

Discussion Paper Series – CRC TR 224

Discussion Paper No. 575
Project C 01

Sequential Choices, Option Values, and the Returns to Education

Manudeep Bhuller¹
Philipp Eisenhauer²
Moritz Mendel³

July 2024

¹University of Oslo, Email: manudeep.bhuller@econ.uio.no

²Amazon, Email: peisen@amazon.com

³University of Bonn, Email: moritz.mendel12@gmail.com

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)
through CRC TR 224 is gratefully acknowledged.

Sequential Choices, Option Values, and the Returns to Education*

Manudeep Bhuller[†] Philipp Eisenhauer[‡] Moritz Mendel[§]

July 1, 2024

Abstract: Using detailed Norwegian data on earnings, education and work histories, we estimate a dynamic structural model of education and sector choices that captures rich life-cycle patterns by ability. We validate the model against variation in education choices induced by a compulsory schooling reform. Our approach allows us to estimate the ex-ante returns to different education tracks across the life-cycle by individual ability and quantify the contribution of option values. We find substantial heterogeneity in returns and establish crucial roles for option values and re-enrollment in determining education choices and the impact of schooling policies.

Keywords: sequential decisions, ex-ante returns, option values, reform evaluation

JEL code: J24, J31, D80

*Acknowledgements: We thank James Heckman for numerous helpful discussions, and seminar participants at University of Chicago (Lifecycle Working Group), Arizona State University, Royal Holloway-University of London, Aarhus University and University of Copenhagen for helpful comments. We thank Annica Gehlen, Emily Schwab, and Leiqiong Wan for their outstanding research assistance. We are grateful to the Social Sciences Computing Service (SSCS) at the University of Chicago for the permission to use their computational resources and Statistics Norway for providing access to micro data and computational resources. Manudeep Bhuller received financial support from the Research Council of Norway through the Young Research Talents Grant 275123. Philipp Eisenhauer is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2126/1- 390838866, the TRA Modelling (University of Bonn) as part of the Excellence Strategy of the federal and state governments. Moritz Mendel gratefully acknowledges funding by the German Research Foundation (DFG) through CRC TR 224 (Project C01) and the Research Training Group "The Macroeconomics of Inequality".

[†]University of Oslo; Statistics Norway; CEPR; CESifo; IZA. manudeep.bhuller@econ.uio.no

[‡]Amazon. peisen@amazon.com

[§]University of Bonn. moritz.mendel12@gmail.com

1 Introduction

Standard models of human capital (Becker, 1964; Mincer, 1958) assume that individuals compare the potential future earnings streams at the beginning of their schooling career, choose the alternative with the highest net benefit, and subsequently complete their desired level of schooling.¹ This view ignores both the sequential nature of human capital investments and the uncertainties embedded in this decision-making. The decision to take an additional year of schooling may open up further schooling opportunities as, for instance, a high school diploma is a stepping stone for a college education. And, individuals make such important decisions often facing considerable uncertainty about the associated costs and gains (see, e.g., Wiswall and Zafar (2015), Attanasio and Kaufmann (2017) and Wiswall and Zafar (2021)).

Our paper is motivated by two stylized facts about educational choices related to their sequential and uncertain nature. Firstly, across a wide range of settings, educational researchers have documented the prevalence of drop-out, re-enrollment, and track switching in educational histories. These patterns are present not only in countries with highly subsidized educational system like the Scandinavian countries and Germany, but also in the U.S., where students face high monetary costs of higher education.² Secondly, a recurrent finding in the literature on compulsory schooling policies is that such policies tend to have so-called “inframarginal” impacts, i.e., educational choices beyond the minimum schooling requirements are impacted. In a seminal study of the compulsory attendance laws in the U.S., Lang and Kropp (1986) documented the prevalence of such impacts, while similar findings have echoed in later studies.³

In this paper, we develop and estimate a dynamic structural model of educational choices in a life-cycle context that can accommodate and explain both of the above-mentioned features of educational careers. Agents in our model are forward-looking and make sequential decisions every period from age 15 onwards on whether to attend school, and if so, the type of educational track to attend, to work in one of three sectors, or to stay at home, while they face uncertainty in terms of their work productivity in each sector and tastes for schooling tracks, work and

¹Heckman et al. (2006) and Card (1999) provide extensive reviews of this literature on returns to schooling based on the Becker-Mincer models of human capital investments.

²For instance, according to data from the National Student Clearinghouse (NSC), thirty-six million Americans held some postsecondary schooling in 2019 without completing a college degree or being currently enrolled (Shapiro et al., 2019).

³For the U.S., Acemoglu and Angrist (2000) provide further evidence on how the compulsory attendance laws affected the distribution of schooling, while Bedard (2001) provides evidence on how better university access increased high school drop-out rates. Relatedly, Meghir and Palme (2005) found evidence on “inframarginal” responses to a compulsory schooling reform in Sweden, while similar evidence for Norway is reported (though not emphasized) in Black et al. (2005). While these responses can reflect equilibrium adjustments to policy reforms in line with models of educational signaling (Spence, 1973), these may also reflect the sequential and uncertain nature of educational decisions, as argued by Altonji (1993) and Heckman et al. (2018).

leisure, and also differ in terms of their ability and a vector of latent types. Our model is able to generate a rich set of educational and work histories that feature (i) interruptions and re-enrollment in educational careers, (ii) persistence in choices across the life-cycle, (iii) costly switching between educational tracks, (iv) sector-specific human capital accumulation, and (v) persistent individual heterogeneity. The model is also able to re-produce “inframarginal” responses to actual and hypothetical compulsory schooling reforms, provide evidence on who is affected and inform about the potential economic mechanisms driving these patterns.

Using our model, we define and quantify two key objects related to educational choices. First, we consider the *ex-ante return* to each schooling track choice for a given state in our model. This object takes into account the sequential and uncertain nature of schooling choices and captures both the immediate wage and non-monetary rewards as well as the discounted lifetime rewards associated with a choice as compared to the best alternative choice. Unlike the monetary wage rewards that have been the focus of much of the returns to schooling literature (see, e.g., Card (1999) for a review), the *ex-ante* returns are the objects that drive the educational choices of agents in our model. To illustrate the role of uncertainty, we further contrast *ex-ante* and *ex-post* returns, where the latter depend on the realizations of productivity and taste shocks, while the former reflect agents’ expectations. Second, we consider the *option value* to each schooling track choice for a given state in our model. The sequential nature of schooling decisions generates this value, for instance, as completing a high school diploma generates the option to enroll in college, and college enrollment generates the option to obtain a college degree (Comay et al., 1973; Weisbrod, 1962). Non-linearities in the wage returns to schooling choices (Hungerford and Solon, 1987) and the sequential resolution of uncertainties embedded in these choices can further exacerbate the importance of option values (Trachter, 2015; Stange, 2012; Altonji, 1993). As we will show, these objects are crucial determinants of schooling decisions, essential in understanding the impacts of policy reforms, and under-appreciated in the existing literature.

We implement our modeling approach and provide evidence on the *ex-ante* returns and the option values of education using Norwegian administrative data with career-long earnings information and education histories. We combine these data with detailed demographic information, including measures on individual ability for males collected as part of the compulsory military recruitment testing. This dataset has several advantages, as this provides (i) complete annual information on educational track choices and earnings histories for selected cohorts across four decades, (ii) allows us to capture several sources of persistent heterogeneity, and (iii) only suffers from natural sample attrition due to either death or out-migration. Our dataset further covers a compulsory schooling reform that increased the minimum requirement schooling from seven to nine years across different municipalities at different points in time (Black et al., 2005). The latter feature of our setting allows us to estimate our model on individuals that were not

exposed to the reform and rely on the reform-induced variation to evaluate our model in an out-of-sample validation and to shed light on “inframarginal” responses to educational policy reforms. Specifically, we compare the predictions about policy impacts based on the model estimated on pre-reform data to the observed impacts post-reform (Todd and Wolpin, 2021).⁴

Our analysis presents several insights. We find substantial heterogeneity in the ex-ante returns by the level of schooling, track choice and individual ability. For low-ability individuals, our estimates of average ex-ante returns range between -3% at the 8th grade in the academic track and 3% at the 9th year in the vocational track.⁵ By comparison, both medium- and high-ability individuals have sizeable and positive returns to academic schooling, while high-ability individuals have negative return at the 8th grade of vocational schooling. Underlying this heterogeneity is a strong pattern of *ability-related sorting* into academic and vocational tracks. Indeed, the structure of the returns reflects the separation of the ability groups in the different schooling tracks at an early stage of the educational pathways. Meanwhile, the relatively low average ex-ante returns for an extra schooling year reflect the presence of *re-enrollment opportunities*, as many individuals who drop out early do re-enroll and complete further education later in their life-cycle. Moving beyond the averages, we document substantial heterogeneity in the *distributions* of ex-ante returns, with more dispersion in returns at the earlier stages of the educational careers, in the academic track and among high-ability individuals.

Our evidence further shows that the option values make up a sizeable fraction of the overall values of educational choices, however, their contribution varies considerably by the stage of educational career, track and individual ability. Option values depend on the likelihood that an individual will continue to pursue schooling further and the rewards they accrue if they do so. A recurring finding is that the *option value contribution* is highest for the year of schooling right before the completion of an academic degree that entails considerable “diploma” effects. Intuitively, completing the schooling year right before the degree awarding year makes it more likely that individual will indeed pursue a degree and thus this choice holds a high option value. We also find important heterogeneity by ability and track. The option value contributions in the academic track tend to be highest for high-ability individuals, as they are also likely to benefit the most from the additional schooling opportunities. Indeed, we find more than 90% of the high-ability individuals facing the choice of attending the 11th year in academic track would not have completed this education had it not triggered the option of continuing further to attain a high school diploma. Our evidence also provides insights on how *sector-*

⁴The recent review by Galiani and Pantano (2021) also emphasizes the need for model validation and provides a structured review of the small literature.

⁵A negative return to a choice alternative in our model reflects that an individual expects to receive a higher reward from another choice alternative; i.e., the respective choice alternative is thus not chosen.

specific returns to academic and vocational schooling influence educational choices for different ability groups. For instance, we find that fewer public sector job opportunities would induce some medium-ability individuals to drop academic schooling and pursue vocational high school.

Finally, we use our model to analyze the impacts of compulsory schooling reforms. Our model predicts that a compulsory schooling reform similar to the one actually implemented in Norway during the 1960s would increase the share of high school graduates by about 3%, a prediction that resembles the actual reform-induced change observed in our data. Option values provide an economic rationale for such *inframarginal responses*. By forcing more schooling on individuals, who prior to the reform would have taken less schooling than the new minimum requirement, we also bring them closer to the margins of schooling choices that hold stronger rewards through diplomas or degrees, and as a result some of these individuals do indeed pursue further education. Another important mechanism that our model brings forth is that of re-enrollment opportunities. Even prior to the reform, a sizeable fractions of individuals would have attained the new minimum schooling requirement, but only after first dropping-out and re-enrolling at a large stage in their careers. Since the reform forces these individuals to take the new minimum schooling requirement in an uninterrupted manner, their educational trajectories are also impacted. Interestingly, some of these individuals now also go on to pursue further education, since they no longer face high re-enrollment costs.

Our paper provides several contributions. We extend the empirical literature that acknowledges the sequential nature of schooling investments and emphasizes the roles of uncertainties and non-linearities. Lee et al. (2017), Eisenhauer et al. (2015), Trachter (2015) and Stange (2012) all study the role of uncertainty and option values in shaping schooling decisions in deliberately simplified settings. However, none of these studies analyze life-cycle decisions, or allow for heterogeneity by ability, re-enrollment, sector and track switching at the same time. Our work is closely related to Heckman et al. (2018), who develop a sequential educational choice model. However, they restrict their attention to ex-post returns of education, and avoid making specific assumptions about individuals' expectations about costs and benefits of schooling. We impose additional structure on the decision process and are able to quantify ex-ante returns and option values. Our paper also relates to a large literature on compulsory schooling reforms (Brunello et al., 2009; Oreopoulos, 2006), providing evidence on the impacts of such reforms along the distribution of schooling attainment and the potential mechanisms driving these patterns.

The structure of our paper is as follows. We outline our structural model in Section 2. Section 3 describes our data and institutional setting, and discusses model implementation and provide evidence on model fit and validation. Section 4 presents our main findings. Section 5 concludes.

2 Model

We now present a model that takes the sequential and uncertain nature of schooling investments into account, besides allowing for nonlinearities in the rewards to such investments. Our model is an example of the Eckstein-Keane-Wolpin (EKW) class of models (Aguirregabiria and Mira, 2010), which are frequently used to study the mechanisms determining human capital investment decisions and to predict the effects of human capital policies (Blundell, 2017; Keane et al., 2011; Low and Meghir, 2017). The model exploits the richness of our data and captures essential features of the Norwegian school system. We start by describing the model setup, and then define our main objects of interest – the ex-ante returns and the option values of schooling.

2.1 Setup

We follow individuals over most of their working life from young adulthood at age 15 to the final period T at age 50.⁶ When entering the model, all individuals have seven years of basic compulsory schooling. Each individual is characterized by one of three levels of measured ability, and within each ability group, by one of several $j \in \mathcal{J}$ latent types (Heckman and Singer, 1984).⁷ Our model treats an individual’s ability group and latent type as predetermined and individuals are assumed to know their ability and latent type. We allow all model parameters to vary across the three ability groups, effectively estimating our model separately for each ability group. In the following, we suppress subscripts for ability groups to ease notation.

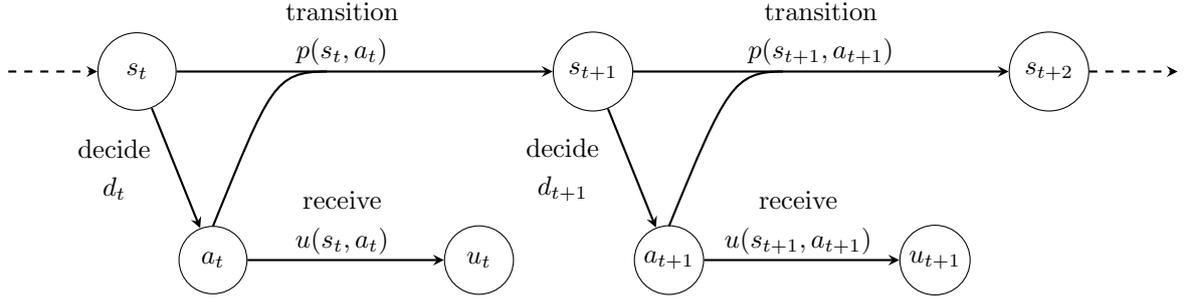
The decision period $t = 15, \dots, 50$ is a school year. Each period individuals observe the state of their choice environment s_t and decide to take action $a_t \in \mathcal{A}$. Individuals can decide whether to work in the private sector ($a_t = PR$), to work in the public sector ($a_t = PU$), to work as self-employed ($a_t = S$), to attend an academic ($a_t = A$) or a vocational ($a_t = V$) schooling track, or to stay at home ($a_t = H$). The decision has two consequences: an individual receives an immediate utility $u(s_t, a_t)$ and the environment is updated to a new state s_{t+1} . The transition from s_t to s_{t+1} is affected by the action but remains partly uncertain. Individuals are forward-looking. Thus, they do not simply choose the alternative with the highest immediate utility. Instead, they take the future consequences of their current action into account.

A policy $\pi = (d_1, \dots, d_T)$ provides the individual with instructions for choosing an action in

⁶In an earlier version of our model, we followed individuals up to age 58. The decision to follow individuals up to age 50 was motivated by computational ease, however, this choice does not comprise our main findings.

⁷To classify individuals in low, medium and high ability groups, we rely on an IQ test score available in our data. See further details in Section 3.1 below. We allow three latent types among medium and high ability groups, and four latent types in the low ability group, so that the total number of distinct ability-by-latent types equals ten. We choose the number of latent types in each ability group based on model diagnostics.

Figure 1: Timing of Events.



any possible future state. It is a sequence of decision rules d_t that specify the planned action at a particular time t for any possible state s_t . The implementation of a policy generates a sequence of utilities that depends on the transition probability distribution $p(s_t, a_t)$ for the evolution of state s_t to s_{t+1} induced by the model. To fix ideas, Figure 1 illustrates the timing of events in the model for two generic periods. At the beginning of period t , an individual fully learns about each alternative's immediate utility, chooses one of the alternatives, and receives its immediate utility. Then, the state evolves from s_t to s_{t+1} and the process is repeated in $t+1$.

Individuals make their decisions facing uncertainty about the future and seek to maximize their expected total discounted utilities over all remaining decision periods. They have rational expectations (Muth, 1961), so their subjective beliefs about the future agree with the objective probabilities for all possible future events determined by the model. Immediate utilities are separable between periods (Kahneman et al., 1997), and individuals discount future over immediate utilities by a discount factor δ (Samuelson, 1937). Equation (1) provides the formal representation of an individual's objective function. Given an initial state s_1 , they implement a policy π^* from the set of all possible policies Π that maximizes the expected total discounted utilities over all decision periods given the information available at the time.

$$\max_{\pi \in \Pi} E_{s_1}^{\pi} \left[\sum_{t=16}^T \delta^{t-16} u(s_t, d_t(s_t)) \right] \quad (1)$$

The immediate utility $u(\cdot)$ of each alternative consists of a non-pecuniary utility $\zeta_a(\cdot)$ and, for the working alternatives, an additional monetary wage component $w_a(\cdot)$. Both depend on an individual's level of human capital as measured by sector-specific work experience $\mathbf{k}_t = (k_{a,t})_{a \in \{PU, PR, S\}}$, years of completed schooling in each track $\mathbf{h}_t = (h_{a,t})_{a \in \{A, V\}}$, and alternative-specific skill endowment $\mathbf{e} = (e_{j,a})_{\mathcal{J} \times \mathcal{A}}$ that depends on the individual's latent type. The immediate utilities are also influenced by the decision a_{t-1} in the previous period, a time trend

t , and alternative-specific shocks $\boldsymbol{\epsilon}_t = (\epsilon_{a,t})_{a \in \mathcal{A}}$. Their general form is given by:

$$u(\cdot) = \begin{cases} \zeta_a(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}, e_{j,a}) + w_a(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}, e_{j,a}, \epsilon_{a,t}) & \text{if } a \in \{PR, PU, S\} \\ \zeta_a(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}, e_{j,a}, \epsilon_{a,t}) & \text{if } a \in \{A, V, H\}. \end{cases}$$

Sector-specific work experience \mathbf{k}_t and years of completed schooling in each track \mathbf{h}_t evolve deterministically. There is no uncertainty about grade completion (Altonji, 1993) and no part-time enrollment. Schooling is defined by time spent in school, not by formal credentials acquired. Once individuals reach a certain amount of schooling, they acquire a degree.

$$\begin{aligned} k_{a,t+1} &= k_{a,t} + \mathbf{I}[a_t = a] & \text{if } a \in \{PU, PR, S\} \\ h_{a,t+1} &= h_{a,t} + \mathbf{I}[a_t = a] & \text{if } a \in \{A, V\} \end{aligned}$$

The productivity and preference shocks $\boldsymbol{\epsilon}_t$ are unknown to the individual in advance, and capture uncertainty about the returns and cost of future schooling. In our model setup, we specify these shocks $\boldsymbol{\epsilon}_t$ to be uncorrelated across time and follow a multivariate normal distribution with mean $\mathbf{0}$ and covariance $\boldsymbol{\Sigma}$. Given the structure of the utility functions and the distribution of the shocks, the state at time t is $s_t = \{\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}, \mathbf{e}, \boldsymbol{\epsilon}_t\}$.⁸

Individuals' ability and latent type are the two sources of persistent heterogeneity in this model. As noted above, we allow all model parameters to vary across three ability groups and further let individuals' alternative-specific skill endowments \mathbf{e} to vary by their latent type. All remaining differences in life-cycle decisions result from differences in the transitory shocks $\boldsymbol{\epsilon}_t$ over time. In each period, individuals can increase their stock of human capital $(\mathbf{k}_t, \mathbf{h}_t)$ by either working in the labor market or enrolling in school. From the beginning, individuals are aware of their level of ability and alternative-specific skill endowments \mathbf{e} . They only learn about the realizations of shocks $\boldsymbol{\epsilon}_t$ at the beginning of each period. As the shocks are distributed independently over time, individuals do not update their prior beliefs about their productivity or alternative-specific tastes (see, e.g. Arcidiacono (2004); Arcidiacono et al. (2016)).

Previous research on the determinants of life-cycle wages and schooling decisions (Keane and Wolpin, 1997; Meghir and Pistaferri, 2011) informs our specification of the immediate utility functions. We specify the wage component $w_a(\cdot) = r x_a(\cdot)$ in the immediate utility from working as the product of the market-equilibrium rental price r and a sector-specific skill level $x_a(\cdot)$ for

⁸While in principle, one could also allow persistence in these shocks over time, the estimation problem becomes computationally more cumbersome as this would increase the state space dramatically. However, as discussed below, since the model includes persistent heterogeneity *and* adjustment costs in moving across states, adding persistence in transitory shocks can also pose challenges for identification (Heckman and Singer, 1984).

$a \in \{PU, PR, S\}$. The skill level is determined by a skill production function, which includes a deterministic component $\Gamma_a(\cdot)$ and a multiplicative stochastic productivity shock $\epsilon_{a,t}$, as follows:

$$x_a(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}, e_{j,a}, \epsilon_{a,t}) = \exp(\Gamma_a(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}, e_{j,a}) \cdot \epsilon_{a,t})$$

The specification above leads to a standard logarithmic wage equation in which the constant term is the skill rental price $\ln(r)$ and wages follow a log-normal distribution. Equation (2) shows the parametrization of the deterministic component $\Gamma_a(\cdot)$ of the skill production function:

$$\begin{aligned} \Gamma_a(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}, e_{j,a}) &= e_{j,a} \\ &+ \underbrace{\delta_{1,a} \cdot h_t^A + \delta_{2,a} \cdot h_t^V + \delta_{3,a} \cdot k_t^W + \delta_{4,a} \cdot (k_t^W)^2 + \delta_{5,a} \cdot k_t^a + \delta_{6,a} \cdot (k_t^a)^2}_{\text{Mincer-inspired returns to schooling and experience}} \\ &+ \underbrace{\sum_{d \in \{9,12,16\}} \gamma_{d,a}^A \cdot \mathbf{I}[h_t^A \geq d] + \sum_{d \in \{9,12\}} \gamma_{d,a}^V \cdot \mathbf{I}[h_t^V \geq d]}_{\text{Non-linear rewards to diplomas or degrees}} \\ &+ \underbrace{\eta_{1,a} \cdot \mathbf{I}[a_{t-1} = W]}_{\text{Skill depreciation}} + \nu_{1,a} \cdot (t - 15) + \nu_{2,a} \cdot \mathbf{I}[t < 17] \end{aligned} \quad (2)$$

The first part of the sectoral skill production function is motivated by Mincer (1974) and hence linear in years of completed schooling by track $(\delta_{1,a}, \delta_{2,a})$, quadratic in past work experience $(\delta_{3,a}, \delta_{4,a}, \delta_{5,a}, \delta_{6,a})$, and separable between the two of them. Both sector-specific experience k_t^a and general work experience $k_t^W = k_t^{PR} + k_t^{PU} + k_t^S$ matter, where we use $W \in \{PU, PR, S\}$ as a generic work choice. We include diploma effects $(\gamma_{d,a}^A, \gamma_{d,a}^V)$ that capture the non-linear rewards associated with a degree d beyond the years of schooling (Hungerford and Solon, 1987; Jaeger and Page, 1996). Skills depreciate by $\eta_{1,a}$ if the individual didn't work in the previous period. Finally, there is an age effect $\nu_{1,a}$ and a penalty when working as a minor $\nu_{2,a}$.

In the following, we discuss how we parametrize the immediate non-pecuniary reward of working in each sector and the immediate utilities of attending academic and vocational schooling, and staying at home, respectively.⁹ Equation (3) shows the non-pecuniary rewards to working in each sector, i.e., for $a \in \{PU, PR, S\}$. We allow the immediate non-pecuniary reward (i.e., disutility) of work to depend on accumulated work experience in the particular sector k_t^a , accumulated general work experience k_t^W and years of completed schooling in each track \mathbf{h}_t , and allow for age and diploma effects as in equation (2). Further, we include indicators for previous period sector-specific work $\mathbf{I}[a_{t-1} = a]$ and previous period private sector work $\mathbf{I}[a_{t-1} = PR]$, which capture the fixed costs of market entry and sector switching.

⁹Note that to enhance model performance, we varied some of these specifications slightly across the three ability groups. We include a more detailed exposition of reward functions by ability groups in the Appendix.

$$\begin{aligned}
\zeta_a(k_t, \mathbf{h}_t, t, a_{t-1}) = & e_{j,a} + \underbrace{\beta_{1,a} \cdot \mathbf{I}[t < 17]}_{\text{Age Effects}} + \underbrace{\beta_{2,a} \cdot k_t^W + \beta_{3,a} \cdot k_t^a + \beta_{4,a} \cdot h_t^A + \beta_{5,a} \cdot h_t^V}_{\text{Experience Effects}} \\
& + \underbrace{\sum_{d \in \{9,12,16\}} \vartheta_{d,a}^A \cdot \mathbf{I}[h_t^A \geq d] + \sum_{d \in \{9,12\}} \vartheta_{d,a}^V \cdot \mathbf{I}[h_t^V \geq d]}_{\text{Degree Effects}} \\
& + \underbrace{\beta_{6,a} \cdot \mathbf{I}[a_{t-1} = PR] + \beta_{7,a} \cdot \mathbf{I}[k_t^a > 0] + \beta_{8,a} \cdot \mathbf{I}[k_t^W > 0] + \beta_{9,a} \cdot \mathbf{I}[a_{t-1} = a]}_{\text{Labor Market Entry and Sector Switching Costs}}
\end{aligned} \tag{3}$$

Next, in Equation (4), we show the immediate rewards of academic and vocational education, i.e., $a \in \{A, V\}$. They include age ($t - 15$), an indicator for whether they have been enrolled in the track in the past period $\mathbf{I}[a_{t-1} = a]$, an indicator for whether they have nonzero experience in the other schooling track $\mathbf{I}[h_t^{a^C} > 0]$ to capture the costs of track switching. Furthermore they include diploma effects to capture changes in the programs as people progress over time.

$$\begin{aligned}
\zeta_a(k_t, \mathbf{h}_t, t, a_{t-1}) = & e_{j,a} + \underbrace{\beta_{1,a} \cdot (t - 15) + \beta_{2,a} \cdot \mathbf{I}[t < 17]}_{\text{Age Effects}} + \underbrace{\beta_{3,a} \cdot \mathbf{I}[h_t^{a^C} > 0]}_{\text{Track Switching Cost}} \\
& + \underbrace{\beta_{4,a} \cdot \mathbf{I}[a_{t-1} = a] + \beta_{5,a} \cdot \mathbf{I}[a_{t-1} \in \{PR, PU, S, H\}] + \beta_{6,A} \cdot \mathbf{I}[a_{t-1} \neq A] \cdot \mathbf{I}[h_t^A \geq 12]}_{\text{Reenrolment Costs}} \\
& + \underbrace{\vartheta_{9,a}^A \cdot \mathbf{I}[h_t^V \geq 9] + \vartheta_{9,a}^A \cdot \mathbf{I}[h_t^A \geq 9]}_{\text{Degree Effects}} + \epsilon_{t,a}
\end{aligned} \tag{4}$$

Finally we discuss the non-pecuniary reward to home production contained in Equation (5), i.e., $a = H$. The utility of staying at home is allowed to depend on whether one is below age 17, which captures different opportunity costs of home production for minors. Furthermore it depends on indicators for having completed a high school or a college degree which captured differences in labor market transitions across education levels.

$$\begin{aligned}
\zeta_H(k_t, \mathbf{h}_t, t, a_{t-1}) = & e_{j,H} + \underbrace{\beta_{1,H} \cdot \mathbf{I}[t < 17]}_{\text{Age Effects}} + \underbrace{\sum_{d \in \{12,16\}} \vartheta_{d,H}^A \cdot \mathbf{I}[h_t^A \geq d] + \vartheta_{12,H}^V \cdot \mathbf{I}[h_t^V \geq 12]}_{\text{Degree Effects}} + \epsilon_{t,H}
\end{aligned} \tag{5}$$

2.2 Objects of Interest

We now define two primary objects of interest in our analysis within the framework of the above model, namely the ex-ante return to schooling and the option value of schooling. While in the empirical analysis, we will present evidence on these objects separately by academic and vo-

cational schooling, we refer to a generic schooling choice $G \in \{A, V\}$ here to ease the exposition.

We define these objects in terms of value functions $v(s_t, a)$. The value functions are alternative- and state-specific and summarize the total value that individuals receive of choosing alternative a for a given state s_t , including the immediate reward and the discounted future rewards, assuming that the optimal policy π^* is followed in the future:

$$v(s_t, a) = u(s_t, a) + \delta E_{s_t, a} [v^{\pi^*}(s_{t+1})] \quad \forall a \in \mathcal{A}$$

Accordingly, the total value of schooling $v(s_t, G)$ in state s_t captures the immediate and expected future benefits from continuing one's education in another period, subject to optimal policy π^* .

Ex-ante Return Next, we describe the elements that are needed to isolate the ex-ante return to an additional year of schooling. The thought-experiment we perform is to compare the total value of schooling $v(s_t, G)$ against the value of choosing the best alternative choice, which can be another schooling track. Notably, the alternative choice can also contain the option of taking more schooling at a later stage in the life-cycle through re-enrollment. Quantifying the ex-ante return requires making comparisons of counterfactuals as a given individual in state s_t in the estimated model only chooses one of the available alternatives. We construct such counterfactuals through model simulations, where we require each individual to make alternative choices in state s_t , but restrict the realizations of shocks in each period to be held fixed across comparisons.

We denote by $ER(s_t)$ the ex-ante return capturing the value of an additional year of schooling against the best alternative choice in state s_t . Formally, we can express this object as follows:

$$ER(s_t) = \frac{v(s_t, G) - \tilde{v}(s_t)}{\tilde{v}(s_t)}, \quad \text{where } \tilde{v}(s_t) = \max_{a \neq G} \{v(s_t, a)\}. \quad (6)$$

In our model, the ex-ante return of an additional year of schooling in a particular educational track is positive for all individuals who appear in state s_t making this choice and negative for those that decide to attend another educational track, work or stay at home instead.

We also consider the ex-ante lifetime earnings returns associated with each choice. The expected lifetime earnings value associated with a choice a at state s_t is the expected lifetime discounted earnings conditional on choosing the optimal policy π^* from the next period onward.

$$W(s_t, a) = \sum_{\tau \geq t} \delta^{\tau-t} E_{s_t, a}^{\pi^*} [w_\tau] \quad (7)$$

The expected ex-ante earnings return is then defined as follows:

$$\text{ERW}(s_t) = \frac{W(s_t, G) - \tilde{W}(s_t)}{\tilde{W}(s_t)}, \quad \text{where} \quad \tilde{W}(s_t) = \max_{a \neq G} \{W(s_t, a)\} \quad (8)$$

Option Value We are further interested in the option value of schooling $OV(s_t)$. This object captures the part of the value of another year of schooling that can be attributed to having an opportunity to pursue further schooling in the future. This component arises due to the sequential nature of schooling investments. To compute the option value component, we perform another counterfactual comparison. We now compare the total value of schooling $v(s_t, G)$ in state s_t to the value of the same alternative but with an optimal policy $\hat{\pi}$ that does not allow one to increase schooling further beyond the next period. We use superscripts π^* and $\hat{\pi}$ to differentiate between the two total values. The latter is by construction a counterfactual scenario, unless one has already reached the maximum level of schooling.

$$OV(s_t) = v^{\pi^*}(s_t, G) - v^{\hat{\pi}}(s_t, G)$$

The option value of schooling is non-negative at all states and zero once an individual attains the maximum schooling level. The option value increases with the future benefits of pursuing higher education and the probability of doing so. To calculate this value, we compare the total value of continued schooling under the scenario that the individual *may* continue to increase their schooling level in the future periods and the counterfactual where it is impossible to do so.

As a measure for the importance of the option value, we compute its contribution to the overall value of a state by taking the following ratio:

$$\text{OVC}(s_t) = \frac{OV(s_t)}{v^{\pi^*}(s_t)}. \quad (9)$$

In the empirical analysis, we will report estimates of the option value contributions based on the above measure, which provides a decomposition of the total value of schooling in a state.

3 Data and Implementation

In this section, we first describe our data sources, then briefly describe the Norwegian education system and the compulsory schooling reform, before we discuss the implementation and estimation of our model on these data, and finally, provide evidence on model fit and validation.

3.1 Data Sources

Our empirical analysis uses several registry databases maintained by Statistics Norway. First, we use the Norwegian National Education Database, a comprehensive population-wide event-history dataset with information on the dates of enrollment, termination and completion of 6-digit educational courses for all residents since 1970. Second, we use a longitudinal dataset containing annual earnings and tax records for all Norwegians for every year from 1967 onwards. Third, we use demographic information (e.g., cohort of birth, gender and childhood municipality of residence) for all individuals ever registered in the Norwegian Central Population Register, established in 1964. Fourth, we are also able to access supplementary demographic data from the Decennial Population Censuses held in 1960 and 1970, and link information on the primary sector of work based on a register of public sector employees established in 1974, and more detailed matched employer-employee data available from 1983. Finally, we received information on IQ test scores from the Norwegian Armed Forces for male conscripts born in 1950 and later. Importantly, each of these datasets include unique personal identifiers which allow us to follow individuals' educational choices and earnings across time, and add a proxy of ability for males.

Sample construction We restrict our sample to Norwegian males born between 1955 and 1960. We can follow each of these individuals' educational choices and earnings from age 15 and up to age 50.¹⁰ Our initial sample consists of 176,804 Norwegian-born males. Dropping individuals with missing information on childhood municipality of residence, exposure to compulsory schooling reform and education enrollment or attainment, we retain 165,171 individuals. Further dropping individuals with missing information on IQ test scores in the Norwegian military records, we retain 136,292 individuals (i.e., around 77% of the initial sample). Using annual information on individuals' educational choices and earnings, we create a weakly-balanced panel of individuals entering our sample across 36 annual observations, which an individual can exit only due to natural attrition (i.e., death or out-migration). Our panel dataset thus consists of 4,820,863 person-year observations.

We further split the analytical sample in two parts, depending on the type of compulsory schooling system each individual was subject to, exploiting variation in the timing of a compulsory schooling reform across different municipalities in Norway (see details in Section 3.4). Specifically, there are 9,156 individuals in our sample (i.e., around 7%) who grew up in one of the 200 (out of 732) municipalities which hadn't implemented the compulsory schooling reform by the year they turned age 14 (threshold age for completely compulsory schooling) and as such were subject to the pre-reform education system. We will utilize of this sample of 9,156

¹⁰We observe educational choices annually for birth cohort 1955 at age 15 and onwards since the National Education Database was established in 1970. Educational histories are partially observed for earlier cohorts.

individuals and 324,157 person-year observations to estimate our structural model, accounting for key features of the pre-reform education system, and refer to this as the *estimation sample*. We will use the remaining sample of 127,136 post-reform individuals and 4,496,706 person-year observations to validate the structural model, and refer to this as the *validation sample*.

Education Our education information is primarily based on the Norwegian National Education Database (NUDB), which in many respects is an ideal dataset to study the enrollment, drop-out and completion behavior of individuals across time. The NUDB is an event-history dataset providing population-wide information on the dates of each enrollment and exit within a 6-digit educational course code across all lower secondary educations to tertiary educations. The detailed classification of educational course codes allows distinguishing educations by the level of attainment, the standard length of each course/degree, and the type of field for secondary (vocational/academic) and tertiary educations. Each entry in this dataset further has information on the outcome of each enrollment, e.g., allowing the researcher to distinguish drop-out/early terminations and successful completions of a course, and whether the individual was enrolled as a part-time or as a full-time attendant in a specific course.

Combining information obtained from the NUDB and Statistics Norway’s Education Register, where the latter also comprises information on compulsory education attainment, we can classify individuals’ educational choices in a detailed manner across all education levels.¹¹ Noteworthy, information in both the NUDB and the Education Register is based on the annual reports submitted by educational establishments for each of their attendants directly to Statistics Norway, which minimizes the chance of misreporting. In comparison, survey-based data that are often used to study enrollment, drop-out and completion behavior may suffer from non-response or other biases due to misreporting.

Earnings and Work Our earnings data are based on annual tax records. Our earnings measure is the sum of labor income (from wages and self-employment) and work-related cash transfers (such as unemployment benefits and short-term sickness benefits). This dataset has several advantages over those available in most countries. First, there is no attrition from the original sample other than natural attrition due to either death or out-migration. Second, our earnings data pertain to all individuals, and are not limited to some sectors or occupations. Third, we can construct long earnings histories that allow us to estimate the returns to education across the life-cycle. Finally, using the annual tax records, we are able to distinguish between individuals

¹¹Earlier studies on the returns to education in Norway have relied on Statistics Norway’s Education Register and used information on an individual’s highest completed education level or the years of schooling corresponding to the highest level of education, see, e.g., Aryal et al. (2022), Bhuller et al. (2017) and Aakvik et al. (2010). Neither of these studies consider the ex-ante returns or the option value of educational choices.

who work as salaried workers and self-employed, respectively. Further, we add information on the primary sector of employment for salaried workers by relying on information recorded in the register of public sector employees from 1974 and matched employer-employee data from 1983 and onwards. This allows us to track individuals' sectoral choices over their careers.¹²

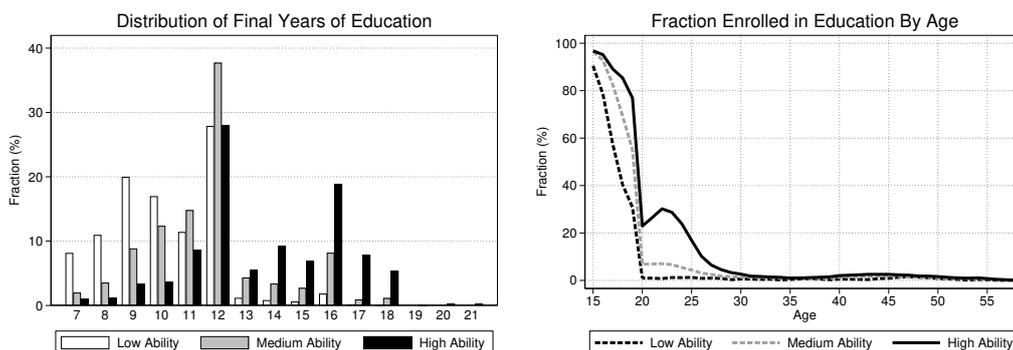
Ability An important measure we exploit in our analysis to capture observational heterogeneity is an IQ test score accessed from the Norwegian Armed Forces. In Norway, military service was compulsory for all able males in the birth cohorts we study. Before each male entered the service, his medical and psychological suitability was assessed. Most eligible Norwegian males in our sample took this test around their 18th birthday. The IQ test score is a composite unweighted mean from three speeded tests--arithmetics, word similarities, and figures (Sundet et al. (2004)). The score is reported in stanine (standard nine) units, a method of standardizing raw scores into a 9-point standard scale with a normal distribution, a mean of 5.8, and a standard deviation of 1.7. This score is strongly related to individuals' actual completed education, with a correlation of around 0.5 with years of schooling in our analytical sample.

Descriptives Figure 2 presents key variables in our dataset. In each panel, we categorize individuals as either low (scores 1-4), medium (scores 5-6) or high (scores 7-9) ability. Panel (a) shows the distributions of final years of education, where we find clear associations between education and ability. Panel (b) shows the fraction of individuals enrolled in education by age. Individuals are more likely to attend education during the early part of their life-cycle, with gradual declines in the enrollment rate up to age 30, and virtually no enrollment beyond age 33. Next, in panels (c)-(d), we consider the conditional exit and re-enrollment rates, focusing on the earlier part of the life-cycle. The exit rate is substantially higher among low- and medium-ability individuals, while high-ability individuals have consistently higher re-enrollment. Panel (e) illustrates the choices of academic and vocational tracks in middle and high school, reflecting clear ability-related differences in track choices. Next, panel (f) shows employment rates by ability. While low-ability individuals reach an employment rate of 96 percent already at age 20, high- and medium-ability individuals gradually increase their employment rates until age 30. Beyond age 30, the employment rates remain relatively stable across all groups, though low-ability have earlier labor market exits. Panels (g)-(h) show age-specific means and standard deviations of annual earnings (in 1000s, 2015-Norwegian Kroner) conditional on working, respectively, by ability. The age-earnings profiles are increasing for all ability groups up to age 45, aside from a temporary drop in earnings at ages 19/20 due to military service. Earnings dispersion is stable in the earlier parts of the life-cycle and relatively similar across ability groups, but increases sharply after age 35.

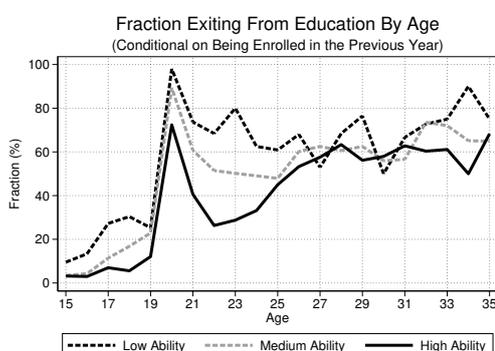
¹²Unfortunately, Norwegian register data contain limited information on occupations prior to 2003.

Figure 2: Descriptive Statistics.

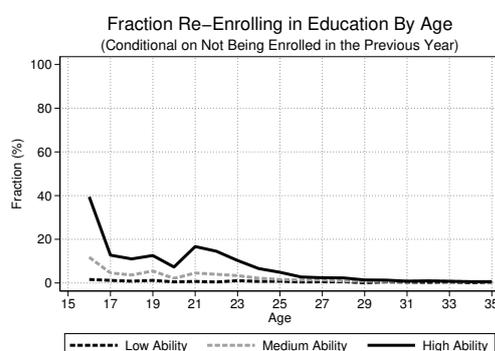
(a) Distribution of Final Years of Education (b) Fraction Enrolled in Education by Age



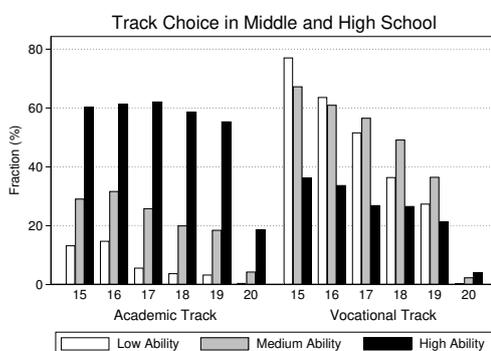
(c) Exit Rate By Age



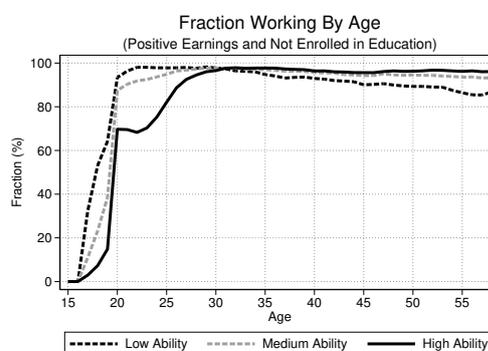
(d) Re-enrollment Rate by Age



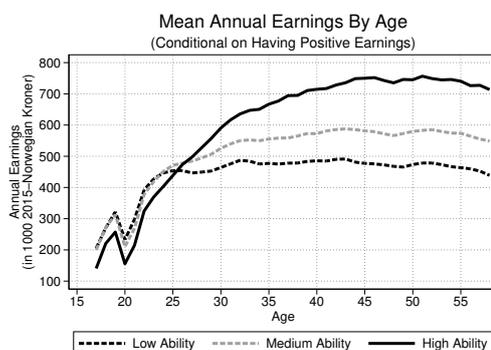
(e) Track Choice in Middle/High School



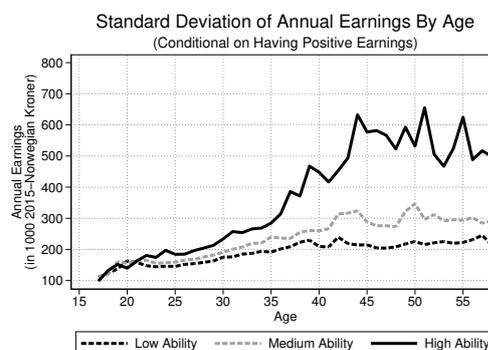
(f) Fraction working by age



(g) Mean Earnings by Age

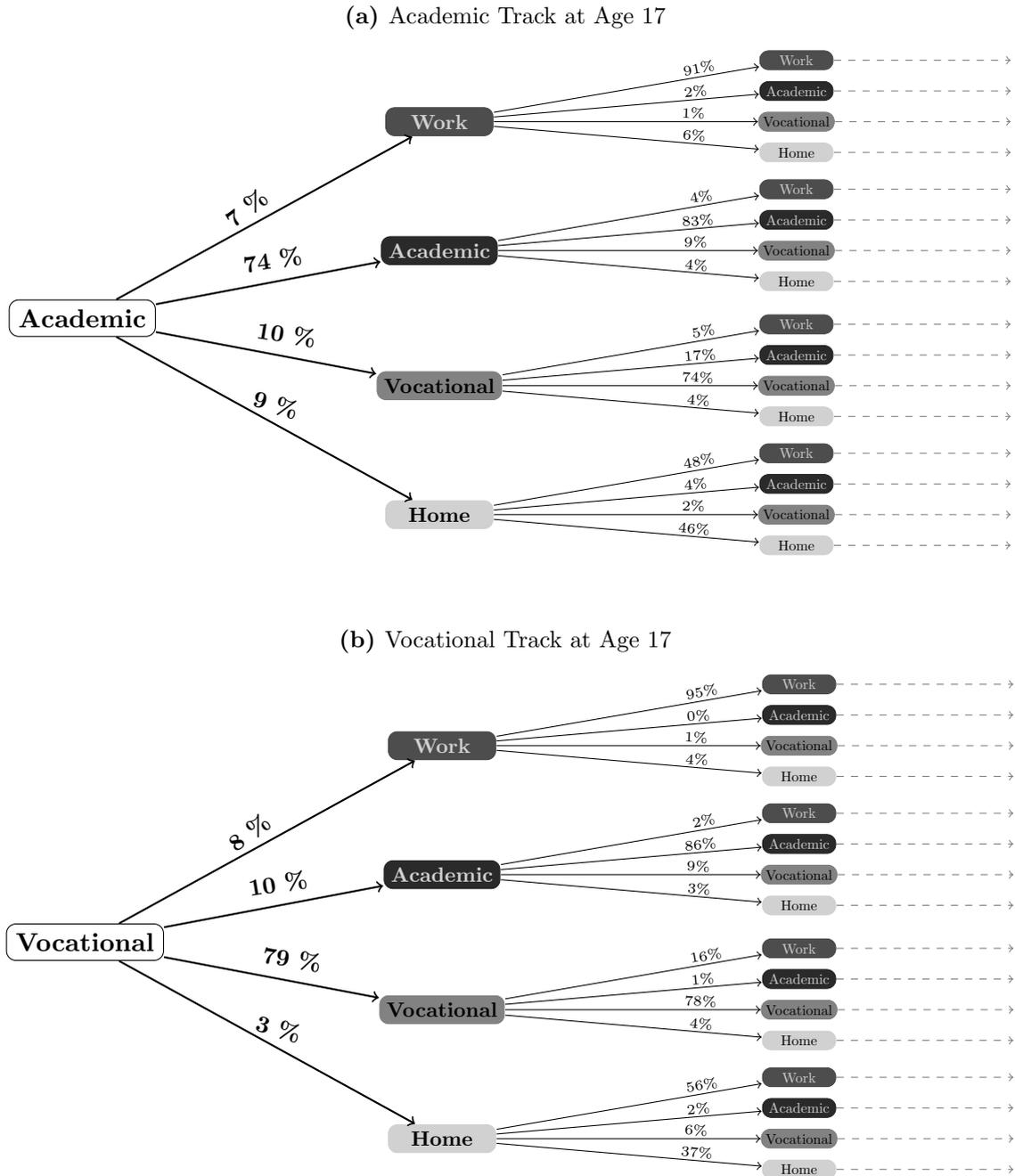


(h) Standard Deviation of Earnings by Age



Note: The sample consists of Norwegian males born 1955-1960, who grew up in a municipality which hadn't implemented the compulsory schooling reform by the year they turned age 14 (see details in Section 3.2). Individuals are followed over ages 15-58, corresponding to calendar years 1970-2018, unless there is natural attrition due to death or out-migration. Individuals' ability is split in three discrete categories, constructed as low (stanine IQ scores 1-4), medium (scores 5-6) and high (scores 7-9). N=9,143.

Figure 3: Illustration of the Decision Tree and Transitions.



Note: This figure shows the decision tree between ages 17 to 19 for those attending the academic (a) or vocational (b) tracks at ages 15–17, after completing compulsory schooling at age 14. This also shows the fraction transiting from one state to another at each age. The corresponding decision tree and transitions between ages 14 to 17 are illustrated in the Appendix Figure A.2.

Besides some of the key moments of our dataset that are illustrated in Figure 2, our model implementation will also rely crucially on the fractions of individuals who transit between different choices over their life-cycle. To illustrate the rich transition patterns present in our data, in Figure 3, panels (a) and (b), we consider individuals who were enrolled in the academic and vocational tracks, respectively, at ages 15 to 17 after having compulsory schooling at age 14, and follow their transition histories over the following three years in our data. Around 74%

(79%) of these individuals continue in the academic (vocational) track for a third year while 10% switch to the other educational track. Less than 10% enter the labor market, while the remaining stay at home. Among those who decide to stay at home for one period, roughly 6-8% re-enroll in an academic or vocational school. These rich patterns of (i) persistence in choices over time, (ii) presence of track switching, and (iii) re-enrollment after spells of work or home stay provide a motivation for the flexible modelling of schooling choices in Section 2.

3.2 The Norwegian Education System

We describe here the structure of the Norwegian education system that existed in the 1960s, which our model is specified to fit. This system had four main stages; see Appendix Figure A.1. The first stage consisted of seven years of compulsory elementary education. The second stage involved tracking, where pupils could attend either a vocational middle school (*framhaldsskole*) or an academic middle school (*realskole*). The vocational middle school could be either one, two or three years, with most attending two years, while the academic middle school could be either two or three years, where the final year was only required for those who wanted to later pursue further academic education. The third stage corresponded to a high school education, which again was track-specific. After attending the academic middle school, students could move on to attend a academic high school (*gymnas*). In contrast, pupils attending the vocational middle school normally didn't qualify for the academic high school, but could rather attend a vocational high school (*yrkesskole*). The academic high school was required to be three years, while the vocational high school could be of varying lengths, depending on the particular vocational field. Finally, the fourth stage involved higher education, leading up to an academic degree at a college or a university, enrollment to which was typically contingent on having completed the academic high school. There existed two main degrees in tertiary education; a 4-year (*cand.mag*) and a 6-year (*hovedfag*), but degrees of other durations also existed.

3.3 Model Implementation and Estimation

The model we described in Section 2 is set up as a standard Markov decision process (MDP), which can be solved by a simple backward induction procedure (Puterman, 1994; White, 1993; Rust, 1994). In the final period T , there is no future to consider, and the optimal action is choosing the alternative with the highest immediate utility in each state. With the decision rule for the final period, we can determine all other optimal decisions recursively.¹³

¹³We use an open-source research code *respy* (Gabler and Raabe, 2020) that allows for the flexible specification, simulation, and estimation of EKW models. Detailed documentation of this software and its numerical components is available at <http://respy.readthedocs.io>.

Table 1: Summary of Moments Used in the Estimation.

Type of Moment	Number
I. Average of Annual Earnings Per Period and Sector	288
II. Standard Deviation of Annual Earnings Per Period and Sector	288
III. Fraction in Academic Schooling Per Period	108
IV. Fraction in Vocational Schooling Per Period	108
V. Fraction Working in Public Sector Per Period	108
VI. Fraction Working in Private Sector Per Period	108
VII. Fraction Working as Self-Employed Per Period	108
VIII. Fraction Staying at Home Per Period	108
IX. Distribution of Final Schooling Levels Per Track	21
X. Distribution of Final Schooling Levels Per Sector	54

Note: This table provides an overview of the 1,083 moments used in the estimation by the type of moment. Moments of type I–II are each calculated between ages 18 and 50 in the private sector and ages 21 and 50 in the public sector and in self employment. Thus there are 96 moments in each ability group and $96 \times 3 = 288$ moments in total. Moments of type III–VII are calculated for each of the 36 periods across all ability groups $36 \times 3 = 108$. Moments of type IX capture the distribution of schooling degrees in each of the two schooling types, which can take four values in academic schooling and three values in vocational schooling, calculated for each of the three ability types, i.e., $7 \times 3 = 21$ unique moments. Moments of type X capture the distribution of final schooling degrees which can take six values in each sector and ability group, i.e., $6 \times 3 \times 3 = 54$ unique moments.

Implementation We use the method of simulated moments (Pakes and Pollard, 1989; Duffie and Singleton, 1993) to estimate the parameters $\hat{\theta}$ of the model for each ability group. Equation (10) shows our criterion function. We select the parameterization for our analysis that minimizes the weighted squared distance between our specified set of moments computed on the observed \mathbf{M}_D and the simulated data $\mathbf{M}_S(\theta)$. We weigh the moments with a diagonal matrix \mathbf{W} that contains the variances of the observed moments (Altonji and Segal, 1996) and use the *estimagic* package (Gabler, 2022) for the optimization of the criterion function.¹⁴ We simulate a sample of 100,000 individuals based on the candidate parameterizations of the model.

$$\hat{\theta} = \arg \min_{\theta \in \Theta} (\mathbf{M}_D - \mathbf{M}_S(\theta)) \mathbf{W}^{-1} (\mathbf{M}_D - \mathbf{M}_S(\theta))' \quad (10)$$

Table 1 provides an overview of the 1,083 empirical moments used in our estimation. These consist of aggregate moments of annual earnings (type I–II), aggregate annual choice proportions in each alternative (type III–VIII), and the distribution of final schooling (type XI–X). Moments of type I–II are each calculated between ages 18 to 50 in the private sector and ages 21 to 50 in the public sector and in self-employment. Thus, there are 96 such moments in

¹⁴We use a global version of the BOBYQA algorithm within the package (Powell, 2009).

each ability group and a total of $96 \times 3 = 288$ unique moments. Moments of type III–VIII are calculated for each of the 36 periods across all ability groups, i.e., there are $36 \times 3 = 108$ unique moments. Moments of type IX capture the distribution of schooling levels in each of the two educational tracks, which can take four values in academic schooling and three values in vocational schooling, calculated for each of the three ability types, i.e., $7 \times 3 = 21$ unique moments. Moments of type X capture the distribution of final schooling levels which can take six values in each sector and ability group, i.e., $6 \times 3 \times 3 = 54$ unique moments.

Identification The moments are selected to determine the various components of our model. In terms of identification, it is useful to distinguish five distinct groups of parameters, for each ability group. These include (i) wage parameters, (ii) non-pecuniary parameters, (iii) latent types, (iv) distribution of shocks, and (v) discount factor. Wage parameters and non-pecuniary parameters are fixed across individuals and vary across states. Latent types are fixed across states but vary across individuals. Shocks vary both across states and individuals but are independent of each other. The discount factor is fixed across individuals and across states.

While all parameter are identified jointly by all the moments, we provide some heuristic arguments for their identification. Wage parameters determine average wage returns to work choices across different states in the model. Changes in average wage moments as the composition of the work force changes over time is an important factor in the identification of wage returns. Intuitively, the movements in aggregate wages across periods where people with more schooling enter the labor market help to identify the wage rewards to additional years of schooling. And, aggregate wages across periods where more high school graduates enter the labor market help to identify the wage rewards to a high school degree. Including the distribution of final schooling level as an additional set of moments further allows recovering non-linearities in the wage-schooling relationship and associated bunching at specific degrees.¹⁵ Similar arguments apply for the identification of wage rewards to additional years of labor market experience.

Non-pecuniary parameters refer to the non-monetary value of choices in the model. Generally they can be identified by differences in choices across states keeping wages and the composition of individuals fixed. As an example take the non-pecuniary reward to working in the private sector with a high school degree. Given wage returns and non-pecuniary schooling returns, the fraction of private sector workers with a high school degree and the variation in sector choices across individuals with different schooling levels are important for identifying these returns.

¹⁵Note that we do not rely on the cross-sectional wage-schooling relationship directly in our estimation, as these moments usually suffer from problems related to endogeneity of schooling and sample selection bias. Using the distribution of final schooling levels and average earnings and choice shares by period, we can nonetheless recover the parameters of the wage-schooling relationship in a relatively flexible manner.

Latent types and shock parameters introduce heterogeneity across individuals. Latent types introduce persistent heterogeneity across individuals whereas shocks introduce temporary and serially uncorrelated heterogeneity across individuals. The model picks the combination of persistent and temporary heterogeneity that leads to the best model fit conditional on wage parameters and non-pecuniary parameters. Consider as an example the case where half of individuals in our data leave school at age 16 and the other half at age 24. This pattern would be much easier to reconcile with persistent differences than with serially uncorrelated taste shocks. In this case the model would load most of the heterogeneity across individuals into two different latent types. On the other hand, if the shares of individuals leaving school varies more continuously across periods, then this is more suggestive of a larger role of independent taste shocks. The key assumption that this representation places on the data generating process is that persistent heterogeneity does not systematically vary across states. If individuals differ in a systematically different way before and after they have obtained a high school degree for example then the approach would be invalid. Note, however, that both latent types and the distribution of shocks can be distinguished from non-pecuniary returns because they do not systematically vary across states. If choices or outcomes vary across states while the composition of individuals across these states does not change, then non-pecuniary parameters are likely to drive these differences. A similar argument holds for identifying wage parameters versus latent types and the distribution of shocks. If the change in average wage after individuals receive high school degree implies very large returns for labor market entrants, then this is not line with the choices of individuals who have not obtained a high school degree. This would rather suggest that some part of the increase is due to selection as opposed to returns.

Finally, conditional on the structure of wages, non-pecuniary rewards, and latent types, the average changes in choices over time provide variation that help identify the discount factor.

3.4 Model Validation

We now demonstrate our model’s credibility by discussing some selected parameter estimates and comparing them to the existing literature. We then report the estimated model’s in-sample model fit and discuss the results from an out-of-sample validation based on a schooling reform.

Parameter Estimates All parameter estimates are reported in Appendix, Section A.2, along with the associated standard errors based on simulation-based inference. In the following, we discuss some of these estimates. Most of our parameter estimates are standard and in line with the previous literature (Eisenhauer et al., 2015; Keane and Wolpin, 1997). The annual discount rate is about 4% for all ability levels. Returns to sector-specific experience are concave. Latent type heterogeneity plays an important role in shaping schooling decisions even within ability

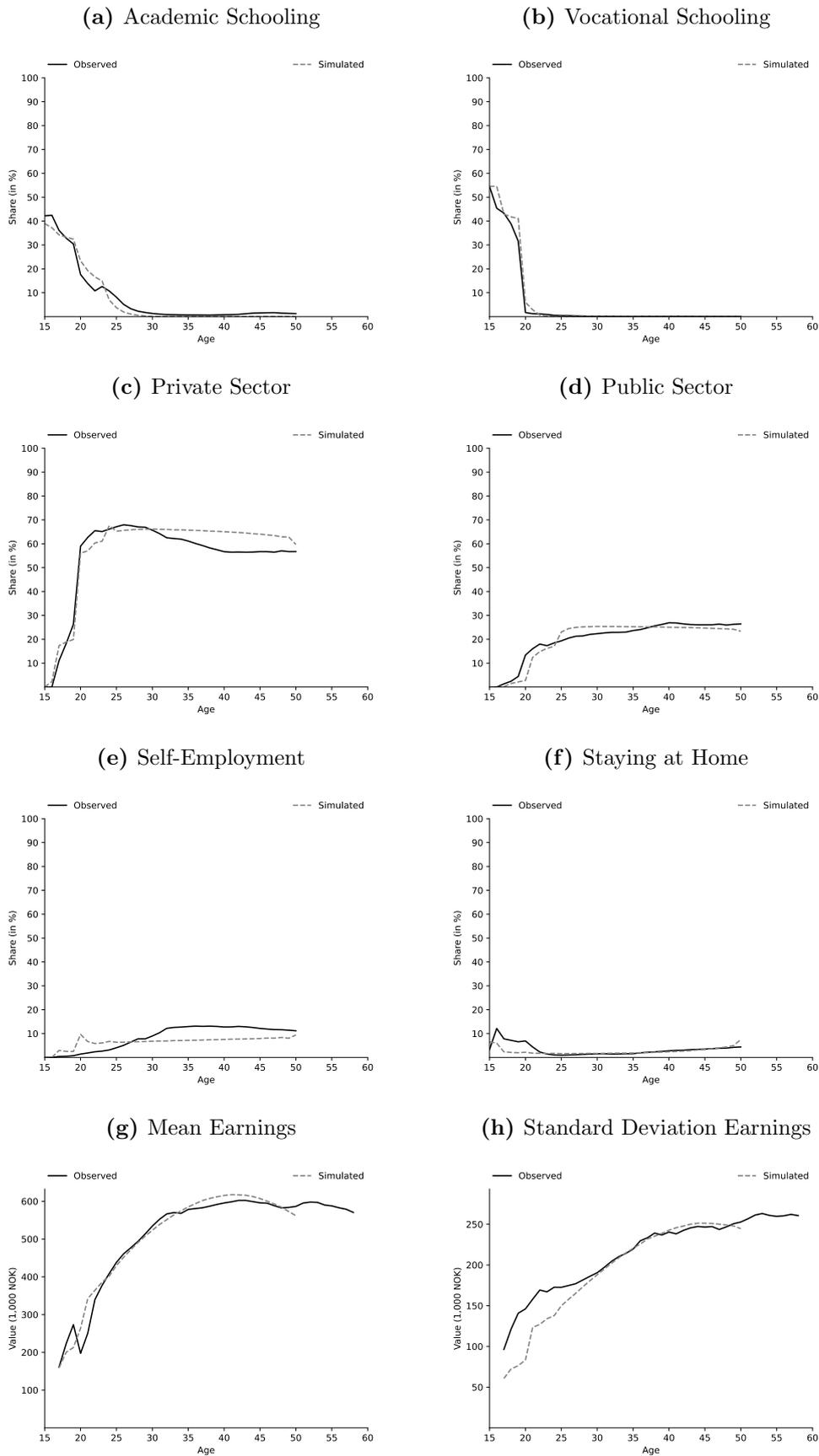
groups. The cost of re-enrollment in school after dropping out is very high. There are interesting differences in the wage returns to academic and vocational schooling by ability and sector. For the academic choice, we find the highest wage returns for high ability individuals, for whom an additional year of schooling increases wages by 16 – 17% in all sectors. By comparison, we find only around 7% return to an additional year of academic schooling for low ability individuals in the public sector. For the vocational choice, however, we find the largest wage returns for low ability individual in all sectors. It should be noted that these returns abstract from the additional degree premiums associated with completing particular years, which are also important elements of our model specification. Further, the associated non-pecuniary rewards reinforce the sorting of high-ability individuals into academic and low-ability individuals into a vocational school. For example, the non-pecuniary benefits from an academic education are negative for low ability individuals but positive for high ability. These patterns likely reflect the heterogeneous drop-out risks across ability groups found in our data.

In-Sample Model Fit We now assess our model’s ability to reproduce the overall patterns of choices and earnings by comparing our simulated sample to the observed data. Figure 4 shows the shares of individuals deciding to either attend academic (panel a) or vocational (panel b) school, private-sector work (panel c), public-sector work (panel d), self-employment (panel e) or stay at home (panel f). The model predictions (dotted gray lines) are in most cases closely aligned with the observed patterns in our data (black solid lines). Next, we show the model fit for the average (panel g) and standard deviation (panel h) of annual earnings by age. Our model does a decent job of reproducing these basic patterns over the life cycle as well.

Out-of-Sample Model Validation We now assess the out-of-sample performance of our model. For this purpose, we rely on variation in schooling choices coming from a compulsory schooling reform. As discussed in Black et al. (2005), since 1959, seven years of elementary education had been compulsory in Norway. However, each municipality—the lowest level of local administration—was allowed to enact nine years of compulsory school, i.e., *two additional years beyond the national minimum requirement*. In subsequent legislation in 1969, nine years of elementary education (*grunnskole*) was made compulsory throughout Norway. Due to the lack of resources some municipalities nevertheless didn’t enforce nine years of compulsory education before 1974. These features led to substantial geographic variation in compulsory education across Norway between 1960 and 1975. For more than a decade, Norwegian schools were divided into two separate systems, where the length of compulsory schooling depended on the birth year and the municipality of residence at age 14, i.e., the childhood municipality.¹⁶

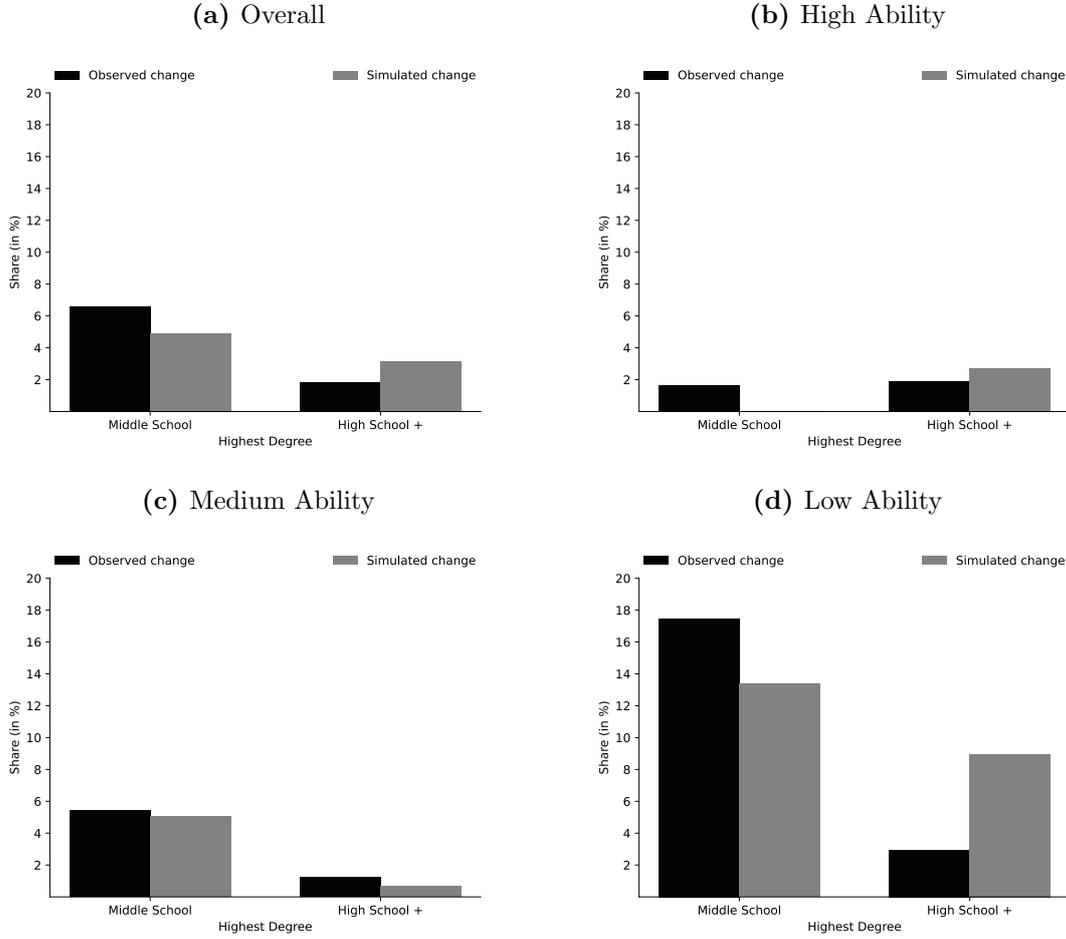
¹⁶The staggered roll-out of the compulsory schooling reform in Norway has led to extensive literature relying on quasi-experimental designs to study the causal effects of schooling, see, e.g., Black et al. (2005), Monstad et al. (2008), Aakvik et al. (2010), Machin et al. (2012), Bhuller et al. (2017), and Aryal et al. (2022). None

Figure 4: Model Fit.



Note: The figure is based on averaging across 100,000 simulated life-cycle profiles using the estimated model.

Figure 5: Out-of-Sample Validation Using Compulsory Schooling Reform.



Note: The figure is based on two samples of 100,000 simulated schooling careers based on the estimated model. We first simulate the model with seven years of compulsory schooling, as in our baseline simulations. Next, we rerun the simulation but impose nine years of compulsory schooling. Throughout, we keep the random realizations of the productivity and taste shocks ϵ_t fixed, and we are thus able to compare the schooling decisions of the same individual under the two different regimes. ‘Middle School’ represents all individuals between 9 and 12 years of academic schooling and 9 and 11 years of vocational schooling. ‘High School +’ represents all individuals with 12 or more years of academic schooling and with 11 or more years of vocational schooling. The black bars *Simulated change* show the percentage points differences in the fraction of individuals who receive a middle school degree and those that receive at least a high school degree, respectively. The gray bars *Observed change* show the same difference in the observed data before and after the reform. We do not distinguish between vocational and academic tracks in this calculation.

We exploit the Norwegian compulsory schooling reform to validate our model in the following manner. First, as described in Section 3.1, our model is estimated solely using data on individuals born 1955–1960 who were subject to the pre-reform education system, and is geared to capture the schooling system that existed pre-reform. Second, for the purposes of model validation, we constraint the choice sets in a modified version of our model where individuals cannot leave schooling before nine years to reflect the implementation of a nine year compulsory schooling. We then compare the predicted changes in education choices in our model to the

of these studies consider the ex-ante returns or the option values of schooling choices.

observed differences in education choices across pre- and post-reform cohorts born 1955–1960, respectively. Finally, in order to provide an insightful and strong validation of our model, we focus on “inframarginal” responses beyond the new minimum schooling requirement. We focus on such responses as the reform is mechanically expected to induce increases in educational attainment up to the new minimum schooling requirement among those who otherwise would have had less than nine years of schooling, both in the observed data and in our model, while it is less clear how educational choices beyond the new minimum schooling requirement are affected.

Figure 5 provides the results of our validation exercise. We compare the predicted changes in the fractions of individuals with nine years of completed schooling or more but less than a high school degree, i.e., middle school, and the fractions of individuals with at least a high school degree, respectively, in model simulations pre- and post-reform to the corresponding changes observed in actual data. The middle school margin is directly affected by the reform as the compulsory schooling reform forces individuals to attend at least nine years of schooling. The high school margin on the other hand is not directly affected by the reform. The validation exercise highlights the main motivation of our modelling approach as there are indeed changes in the distribution of education attainment across margins that are not directly affected by the policy reform, and the model does predict that there will be such “inframarginal” responses. In panel (a), our model predicts an increase in the share completing middle school, which is the margin that is directly affected by the reform, of around 5%. Additionally we predict that the share completing a high school degree or more by around 3%. In the validation sample, the share of individuals with a middle school degree increases by 6.5% and the share of individuals with at least a high school degree increases by around 2%. Thus, our model predictions slightly overestimate the “inframarginal” responses, but is capable of reproducing the general patterns found in the validation sample. Considering the different ability groups separately, we show that we are able to predict changes for medium and high ability groups quite well, while our model overestimates the “inframarginal” responses for low ability individuals. Notably, the low ability group also has the fewest individuals who proceed until high school in the observed data, which renders the extrapolation that the reform induces in our baseline model more challenging for this group. In a nutshell, the predictions from our model line up with the observed changes. We return to these findings in Section 4.4, where we provide additional evidence on economic mechanisms from our model that shed more light on the nature of these responses.

4 Empirical Results

We now present our empirical findings based on the estimated model. We first document heterogeneity in ex-ante returns by year of schooling, choices of academic and vocational track, and ability, before we consider the importance of option values, wage risk and sectoral choices,

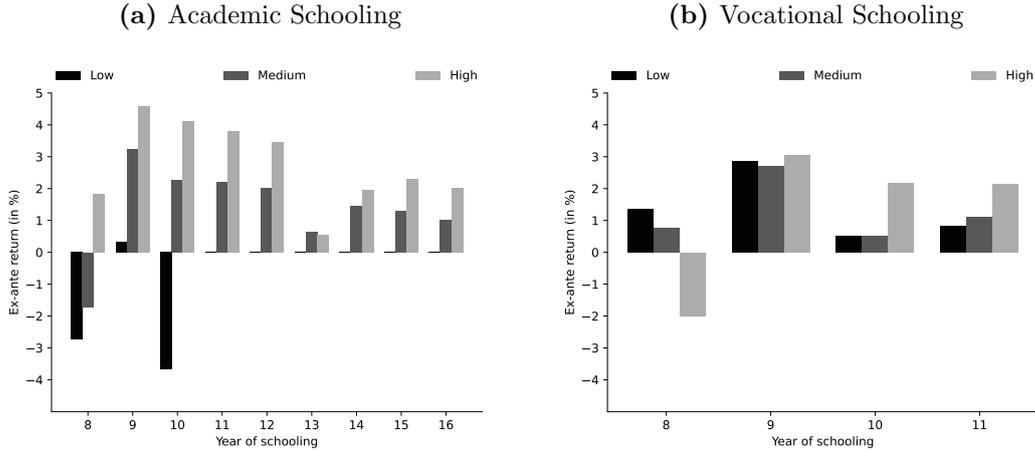
and finally the impacts of alternative education reforms. In these calculations, we simulate life-cycle histories for 5,000 individuals for each ability group using the estimated model.

4.1 Evidence on Ex-ante Returns

To construct measures of ex-ante returns to schooling, we will compare the discounted lifetime value of attending schooling in a particular track in a given period in our model to the corresponding value associated with the best alternative choice. Notably, since our model allows for re-enrollment in a flexible manner, the best alternative choice can include the possibility of attending more education at a later stage in the life-cycle. Individuals can thus reach a given level of final schooling at the end of their life-cycle through many different paths due to the opportunities of track switching and re-enrollment. To ease interpretation and tractability of our findings, we will therefore focus on individuals who in our model have had uninterrupted schooling careers in given track up to the period where we explore the ex-ante returns associated with different schooling choices. To avoid our results to be driven by a small number of individuals, we will drop particular transitions in our expositions for groups who have less than 0.5% chance of reaching such transitions (e.g., low ability individuals attending college).

Ex-ante Returns Figure 6 shows our main evidence on the ex-ante returns to continue schooling in academic or vocational track for an extra year by individual ability type. The returns are shown for each year of schooling along the horizontal axis and are computed for individuals that have reached this particular stage and are faced with the decision to continue schooling, i.e., the bars at 10 years illustrate the average ex-ante returns of attending the 10th grade by track and ability for individuals who have completed 9 years of schooling in this track in an uninterrupted spell. As we move towards right in each panel, the illustrate returns pertain to only individuals who actually reach those stages in our model. At each stage, these returns capture both the immediate rewards and the discounted future rewards. In this sense, the model allows us to estimate the average treatment effect for those facing a particular treatment choice, as compared to their best alternative choice (Heckman et al., 2016; Heckman and Navarro, 2007). We find substantial heterogeneity in the average ex-ante returns by ability, track choice and year of schooling. In the academic track, we estimate negative returns among low-ability individuals of -2.75% and -3.5% , respectively, at the 8th and 10th grades, and positive returns among high-ability individuals at all years. By contrast, we estimate a negative return for high-ability individuals at the 8th grade in vocational track. Underlying this heterogeneity is a strong pattern of sorting into academic and vocational tracks by ability, which in our model is both caused by differences in wage rewards and preference heterogeneity. Indeed, the structure of the 8th grade returns reflects the separation of the ability groups in the different schooling tracks at an early stage of the educational pathways, consistent with the descriptive evidence in Figure 2, panel (e). As the average return to vocational track at the 8th year is negative

Figure 6: Average Ex-ante Returns to Academic and Vocational Schooling.



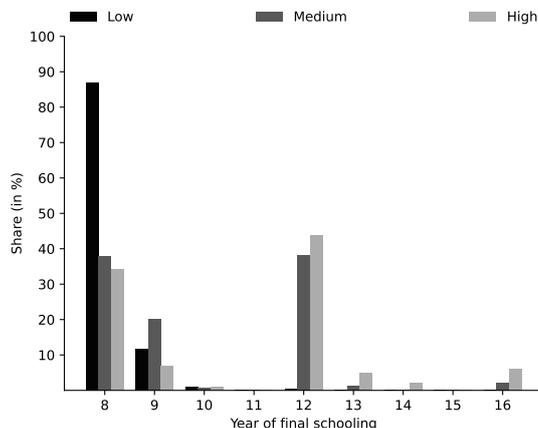
Note: The figure is based on samples of 5,000 simulated schooling careers for each ability group based on the estimated model. This figure contains average ex-ante returns in utility terms for each schooling year-track-ability cell. The left figure shows how ex-ante returns to academic schooling develop over time for each group while the right figure shows the same for the vocational track. Each bar shows the average ex-ante return to a particular year of schooling for the individuals in the respective ability group that reaches the relevant transition in our model. For instance, the bar for the high-ability group in the academic panel in year 11 shows the average ex-ante return of the 11th year overall high-ability individuals that have had an uninterrupted academic schooling career until the 10th year. We compute the ex-ante return as defined in Equation (6). Whenever there are only a few people of a particular ability group that reach a particular transition we do omit this group from the calculation.

for high-ability individuals, the majority of them never enter vocational school. The opposite is true for low-ability individuals, for whom the returns to an academic track are negative initially, pushing them into vocational schooling instead. Indeed, most low-ability individuals never attend an academic school. As we move towards right in panel (a), we notice that the returns to academic schooling remain consistently positive only for high-ability individuals. In each track, we find the highest returns for transitions that entail a middle school degree at the 9th year, with gradually decreasing returns as we progress further.

Earnings Returns The above evidence on ex-ante average *true* returns in terms of *utility* may contrast ex-ante average *lifetime earnings* returns associated with different schooling choices, as discussed in Section 2.2. To illustrate this, we also provide evidence on ex-ante earnings returns in the Appendix Figure A.3 for each schooling year-track-ability cell, using our estimated model. Consistent with the notion of strong ability-related sorting in our data, we again find that the ex-ante earnings returns to academic (vocational) schooling are negative for low-ability (high-ability) individuals. However, we find different patterns across the schooling careers. Unlike the true returns, earnings returns in the academic track peak at the 13th year of schooling, reflecting that once an individual enters college their probability of receiving a college degree increases substantially. Notably, however, earnings returns may not align with individual choices, as opposed to the true (utility) returns, which also capture expected non-pecuniary rewards.

The Role of Re-enrollment A salient aspect of our model—crucial in interpreting the results above—is the individuals’ ability to re-enroll in school after having exited and undergone a non-schooling spell. Indeed, the relatively low average ex-ante returns in Figure 6 can be partly attributed to the fact that many individuals come back to school to pursue more education. In an attempt to illustrate this feature of our model, we show the final schooling level for individuals that initially leave school after the 8th grade in either track by ability in Figure 7. Among low ability individuals, dropping out at that stage determines the final schooling level for about 90%. Only 10% do still acquire their middle school degree at a later stage. For high-ability individuals, however, the large majority do continue their schooling at some point. Roughly 85% do eventually end up with at least a middle school degree, and 30% even continue to obtain at least a high school degree. We need to keep this feature of our model in mind when interpreting average ex-ante returns at each transition that are presented above.

Figure 7: The Final Years of Schooling for those Exiting School at the 8th Grade.



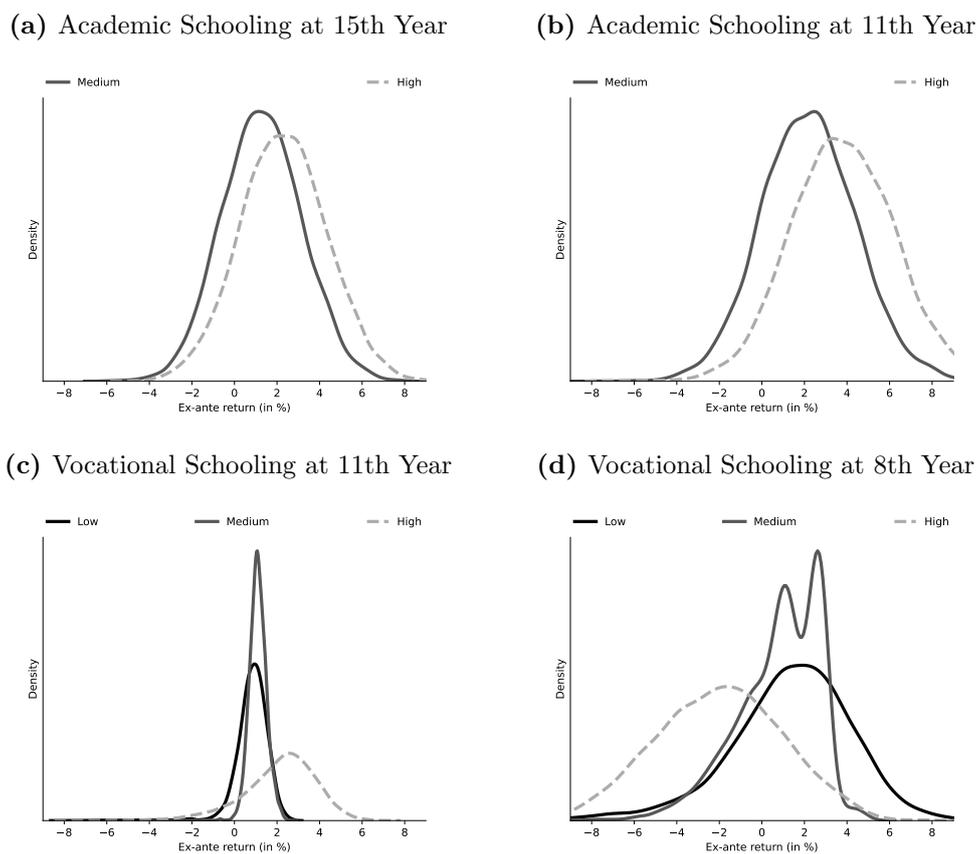
Note: The figure is based on samples of 100,000 simulated schooling careers for each ability group based on the estimated model. We further restrict the sample to individuals who continue their schooling for only one additional year and then initially drop out of school at the 8th year of schooling, i.e., early drop-outs. We determine an individual’s final schooling level as the sum of the years spent in academic and vocational school.

The Distributions of Ex-ante Returns The average returns in Figure 6 can mask considerable heterogeneity in returns by track choice across individuals with the same ability facing identical choices. For instance, Wiswall and Zafar (2015) and Attanasio and Kaufmann (2017) document substantial heterogeneity in ex-ante returns using survey data on subjective expectations. In our model, heterogeneous returns across observationally similar individuals, i.e., those from the same ability group who have reached the same stage of the decision tree after an uninterrupted spell, are represented through the presence of heterogeneous latent types and different realizations of shocks to productivity and tastes associated with different choices.

To illustrate heterogeneity in ex-ante returns, we now focus in Figure 8 on individuals in each

ability group who in our model faced the choices of 11th and 15th year of academic schooling and 8th and 11th year of vocational schooling, respectively. In panel (b), we see that the average return of the 11th year of academic schooling is positive for both high- and medium-ability groups, but also that a considerable share within each group has negative returns. In our model, all individuals with negative returns to academic schooling would decide not to attend academic schooling as this choice is (in expectation) dominated by the best alternative choice. They might, however, pursue academic school later as this does not rule out the possibility of re-enrollment. In panel (b), we also see that the distributions overlap, so there are many individuals of medium-ability for whom returns to the 11th year of academic schooling are higher than for high-ability individuals. In panel (a), we see that the distributions of returns to the 15th year of academic schooling are shifted towards left as compared to panel (b), reflecting that fewer go on to attend the 15th year as compared to the 11th year.

Figure 8: Distributions of Ex-Ante Returns to Academic and Vocational Schooling.



Note: The figure is based on samples of 5,000 simulated schooling careers for each ability group based on the estimated model. Each panel shows the distribution of ex-ante returns to a particular schooling choice by ability who those with an uninterrupted schooling career up to that point who have reached this transition. For instance, in panel (a), we show the distribution of ex-ante returns to the 15th year of academic schooling for all individuals with uninterrupted schooling careers up to that point. We compute the ex-ante return as defined in Equation (6). The heterogeneity in returns follow from the permanent differences across latent types \mathbf{e} and the realizations of transitory shocks $\boldsymbol{\epsilon}_t$. Whenever there are only a few people of a particular ability group that reach a particular transition we do omit this group from the calculation.

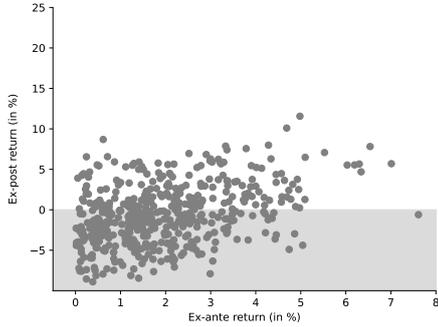
Next, in Figure 8, panels (c)-(d), we consider the distributions of returns to 11th and 8th year of vocational schooling, respectively. In these plots, we include all three ability groups as sufficiently many from each group are present at these transitions in our model. We find interesting ability-related patterns in panel (d), where the majority of low-ability individuals have positive returns to 8th year of vocational schooling, while the majority of high-ability individuals have negative returns. These results reflect the strong patterns of ability-related sorting at the 8th year schooling tracks. By contrast, the distributions of returns are much more similar across low- and medium-ability groups in panel (c), reflecting that conditional on having reached the 10th year of vocational schooling, the transition rates to 11th year of vocational school do not differ substantially across these groups. This finding may reflect that there are latent types in all ability groups with a high propensity to attend vocational schooling, and these types stay in vocational schooling through-out middle and high school, conditional on having enrolled.

Ex-ante and Ex-post Returns The preceding analysis have been focused on ex-ante returns, i.e., the relative rewards that agents in our model base their decisions on. We now contrast our estimates of ex-ante returns to another set of objects which we refer to as ex-post returns. The latter returns are based on our baseline model with taste and productivity shocks turned on, but where we use the actual realizations of shocks rather than the agents' expectations about these to calculate their returns to difference choices. Since our model assumes rational expectations, on average, ex-ante and ex-post returns must agree. However, there does exist a non-degenerate joint distribution of ex-ante and ex-post returns across agents. To construct measures of ex-post returns, we limit attention to individuals in our model who ended up selecting specific schooling choices and retain the realization of shocks they were exposed to as they traversed through their decision trees. By construction, the ex-ante returns for these individuals for the set of schooling choices they ended up making are strictly positive. To provide a comparison of the ex-ante and ex-post returns, we compare both the expected and the realized utility flows to the expected values of the next best alternatives the individuals faced.

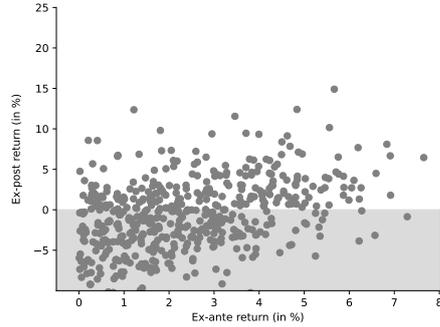
Figure 9 presents the joint distribution of ex-ante and ex-post returns for a random set of 500 individuals from our model at each transition. As expected, the ex-ante return are always positive by construction in each panel. However, the shaded areas indicate that the ex-post returns from pursuing an education were negative for some individuals, i.e., they faced regret due to the actual shock realizations. We also note that the ex-post returns to the 8th year of vocational schooling are relatively dispersed. As this decision is made early in the life-cycle, agents face very different life-time trajectories subject to the future shock realizations and choices. By contrast, the ex-post returns to the 15th year of academic schooling are more compressed. As most individuals would go on to attend the 16th year to attain a college degree

Figure 9: Joint Distributions of Ex-Ante and Ex-post Returns.

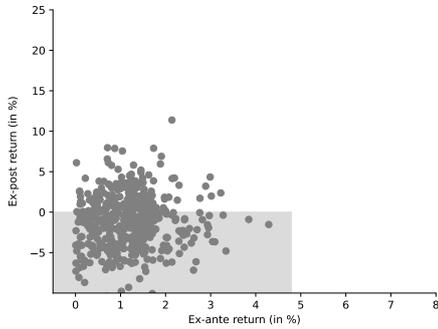
(a) Academic Schooling at 15th Year



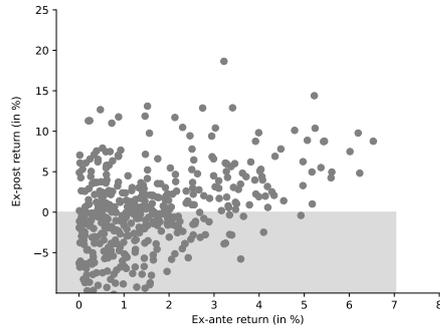
(b) Academic Schooling at 11th Year



(c) Vocational Schooling at 11th Year



(d) Vocational Schooling at 8th Year



Note: The figure is based on samples of 5,000 simulated schooling careers for each ability group based on the estimated model. We then restrict the sample to 500 random individuals with uninterrupted careers in the respective period for a particular track. Each panel shows the joint distributions of ex-ante and ex-post returns to a particular schooling choice in either academic or vocational track. Ex-post returns are the realized total discounted utilities over the remaining decision periods relative to the value function of the best alternative. The gray area shows all points where the realized return is smaller than the expected return from the second best option. Ex-ante return as defined in Equation (6). Whenever there are only a few people of a particular ability group that reaches a particular transition we do omit this group from the calculation.

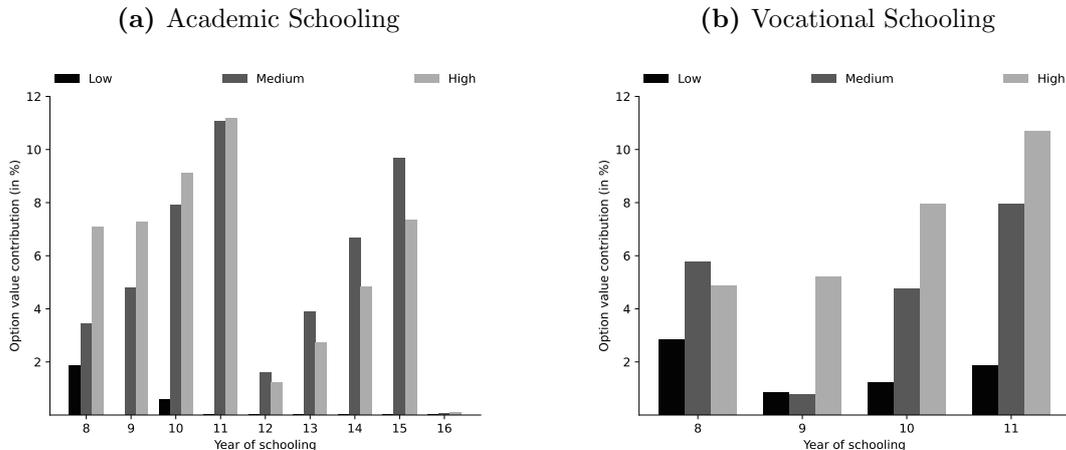
and then enter the labor market, these ex-post returns are associated with more similar lifetime trajectories. These findings also demonstrate how uncertainty is highest at the beginning of the life-cycle and the choices made early on can be more consequential in a dynamic setting.

4.2 Evidence on Option Values

Part of the overall value to a schooling choice is the option to continue schooling further. We now provide quantitative evidence on such option values based on our estimated model. To construct measures of option values to schooling, we will compare the discounted lifetime value of attending schooling in a particular track in a given period in our model to the corresponding value of the same schooling track under a counterfactual policy where the individual is prohibited from making a schooling choice in any future period. As earlier, we will focus on individuals who in our model have had uninterrupted schooling careers in given track up to the period where we explore the option values associated with a schooling track choice.

Option Value Contributions Figure 10 shows the contribution of the option value to the overall value of a schooling track by the year of schooling and ability. The option value contributions in the academic track range from 11% for the 11th year to almost zero beyond the 15th year of schooling. A recurring pattern is that the option value contribution is always the highest for the year of schooling right before the completion of an academic degree that entails considerable “diploma” effects. We also find sizeable heterogeneity by ability level. The option value contributions in the academic track tend to be the highest for medium- and high-ability individuals, as they are also likely to benefit the most from the additional schooling opportunities that open up from taking an extra year of schooling. By contrast, in the vocational track, the option value contributions for low-ability individuals are the highest at the 8th year of schooling and decline in later years. This pattern may reflect that most low-ability individuals attending this track go on to complete the 9th year in vocational school. Among those who progress further in vocational schooling, we again find a strong ability gradient. This likely reflects that among this group, also the medium- and high-ability individuals have the highest gain from attending the 10th and 11th year, and reach a vocational high school degree.

Figure 10: The Option Value Contributions of Academic and Vocational Schooling.

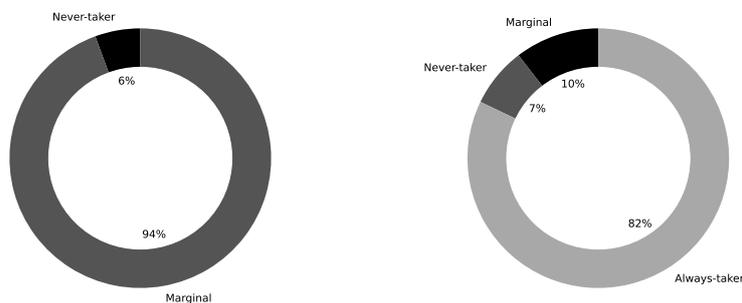


Note: The figure is based on samples of 5,000 simulated schooling careers for each ability group based on the estimated model. We restrict the sample to individuals with uninterrupted schooling careers up the relevant transition using the estimated model. The left figure shows how option value contributions for academic schooling develop over time for each group while the right figure shows the same for the vocational track. The option value contribution is defined in Equation (9). Whenever there are only a few people of a particular ability group that reaches a particular transition we do omit this group from the calculation.

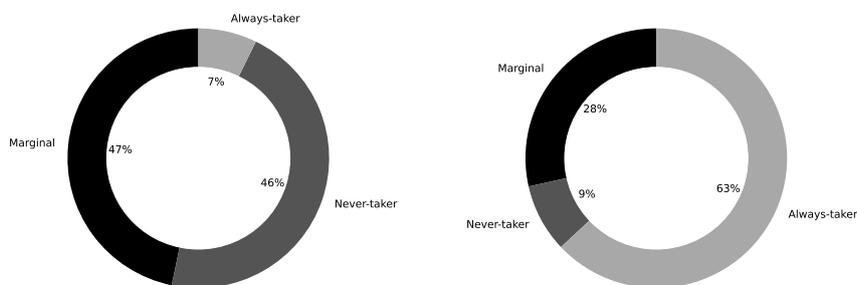
While the previous illustration provides evidence on the option value contributions, measured as a fraction of the overall value of a choice, we now provide more direct evidence on how the option value channel can play a crucial role in shaping schooling careers. To get at this, we perform counterfactual experiments based on our model where we turn on and off the option value of a choice and characterize the schooling decisions made by the agents in our model under each scenario. Based on these comparisons and inspired by the IV/LATE complier char-

Figure 11: ‘Complier’ Characterization – Switching Off the Option Value.

(a) Academic at 11th Year (High) (b) Academic at 12th Year (High)



(c) Vocational at 8th Year (All) (d) Vocational at 9th Year (All)



Note: The figure is based on samples of 5,000 simulated schooling careers for each ability group based on the estimated model. Panels (a)-(b) are restricted to high-ability individuals, while panels (c)-(d) average across all ability groups. Each panel provides a complier characterization based on model simulations where we turn off the option value of a particular schooling choice. This calculation compares the total value of schooling and the value of schooling net of the option value contribution to the next best alternative. The option value contribution is defined in Equation (9). Always-takers (never-takers) always (never) choose to continue with another year of schooling, irrespective of the option value contribution, while marginal individuals take the additional year only because of the option value contribution. Whenever there are only a few people of a particular ability group that reaches a particular transition we do omit this group from the calculation.

acterizations done in the program evaluation literature (e.g., Angrist et al. (1996)), we perform a characterization of agents into three groups; always-takers, never-takers and marginal agents (or compliers). Agents that always (never) decide to take another year of schooling when faced with this choice, irrespective of the option value, are characterized as always-takers (never-takers). While, agents that decide to take another year of schooling when the option value is turned on but not when it is off are characterized as marginal (i.e., compliers). By construction, monotonicity is satisfied, as for given shock realizations, agents in our model are never more likely to take more schooling when the option value is turned off compared to when it is on.

Figure 11 illustrates our evidence based on the complier characterizations described above. In panels (a)-(b), we focus on high-ability individuals who faced the decision to continue their

schooling for the 11th and 12th year in the academic track. These figures provide interesting illustrations of how important option values can be close to the degree-rewarding schooling choices. Among those facing the decision to continue their academic schooling in the 11th year, we find that a large majority at 94% among the high-ability individuals consists of marginal compliers, i.e., individuals who continue for another year of schooling only because of the option value stemming from being able to complete a high school degree right afterwards. By contrast, there are hardly any always-takers, who move ahead with their schooling even when no future schooling opportunities are available, and about 6% are never-takers, who drop out regardless. The picture is somewhat different at the 12th year of academic schooling. The fraction of always-takers rises drastically to 82% as completing the high school degree provides immediate considerable wage rewards. For 10% the option value of the high school diploma is crucial to complete the 12th year, as receiving this diploma opens up the possibility of attending college. Still, 7% drop-out and do not complete high school regardless of the option value.

Next, in panels (c)-(d) of Figure 11, we consider all individuals (irrespective of ability) who faced the decision to continue their schooling for the 8th and 9th year in the vocational track. Option values in the vocational track are highest at the 8th year of schooling as this gives the option to continue later on with the 9th year of vocational schooling, i.e., receive a two-year vocational diploma. At the 9th year of vocational schooling, a majority at 63% of individuals facing this choice are characterized as always-takers, who attend this year due to the immediate wage gains associated with this choice, while the option value associated with the possibility to continue with a vocational high school degree drive this choice for 28% of individuals.

Uncertainty and Option Values In our model, transitory shocks to productivity (i.e., wage risk) and tastes give rise to uncertainty in agents' decision-making. To shed light on these aspects through the lens of our model, we perform a series of comparative statics, where we recompute option values contributions by ability and track shutting off various sources of shocks and compare the findings to our baseline model.¹⁷ We now consider how these sources of uncertainty contribute to the option values associated with different educational choices.¹⁸

¹⁷We summarize our findings on ex-ante returns from these exercises in Appendix Figure A.4. To facilitate comparison, panel A reports estimates of average ex-ante returns by ability and track from our baseline model (as in Figure 6), while panel B shows the corresponding estimates from a version of the model where we remove shocks to productivity (i.e., no wage risk), while in panel C we further also remove unobserved transitory shocks related to tastes for schooling, working or staying at home. While the ex-ante returns are broadly similar across panels A-B, we find substantially larger returns for low-ability (high-ability) individuals in vocational (academic) tracks in panel C. Thus, removing taste shocks would further strengthen the patterns of ability-related sorting into tracks in the early stage of educational careers in our model.

¹⁸The existing literature on learning and educational choices in dynamic settings actually emphasizes uncertainty as the primary source of option values, i.e., individual learn about their own ability and preferences as they progress in their schooling career (Arcidiacono et al., 2016; Trachter, 2015; Stinebrickner and Stinebrickner, 2014; Stange, 2012). As the shocks in our model are distributed independently over time, individuals do

We present in Appendix Figure A.9 two different scenarios where the presence of transitory shocks has opposite signed effects on the option value contribution in our model. In panel (a), we consider the option value of the 11th year in the academic track for individuals of medium ability under different sources of uncertainty. We first turn off the productivity shocks and then also turn off the taste shocks. In the baseline model, the option value contribution amounts to about 11%. When we turn off the productivity shocks alone, the option value contribution increases to 12%, and decreases to about 10% in a scenario without any uncertainty. The latter pattern reflects that the continuation of schooling becomes less likely when we reduce the extent of uncertainty for some individuals. Specifically, some medium-ability individuals pursue the 11th year of academic schooling solely due to the realization of taste shocks. Once we remove these realizations, the option value of the 11th academic year declines for these individuals.

In panel (b) of Appendix Figure A.9, we consider another scenario showing how the presence of transitory shocks affects the option value contribution for the 8th year in the vocational track for low-ability individuals. In the baseline model, the option value contribution is at 3%. When we turn off the productivity shocks alone, the option value remains almost unchanged, but when we remove taste shocks this value increase to almost 6%. This pattern reflects that some among the low-ability individuals in our baseline model who drop-out at the 8th year of vocational schooling do so due to the realization of taste shocks. Once we remove these shocks from our model, their likelihood of continuing beyond the 8th year increases even further.

4.3 Evidence on Sectoral Choices

An important feature of our model setup is that we allow individuals to work in one of three sectors—private, public and self-employment—and further allow sector-specific returns to schooling in academic and vocational tracks and to work experience. In the following, we use the estimated model to perform a series of comparative statics where we alter sectoral job opportunities, with the goal of informing how these options influence educational choices.

Sectoral Opportunities As noted above, we allow the returns to schooling tracks and work experience to differ across sectors. Consequentially, the ex-ante returns depend on the availability of job opportunities in different sectors. In Figure 12, we compares ex-ante returns to each year of academic and vocational schooling by ability under different assumptions on sectoral

not update their prior beliefs about their productivity or alternative-specific tastes, i.e., our model does not feature learning over time. We rather focus on the overall value attached to a schooling choice stemming from the possibility of pursuing further education, and not only the value associated with resolution of uncertainty and learning. The presence of transitory shocks may nonetheless affect individuals' schooling choices and alter the likelihood of attending further schooling, i.e., option values can depend on the presence of shocks.

job opportunities. For comparison, panel A shows the returns estimated in our baseline model. In panel B, we remove the option of public sector choice, and recompute the ex-ante returns associated with each educational choice. Notably, removing the public sector job opportunities decreases ex-ante returns to academic schooling considerably for high- and medium-ability individuals, who are most likely to benefit from these opportunities.

The most noticeable decline in the ex-ante return is observed for the 9th year of academic schooling. This change implies that removing the public sector opportunities alters the schooling track choice of some individuals. Indeed, in Appendix Figure A.5 (A.6), we show that removing public sector opportunities induces some (medium-ability) individuals towards a vocational high school degree instead. Further, in Figure 12, panel C, we further remove the option to work as self-employment. The latter change, however, does not substantially alter returns for low- and medium-ability individuals, as these groups alternatively move into private sector jobs and retain similar premiums. More surprisingly, the return to academic schooling decreases for high-ability individuals once we remove self-employment choice. This phenomenon may reflect the presence of sizeable returns to academic college education also in the self-employment sector (e.g., high-paying consultancy work).

Sectoral Opportunities and the Option Value Contributions Sectoral opportunities shapes educational choices and thus also the size of option value contributions. In Figure A.10, we compare option values across model specifications with different sectoral opportunities for medium- and low-ability individuals. Panel (a) shows option value contributions for medium-ability individuals who consider enrolling in the 8th year of academic schooling. Removing the option to enter the public sector substantially reduces the option value of academic schooling. The return to academic degrees is relatively large in the public sector which leads to a reduction in option values if this job opportunity is removed. However, removing the option to work as self-employed does not change option values to academic schooling, as degree effects for self-employed are substantially smaller. Next, panel (b) shows how the option value of the 8th year of vocational schooling changes for low-ability individuals when sectoral opportunities are constrained. Notably, the option values for vocational schooling do not react to removing sectoral opportunities. Low-ability individuals who choose vocational schooling mostly work in the private sector, so the option values remain constant once the options to work in the public sector and/or be self-employed are removed. Thus, sectoral opportunities influence educational choices, partly by altering the composition of returns to different educational options.

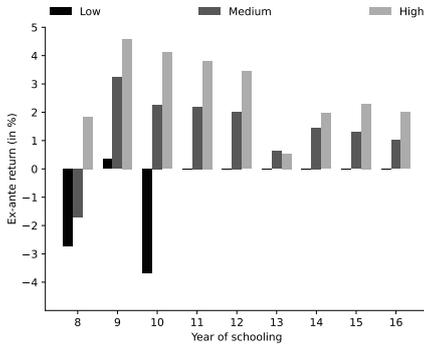
4.4 Educational Policy Evaluation

We now use our model to analyze the impacts of compulsory schooling reforms. First, we provide further evidence on compliance to the Norwegian compulsory schooling reform that we

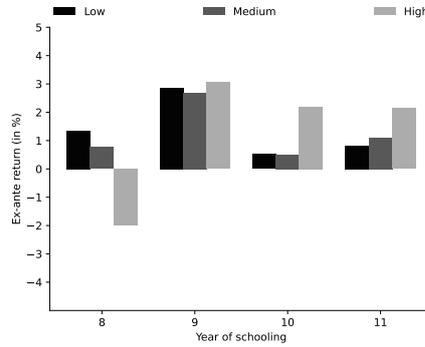
Figure 12: Average Ex-ante Returns – The Role of Sectoral Job Opportunities.

Panel A. Baseline Model

(A-1) Academic Schooling

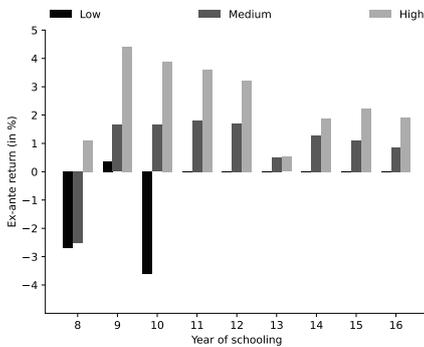


(A-2) Vocational Schooling

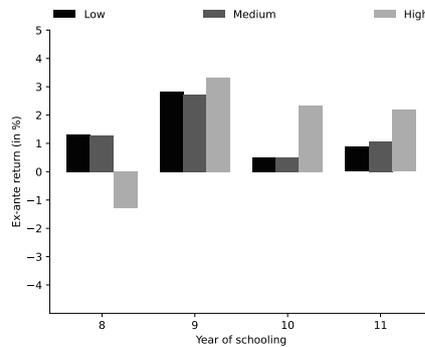


Panel B. No Public Sector Choice

(B-1) Academic Schooling

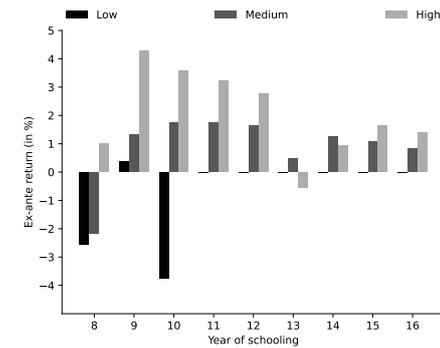


(B-2) Vocational Schooling

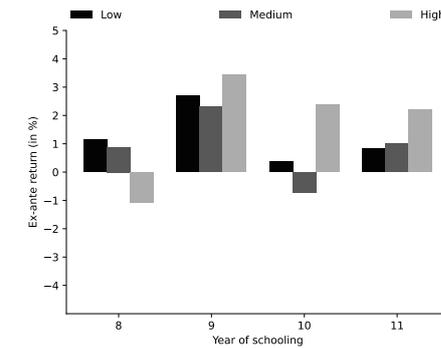


Panel C. Only Private Sector Choice

(C-1) Academic Schooling



(C-2) Vocational Schooling

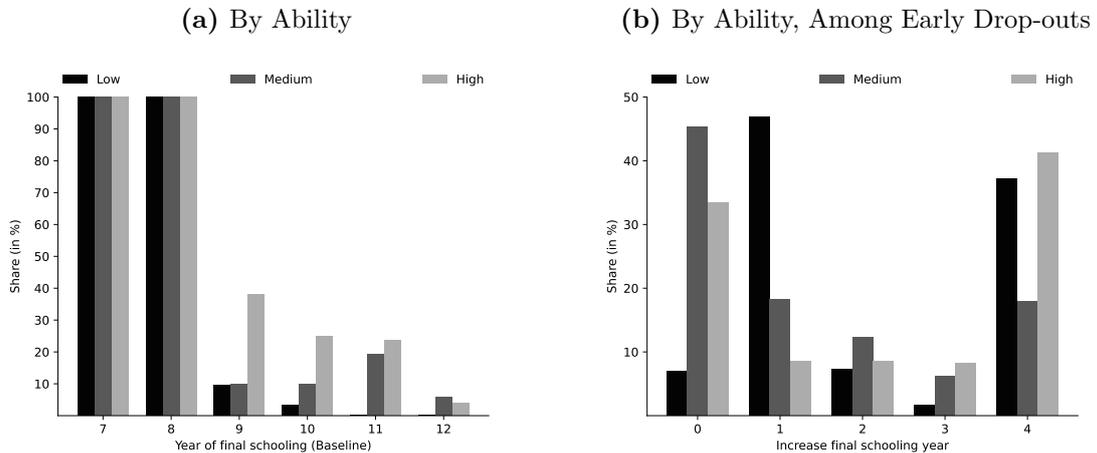


Note: The figure is based on samples of 5,000 simulated schooling careers for each ability group based on the estimated model under different specifications of transitory shocks. Each panel contains average ex-ante returns for each year-track-ability cell for alternative model specifications. Panel A shows results based on the baseline model, while panel B removes the public sector and panel C further also removes the option to work as self employed. In each panel, the left figure shows how ex-ante returns to academic schooling develop over time for each group while the right figure shows the same for the vocational track. Each bar shows the average ex-ante return to a particular year of schooling for the subset of the respective ability group that has reached the relevant transition in our model. For instance, the bar for the high-ability group in the academic panel in year 11 shows the average ex-ante return of the 11th year for those that have had an uninterrupted academic schooling career until the 10th year. We compute the ex-ante return as defined in Equation (6). Whenever there are only a few people of a particular ability group that reaches a particular transition we do omit this group from the calculation.

earlier used to validate our model. We show who is affected by the policy along the distribution of schooling by ability and by early drop-out status. Second, we investigate the impacts of a high school enrollment policy, which requires everyone to attend ten years of schooling. Finally, we consider the impacts of introducing a tuition fees policy.

The Norwegian Compulsory Schooling Reform As described in Section 3.4, the Norwegian compulsory schooling reform increased the minimum schooling requirement from seven to nine years, and was gradually introduced in different municipalities in different years. In our analysis thus far we used individuals born 1955-1960 who were not exposed to the reform and relied on the reform variation in an out-of-sample validation of our model. We now use our estimated model to shed light on the compliance to this reform by ability and early drop-out status. In panel (a) of Figure 13, we show the fractions of individuals by their final year of schooling in the baseline scenario (i.e., pre-reform) along the horizontal axis that change their schooling choices due to the reform. By construction, since the post-reform compulsory schooling is nine years, all individuals that previously decided to stop after seven or eight years are affected. Notably, as discussed in Section 3.4, some of these individuals even increase their schooling beyond the new minimum requirement. Such “inframarginal” responses in our model can be explained by the presence of option values; by forcing individuals to attend nine years of schooling, we also bring them closer to transitions that make a high school diploma within reach.¹⁹

Figure 13: Compliance to the Norwegian Compulsory Schooling Reform.



Note: The figure is based on two samples of 100,000 simulated schooling careers for each ability group under alternative scenarios. Using the point estimates, we first simulate the model with the original seven years of compulsory schooling. Next, we rerun the simulation but impose nine years of compulsory schooling. Throughout, we keep the random realizations of the productivity and taste shocks ϵ_t fixed, and we are thus able to compare the schooling decisions of the same individual under the two different regimes. In panel (a), we plot the fractions of individuals who change their schooling decisions for varying levels of final schooling in the baseline scenario along the horizontal axis. In panel (b), we restrict our sample to individuals who initially dropped out after the 8th year of uninterrupted schooling in the baseline scenario and then illustrate the distribution of observed increases in their final schooling due to the policy reform.

¹⁹Indeed, we relied on the extent of such “inframarginal” responses for the model validation in Section 3.4.

More interestingly, our model also predicts alterations in the educational trajectories among those who in the baseline scenario actually had attended nine or more years of schooling. While the presence of option values can trigger the “inframarginal” responses discussed above among those with less than nine years of schooling, an additional mechanism of re-enrollment possibilities is at play when we consider those having nine or more years of schooling pre-reform. Prior of the reform, 10% of individuals had dropped out either after the 7th or 8th grade but then re-enrolled at a later time. Since the reform rules out any interruptions between 7th and 9th year of schooling, the educational trajectories individuals who earlier dropped out after the 7th or 8th grade and re-enrolled are also affected. Indeed, about 40% of high-ability individuals who end up with only nine years of schooling in the baseline scenario increase their final schooling level after the reform. They do so because they no longer face the considerable re-enrollment costs they had to incur in the baseline scenario where they dropped out after the 7th or 8th year.

In panel (b) of Figure 13, we show that the compulsory schooling reform affects individuals who initially dropped out after the 8th year of schooling in our model. Around 40% of the early drop-outs with high- and medium-ability do not increase their final years of schooling post reform; these individuals all re-enrolled even in the baseline scenario and attained at least nine years of schooling in the end. However, among low-ability individuals, around 45% increase their schooling by one year and thus only meet the new requirement, while around 35% increase their schooling level by four years and thus attain a high school degree after the reform. Thus, option values and re-enrollment are important channels that explain these compliance patterns.

Compulsory High School Enrollment Policy Another policy we consider next is the introduction of compulsory high school enrollment, which requires all individuals to attend ten years of schooling, i.e., one more year than the Norwegian compulsory schooling reform. Appendix Figure A.7 compares the distributions of final years of schooling in our model between the simulated reform with the nine year of compulsory schooling (‘Reform 9’) and the compulsory high school enrollment policy (‘Reform 10’). Interestingly, we again find evidence of strong “inframarginal” responses; most individuals that are induced to change their schooling level by the compulsory high school enrollment policy indeed go on to complete a high school degree. Overall, the fraction of individuals who at least hold a high school degree increases from 82% to more than 95%. Most of these increases are stem from low- and medium-ability individuals, who increase their high school shares from about 78% to 97% and 52% to 82%, respectively.

College Tuition Fees Policy The final policy we consider next is the introduction of a tuition fees corresponding to 50,000 Norwegian Kroner (2018) per year of college attendance, corresponding to around USD 6,000. As illustrated in Appendix Figure A.8, our model predicts that the introduction of this tuition fees would drive out virtually all medium-ability individuals

from college, while this would apply to two-thirds of high-ability individuals who would attend college in the baseline. Overall college graduation share declines from 17.5 to 3.5 percent, with a corresponding increase in high school attendance. In particular, medium-ability individuals drawn away from college by this policy would instead pursue a vocational high school degree.

5 Conclusion

This paper has attempted to provide evidence on the ex-ante returns and option values to educational choices. To achieve this, we devised a dynamic model of education and work decisions in a life-cycle context that acknowledges uncertainty and sequential nature of schooling decisions. We estimated this model using Norwegian population panel data with long earnings histories, and validated this against variation in schooling choices induced by a compulsory schooling reform. Finally, we used the structure of our model to learn about the rich patterns of compliance observed in our data and the potential economic mechanisms driving these.

Our analysis gave several interesting insights. The ex-ante returns to schooling vary across the different stages of educational careers, depend on the choice of track and the ability of individuals. Underlying these heterogeneities is a strong pattern of ability-related sorting into different educational tracks. We find that option values play a dominant role in shaping schooling decisions at several points in the educational career. We also document how the presence of option values and re-enrollment opportunities could explain the “inframarginal” impacts of compulsory schooling reforms across the distribution of schooling attainment. Further, we provide insights on how sectoral opportunities influence individuals’ educational choices, reflecting differences in sector-specific returns to alternative educational tracks.

While our paper provides several insights, one shortcoming is worth mentioning. Recent studies emphasize the role of “experimentation” in educational decisions, where individuals make such decisions in view of the returns generated through the subsequent resolution of uncertainty that they are initially faced with (see, e.g., Arcidiacono et al. (2016) and references therein). We regard this as an important stream of research, which highlights another channel for why option values can matter in educational decisions. Our model however does not feature learning and updating of individuals’ prior beliefs, but instead has focused on the analyzing educational decisions in a life-cycle context. We leave it for future work to develop a modelling framework for educational choices with learning in a life-cycle context with many periods.

References

- Aakvik, A., Salvanes, K. G., and Vaage, K. (2010). Measuring heterogeneity in the returns to education in Norway using educational reforms. *European Economic Review*, 54(4):483–500.
- Acemoglu, D. and Angrist, J. (2000). How large are human-capital externalities? Evidence from compulsory schooling laws. *NBER Macroeconomics Annual*, 15:9–59.
- Aguirregabiria, V. and Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics*, 156(1):38–67.
- Altonji, J. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics*, 11(1):48–83.
- Altonji, J. G. and Segal, L. M. (1996). Small-sample bias in GMM estimation of covariance structures. *Journal of Business & Economic Statistics*, 14(3):353–366.
- Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455.
- Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1):343–375.
- Arcidiacono, P., Aucejo, E., Maurel, A., and Ransom, T. (2016). College attrition and the dynamics of information revelation. Technical report, National Bureau of Economic Research.
- Aryal, G., Bhuller, M., and Lange, F. (2022). Signaling and employer learning with instruments. *American Economic Review*, 112(5):1669–1702.
- Attanasio, O. P. and Kaufmann, K. M. (2017). Education choices and returns on the labor and marriage markets: Evidence from data on subjective expectations. *Journal of Economic Behavior & Organization*, 140:35–55.
- Becker, G. S. (1964). *Human capital*. Columbia University Press, New York City, NY.
- Bedard, K. (2001). Human capital versus signaling models: University access and high school dropouts. *Journal of Political Economy*, 109(4):749–775.
- Bertrand, M., Mogstad, M., and Mountjoy, J. (2021). Improving educational pathways to social mobility: Evidence from Norway’s “Reform 94”. *Journal of Labor Economics*, 39(4):965–1010.
- Bhuller, M., Mogstad, M., and Salvanes, K. G. (2017). Life cycle earnings, education premiums and internal rates of return. *Journal of Labor Economics*, 35(4):993–1030.

- Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American Economic Review*, 95(1):437–449.
- Blundell, R. (2017). What have we learned from structural models? *American Economic Review*, 107(5):287–92.
- Brunello, G., Fort, M., and Weber, G. (2009). Changes in compulsory schooling, education and the distribution of wages in europe. *The Economic Journal*, 119(536):516–539.
- Card, D. (1999). The causal effect of education on earnings. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 5, pages 1081–1863. North-Holland Publishing Company, New York City, NY.
- Comay, Y., Melnik, A., and Pollatschek, M. A. (1973). The option value of education and the optimal path for investment in human capital. *International Economic Review*, 14(2):421–435.
- Duffie, D. and Singleton, K. J. (1993). Simulated moments estimation of Markov models of asset prices. *Econometrica*, 61(4):929–952.
- Eisenhauer, P., Heckman, J. J., and Mosso, S. (2015). Estimation of dynamic discrete choice models by maximum likelihood and the simulated method of moments. *International Economic Review*, 56(2):331–357.
- Gabler, J. (2022). *A python tool for the estimation of large scale scientific models*.
- Gabler, J. and Raabe, T. (2020). respy - A framework for the simulation and estimation of Eckstein-Keane-Wolpin models.
- Galiani, S. and Pantano, J. (2021). Structural models: Inception and frontier. Technical report, National Bureau of Economic Research.
- Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica: Journal of the Econometric Society*, pages 271–320.
- Heckman, J. J., Humphries, J. E., and Veramendi, G. (2016). Dynamic treatment effects. *Journal of Econometrics*, 191(2):276–292.
- Heckman, J. J., Humphries, J. E., and Veramendi, G. (2018). Returns to education: The causal effects of education on earnings, health, and smoking. *Journal of Political Economy*, 126(1):197–246.

- Heckman, J. J., Lochner, L. J., and Todd, P. E. (2006). Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. In Hanushek, E. A. and Welch, F., editors, *Handbook of the Economics of Education*, volume 1, pages 307–458. North-Holland Publishing Company, Amsterdam, Netherlands.
- Heckman, J. J. and Navarro, S. (2007). Dynamic discrete choice and dynamic treatment effects. *Journal of Econometrics*, 136(2):341–396.
- Hungerford, T. and Solon, G. (1987). Sheepskin effects in the returns to education. *Review of Economics and Statistics*, 69(1):175–177.
- Jaeger, D. A. and Page, M. E. (1996). Degrees matter: New evidence on sheepskin effects in the returns to education. *Review of Economics and Statistics*, 78(4):733–740.
- Kahneman, D., Wakker, P. P., and Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics*, 112(2):375–406.
- Keane, M. P., Todd, P. E., and Wolpin, K. I. (2011). The structural estimation of behavioral models: Discrete choice dynamic programming methods and applications. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, pages 331–461. Elsevier Science, Amsterdam, Netherlands.
- Keane, M. P. and Wolpin, K. I. (1997). The career decisions of young men. *Journal of Political Economy*, 105(3):473–522.
- Lang, K. and Kropp, D. (1986). Human capital versus sorting: the effects of compulsory attendance laws. *The Quarterly Journal of Economics*, 101(3):609–624.
- Lee, S. Y. T., Shin, Y., and Lee, D. (2017). The option value of human capital: Higher education and wage inequality. *NBER Working Paper*.
- Low, H. and Meghir, C. (2017). The use of structural models in econometrics. *Journal of Economic Perspectives*, 31(2):33–58.
- Machin, S., Salvanes, K. G., and Pelkonen, P. (2012). Education and mobility. *Journal of the European Economic Association*, 10(2):417–450.
- Meghir, C. and Palme, M. (2005). Educational reform, ability, and family background. *American Economic Review*, 95(1):414–424.
- Meghir, C. and Pistaferri, L. (2011). Earnings, consumption and life cycle choices. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4B, pages 773–854. Elsevier Science, Amsterdam, Netherlands.

- Miller, R. A. (1984). Job matching and occupational choice. *The Journal of Political Economy*, 92(6):1086–1120.
- Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, 66(4):281–302.
- Mincer, J. (1974). *Schooling, experience and earnings*. National Bureau of Economic Research, New York City, NY.
- Monstad, K., Propper, C., and Salvanes, K. G. (2008). Education and fertility: Evidence from a natural experiment. *Scandinavian Journal of Economics*, 110(4):827–852.
- Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, 29(3):315–335.
- Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review*, 96(1):152–175.
- Pakes, A. and Pollard, D. (1989). Simulation and the asymptotics of optimization estimators. *Econometrica*, 57(5):1027–1057.
- Powell, M. J. D. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. Technical report, University of Cambridge, Cambridge, United Kingdom.
- Puterman, M. L. (1994). *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons, New York City, NY.
- Rust, J. (1994). Structural estimation of Markov decision processes. In Engle, R. and McFadden, D., editors, *Handbook of Econometrics*, pages 3081–3143. North-Holland Publishing Company, Amsterdam, Netherlands.
- Samuelson, P. A. (1937). A note on measurement of utility. *Review of Economic Studies*, 4(2):155–161.
- Shapiro, D., Ryu, M., Huie, F., and Liu, Q. (2019). *Some college, no degree*. National Student Clearinghouse Research Center, Herndon, VA. A 2019 Snapshot for the Nation and 50 States. Signature Report No. 17.
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374.
- Stange, K. M. (2012). An empirical investigation of the option value of college enrollment. *American Economic Journal: Applied Economics*, 4(1):49–84.
- Stinebrickner, R. and Stinebrickner, T. (2014). Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics*, 32(3):601–644.

- Sundet, J. M., Barlaug, D. G., and Torjussen, T. M. (2004). The end of the Flynn effect? A study of secular trends in mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence*, 32:349–362.
- Todd, P. E. and Wolpin, K. I. (2021). The best of both worlds: Combining RCTs with structural modeling. *Journal of Economic Literature*.
- Trachter, N. (2015). Stepping stone and option value in a model of postsecondary education. *Quantitative Economics*, 6(1):223–256.
- Weisbrod, B. A. (1962). Education and investment in human capital. *Journal of Political Economy*, 70(5, Part 2):106–123.
- White, D. J. (1993). *Markov decision processes*. John Wiley & Sons, New York City, NY.
- Wiswall, M. and Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2):791–824.
- Wiswall, M. and Zafar, B. (2021). Human capital investments and expectations about career and family. *Journal of Political Economy*, 129(5):1361–1424.

Appendix

A.1 Specifications of Immediate Reward Functions

We present here the parametrizations of immediate reward functions in our model. In the estimation, all parameters are allowed to vary freely across three observed ability types. The specification of non-pecuniary utilities varies slightly across ability groups. Given the observed patterns in data, some parameters are well identified for one group but not for the other groups.

Work Choice: Wages and Nonpecuniary Utility

The immediate reward of choosing the work alternative, i.e., $a \in \{PR, PU, S\}$, consists a wage and a non-pecuniary component, where all parameters are allowed to vary flexibly across the private (PR), public (PU) and self-employment (S) sectors, as follows:

$$\underbrace{\zeta_a(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}, e_{j,W})}_{\text{Non-Pecuniary Component}} + \underbrace{w_a(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}, e_{j,W}, \epsilon_{W,t})}_{\text{Wage Component}}$$

The wage reward is defined as follows:

$$x_a(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}, e_{j,a}, \epsilon_{a,t}) = \exp(\Gamma_a(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}, e_{j,a}) \cdot \epsilon_{a,t})$$

Wage Component

$$\begin{aligned} \Gamma_a(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}, e_{j,a}) &= e_{j,a} \\ &+ \delta_{1,a} \cdot h_t^A + \delta_{2,a} \cdot h_t^V + \delta_{3,a} \cdot k_t^W + \delta_{4,a} \cdot (k_t^W)^2 + \delta_{5,a} \cdot k_t^a + \delta_{6,a} \cdot (k_t^a)^2 \\ &+ \sum_{d \in \{9,12,16\}} \gamma_{d,a}^A \cdot \mathbf{I}[h_t^A \geq d] + \sum_{d \in \{9,12\}} \gamma_{d,a}^V \cdot \mathbf{I}[h_t^V \geq d] \\ &+ \eta_{1,a} \cdot \mathbf{I}[a_{t-1} = W] + \nu_{1,a} \cdot (t - 15) + \nu_{2,a} \cdot \mathbf{I}[t < 17] \end{aligned}$$

Nonpecuniary Component

$\beta_{9,a}$ is only included in the reward equations for self-employment and the public sector for low-ability individuals. $\beta_{6,a}$ is included in the self-employment and public sector rewards for medium- and high-ability individuals. The choice is motivated by the fact that low-ability individuals tend to have high incentives to not work at all and the term counteracts this. $\beta_{8,a}$, $\beta_{2,a}$ account for the importance of general work experience. Since low-ability individuals work primarily in the private sector, these terms are not included for this group.

$$\begin{aligned}
\zeta_a(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}) &= e_{j,a} + \beta_{1,a} \cdot \mathbf{I}[t < 17] + \beta_{2,a} \cdot k_t^W + \beta_{3,a} \cdot k_t^a + \beta_{4,a} \cdot h_t^A + \beta_{5,a} \cdot h_t^V \\
&+ \sum_{d \in \{9,12,16\}} \vartheta_{d,a}^A \cdot \mathbf{I}[h_t^A \geq d] + \sum_{d \in \{9,12\}} \vartheta_{d,a}^V \cdot \mathbf{I}[h_t^V \geq d] \\
&+ \beta_{6,a} \cdot \mathbf{I}[a_{t-1} = PR] + \beta_{7,a} \cdot \mathbf{I}[k_t^a > 0] + \beta_{8,a} \cdot \mathbf{I}[k_t^W > 0] + \beta_{9,a} \cdot \mathbf{I}[a_{t-1} = a]
\end{aligned}$$

Academic Choice: Nonpecuniary Utility

$$\begin{aligned}
\zeta_A(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}) &= e_{j,A} + \beta_{1,A} \cdot (t - 15) + \beta_{4,A} \cdot \mathbf{I}[h_t^V > 0] \\
&+ \beta_{4,A} \cdot \mathbf{I}[a_{t-1} = A] + \beta_{5,A} \cdot \mathbf{I}[a_{t-1} \in \{PR, PU, S, H\}] + \beta_{6,A} \cdot \mathbf{I}[a_{t-1} = A] \cdot \mathbf{I}[h_t^A \geq 12] \\
&+ \sum_{d \in \{9,12,16\}} \vartheta_{d,A}^A \cdot \mathbf{I}[h_t^A \geq d] + \vartheta_{9,A}^A \cdot \mathbf{I}[h_t^V \geq 9] + \epsilon_{t,A}
\end{aligned}$$

The coefficient $\beta_{4,A}$ is only included for high IQ individuals. The coefficient is included to avoid excessive dropping out and re-enrolling. The coefficient $\vartheta_{9,A}^A$ is mainly relevant in settings where individuals dropout early and potentially re-enroll later. Since this does not happen frequently for high IQ individuals the coefficient is only included for medium and low IQ individuals.

Vocational Choice: Nonpecuniary Utility

$$\begin{aligned}
\zeta_V(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}) &= e_{j,a} + \beta_{1,a} \cdot (t - 15) + \beta_{2,a} \cdot \mathbf{I}[t < 17] + \beta_{3,a} \cdot \mathbf{I}[h_t^A > 0] \\
&+ \beta_{4,a} \cdot \mathbf{I}[a_{t-1} = a] + \beta_{5,a} \cdot \mathbf{I}[a_{t-1} \in \{PR, PU, S, H\}] \\
&+ \vartheta_{9,a}^A \cdot \mathbf{I}[h_t^V \geq 9] + \vartheta_{9,a}^A \cdot \mathbf{I}[h_t^A \geq 9] + \epsilon_{t,V}
\end{aligned}$$

The coefficient $\beta_{5,V}$ is only included for high IQ individuals. The coefficient is included to avoid excessive dropping out and re-enrolling. The coefficient $\vartheta_{9,V}^A$ is mainly relevant in settings where individuals dropout early and potentially re-enroll later. Since this does not happen frequently for high IQ individuals the coefficient is only included for medium and low IQ individuals.

Home Choice: Nonpecuniary Utility

$$\zeta_H(\mathbf{k}_t, \mathbf{h}_t, t, a_{t-1}) = e_{j,H} + \beta_{1,H} \cdot \mathbf{I}[t < 17] + \sum_{d \in \{12,16\}} \vartheta_{d,H}^A \cdot \mathbf{I}[h_t^A \geq d] + \vartheta_{12,H}^V \cdot \mathbf{I}[h_t^V \geq 12] + \epsilon_{t,H}$$

A.2 Estimation Results

This section presents the estimated parameters and their corresponding standard errors. Whenever a parameter has no standard error, it means that it is not well-identified or irrelevant and that we have taken this parameter out of the corresponding specification. See Section A.1 for a more detailed explanation. In some cases, a particular unobserved latent type is not well identified for some choice options since the corresponding type is far from ever choosing this option. Another special case relates to the penalty for minors working, which are set so that people below 17 do not work as self-employed or in the public sector.

Choice Alternative: Private Sector

Table A.1 presents the point estimates and the standard errors for the parameters in the wage component, while Table A.2 presents the point estimates and the standard errors for the parameters in the specification of non-pecuniary component.

Table A.1: Choice Alternative: Private Sector – Wage Component.

		Low IQ	Medium IQ	High IQ
Constant	$e_{1,PR}$	12.1 (0.00469)	11.5 (0.01962)	11.5 (0.02496)
Type 1	$e_{2,PR}$	0.17628 (0.02692)	0.07283 (0.02309)	0.25511 (0.05611)
Type 2	$e_{3,PR}$	-0.05007 (0.02414)	0.50311 (0.01856)	0.37545 (0.02000)
Type 3	$e_{4,PR}$	0.09268 (0.04047)	- (-)	- (-)
Reward Years of Academic Schooling	$\delta_{1,PR}$	0.10764 (0.00136)	0.14816 (0.00078)	0.17273 (0.00042)
Reward Years of Vocational Schooling	$\delta_{2,PR}$	0.16961 (0.00109)	0.12013 (0.00250)	0.15994 (0.00032)
Reward Experience General Work	$\delta_{3,PR}$	- (-)	0.10525 (0.00030)	0.11110 (0.00301)
Reward Experience General Work Square	$\delta_{4,PR}$	- (-)	-0.10907 (0.01408)	-0.09934 (0.00247)
Reward Experience Work Private	$\delta_{5,PR}$	0.10496 (0.00145)	0.08750 (0.00029)	0.06362 (0.00336)
Reward Experience Work Private Square	$\delta_{6,PR}$	-0.05487 (0.00309)	-0.04357 (0.01399)	-0.03549 (0.00221)
Reward Vocational High School Degree	$\gamma_{12,PR}^V$	0.05735 (0.00603)	0.33053 (0.00819)	0.35622 (0.01879)
Reward Vocational Middle School Degree	$\gamma_{9,PR}^V$	0.00600 (0.00731)	0.24167 (0.00914)	0.02969 (0.01233)
Reward Middle School Degree	$\gamma_{9,PR}^A$	0.10091 (0.00554)	-0.00008 (0.04361)	0.05080 (0.01109)
Reward College Degree	$\gamma_{15,PR}^A$	0.02324 (-)	0.28610 (0.00993)	0.14156 (0.00672)
Reward High School Degree	$\gamma_{12,PR}^A$	0.05448 (0.09879)	0.31063 (0.04340)	0.27187 (0.01993)
Period	$\nu_{1,PR}$	-0.07769 (0.00071)	-0.13118 (0.00015)	-0.11556 (0.00045)
Minor Age	$\nu_{2,PR}$	- (-)	-2.0 (2.5)	-1.6 (0.65087)
Work Lagged	$\eta_{1,PR}$	0.34226 (0.00565)	0.36983 (0.02131)	0.25920 (0.01372)

Table A.2: Choice Alternative: Private Sector – Non-Pecuniary Component.

		Low IQ	Medium IQ	High IQ
Constant	$e_{1,PR}$	240864.3 (8270.5)	302097.3 (6917.2)	280533.2 (15478.1)
Type 1	$e_{2,PR}$	-8427.4 (10530.6)	-654.3 (32282.4)	-11050.5 (12603.6)
Type 2	$e_{3,PR}$	-8005.3 (6211.8)	24773.1 (9107.2)	130.8 (11889.3)
Type 3	$e_{4,PR}$	-1410.1 (8725.7)	- (-)	- (-)
Reward Experience General Work	$\beta_{2,PR}$	- (-)	3218.4 (7597.3)	2471.5 (850.0)
Reward Experience Work Private	$\beta_{3,PR}$	11547.3 (198.9)	3763.4 (7593.3)	3195.7 (853.8)
Reward Years of Academic Schooling	$\beta_{4,PR}$	1687.0 (1704.2)	11947.7 (1409.3)	14105.7 (1065.5)
Reward Years of Vocational Schooling	$\beta_{5,PR}$	11593.5 (933.2)	2753.6 (823.3)	13330.3 (1322.7)
Any Private Work Experience	$\beta_{7,PR}$	137907.6 (8478.6)	32540.4 (13397.0)	33311.2 (55400.4)
Any Work Experience	$\beta_{8,PR}$	- (-)	57641.4 (12362.3)	50964.2 (55540.5)
Reward Vocational High School Degree	$\vartheta_{12,PR}^V$	-2896.1 (3909.6)	12191.0 (3097.0)	13776.8 (7457.6)
Reward Vocational Middle School Degree	$\vartheta_{9,PR}^V$	-3005.7 (2861.5)	16408.3 (4014.0)	13558.2 (7418.2)
Reward Middle School Degree	$\vartheta_{9,PR}^A$	4049.9 (3725.5)	-4154.2 (7603.9)	15394.2 (5103.6)
Reward College Degree	$\vartheta_{16,PR}^A$	- (-)	24647.0 (6302.7)	15457.1 (4607.5)
Reward High School Degree	$\vartheta_{12,PR}^A$	38683.3 (10713.2)	12760.1 (7812.6)	35410.9 (5174.1)

Choice Alternative: Public Sector

Table A.3 presents the point estimates and the standard errors for the parameters in the wage component, while Table A.4 presents the point estimates and the standard errors for the parameters in the specification of non-pecuniary component.

Table A.3: Choice Alternative: Public Sector – Wage Component.

		Low IQ	Medium IQ	High IQ
Constant	$e_{1,PU}$	11.8 (0.02055)	11.1 (0.00478)	10.9 (0.07144)
Type 1	$e_{2,PR}$	- (-)	0.33312 (0.02139)	0.25268 (0.02767)
Type 2	$e_{3,PR}$	0.63879 (0.03723)	-0.01928 (-)	0.00786 (0.02514)
Reward Years of Academic Schooling	$\delta_{1,PU}$	0.07376 (0.00285)	0.14692 (0.00053)	0.17320 (0.00018)
Reward Years of Vocational Schooling	$\delta_{2,PU}$	0.15731 (0.00610)	0.10570 (0.00793)	0.15139 (0.00163)
Reward Experience General Work	$\delta_{3,PU}$	-0.00313 (0.00315)	0.10402 (0.00028)	0.11290 (0.00018)
Reward Experience General Work Square	$\delta_{4,PU}$	-0.00199 (0.09412)	-0.12492 (0.00414)	-0.13500 (0.00405)
Reward Experience Public Work	$\delta_{5,PU}$	0.09588 (0.00302)	0.04119 (0.00031)	0.04733 (0.00029)
Reward Experience Public Work Square	$\delta_{6,PU}$	-0.06687 (0.10340)	0.00276 (0.00513)	-0.01695 (0.00512)
Reward Vocational High School Degree	$\gamma_{12,PU}^V$	0.02690 (0.01088)	0.52160 (1.8)	0.10028 (1.5)
Reward Vocational Middle School Degree	$\gamma_{9,PU}^V$	0.00773 (0.01425)	0.03469 (1.8)	0.08647 (0.00874)
Reward Middle School Degree	$\gamma_{9,PU}^A$	0.10618 (0.01650)	0.02698 (0.13418)	0.02118 (0.09915)
Reward College Degree	$\gamma_{15,PU}^A$	- (-)	0.17475 (0.01211)	0.11769 (0.00889)
Reward High School Degree	$\gamma_{12,PU}^A$	0.04767 (0.58908)	0.47086 (0.12628)	0.50295 (0.08350)
Period	$\nu_{1,PU}$	-0.06696 (0.00033)	-0.09688 (0.00051)	-0.10288 (0.00015)
Work Lagged	$\eta_{1,PU}$	0.32174 (0.01761)	0.60734 (0.01458)	0.65906 (0.06599)

Note: Type 3 is missing in this table as it is not well identified in any of the three specifications. Type 3 is only included for low IQ individuals and the third type always works in the private sector in this specification.

Table A.4: Choice Alternative: Public Sector – Non-Pecuniary Component.

		Low IQ	Medium IQ	High IQ
Constant	$e_{1,PU}$	264882.5 (5465.2)	238795.9 (14060.6)	308910.4 (12454.8)
Type 1	$e_{2,PR}$	- (-)	4911.4 (14060.6)	36709.8 (9129.3)
Type 2	$e_{3,PR}$	-8241.0 (5465.2)	- (-)	- (-)
Reward Experience Public Work	$\beta_{2,PU}$	5990.7 (280.8)	3538.9 (4247.8)	3723.8 (783.8)
Reward Experience General Work	$\beta_{3,PU}$	- (-)	3993.8 (3978.4)	3968.8 (811.6)
Reward Years of Academic Schooling	$\beta_{4,PU}$	2535.7 (2850.4)	2427.6 (1618.7)	9569.9 (999.9)
Reward Years of Vocational Schooling	$\beta_{5,PU}$	10258.5 (754.8)	12569.4 (1886.2)	13465.2 (3871.4)
Work Lagged	$\beta_{6,PU}$	- (-)	- (-)	94855.7 (8329.8)
Any Public Work Experience	$\beta_{7,PU}$	125227.1 (10162.9)	88779.2 (12429.7)	5227.3 (3458.3)
Any Work Experience	$\beta_{8,PU}$	- (-)	23630.2 (12429.7)	114214.1 (11984.3)
Public Work Lagged	$\beta_{9,PU}$	1644.7 (4891.2)	- (-)	- (-)
Reward Vocational High School Degree	$\vartheta_{12,PU}^V$	-2905.2 (2525.8)	21144.7 (6018.8)	- (-)
Reward Vocational Middle School Degree	$\vartheta_{9,PU}^V$	-4699.7 (3369.8)	10122.4 (6018.8)	13544.8 (6406.3)
Reward Middle School Degree	$\vartheta_{9,PU}^A$	9307.8 (5019.9)	842.7 (20783.7)	-4233.7 (15409.8)
Reward College Degree	$\vartheta_{16,PU}^A$	- (-)	20541.4 (6433.7)	74809.3 (3953.1)
Reward High School Degree	$\vartheta_{12,PU}^A$	38956.0 (4327.4)	425.0 (14547.8)	17722.6 (14975.3)

Note: Type 3 is missing in this table as it is not well identified in any of the three specifications. Type 3 is only included for low IQ individuals and the third type always works in the private sector in this specification.

Choice Alternative: Self Employment

Table A.5 presents the point estimates and the standard errors for the parameters in the wage component, while Table A.6 presents the point estimates and the standard errors for the parameters in the specification of non-pecuniary component.

Table A.5: Choice Alternative: Self Employment – Wage Component.

		Low IQ	Medium IQ	High IQ
Constant	$e_{1,S}$	11.9 (0.00259)	11.5 (0.02111)	10.9 (0.09371)
Type 1	$e_{2,PR}$	0.11961 (0.06508)	0.14859 (0.00475)	0.17965 (0.02237)
Type 2	$e_{3,PR}$	-0.00542 (0.04897)	0.38101 (0.02913)	0.06762 (0.01721)
Type 3	$e_{4,PR}$	0.00229 (0.08717)	- (-)	- (-)
Reward Years of Academic Schooling	$\delta_{1,S}$	0.11141 (0.00033)	0.12539 (0.00048)	0.16501 (0.00013)
Reward Years of Vocational Schooling	$\delta_{2,S}$	0.16222 (0.00268)	0.11082 (0.00026)	0.11707 (0.00052)
Reward Experience General Work	$\delta_{3,S}$	-0.00098 (0.00450)	0.08185 (0.00089)	0.10290 (0.00026)
Reward Experience General Work Square	$\delta_{4,S}$	-0.00024 (0.16097)	-0.06265 (0.00022)	-0.12382 (0.00264)
Reward Experience Self Employed	$\delta_{5,S}$	0.10448 (0.00459)	0.04706 (0.00063)	0.03522 (0.00082)
Reward Self Employed Square	$\delta_{6,S}$	-0.05367 (0.16061)	-0.02710 (0.00091)	-0.06122 (0.00278)
Reward Vocational High School Degree	$\gamma_{12,S}^V$	0.10238 (0.01694)	0.14908 (0.01239)	0.38303 (0.08259)
Reward Vocational Middle School Degree	$\gamma_{9,S}^V$	0.01859 (0.01575)	0.04600 (0.01266)	0.03953 (0.08822)
Reward Middle School Degree	$\gamma_{9,S}^A$	0.09662 (0.01785)	0.07674 (0.00982)	0.09875 (0.03282)
Reward College Degree	$\gamma_{15,S}^A$	0.01803 (-)	0.13065 (0.10580)	0.16817 (0.02567)
Reward High School Degree	$\gamma_{12,S}^A$	0.25517 (0.02935)	0.19684 (0.01131)	0.19378 (0.05081)
Period	$\nu_{1,S}$	-0.07432 (0.00014)	-0.08945 (0.00038)	-0.06156 (0.00021)
Work Lagged	$\eta_{1,S}$	0.34353 (0.00312)	0.54149 (0.02402)	0.56502 (0.08108)

Note: Type 3 is missing in this table as it is not well identified in any of the three specifications. Type 3 is only included for low IQ individuals and the third type always works in the private sector in this specification.

Table A.6: Choice Alternative: Self Employment – Non-Pecuniary Component.

		Low IQ	Medium IQ	High IQ
Constant	$e_{1,S}$	295981.5 (9347.0)	180293.9 (9505.6)	230336.8 (31204.5)
Type 1	$e_{2,PR}$	- (-)	2221.2 (12988.7)	3319.5 (12468.2)
Type 2	$e_{3,PR}$	-3809.4 (9340.8)	73325.3 (9990.1)	-880.2 (21064.2)
Reward Experience Self Employed	$\beta_{2,S}$	12380.7 (277.3)	3998.5 (4407.8)	2366.1 (1256.7)
Reward Experience General Work	$\beta_{3,S}$	- (-)	4753.4 (4416.7)	3060.0 (641.9)
Reward Years of Academic Schooling	$\beta_{4,S}$	1667.3 (2882.2)	4520.6 (1645.4)	1835.7 (1110.8)
Reward Years of Vocational Schooling	$\beta_{5,S}$	12887.7 (778.1)	13212.8 (831.7)	12496.1 (1888.6)
Work Lagged	$\beta_{6,S}$	- (-)	50218.3 (38470.0)	200240.3 (3527.5)
Any Self Employed Experience	$\beta_{7,S}$	84565.9 (9372.8)	40700.5 (23442.9)	46430.6 (9830.4)
Any Work Experience	$\beta_{8,S}$	- (-)	105147.0 (23733.8)	95528.2 (28781.1)
Self Employed Lagged	$\beta_{9,S}$	-119.1 (6091.9)	- (-)	- (-)
Reward Vocational High School Degree	$\vartheta_{12,S}^V$	308.8 (3169.3)	48114.5 (3586.9)	18121.9 (20039.2)
Reward Vocational Middle School Degree	$\vartheta_{9,S}^V$	2565.6 (2902.6)	29776.4 (4919.8)	12126.4 (19452.7)
Reward Middle School Degree	$\vartheta_{9,S}^A$	2731.2 (5757.4)	-5594.5 (6172.8)	-6087.2 (8189.4)
Reward College Degree	$\vartheta_{16,S}^A$	- (-)	25394.5 (23798.4)	44290.8 (6371.5)
Reward High School Degree	$\vartheta_{12,S}^A$	-28042.9 (10708.8)	14707.1 (6758.2)	40264.8 (7201.0)

Note: Type 3 is missing in this table as it is not well identified in any of the three specifications. Type 3 is only included for low IQ individuals and the third type always works in the private sector in this specification.

Choice Alternative: Academic Schooling

Table A.7 presents the point estimates and the standard errors for the parameters determining the immediate utility from academic schooling.

Table A.7: Choice Alternative: Academic Schooling.

		Low IQ	Medium IQ	High IQ
Constant	$e_{1,A}$	-98271.5 (9596.7)	-40476.2 (14339.9)	57904.7 (9465.8)
Type 1	$e_{2,PR}$	2024.4 (10347.7)	16843.7 (16143.9)	-4009.4 (10135.0)
Type 2	$e_{3,PR}$	19937.8 (9580.2)	41294.3 (15131.9)	60647.6 (11910.4)
Type 3	$e_{4,PR}$	-59833.1 (9486.5)	- (-)	- (-)
Period	$\beta_{1,A}$	-14372.3 (3861.3)	-39389.8 (2568.7)	-45497.9 (1470.2)
Any Vocational Experience	$\beta_{3,A}$	-97028.4 (10997.0)	-120161.0 (33595.5)	-114535.1 (9797.8)
Academic Lagged	$\beta_{4,A}$	53576.5 (9334.5)	83838.9 (10075.2)	74942.5 (7399.6)
Lagged Choice not Education	$\beta_{5,A}$	- (-)	- (-)	-269762.2 (10205.2)
Post Degree Return	$\beta_{6,A}$	- (-)	40926.1 (10992.6)	71997.1 (14822.9)
Reward High School Degree	$\vartheta_{12,A}^A$	- (-)	146293.7 (12784.7)	46283.7 (7794.7)
Reward College Degree	$\vartheta_{16,A}^A$	- (-)	145380.5 (34924.4)	147851.6 (18013.2)
Reward Middle School Degree	$\vartheta_{9,A}^A$	79305.3 (10867.9)	83377.3 (10443.0)	92611.7 (9277.5)
Reward Vocational Middle School Degree	$\vartheta_{9,V}^A$	- (-)	-247748.4 (18167.5)	- (-)

Choice Alternative: Vocational Schooling

Table A.8 presents the point estimates and the standard errors for the parameters determining the immediate utility from vocational schooling.

Table A.8: Choice Alternative: Vocational Schooling.

		Low IQ	Medium IQ	High IQ
Constant	$e_{1,V}$	-80064.2 (10198.2)	-5157.2 (13003.6)	76784.2 (11992.3)
Type 1	$e_{2,PR}$	12692.6 (9187.6)	3365.8 (16773.2)	-10485.6 (17157.2)
Type 2	$e_{3,PR}$	-5714.3 (8374.0)	43587.5 (13389.6)	75185.9 (15083.3)
Type 3	$e_{4,PR}$	-59483.1 (10103.0)	- (-)	- (-)
Period	$\beta_{1,V}$	-28210.9 (3265.9)	-15538.9 (3563.6)	-15306.9 (4470.3)
Minor Age	$\beta_{2,V}$	- (-)	68987.5 (13400.4)	91075.1 (10623.9)
Any Academic Experience	$\beta_{3,V}$	-101719.8 (11000.8)	-100112.4 (21400.8)	-115946.1 (9184.2)
Vocational Lagged	$\beta_{4,V}$	1606.0 (6611.7)	8723.5 (9392.2)	47514.7 (9853.1)
Lagged Choice not Education	$\beta_{5,V}$	- (-)	- (-)	-260018.9 (24414.3)
Reward Vocational Middle School Degree	$\vartheta_{9,V}^V$	168603.7 (10419.5)	75011.9 (11311.6)	71910.2 (12826.3)

Choice Alternative: Staying at Home

Table A.9 presents the point estimates and the standard errors for the parameters determining the immediate utility of staying at home.

Table A.9: Choice Alternative: Staying at Home.

		Low IQ	Medium IQ	High IQ
Constant	$e_{1,H}$	-61993.5 (7897.9)	39523.1 (12138.7)	134951.0 (8204.2)
Type 1	$e_{2,PR}$	-109584.5 (11445.5)	14010.0 (19666.1)	80035.0 (11524.8)
Type 2	$e_{3,PR}$	-136231.4 (7848.4)	13232.0 (17915.7)	- (-)
Type 3	$e_{4,PR}$	35911.6 (8541.0)	- (-)	- (-)
Minor Age	$\beta_{1,H}$	205531.2 (10388.8)	418226.6 (15531.8)	111000.0 (15939.4)
Period	$\beta_{2,H}$	-2789.1 (550.9)	- (-)	- (-)
Reward High School Degree	$\vartheta_{12,H}^A$	153457.4 (6781.8)	79111.3 (29987.2)	67771.2 (10681.0)
Reward College Degree	$\vartheta_{16,H}^A$	- (-)	63060.1 (36040.7)	67147.8 (27959.3)
Reward Vocational High School Degree	$\vartheta_{12,H}^V$	4752.2 (8417.2)	48480.3 (13685.5)	47196.8 (10132.1)

Time Preferences and the Distribution of Shocks

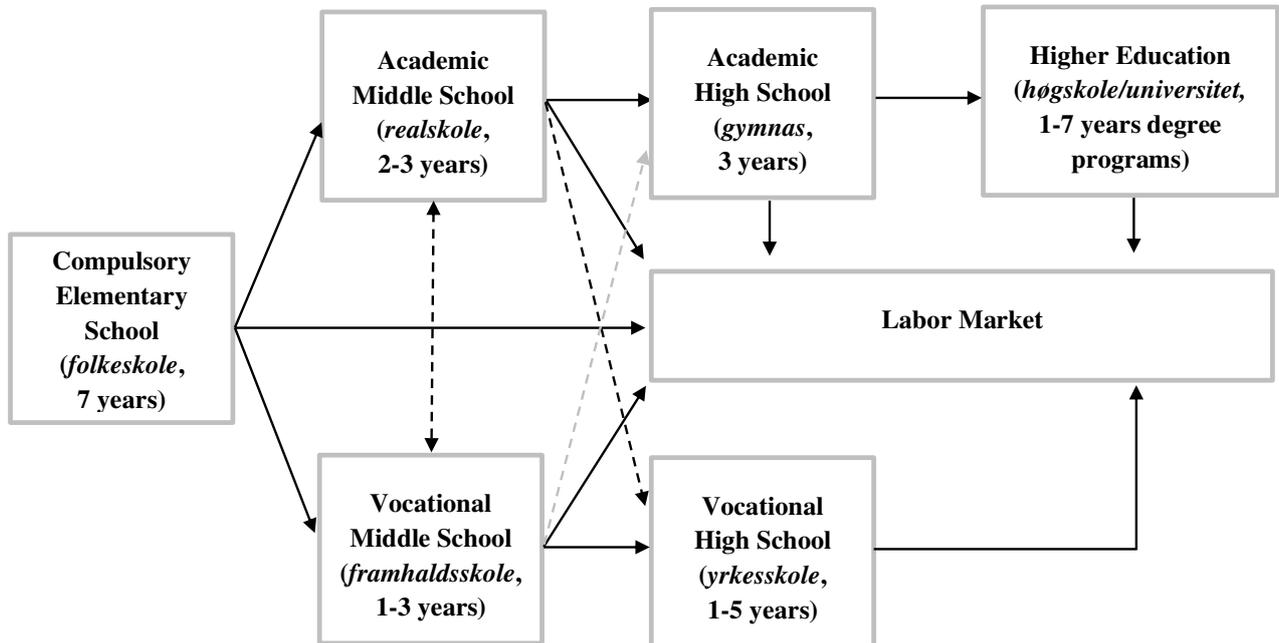
Table A.10 presents the point estimates and the standard errors for the parameters determining the discount rate and distribution of the shocks.

Table A.10: Time Preferences and Distribution of Shocks.

		Low IQ	Medium IQ	High IQ
Discount Rate	Discount Rate	0.96391 (0.00060)	0.96120 (0.00082)	0.95399 (0.00045)
Shock SD Work	Shock SD Work	0.19820 (0.00397)	0.22810 (0.01502)	0.25433 (0.00728)
Shock SD Public	Shock SD Public	0.14752 (0.01042)	0.20060 (0.00901)	0.22360 (0.00847)
Shock SD Self	Shock SD Self	0.22307 (0.02342)	0.32006 (0.01320)	0.31039 (0.01087)
Shock SD Academic	Shock SD Academic	351343.2 (5547.8)	357454.0 (1402.4)	372697.7 (1009.1)
Shock SD Vocational	Shock SD Vocational	99114.8 (3365.6)	42503.6 (1751.0)	173032.0 (1293.4)
Shock SD Home	Shock SD Home	1041070.1 (3525.2)	652166.2 (866.5)	736163.6 (839.7)

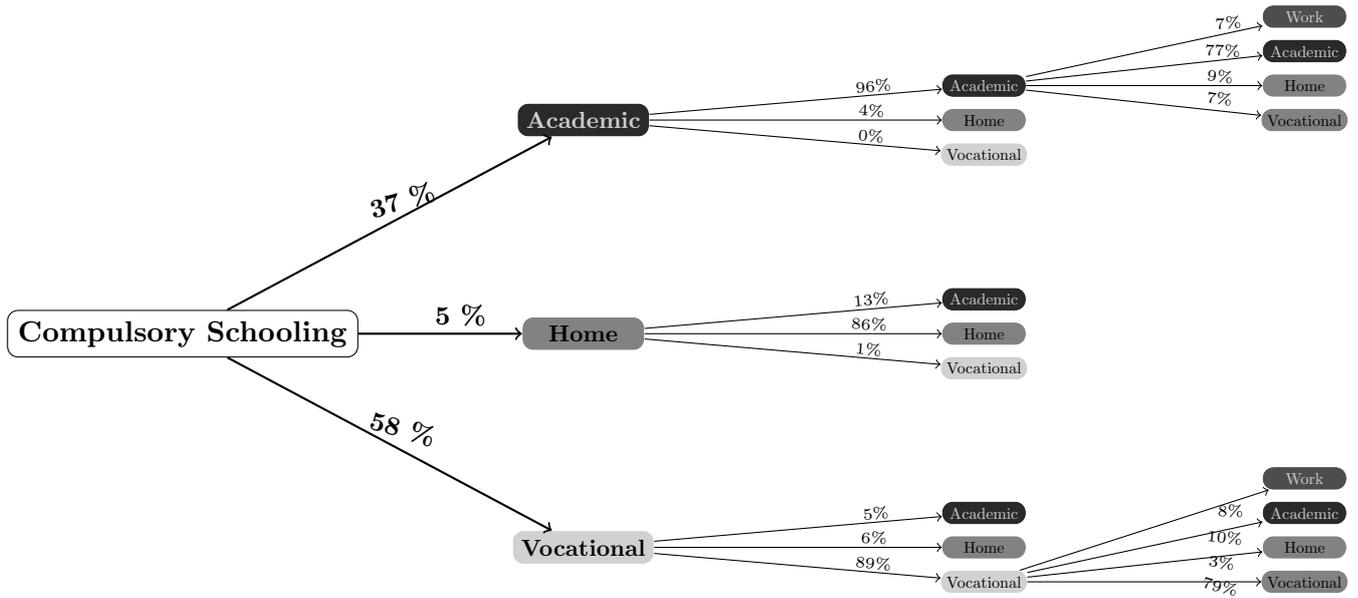
A.3 Additional Figures and Tables

Figure A.1: Illustration of the Norwegian Education System in 1960s.



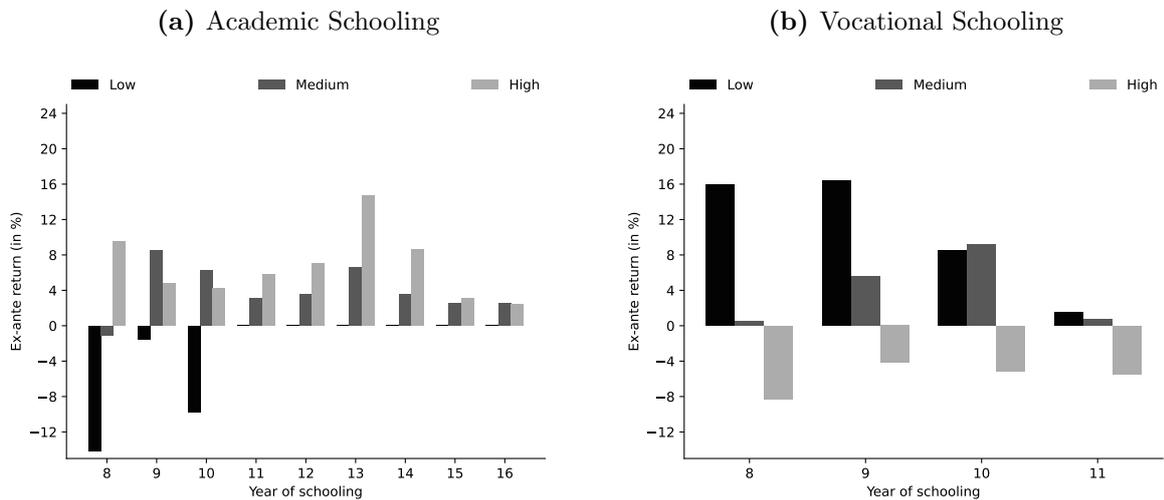
Note: This figure illustrates the Norwegian education system following the 1959 legislation (*Lov om folkeskolen 1959*), when only seven years of elementary schooling was compulsory. Solid black arrows indicate the typical paths that pupils could take as they traverse through the school system, while dotted black (gray) arrows indicate switching between tracks (considered particularly difficult or associated with additional requirements). Following the subsequent legislation in 1969 (*Lov om grunnskolen 1969*), the compulsory schooling was extended to nine years, which was rolled-out in a staggered manner between 1960 and 1975 (see details in Section 3.4). Our baseline analysis uses individuals born between 1955-1960, who faced the pre-reform education system. In 1974, a new type of comprehensive high school (*videregående skole*) was introduced, which made track switching easier. The latter system remained in place up to the mid-1990s, when two reforms (Reform 94 and Reform 97) were enacted, which altered the structure of high school education and lowered the school starting age to six, respectively. See further details in Bertrand et al. (2021).

Figure A.2: Illustration of the Decision Tree and Transitions at Ages 14–17.



Note: This figure shows the decision tree between ages 14 to 17, along with the fraction transiting from one state to another at each age. All individuals start with compulsory schooling at age 14, and transit into either academic schooling, vocational schooling or home at ages 15–16. From age 16 to 17, individuals may also transit into work. To ease the illustration of possible decision paths at age 17, the figure plots only the transitions from the academic–academic and vocational–vocational branches at age 15–16.

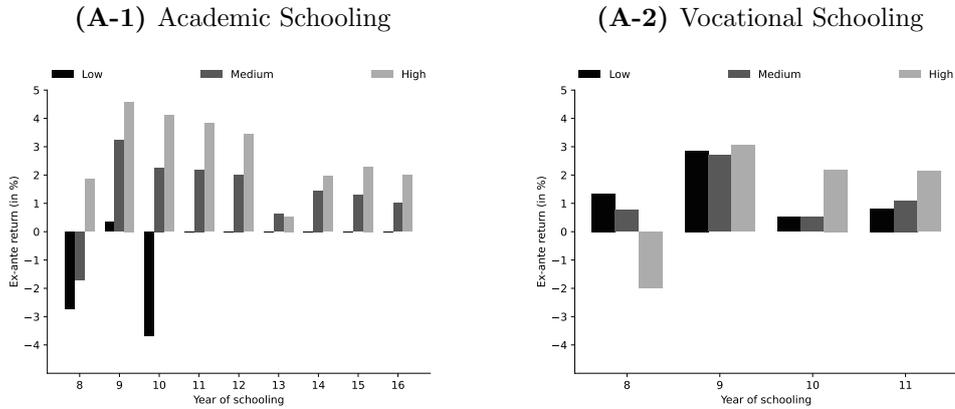
Figure A.3: Average Ex-ante Earnings Returns to Academic and Vocational Schooling.



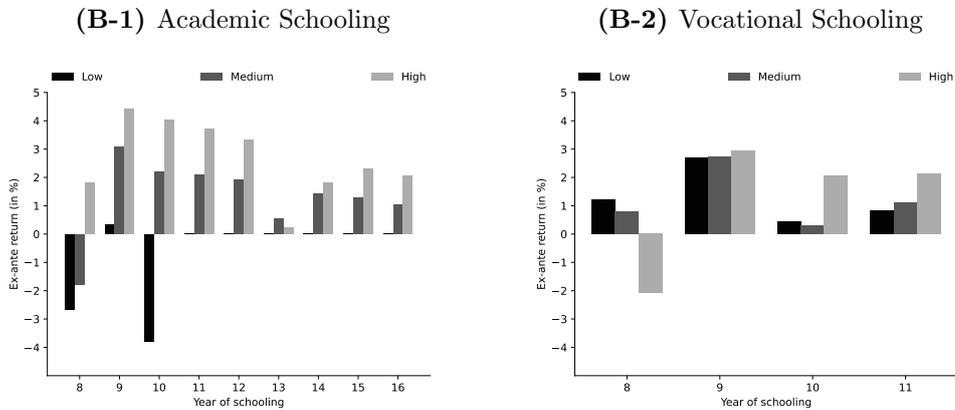
Note: The figure is based on samples of 5,000 simulated schooling careers for each ability group based on the estimated model. This figure contains average ex-ante lifetime earnings returns for each schooling year-track-ability cell. The left figure shows how ex-ante returns to academic schooling develop over time for each group while the right figure shows the same for the vocational track. Each bar shows the average ex-ante return to a particular year of schooling for the individuals in the respective ability group that reaches the relevant transition in our model. For instance, the bar for the high-ability group in the academic panel in year 11 shows the average ex-ante return of the 11th year overall high-ability individuals that have had an uninterrupted academic schooling career until the 10th year. We compute the ex-ante return as defined in Equation (8). Whenever there are only a few people of a particular ability group that reach a particular transition we do omit this group from the calculation.

Figure A.4: Average Ex-ante Returns – The Role of Shocks to Productivity and Tastes.

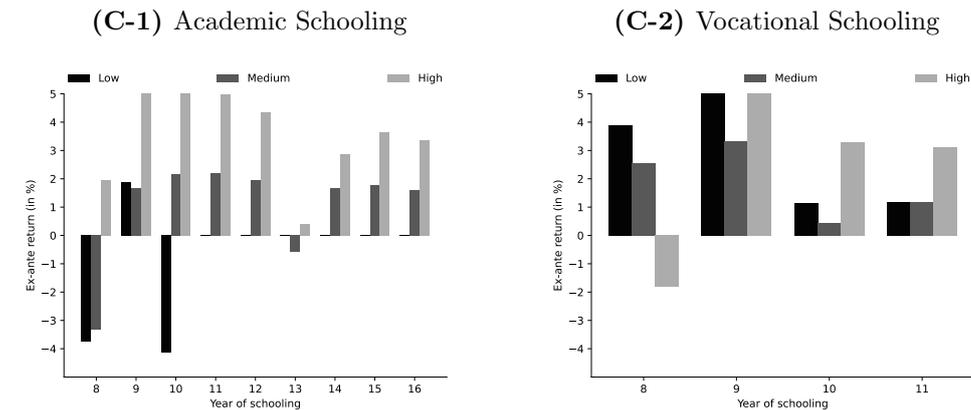
Panel A. Baseline Model



Panel B. No Shocks to Productivity

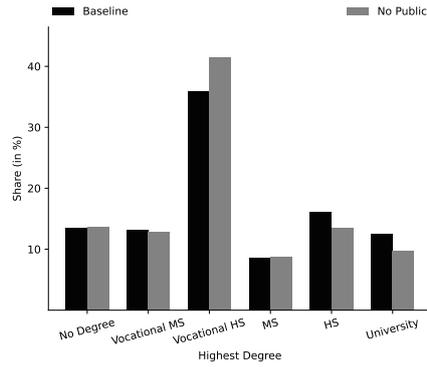


Panel C. No Shocks to Productivity and Tastes



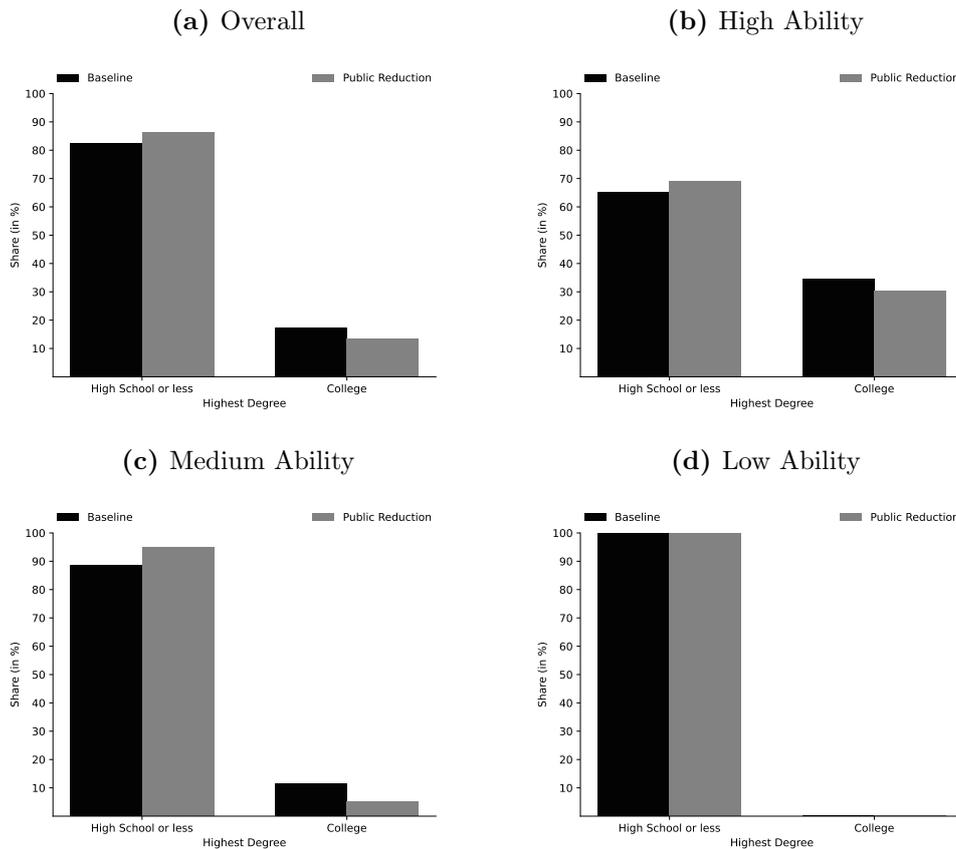
Note: The figure is based on samples of 5,000 simulated schooling careers for each ability group based on the estimated model under different specifications of transitory shocks. Each panel contains average ex-ante returns for each year-track-ability cell for alternative model specifications. Panel A shows results based on the baseline model, while panel B removes shocks to productivity (i.e., no wage risk) and panel C further also removes taste shocks. In each panel, the left figure shows how ex-ante returns to academic schooling develop over time for each group while the right figure shows the same for the vocational track. Each bar shows the average ex-ante return to a particular year of schooling for the subset of the respective ability group that has reached the relevant transition in our model. For instance, the bar for the high-ability group in the academic panel in year 11 shows the average ex-ante return of the 11th year for those that have had an uninterrupted academic schooling career until the 10th year. We compute the ex-ante return as defined in Equation (6). Whenever there are only a few people of a particular ability group that reaches a particular transition we do omit this group from the calculation.

Figure A.5: Change in Final Schooling without Public Sector Job Opportunities.



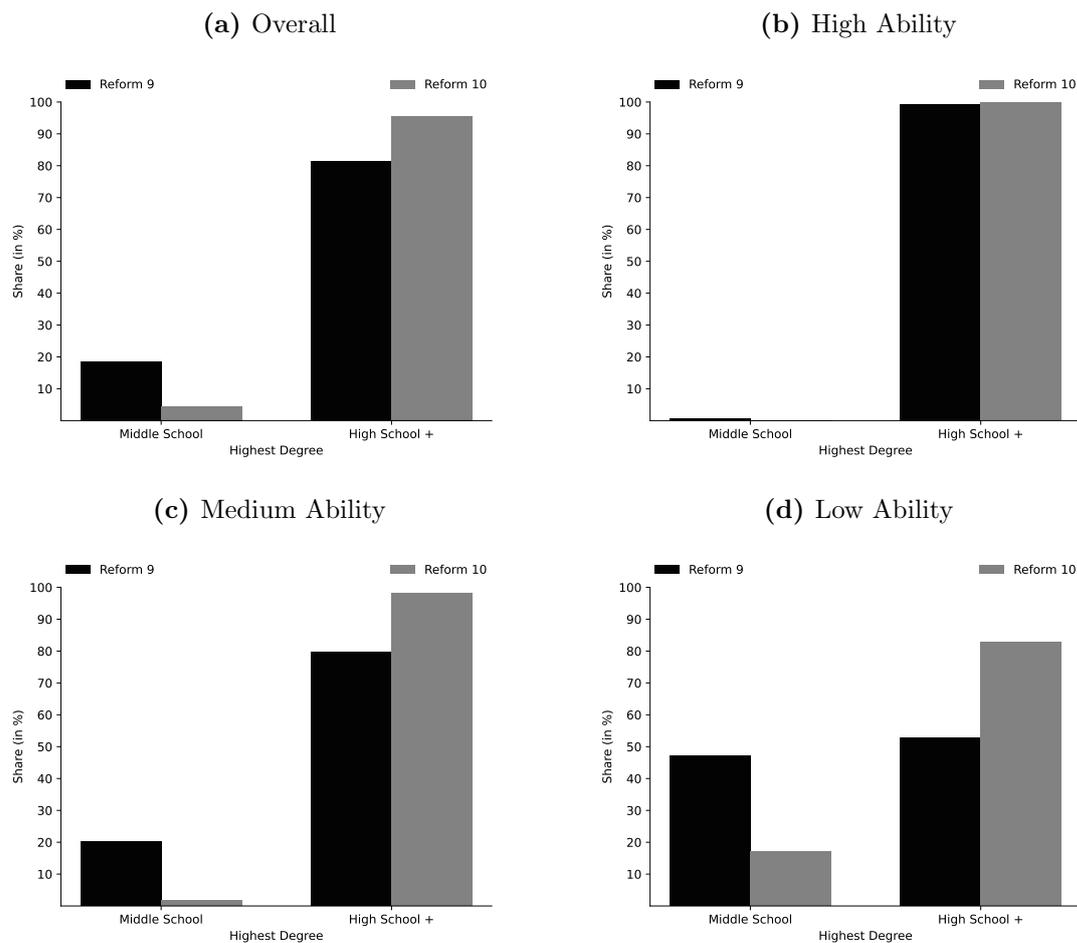
Note: The figure is based on two samples of 100,000 simulated schooling careers under alternative policies. The first bar shows the distribution of final schooling degrees in the baseline model. The second bar shows the distribution of final schooling in an alternative model without a public sector. Throughout, we keep the random realizations of the productivity and taste shocks ϵ_t fixed, and we are thus able to compare the schooling decisions of the same individual under the two different regimes. Finally, we illustrate the distributions of final years of schooling under each policy simulation. The labels on the x-axis refer to different degrees. The abbreviation MS refers to a middle school degree. The abbreviation HS refers to a high school degree. Both of these can be obtained in an academic and a vocational school.

Figure A.6: Change in Final Schooling with Reduction of Public Sector Jobs.



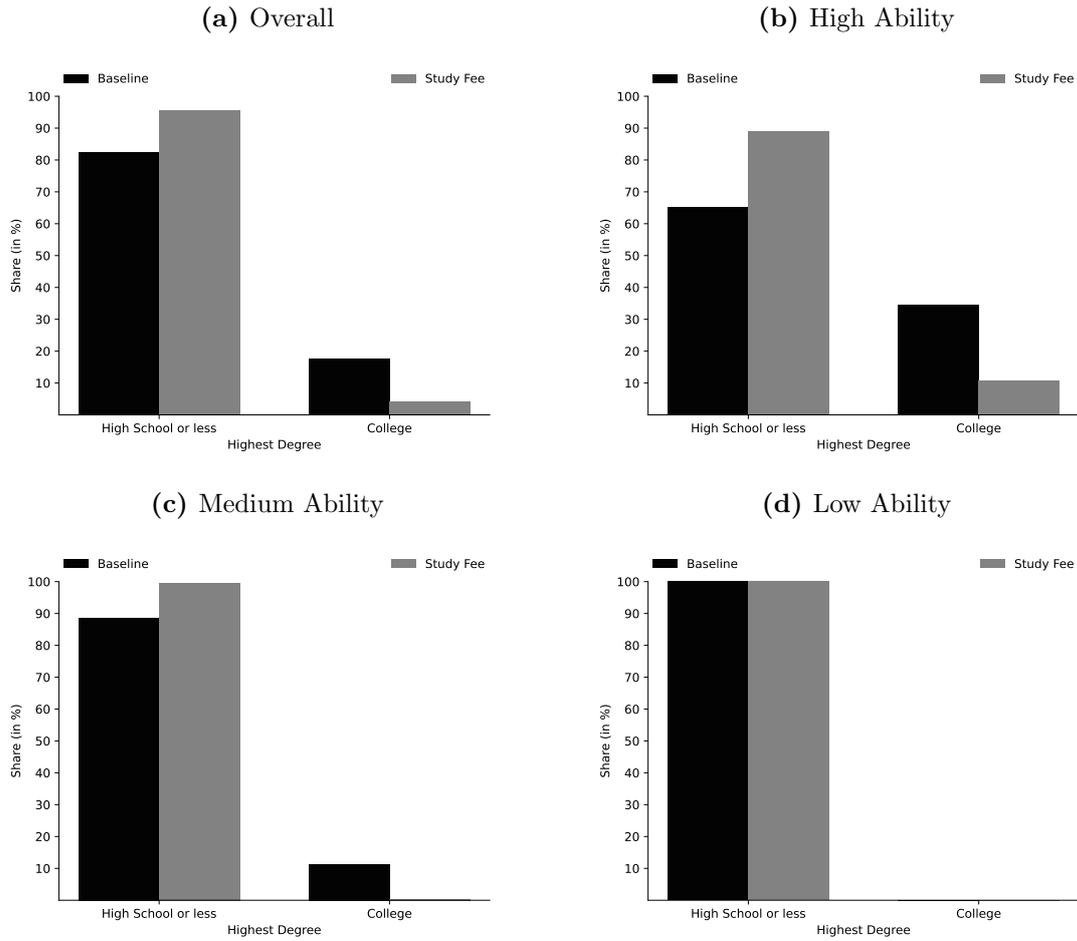
Note: The figure is based on two samples of 100,000 simulated schooling careers under alternative policies. Using the point estimates, we first simulate the baseline model. Next, we rerun the simulation but reduce the baseline nonpecuniary utility associated with the public sector by a half. Throughout, we keep the random realizations of the productivity and taste shocks ϵ_t fixed, and we are thus able to compare the schooling decisions of the same individual under the two different regimes. Finally, we illustrate the distributions of final schooling under each model. 'College' represents all individuals with at least 16 years of academic schooling and 'High School or Less' represents all individuals with less than sixteen years of academic schooling.

Figure A.7: The Impacts of Compulsory High School Enrollment Policy (‘Reform 10’).



Note: The figure is based on two samples of 100,000 simulated schooling careers under alternative policies. Using the point estimates, we first simulate the model with the nine years of compulsory schooling (‘Reform 9’). Next, we rerun the simulation but impose ten years of compulsory schooling to illustrate the compulsory high school enrollment policy (‘Reform 10’). Throughout, we keep the random realizations of the productivity and taste shocks ϵ_t fixed, and we are thus able to compare the schooling decisions of the same individual under the two different regimes. Finally, we illustrate the distributions of final schooling under each model. ‘Middle School’ represents all individuals between 9–12 years of academic schooling and 9–11 years of vocational schooling. ‘High School +’ represents all individuals with 12 or more years of academic schooling and with 11 or more years of vocational schooling.

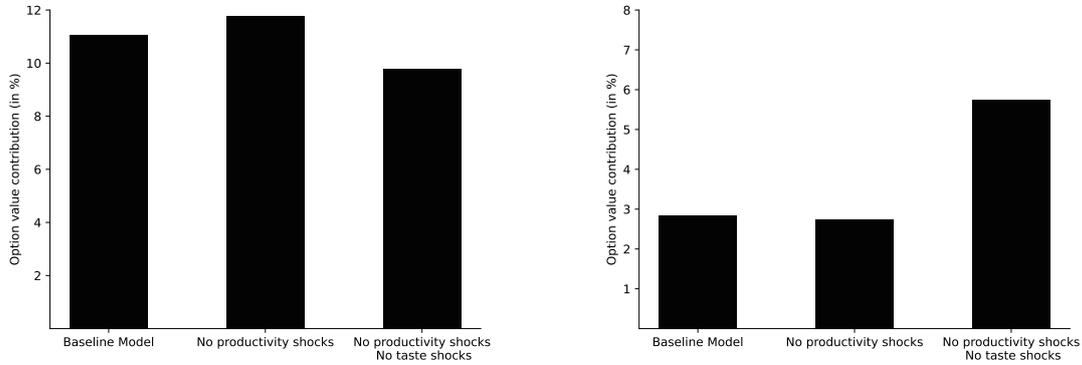
Figure A.8: The Impacts of Tuition Fees Policy.



Note: The figure is based on two samples of 100,000 simulated schooling careers under alternative policies. Using the point estimates, we first simulate the baseline model. Next, we rerun the simulation but impose a college tuition fees of 50,000 Norwegian Kroner (2018). Throughout, we keep the random realizations of the productivity and taste shocks ϵ_t fixed, and we are thus able to compare the schooling decisions of the same individual under the two different regimes. Finally, we illustrate the distributions of final schooling under each model. ‘College’ represents all individuals with at least sixteen years of academic schooling and ‘High School or Less’ represents all individuals with less than sixteen years of academic schooling.

Figure A.9: Option Value Contributions – The Role of Shocks to Productivity and Tastes.

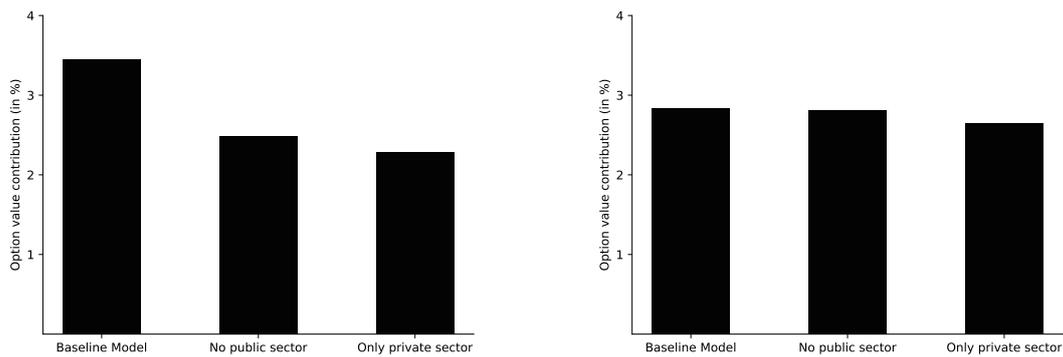
(a) Academic Schooling at 11th Year (Medium) (b) Vocational Schooling at 8th Year (Low)



Note: The figure is based on samples of 5,000 simulated schooling careers for each ability group based on alternative model specifications. The left panel shows the different option value contributions for medium-ability individuals for the 11th year of academic schooling, while the right panel shows the option value contributions for low-ability individuals for the 8th year of vocational schooling. The option value contribution is defined in Equation (9). The bars correspond to different model specification; the first bar corresponds to the estimated model, the medium bar corresponds to an adapted version of the estimated model where productivity shocks (i.e., wage risk) is turned off and the final bar to a model where both productivity and taste shocks are turned off. Whenever there are only a few people of a particular ability group that reaches a particular transition we do omit this group from the calculation.

Figure A.10: Option Value Contributions – The Role of Sectoral Opportunities.

(a) Academic Schooling at 8th Year (Medium) (b) Vocational Schooling at 8th Year (Low)



Note: The figure is based on samples of 5,000 simulated schooling careers for each ability group based on alternative model specifications. The left panel shows the different option value contributions for medium-ability individuals for the 8th year of academic schooling, while the right panel shows the option value contributions for low-ability individuals for the 8th year of vocational schooling. The option value contribution is defined in Equation (9). The bars correspond to different model specification; the first bar corresponds to the estimated model, the medium bar corresponds to an adapted version of the estimated model where individuals can only work as self employed and in the private sector and the final bar to a model where agents can only work in the private sector. Whenever there are only a few people of a particular ability group that reaches a particular transition we do omit this group from the calculation.