). The study contains information on high school transcripts in 2013-2014 and postsecondary transcripts in 2017-2018. It also includes a student assessment in algebraic skills, reasoning, and problem solving for 9th and 11th grades.

- **Scope:** By beginning with a 9th-grade cohort and including information collected from parents and school counselors, the HSLS09 aims to provide new data for study of how incoming 9th-graders are led into early math and science courses that can affect future coursework necessary for STEM career progress.

- **Relevant content:** Among others, students' expected number of years of math and of science they will study during high school, students' educational plans and attainment expectations (asked also of parents), and students' expectations for their job or occupation at 30.

- **Documentation:** Survey instruments and data are available at the following links: HSLS09 Questionnaires and HSLS09 Data.

- **Relevant research:** Holian (2020), Holian and Kelly (2020).

### C.16  Kerr-Pekkarinen-Sarvimäki-Uusitalo's Finnish Study

- **Link to study:** -

- **Institution(s):** -

- **Principal investigator(s):** Sari Pekkala Kerr, Tuomas Pekkarinen, Matti Sarvimäki, and Roope Uusitalo.

- **Study population:** Final-year high school students in Finland in school year 2011/2012, before application to post-secondary education.

- **Study type:** Cross-sectional study, with a survey and an experimental component. The study was conducted between September 2011 and January 2012 and consisted of a paper&pencil survey, plus a randomized information package and lecture material regarding the differences in earnings distributions and employment rates following graduation from different post-secondary degrees.

- **Scope:** To evaluate the impact on post-secondary education applications and enrollment of a randomized intervention providing high school students with information about labor market outcomes of graduates from alternative post-schooling degrees.

- **Relevant content:** The survey elicits students' preferences and expectations about education and labor market outcomes, as well as their impressions about the information treatment received. The information treatment includes a booklet reporting the distribution of earnings, employment rates, and the most common occupations among the current population of 30–34-year-old persons by 104 most common post-secondary degrees. Students are also given the supplementary material to consult at home. The experiment was implemented 5–6 months before the students applied to post-secondary programs.

- **Documentation:** The data used for Kerr et al. (2020)'s analysis is available as supplementary material on the Labour Economics website. A description of the main survey questions is provided by Kerr et al. (2020).

- **Relevant research:** Kerr et al. (2020).

### C.17  Saarland University Study

- **Link to study:** -

- **Institution(s):** -

- **Principal investigator(s):** Aderonke Osikominu, Gregor Pfeifer, and others.

- **Study population:** Students in the process of applying to Saarland University (Germany) in 2011 and 2012.

- **Study type:** Multi-cohort study, whose surveys were fielded through an online URL in 2011 and 2012.

- **Scope:** To elicit students' expectations of future gross salaries for themselves and other students in multiple ways and to evaluate the determinants of students' estimation errors, especially but not exclusively in relation to students' understanding of progressive income taxation.

- **Relevant content:** Students' prospective field of study, whether they aspire to obtain an additional degree after university (master, second state examination, or a doctoral degree), and with which of these degrees they aim at earning their first salary. Two different types of questions about students' expectations for monthly gross salaries. Students' potential future income and profession. Their personal and family background. Students' estimate of net equivalents of given gross salaries (1,500, 3,000, and 4,500 Euros). The first type of salary question asks students to estimate future monthly salary for themselves and for other (average) students within the same field (corresponding to the degree with which the student intends to earn their first salary). These expectations are asked with reference to salaries at labor market entry and after five years on the job. In the second type of questions, students are asked to estimate average monthly gross salaries for other students in different fields of study (business studies, education, humanities, law studies, mathematics/computer science, medicine, and natural sciences). In the latter case, expectations are elicited without conditioning on a degree (i.e., bachelor, master, etc.). Before reporting their expectations, students are provided a detailed explanation about the difference between gross and net salaries. Students' estimates of net equivalents of given gross salaries are used to measure students' understanding of the income tax system, to correct their salary estimates, and to reevaluate their knowledge of future wages.

- **Documentation:** -

- **Relevant research:** Osikominu and Pfeifer (2018), Klossner and Pfeifer (2019), Briel et al. (2022).

## C.18   Peter-Zambre Study on the Berliner-Studienberechtigten Panel (Best Up)

- **Link to studys:** Best Up webpage, Best Up on DIW webpage, and Best Up on WZB webpage.

- **Institution(s):** German Institute of Economic Relevant Research (DIW) and Berlin Social Science Center (WZB).

- **Principal investigator(s):** Of Best Up, C. Katharina Spieß and Heike Solga. Of the current study, Frauke H. Peter and Vaishali Zambre.

- **Study population:** Students in upper secondary schools of Berlin (Germany) with a high share of nonacademic family backgrounds.

- **Study type:** Best Up is a longitudinal study with five survey waves, conducted yearly between 2012 and 2016. Peter and Zambre's project consists of a randomized field experiment conducted as a part of Best Up's wave 1 (2012) and wave 2 (2013).

- **Scope:** To investigate the causal relationship between information and educational expectations. The intervention provided high school students with randomized information on the benefits of university education and on different funding possibilities.

- **Relevant content:** Expectations and realizations about education and labor market outcomes.

- **Documentation:** -

- **Relevant research:** Ehlert, Peter, Finger, Rusconi, Solga, Spiess and Zambre (2017); Ehlert, Finger, Rusconi and Solga (2017), Peter et al. (2017), Zambre (2018), Peter et al. (2021), Berkes et al. (2022), Leibing et al. (2022).

## C.19    Hastings-Neilson-Ramirez-Zimmerman-Mineduc's Studies

- **Link to study:** -

- **Institution(s):** Chilean Ministry of Education (Mineduc).

- **Principal investigator(s):**  Justine S. Hastings, Christopher A. Neilson, Anely Ramirez, and Seth D. Zimmerman.

- **Study population:**  The study encompasses two sets of surveys, the Proyecto 3E pilot survey and the Tarjeta Nacional Estudantil (TNE) survey. The sample for the Proyecto 3E pilot survey was drawn from the population of the 2012 *Prueba Seleccion Universitaria* (PSU) registrants (i.e., college applicants) who graduated from high school in 2010 or later. The sample for the first wave of the TNE survey was drawn from the population of first-year postsecondary students in 2012. The two subsequent waves of the TNE survey further drew participants from the population of students who graduated from high school in 2011 or earlier and either registered for the PSU or applied for state-subsidized financial aid in 2012.

- **Study type:** Multi-survey study encompassing administrative-data and experimental components.  Participants in the Proyecto 3E pilot survey were recruited via email invitations sent from the Chilean testing authority (DEMRE) in two waves, the first before the PSU exams in early December 2011 and the second after the PSU exams.  Proyecto 3E was fielded as a test run for an informational intervention conducted the following year.  The intervention was a randomized information treatment disclosing information about college costs and monetary returns to college education. Participants in the Tarjeta Nacional Estudantil (TNE) survey were initially recruited from the population of first-year postsecondary students in 2012 via email invitations from school administrators. In the two subsequent waves, additional respondents from the population of students who graduated from high school in 2011 or earlier and either registered for the PSU or applied for state-subsidized financial aid in 2012 were recruited via emails sent through DEMRE. The first wave of TNE was conducted between March and December 2012. The administrative data include education outcomes from administrative records and labor market outcomes from tax returns.

- **Scope:** To evaluate students' beliefs about college costs and monetary returns to college against observed benchmarks; to investigate the determinants of students' degree program choice; to track enrollment choices in the years following the intervention.

- **Relevant content:** The Proyecto 3E pilot survey includes questions about students' institution and career choices, the main factors students considered when making these choices, their sources of information about degree programs, and their expectations about college costs and major-and-institution-specific earnings for themselves, for graduates with similar characteristics to their own, and for the average graduate. The TNE survey questions deals with informed decision-making, resources used to finance studies, knowledge of loan terms, and financial literacy.

- **Documentation:**  The survey instruments are provided in Online Appendix B of Hastings et al. (2015)'s NBER working paper.

- **Relevant research:** Hastings et al. (2015, 2016).


## C.20    Tincani-Kosse's Study

- **Link to study:** -

- **Institution(s):** University College London, the Karta Initiative, and the Focus Data Collection Agency, with the external support of the Chilean Ministry of Education (Mineduc).

- **Principal investigator(s):** Michela Tincani and Fabian Kosse.
- **Study population:** Chilean 12th graders attending 128 disadvantaged schools, 64 assigned to treatment and 64 assigned to control within the Chilean Government PACE program.
- **Study type:** Multi-component study, including an experimental component (the 2016 extension of the PACE program), an administrative data component, and a survey component. The survey was fielded between September and November 2017 when participants attended 12th grade.
- **Scope:** To study the role of belief biases in students' college admission and enrollment distortions in Chile and how those interact with preferential admission policies.

- **Relevant content:** The survey elicits students' pre-college achievement (via a 20-minute mathematics achievement test); students' study effort and subjective beliefs about future outcomes and returns to effort (via survey questions); additional information from mathematics and Spanish teachers as well as from school principles (via survey questions).
- **Documentation:** Descriptions of the fieldwork and of selected survey questions are provided in Tincani et al. (2021)'s appendix.
- **Relevant research:** Tincani et al. (2021).

# References

Abdulkadiroglu, A. and T. Sonmetz (2003), 'School Choice: A mechanism design approach', *American Economic Review* **93**(3), 729–747.

Agarwal, N. and P. Somaini (2018), 'Demand Analysis Using Strategic Reports: An application to a school choice mechanism', *Econometrica* **86**(2), 391–444.

Agopsowicz, A., C. Robinson, T. Stinebrickner and R. Stinebrickner (2020), 'Careers and Mismatch for College Graduates: College and Non-college Jobs', *Journal of Human Resources* **55**(4), 1194–1221.

Altonji, J.G. (1993), 'The Demand for and Return to Education When Education Outcomes are Uncertain', *Journal of Labor Economics* **11**(1), 48–83.

Altonji, J.G., P. Arcidiacono and A. Maurel (2016), The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects, *in* E.Hanushek, S.Machin and L.Woessmann, eds, 'Handbook of Economics of Education', Vol. 5, Elsevier B.V., chapter 7, pp. 305–396.

An, Y., Y. Hu and R. Xiao (2021), 'Dynamic Decisions under Subjective Expectations: A structural analysis', *Journal of Econometrics* **222**(1C), 645–675.

Arcidiacono, P. (2004), 'Ability Sorting and the Returns to College Major', *Journal of Econometrics* **121**(1-2), 343–375.

Arcidiacono, P., V.J. Hotz, A. Maurel and T. Romano (2020), 'Ex Ante Returns and Occupational Choice', *Journal of Political Economy* **128**, 4475–4522.

Arcidiacono, P., V.J. Hotz and S. Kang (2012), 'Modeling College Choices Using Elicited Measures of Expectations and Counterfactual', *Journal of Econometrics* **166**(1), 3–16.

Ariga, K., G. Brunello, R. Iwahashi and L. Rocco (2005), Why Is the Timing of School Tracking So Heterogeneous?, Discussion Paper 1854, IZA.

Arteaga, F., A.J. Kapor, C.A. Neilson and S.D. Zimmerman (2021), Smart Matching Platforms and Heterogeneous Beliefs in Centralized School Choice, Working Paper 28946, NBER.

Arteaga, F., A.J. Kapor, C.A. Neilson and S.D. Zimmerman (2022), 'Smart Matching Platforms and Heterogeneous Beliefs in Centralized School Choice', *Quarterly Journal of Economics* **137**(3), 1791–1848.

Attanasio, O. and K.M. Kaufmann (2014), 'Education Choices and Returns to Schooling: Intrahousehold Decision Making, Gender and Subjective Expectations', *Journal of Development Economics* **109**, 203–216.

Attanasio, O., T. Boneva and C. Rauh (2020), 'Parental Beliefs about Returns to Different Types of Investments in School Children', *Journal of Human Resources* **Forthcoming**.

Aucejo, E., J. French and B. Zafar (2021), Estimating Students' Valuation for College Experiences, Working paper, Arizona State University.

Aucejo, E., J. French, M.P. Ugalde Araya and B. Zafar (2020), 'The Impact of COVID-19 on Student Experiences and Expectations: Evidence from a Survey', *Journal of Public Economics* **191**, 104271.

Azmat, G. and K.M. Kaufmann (2021), Formation of College Plans: Expected Returns, Preferences, and Adjustment Process, Working paper.

Baker, R., E. Bettinger, B. Jacob and I. Marinescu (2017), The Effect of Labor Market Information on Community College Students' Major Choice, Working Paper 23333, NBER.

Baker, R., E. Bettinger, B. Jacob and I. Marinescu (2018), 'The Effect of Labor Market Information on Community College Students' Major Choice', *Economics of Education Review* **65**, 18–30.

Barrows, S., M.R. Busemeyer, M.A. Cattaneo, P. Dolton, J.L. Garritzmann, M.B. Henderson, A. Kilibarda, P. Lergetporer, O. McClellan, P.E. Peterson, R.Y. Shapiro, S. Sinozich K. Werner, M.R. West, L. Woessmann and S.C. Wolter (2021), *Public Opinion and the Political Economy of Education Policy around the World*, CESifo Seminar Series, MIT Press.

Becker, S.G. (1962), 'Investment in human capital: A theoretical analysis', *Journal of Political Economy* **70**, 9–49.

Beffy, M., D. Fougere and A. Maurel (2012), 'Choosing the Field of Study in Postsecondary Education: Do expected earnings matter?', *Review of Economics and Statistics* **94**, 334–347.

Belfield, C., T. Boneva, C. Rauh and J. Shaw (2020), 'What Drives Enrollment Gaps in Further Education? The Role of Beliefs in Sequential Schooling Decisions', *Economica* **87**, 490–529.

Bellemare, C., L. Bissonnette and S. Kröger (2012), 'Flexible Approximation of Subjective Expectations using Probability Questions', *Journal of Business and Economic Statistics* **30**(1), 125–131.

Ben-Akiva, M., M. Bradley, T. Morikawa, J. Benjamin, T. Novak, H. Oppewal and V. Rao (1994), 'Combining Revealed and Stated Preference Data', *Marketing Letters* **5**(4), 335–350.

Benabou, R. and J. Tirole (2016), 'Mindful Economics: The Production, Consumption, and Value of Beliefs', *Journal of Economic Perspectives* **30**(3), 141–164.

Berkes, J., F. Peter, K.C. Spieß and F. Weinhardt (2022), 'Information Provision and Postgraduate Studies', *Economica* **Forthcoming**.

Bettinger, E.P., B. Terry Long, P. Oreopoulos and L. Sanbonmatsu (2012), 'The Role of Application Assistance and Information in College Decisions: Results from the H&R block FAFSA experiment', *Quarterly Journal of Economics* **127**(3), 1205–1242.

Betts, J.R. (1996), 'What Do Students Know about Wages? Evidence from a Survey of Undergraduates', *Journal of Human Resources* **31**, 27–56.

Betts, J.R. (2011), The Economics of Tracking in Education, *in* E.Hanushek, S.Machin and L.Woessmann, eds, 'Handbook of Economics of Education', Vol. 3, Elsevier B.V., chapter 7, pp. 341–381.

Blass, A.A., S. Lach and C.F. Manski (2010), 'Using Elicited Choice Probabilities to Estimate Random Utility Models: Preferences for Electricity Reliability', *International Economic Review* **51**(2), 421–440.

Bleemer, Z. and B. Zafar (2018), 'Intended College Attendance: Evidence from an Experiment on College Returns and Costs', *Journal of Public Economics* **157**, 184–211.

Boneva, T. and C. Rauh (2018), 'Parental Beliefs about Returns to Educational Investments – The Later the Better?', *Journal of the European Economic Association* **16**, 1669–1711.

Boneva, T. and C. Rauh (2021), Socio-economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns, Working paper.

Boneva, T., M. Golin and C. Rauh (2022), 'Can Perceived Returns Explain Enrollment Gaps in Postgraduate Education?', *Labour Economics* **77**, 1–17.

Boneva, T., T. Buser, A. Falk and F. Kosse (2022), The Origins of Gender Differences in Competitiveness and Earnings Expectations: Causal evidence from a mentoring intervention, Working Paper 9509, CESifo.

Botelho, A. and L. Costa Pinto (2004), 'Students' Expectations of the Economic Returns to College Education: Results of a controlled experiment', *Economics of Education Review* **23**, 645–653.

Bramoullé, Y., H. Djebbari and B. Fortin (2020), 'Peer Effects in Networks: A survey', *Annual Review of Economics* **12**, 603–629.

Briel, S., A. Osikominu, G. Pfeifer, M. Reutter and S. Satlukal (2022), 'Gender Differences in Wage Expectations: The role of biased beliefs', *Empirical Economics* **62**, 187–212.

Brock, W. and S. Durlauf (2001*a*), 'Discrete Choice with Social Interactions', *Review of Economic Studies* **68**, 235–260.

Brock, W. and S. Durlauf (2001*b*), Interactions-based Models, *in* J.Heckman and E.Leamer, eds, 'Handbook of Econometrics', Vol. 5, Elsevier, chapter 54, pp. 3297–3380.

Bruine de Bruin, W., A. Chin, J. Dominitz and W. van der Klaauw (2022), Household Surveys and Probabilistic Questions, *in* 'Handbook of Economic Expectations', Bachman, R., G. Topa, and W. van der Klaauw edn, Elsevier, chapter 1.

Brunello, G., C. Lucifora and R. Winter-Ebmer (2004), 'The Wage Expectations of European Business and Economics Students', *Journal of Human Resources* **39**, 1116–1142.

Calsamiglia, C. and M. Güell (2018), 'Priorities in School Choice: The case of the Boston Mechanism in Barcelona', *Journal of Public Economics* **163**, 20–36.

Card, D. (1999), The Causal Effects of Education on Earnings, *in* O.Ashenfelter and D.Card, eds, 'Handbook of Economics of Labor Economics', Vol. 3, Elsevier B.V., chapter 30, pp. 1801–1863.

Carlana, M. (2019), 'Implicit Stereotypes: Evidence from teachers' gender bias', *Quarterly Journal of Economics* **134**(3), 1163–1224.

Carlana, M., E. La Ferrara and P. Pinotti (Forthcoming), 'Goals and Gaps: Educational careers of immigrant children', *Econometrica* .

Carter, M. and R. Maddock (1984), Expectations in Economics, *in* 'Rational Expectations. Macroeconomics for the 1980s?', Palgrave Macmillan, chapter 2, pp. 12–42.

Carvajal, M.J., D. Bendana, A. Bozorgmanesh, M.A. Castillo, K. Pourmasiha, P. Rao and J.A. Torres (2000), 'Inter-gender Differentials between College Students' Earnings Expectations and the Experience of Recent Graduates', *Economics of Education Review* **19**, 229–243.

Chiswick, B. (1973), Schooling, Screening, and Income, *in* 'Does College Matter?', Academic Press, New York, pp. 151–158.

Comay, Y., A. Melnik and M.A. Pollatschek (1973), 'The Option Value of Education and the Optimal Path for Investment in Human Capital', *International Economic Review* **14**(2), 421–435.

Comay, Y., A. Melnik and M.A. Pollatschek (1976), 'Dropout Risks, Option values, and Returns to Investment in Schooling', *The Canadian Journal of Economics* **9**, 45–56.

Congdon-Hohman, J., N. Anil and J. Svec (2017), Student Uncertainty and Major Choice: A robust control approach, Working paper, College of the Holy Cross.

Conley, T., N. Mehta, R. Stinebrickner and T. Stinebrickner (2017), Social Interactions, Mechanisms, and Equilibrium: Evidence from a model of study time and academic achievement, Working Paper 7, Centre for Human Capital and Productivity.

Conlon, J.J. (2021), 'Major Malfunction: A Field Experiment Correcting Undergraduates' Beliefs about Salaries', *Journal of Human Resources* **56**(3), 922–939.

Cowan, B.W. (2017), 'Sources of Bias in Teenagers' College Expectations', *Social Science Quarterly* **99**(1), 136–153.

Crossley, T., Y. Gong, T. Stinebrickner and R. Stinebrickner (2021), Examining Income Expectations in the College and Early Post-college Periods: New Distributional Tests of Rational Expectations, Working Paper 01, Institute for Fiscal Studies.

Crossley, T., Y. Gong, T. Stinebrickner and R. Stinebrickner (2022), 'The Ex Post Accuracy of Subjective Beliefs: A new measure and decomposition', *Economics Letters* **210**, 110160.

Cunha, F., I. Elo and J. Culhane (2013), Eliciting Maternal Beliefs about the Technology of Skill Formation, Working Paper 19144, NBER.

Cunha, F., I. Elo and J. Culhane (2021), Maternal Subjective Expectations about the Technology of Skill Formation Predict Investments in Children One Year Later, Working Paper 18, HCEO.

Dauphin, A., A.-R. El Lahga, B. Fortin and G. Lacroix (2011), 'Are Children Decision-Makers Within the Household?', *The Economic Journal* **121**, 871–903.

de Bresser, J.R. and A.H.O. van Soest (2013), 'Survey Response in Probabilistic Questions and Its Impact on Inference', *Journal of Economic Behavior & Organization* **96**, 65–84.

Delavande, A. (2022), Expectations in Development, *in* 'Handbook of Economic Expectations', Bachman, R., G. Topa, and W. van der Klaauw edn, Elsevier, chapter 9.

Delavande, A., E. Del Bono and A. Holford (2020), 'Academic and Non-Academic Investments at University: The role of expectations, preferences and constraints', *Journal of Econometrics* **In Press**.

Delavande, A., E. Del Bono and A. Holford (2021), Ambiguity about COVID-Related Health Risks and Protective Health Behaviors Among Young Adults, Working paper.

D'Haultfoeuille, X., C. Gaillac and A. Maurel (2021), 'Rationalizing Rational Expectations: Characterizations and Tests', *Quantitative Economics* **12**(3), 817–842.

Diaz-Serrano, L., J. Hartog, W. Nilsson, H. van Ophem and P. Yang (2016), Student Earnings Expectations: Heterogeneity or Noise?, Working Paper 6110, CESifo.

Diaz-Serrano, L. and W. Nilsson (2022), 'The Reliability of Students' Earnings Expectations', *Labour Economics* **76**, 102182.

Doepke, M. and F. Zilibotti (2017), 'Parenting with Style: Altruism and paternalism in intergenerational preference transmission', *Econometrica* **85**(5), 1331–1371.

Doepke, M., F. Zilibotti and G. Sorrenti (2019), 'The Economics of Parenting', *Annual Review of Economics* **11**, 55–84.

Domina, T., A. Conley and G. Farkas (2011), 'The Link Between Educational Expectations and Effort in the College-for-all Era', *Sociology of Education* **84**, 93–112.

Dominitz, J. (1997), The Role of Conditioning Information in Reports of Subjective Phenomena, Draft prepared for presentation at EML/NSF Symposium on Preference Elicitation, California Institute of Technology.

Dominitz, J. (1998), 'Earning Expectations, Revisions, and Realizations', *Review of Economics and Statistics* **80**, 374–388.

Dominitz, J. (2001), 'Estimation of Income Expectations Models Using Expectations and Realization Data', *Journal of Econometrics* **102**, 165–195.

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

Dominitz, J. and C.F. Manski (1997), 'Using Expectations Data to Study Subjective Income Distribution', *Journal of the American Statistical Association* **87**, 855–867.

Dominitz, J. and C.F. Manski (1999), The Several Cultures of Research on Subjective Expectations, *in* J.Smith and R.Willis, eds, 'Wealth, Work and Health: Innovations in Measurement in the Social Sciences', The University of Michigan Press, chapter 1, pp. 15–33.

Dominitz, J., C.F. Manski and B. Fischhoff (2001), *Who Are the Youth "At Risk"? Expectations Evidence in the NLSY97*, Social Awakening, R. Michael edn, New York: Russel Sage Foundation, Sixth World Congress 8, pp. 230–257.

Dongili, P. and P. Giustinelli (2009), 'Economia e Scelta del Percorso Scolastico', *Formazione e Insegnamento* **1-2**.

Eckstein, Z. and K.I. Wolpin (1999), 'Why Youths Drop Out of High School: The impact of preferences, opportunities, and abilities', *Econometrica* **67**, 1295–1339.

Ehlert, M., C. Finger, A. Rusconi and H. Solga (2017), 'Intended College Enrollment and Educational Inequality: Do Students Lack Information?', *Economics of Education Review* **60**, 125–141.

Ehlert, M., F. Peter, C. Finger, A. Rusconi, H. Solga, C.K. Spiess and V. Zambre (2017), The Berliner-Studienberechtigen-Panel (Best Up): Methodological and Data Report, DIW Data Documentation 90, DIW Berlin.

Ehrmantraut, L., P. Pinger and R. Stans (2020), The Expected (Signaling) Value of Higher Education, Discussion Paper 12522, IZA.

Entwisle, D.R., K.L. Alexander and L.S. Olson (2002), Baltimore Beginning School Study in Perspective, *in* E.Phelps, J.F.F. Furstenberg and A.Colby, eds, 'Looking at Lives: American Longitudinal Studies of the Twentieth Century', Russell Sage Foundation, p. 167–193.

Epple, D. and R.E. Romano (2011), Peer Effects in Education: A survey of the theory and evidence, *in* J.Benhabib, A.Bisin and M.Jackson, eds, 'Handbook of Social Economics', Vol. 1, Elsevier B.V., chapter 20, pp. 1053–1163.

Falk, A. and F. Kosse (2021), The briq Family P, Working paper.

Fischhoff, B., A.M. Parker, W. Bruine de Bruin, J. Downs, C. Palmgren, R. Dawes and C.F. Manski (2000), 'Teen Expectations for Significant Life Eevents', *Public Opinion Quarterly* **64**, 189–205.

Fuster, A. and B. Zafar (2022), Fields Experiments in Surveys, *in* 'Handbook of Economic Expectations', Bachman, R., G. Topa, and W. van der Klaauw edn, Elsevier, chapter 4.

Garbin, F. (2021), The Role of Expectations on Peers in Educational Choices, Phd thesis chapter, Bocconi University.

Gershenson, S., S.B. Holt and N.W. Papageorge (2016), 'Who Believes in Me? The effect of student–teacher demographic match on teacher expectations', *Economics of Education Review* **52**, 209–224.

Gilboa, I. and D. Schmeidler (2004), 'Subjective Distributions', *Theory and Decisions* **56**, 345–357.

Giustinelli, P. (2016), 'Group Decision Making with Uncertain Outcomes: Unpacking Child-Parent Choice of the High School Track', *Internation Economic Review* **57**, 573–602.

Giustinelli, P. and C.F. Manski (2018), 'Survey Measures of Family Decision Processes for Econometric Analysis of Schooling Decisions', *Economic Inquiry* **56**(1), 81–99.

Giustinelli, P., C.F. Manski and F. Molinari (2020), 'Tail and Center Rounding of Probabilistic Expectations in the Health and Retirement Study', *Journal of Econometrics* **In Press**.

Giustinelli, P., C.F. Manski and F. Molinari (2022), 'Precise or Imprecise Probabilities? Evidence from survey response related to late-onset dementia', *Journal of the European Economic Association* **20**(1), 187–221.

Giustinelli, P. and M.D. Shapiro (2019), S*ea*TE: Subjective *ex ante* Treatment Effect of Health on Retirement, Working Paper 26087, NBER.

Giustinelli, P. and M.D. Shapiro (2021), S*ea*TE: Subjective *ex ante* Treatment Effect of Health on Retirement, Working paper.

Giustinelli, P. and N. Pavoni (2017), 'The Evolution of Awareness and Belief Ambiguity In the Process of High School Track Choice', *Review of Economic Dynamics* **25**, 93–120.

Giustinelli, P. and N. Pavoni (2019), Skill Mismatch, Family Information, and High School Track Choice in Italy, Policy report, JPMorgan.

Gong, Y., L. Lochner, T. Stinebrickner and R. Stinebrickner (2021), The Consumption Value of College, Working paper.

Gong, Y., T. Stinebrickner and R. Stinebrickner (2019), 'Uncertainty about Future Income: Initial beliefs and resolution during college', *Quantitative Economics* **10**(2), 607–641.

Gong, Y., T. Stinebrickner and R. Stinebrickner (2020*a*), 'Marriage, Children, and Labor Supply: Beliefs and Outcomes', *Journal of Econometrics* **In Press**.

Gong, Y., T. Stinebrickner and R. Stinebrickner (2020*b*), 'Perceived and Actual Option Values of College Enrollment', *Journal of Applied Econometrics* **Forthcoming**.

Groot, W. and H. Oosterbeek (1992), 'Optimal Investment in Human Capital Under Uncertainty', *Economics of Education Review* **11**(I), 41–49.

Grossman, M. (2006), Education and Nonmarket Outcomes, *in* E.Hanushek and F.Welch, eds, 'Handbook of Economics of Education', Vol. 1, Elsevier B.V., chapter 10, pp. 578–633.

Hakimov, R. and D. Kübler (2021), 'Experiments on Centralized School Choice and College Admissions: A survey', *Experimental Economics* **24**, 434–488.

Hartog, J. and L. Diaz-Serrano (2014), 'Schooling As a Risky Investment. A Survey of Theory and Evidence', *Foundations and Trends in Microeconomics* **9**(3-4), 1–176.

Hartog, J., X. Ding and J. Liao (2014), 'Is Earnings Uncertainty Relevant for Educational Choice? An empirical analysis for China', *Education Economics* **22**, 471–483.

Hastings, J., C. Neilson and S. Zimmerman (2015), The Effects Earnings Disclosure on College Enrollment Decisions, Working Paper 21300, NBER.

Hastings, J.S., C.A. Neilson, A. Ramirez and S.D. Zimmerman (2016), '(Un)information, College and Major Choice: Evidence from Linked Survey and Survey', *Economics of Education Review* **51**, 136–151.

Heckman, J.J., L. Lochner and P.E. Todd (2006), Earnings Functions, Rates of Return and Treatment Effects: The Mincer equation and beyond, *in* E.Hanushek and F.Welch, eds, 'Handbook of Economics of Education', Vol. 1, Elsevier B.V., chapter 7, pp. 307–458.

Holian, L. (2020), Male and Female High School Students' Expectations for Working In a Health-Related Field, Report 82, NCES.

Holian, L. and E. Kelly (2020), STEM Occupational Intentions Stability and Change Through High School, Report 167, NCES.

Hotz, V.J., E. Wiemers, J. Rasmussen and K. Maxwell Koegel (Forthcoming), 'The Role of Parental Wealth and Income in Financing Children's College Attendance and Its Consequences', *Journal of Human Resources* .

Hoxby, C.M. and S. Turner (2013), Expanding college opportunities for high-achieving, low income students, Working Paper 12-014, SIEPR.

Hudomiet, P., M.D. Hurd and S. Rohwedder (2021), The Causal Effects of Wages, Wealth, and Health on Retirement: Estimates Based on Subjective Conditional Probabilities, Working paper, RAND.

Huntington-Klein, N. (2015), 'Subjective and Projected Returns to Education', *Journal of Economic Behavior & Organization* **117**, 10–25.

Huntington-Klein, N. (2016a), Consumption Value and the Demand for College Education, Working paper, California State University Fullerton.

Huntington-Klein, N. (2016b), "'(Un)informed College and Major Choice": Verification in an alternate setting', *Economics of Education Review* **53**, 159–163.

Ilut, C.L. and M. Schneider (2022), Modeling Uncertainty as Ambiguity: a Review, *in* 'Handbook of Economic Expectations', Bachman, R., G. Topa, and W. van der Klaauw edn, Elsevier, chapter 24.

Imbens, G. (2020), 'Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics', *Journal of Economic Literature* **58**(4), 1129–1179.

Jacob, B. and T. Wilder (2010), Educational Expectation and Attainment, Working Paper 15683, NBER.

Jacob, B.A., B. McCall and K.M. Stange (2018), 'College as Country Club: Do colleges cater to students' preferences for consumption?', *Journal of Labor Economics* **36**(2), 309–348.

Jacob, B.A. and L. Lefgren (2007), 'What Do Parents Value in Education? An empirical investigation of parents' revealed preferences for teachers', *Journal of Labor Economics* **122**(4), 1603–1637.

Jerrim, J. (2015), 'Do College Students Make Better Predictions of Their Future Income Than Young Adults in the Labor Force?', *Education Economics* **23**, 162–179.

Juster, F.T. (1966), 'Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design', *Journal of the American Statistical Association* **61**, 658–696.

Kalenkoski, C. (2008), 'Parent-Child Bargaining, Parental Transfers, and the Post-Secondary Education Decision', *Applied Economics* **40**(4), 413–436.

Kapor, A.J., C.A. Neilson and S.D. Zimmerman (2020), 'Heterogeneous Beliefs and School Choice Mechanisms', *American Economic Review* **110**, 1274–1315.

Kerr, S.P., T. Pekkarinen, M. Sarvimaki and R. Uusitalo (2020), 'Post-secondary Education and Information on Labor Market Prospects: A randomized field experiment', *Labour Economics* **66**, 101888.

Kiessling, L., P. Pinger, P. Seegers and J. Bergerhoff (2020), Gender Differences in Wage Expectations: Sorting, Children, and Negotiation styles, Discussion Paper 13729, IZA.

Kleinjans, K.J. and A.H.O. van Soest (2014), 'Rounding, Focal Point Answers and Nonresponse to Subjective Probability Questions', *Journal of Applied Econometrics* **29**, 567–585.

Klossner, S. and G. Pfeifer (2019), 'The Importance of Tax Adjustments When Evaluating Wage Expectations', *The Scandinavian Journal of Economics* **121**, 578–605.

Koşar, G. and C. O'Dea (2022), Expectations Data in Structural Micro Models, *in* 'Handbook of Economic Expectations', Bachman, R., G. Topa, and W. van der Klaauw edn, Elsevier, chapter 21.

Kuhnke, R. (1997), Anlage und Weiterführung der dritten Leipziger Längsschnittsstudie, *in* E.Brislinger, B.Hausstein and E.Riedel, eds, 'Jugend im Osten (Youth in the East)', Trafo, pp. 125–132.

Kunz, J.S. and K.E. Staub (2020), 'Early Subjective Completion Beliefs and the Demand for Post-secondary Education', *Journal of Economic Behavior & Organization* **177**, 34–55.

Lange, F. and R. Topel (2006), The Social Value of Education and Human Capital, *in* E.Hanushek and F.Welch, eds, 'Handbook of Economics of Education', Vol. 1, Elsevier B.V., chapter 8, p. 459–509.

Larroucau, T., M. Martinez, C. Neilson and I. Rios (2021), Application Mistakes and Information Frictions in College Admissions, Working paper, Princeton University.

Lazear, E. (1977), 'Education: Consumption or Production?', *Journal of Political Economy* **85**(3), 569–597.

Lee, L.-F., J. Li and X. Lin (2014), 'Binary Choice Models With Social Network Under Heterogeneous Rational Expectations', *The Review of Economics and Statistics* **93**(3), 402–417.

Lee, S.Y., Y. Shin and D. Lee (2015), The Option Value of Human Capital: Higher Education and Wage Inequality, Working Paper 21724, NBER.

Leibing, A., F. Peter, S. Waights and C.K. Spieß (2022), Gender Gaps in Early Wage Expectations, Discussion Paper 15281, IZA.

Lergetporer, P., G. Schwerdt, K. Werner, M.R. West and L. Woessmann (2018), 'How Information Affects Support for Education Spending: Evidence from survey experiments in germany and the united states', *Journal of Public Economics* **167**, 138–157.

Lergetporer, P., K. Werner and L. Woessmann (2020), 'Educational Inequality and Public Policy Preferences: Evidence from representative survey experiments', *Journal of Public Economics* **188**, 104–226.

Lergetporer, P., K. Werner and L. Woessmann (2021), 'Does Ignorance of Economic Returns and Costs Explain the Educational Aspiration Gap? Representative evidence from adults and adolescents', *Economica* **Forthcoming**.

Lergetporer, P. and L. Woessmann (2019), The Political Economy of Higher Education Finance: How information and design affect public preferences for tuition, Working Paper 7536, CESifo.

Levhari, D. and Y. Weiss (1974), 'The Effect of Risk on the Investment in Human Capital', *American Economic Review* **64**, 950–963.

Lochner, L. (2011), Education and Nonmarket Outcomes, *in* E.Hanushek, S.Machin and L.Woessmann, eds, 'Handbook of Economics of Education', Vol. 4, Elsevier B.V., chapter 10, pp. 183–282.

Lochner, L. and A. Monge-Naranjo (2012), 'Credit Constraints in Education', *Annual Review of Economics* **4**, 225–256.

Lochner, L. and A. Monge-Naranjo (2016), Student Loans and Repayment: Theory, Evidence and Policy, *in* E.Hanushek, S.Machin and L.Woessmann, eds, 'Handbook of Education Economics', Vol. 5, Elsevier B.V., chapter 8, pp. 397–478.

Lundberg, S., J. Romich and K.P. Tsang (2009), 'Decision-Making by Children', *Review of Economics of the Household* **7**, 1–30.

Madeira, C. (2018), 'Testing the Rationality of Expectations of Qualitative Outcomes', *Journal of Applied Econometrics* **33**(6), 837–852.

Madeira, C. (2020*a*), 'Learning Your Own Ability', *Journal of Economic Dynamics and Control* **121**, 104026.

Madeira, C. (2020*b*), Measuring the Perceived Value of an MBA Degree, Working Paper 876, Central Bank of Chile.

Maestas, N., K.J. Mullen, D. Powell, T. von Wachter and J. Wenger (2018), The Value of Working Conditions in the United States and Implications for the Structure of Wages, Working Paper 25204, NBER.

Manski, C.F. (1989), 'Schooling as Experimentation: A Reappraisal of the Postsecondary Dropout Phenomenon', *Economics of Education Review* **8**(4), 305–312.

Manski, C.F. (1990), 'The Use of Intentions Data to Predict Behavior: A Best Case Analysis', *Journal of the American Statistical Association* **85**(412), 934–940.

Manski, C.F. (1993), Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?, *in* 'Studies of Supply and Demand in Higher Education', Clotfelter C.T. and Rothschild M. edn, University of Chicago Press, pp. 43–60.

Manski, C.F. (1999), 'Analysis of Choice Expectations in Incomplete Scenarios', *Journal of Risk and Uncertainty* **19**(1-3), 49–65.

Manski, C.F. (2000), 'Economic Analysis of Social Interactions', *The Journal of Economic Perspectives* **14**(3), 115–136.

Manski, C.F. (2004), 'Measuring Expectations', *Econometrica* **72**(5), 1329–1376.

Manski, C.F. (2016), 'Interpreting Point Predictions: Some Logical Issues', *Foundations and Trends in Accounting* **10**(2-4), 238–261.

Manski, C.F. (2018), 'Survey Measurement of Probabilistic Macroeconomic Expectations: Progress and Promise', *NBER Macroeconomics Annual* **32**, 411–471.

Manski, C.F. (2022), Looking Ahead to Research Enhancing Measurement of Expectations, *in* 'Handbook of Economic Expectations', Bachman, R., G. Topa, and W. van der Klaauw edn, Elsevier, chapter 26.

Manski, C.F. and F. Molinari (2010), 'Rounding Probabilistic Expectations in Surveys', *Journal of Business and Economic Statistics* **28**(2), 219–231.

Mazza, J. and J. Hartog (2011), Do They Understand the Benefits from Education? Evidence on Dutch High School Students' Earnings Expectations, Discussion Paper 5714, IZA.

McFadden, D.L. (1973), Conditional Logit Analysis of Qualitative Choice Behavior, *in* P.Zarembka, ed., 'Avances in Econometrics', New York Academic Press.

Menon Eliophotou, M. (1997*a*), 'Perceived Economic Benefits of Higher Education: The Case of Cyprus', *Education Economics* **5**, 53–61.

Menon Eliophotou, M. (1997*b*), 'Perceived Rates of Return to Higher Education in Cyprus', *Economics of Education Review* **16**, 425–430.

Menon Eliophotou, M. (2008*a*), 'Perceived Rates of Return to Higher Education: Further evidence from Cyprus', *Economics of Education Review* **27**, 39–47.

Menon Eliophotou, M. (2008*b*), 'The Economic Benefits of Higher Education in Cyprus: The expectations of prospective students', *International Journal of Educational Development* **28**, 259–267.

Menon Eliophotou, M. and N. Pashourtidou (2017), 'Low Socioeconomic Status Students in Higher Education: Entry, academic attainment and earnings expectations', *Cyprus Economic Policy Review* **11**, 15–40.

Mincer, J. (1974), *Schooling, Experience and Earnings*, Columbia University Press.

Montmarquette, C., K. Cannings and S. Mahseredjian (2002), 'How Do Young People Choose College Majors?', *Economics of Education Review* **21**, 543–556.

Mueller, A.I. and J. Spinnewijn (2022), Expectations Data, Labor Market and Job Search, *in* 'Handbook of Economic Expectations', Bachman, R., G. Topa, and W. van der Klaauw edn, Elsevier, chapter 22.

Oreopoulos, P. and K.G. Salvanes (2011), 'Priceless: The Nonpecuniary Benefits of Schooling', *Journal of Economic Perspectives* **25**(1), 159–184.

Osikominu, A. and G. Pfeifer (2018), Perceived Wages and the Gender Gap in STEM Fields, Discussion Paper 12719, CEPR.

Papageorge, N.W., S. Gershenson and K. Min Kang (2020), 'Teacher Expectations Matter', *Review of Economics and Statistics* **102**(2), 234–251.

Patnaik, A., J. Venator, M.J. Wiswall and B. Zafar (2020), 'The Role of Heterogeneous Risk Preferences, Discount Rates, and Earnings Expectations in College Major Choice', *Journal of Econometrics* **In Press**.

Patnaik, A., M.J. Wiswall and B. Zafar (2020), College Majors, Working paper.

Peter, F., C.K. Spieß and V. Zambre (2017), 'Intended College Enrollment and Educational Inequality: Do Students Lack Information?', *Economics of Education Review* **60**, 125–141.

Peter, F., C.K. Spieß and V. Zambre (2021), 'Informing Students About College: Increasing enrollment using a behavioral intervention?', *Journal of Economic Behavior & Organization* **190**, 524–549.

Psacharopoulos, G. and H.A. Patrinos (2018), 'Returns to Investment in Education: A decennial review of the global literature', *Education Economics* **26**(5), 445–458.

Reuben, E., M.J. Wiswall and B. Zafar (2017), 'Preferences and Biases in Education Choices and Labor Market Expectations: Shrinking the Black Box of Gender', *Economic Journal* **604**, 2153–2186.

Reynolds, J.R. and J. Pemberton (2001), 'Rising College Expectations among Youth in the United States. a Comparison of the 1979 and 1997 NLSY', *Journal of Human Resources* **36**, 703–726.

Rubin, D. (1974), 'Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies', *Journal of Educational Psychology* **66**(5), 688–701.

Ruder, A.I. and M. van Noy (2017*a*), 'Knowledge of Earnings Risk and Major Choice: Evidence from an information experiment', *Economics of Education Review* **57**, 80–90.

Ruder, A.I. and M. van Noy (2017*b*), Labor Market Expectations and Major Choice for Low-Income, First-Generation College Students: Evidence from an information experiment, Working paper.

Sacerdote, B. (2011), Peer Effects in Education: How might they work, how big are they and how much do we know thus far?, *in* E.Hanushek, S.Machin and L.Woessmann, eds, 'Handbook of Economics of Education', Vol. 3, Elsevier B.V., chapter 4, pp. 249–277.

Sattin-Bajaj, C. (2014), *Unaccompanied Minors: Immigrant Youth, School Choice, and the Pursuit of Equity*, Harvard Education Press.

Savage, L.J. (1954), *The Foundations of Statistics*, 2nd Revised Edition (1972) edn, Dover, New York.

Schultz, T. (1963), *The Economic Value of Education*, Columbia University Press.

Schweri, J. and J. Hartog (2017), 'Do Wage Expectations Predict College Enrollment? Evidence from healthcare', *Journal of Economic Behavior & Organization* **141**, 135–150.

Schweri, J., J. Hartog and S.C. Wolter (2011), 'Do Students Expect Compensation for Wage Risk?', *Economics of Education Review* **30**, 215–227.

Son, S.J. (2020), Distributional Impacts of Centralized School Choice, Job market paper, Yale University.

Spence, M. (1973), 'Job Market Signalling', *Quarterly Journal of Economics* **87**, 355–374.

Stange, K. (2012), 'An Empirical Investigation of the Option Value of College Enrollment', *American Economic Journal: Applied Economics* **4**(1), 49–84.

Stiglitz, J. and A. Weiss (1990), Sorting Out the Differences Between Signalling and Screening Models, *in* M.Bacharach, M.Dempster and J.Enos, eds, 'Mathematical Models in Economics', University of Oxford, chapter 4, pp. 1–34.

Stinebrickner, T. and R. Stinebrickner (2003*a*), 'Understanding Educational Outcomes of Students from Low-Income Families: Evidence from a liberal arts college with a full tuition subsidy program', *Journal of Human Resources* **38**, 591–617.

Stinebrickner, T. and R. Stinebrickner (2003*b*), 'Working During School and Academic Performance', *Journal of Labor Economics* **21**(2), 473–491.

Stinebrickner, T. and R. Stinebrickner (2004), 'Time-Use and College Outcomes', *Journal of Econometrics* **121**(1-2), 243–269.

Stinebrickner, T. and R. Stinebrickner (2008*a*), 'The Causal Effect of Studying on Academic Performance', *B.E. Journal of Economic Analysis and Policy (Frontiers)* **8**(1), 1–55.

Stinebrickner, T. and R. Stinebrickner (2008*b*), 'The Effect of Credit Constraints on the College Dropout Decision: A direct approach using a new panel study', *American Economic Review* **98**(5), 2163–2184.

Stinebrickner, T. and R. Stinebrickner (2012), 'Learning About Academic Ability and the College Drop-Out Decision', *Journal of Labor Economics* **30**(4), 707–748.

Stinebrickner, T. and R. Stinebrickner (2014*a*), 'A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout', *Review of Economic Studies* **81**(1), 426–472.

Stinebrickner, T. and R. Stinebrickner (2014*b*), 'Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model', *Journal of Labor Economics* **32**(3), 601–644.

Tincani, M.M., F. Kosse and E. Miglino (2021), Subjective Beliefs and Inclusion Policies: Evidence from College Admissions, Working paper.

Train, K. (2009), *Discrete Choice Methods with Simulation*, 2nd edition edn, Cambridge University Press.

Trede, I. and J. Schweri (2014), 'Work Values and Intention to Become a Registered Nurse Among Healthcare Assistants', *Nurse Education Today* **34**, 948–953.

van der Klaauw, W. (2012), 'On the Use of Expectations Data in Estimating Structural Dynamic Models', *Journal of Labor Economics* **30**(3), 521–554.

Walker, J. (2001), *Adolescents' Expectations Regarding Birth Outcomes: A Comparison of the NLSY79 and NLSY97 Cohorts*, Social Awakening, R. Michael edn, New York: Russel Sage Foundation, Sixth World Congress 7, p. 201–229.

Webbink, D. and J. Hartog (2004), 'Can Students Predict Starting Salaries? Yes!', *Economics of Education Review* **23**, 103–113.

Weisbrod, B.A. (1962), 'Education and Investmant in Human Capital', *Journal of Political Economy* **70**, 106–123.

Weiss, Y. (1972), 'The Risk Element in Occupational and Educational Choices', *Journal of Political Economy* **80**, 1203–1213.

Winkler, R. and A. Murphy (1970), 'Nonlinear Utility and the Probability Score', *Journal of Applied Meteorology* **9**(1), 143–148.

Wiswall, M.J. and B. Zafar (2015*a*), 'Belief Updating among College Students: Evidence from Experimental Variation in Information', *Journal of Human Capital* **9**(2), 117–169.

Wiswall, M.J. and B. Zafar (2015*b*), 'Determinants of College Major Choice: Identification Using an Information Experiment', *Review of Economic Studies* **82**(2), 791–824.

Wiswall, M.J. and B. Zafar (2018), 'Preference for the Workplace, Investment in Human Capital, and Gender', *Quarterly Journal of Economics* **133**(1), 457–507.

Wiswall, M.J. and B. Zafar (2020), 'Human Capital and Expectations about Career and Family', *Journal of Political Economy* **129**(5), 1361–1424.

Wolter, S. and A. Zbinden (2005), 'Labour Market Expectations of Swiss University Students', *International Journal of Manpower* **23**(5), 458–470.

Wolter, S.C. (2000), 'Wage Expectations: A Comparison of Swiss and US Students', *Kyklos* **53**, 51–69.

Zafar, B. (2011*a*), 'Can Subjective Expectations Data Be Used in Choice Models? Evidence on Cognitive Biases', *Journal of Applied Econometrics* **26**(3), 520–544.

Zafar, B. (2011*b*), 'How Do College Students Form Expectations?', *Journal of Labor Economics* **29**(2), 301–348.

Zafar, B. (2012), 'Double Majors: One For Me, One For the Parents?', *Economic Inquiry* **50**(2), 287–308.

Zafar, B. (2013), 'College Major Choice and the Gender Gap', *Journal of Human Resources* **48**(3), 545–595.

Zambre, V. (2018), The Gender Gap in Wage Expectations: Do Young Women Trade Off Higher Wages for Lower Wage Risk?, Discussion Paper 1742, DIW Berlin.