The selection effect of childhood abilities on educational decisions

Zuoqi Zhang*[a], Andrew M. Jones[a], and Nigel Rice[a,b]

[a]Department of Economics and Related Studies, University of York, York, UK
[b]Centre for Health Economics, University of York, York, UK

Abstract

Do people with stronger abilities have a greater probability of progressing to higher levels of education? We address this question by examining the influence of childhood cognitive and non-cognitive abilities on three sequential educational decisions made following completion of compulsory education. Using data from the 1970 British Cohort Study, we specify a structural model which combines a sequential decision model with a cognitive development model, and apply confirmatory factor analysis in a measurement model for latent abilities. Estimation follows a structural equation modelling approach. We find that both cognitive and non-cognitive abilities have positive selection effects on encouraging people to progress to the next stage of education irrespective of the level completed. For females preschool cognitive ability plays a more important role in determining educational decisions than it does for men.

Keywords: Early cognitive ability, Educational decisions, British Cohort Study, Structural equation modelling
JEL classification: C31; I21; I23.

*Corresponding author. *Email address:* `zuoqizhang106@gmail.com`.

1 Introduction

At its core education constitutes an investment by individuals and society in human capital intended to enhance future productivity and income. The decision to invest in education depends not only on external opportunities and resources (Favaro and Sciulli, 2022), but also on the influence of individual endowments in determining learning efficiency and costs (Almlund et al., 2011; Harris, 1940). A substantial body of evidence has identified the potential determinants of educational attainment (see e.g., Ermisch and Francesconi, 2001; Glick and Sahn, 2010; Ganzach, 2000), and indicates that cognitive and non-cognitive abilities are significant predictors of academic *achievement* (see e.g., Harris, 1940; Glewwe et al., 2017; Heckman and Vytlacil, 2001). However, comparatively few studies have explored the extent to which these abilities influence educational *preferences*. Weisbrod (1962) demonstrates that the effect of abilities is not restricted to final educational achievement but also extends to educational choices. Their findings reveal that receiving an extra year of schooling opens up options for additional schooling and provides opportunities for learning about personal abilities. On that basis, Heckman et al. (2018) find strong evidence that both cognitive and non-cognitive endowments influence educational choices and outcomes. Evidence from the UK where educational is predominantly publicly funded is, however, scarce. We seek to address this gap in the literature by examining how childhood cognitive and non-cognitive abilities shape educational choices in the context of the UK educational system.

Cognitive ability, also known as intelligence, is needed in the acquisition of knowledge, manipulation of information, and reasoning. It shapes individual memory, learning, decision-making and language abilities (Michalos, 2014). Frederick (2005) demonstrates that the decision-making of an individual is causally determined by general intelligence or various specific cognitive abilities, while Dohmen et al. (2018) note that decisions made under any given task involving risk and uncertainty are, at least in part, the result of a conscious process of mental deliberation, which consequently requires cognitive abilities such as processing probabilities and stakes and evaluating alternative options. Individuals with higher cognitive ability tend to achieve stronger academic performance and gain access to higher-quality educational opportunities. They are better at analysing the potential benefits of these educational opportunities and making informed choices. Non-cognitive skills are defined as abilities to engage constructively with others, for example perseverance, self-control, conscientiousness, resilience, and empathy. These skills are highly valued in educational learning, the labour market, and society more broadly (Kautz et al., 2014). Indeed Brunello and Schlotter (2011) suggest that non-cognitive skills have an impact on academic achievement and labour market performance that is as important as that of cognitive skills. Those with stronger non-cognitive abilities such as perseverance and conscientiousness are more likely to complete their studies and avoid early school leaving (Ryberg, 2018; Heckman et al., 2006). Individuals with greater patience are also more inclined to undertake long-term investments, including additional years of schooling (Coneus and Laucht, 2014). In addition, people with higher abilities tend to benefit more from education than those with lower abilities. Therefore, they are more motivated to pursue long-term education (Heckman et al., 2018). Collectively, these factors point to a 'selection by ability' hypothesis: individuals with higher cognitive and non-cognitive abilities are more likely to choose to continue their studies at each stage of the educational decision-making process.

This paper contributes to the literature in four ways. First, there is little evidence from the UK. By using data from the 1970 British Cohort Study (BCS70), we contribute to the literature by developing the multistage sequential decision model in the British educational context where, at the time, educational access and provision were largely state funded. Secondly, in addition to non-cognitive ability at age 10, two early cognitive abilities are considered: preschool cognitive ability measured at age 5 and post-compulsory school cognitive ability measured at age 16. These ages coincide respectively with children starting primary school and their final year of compulsory education.[1] The timing of these measures enable us to explore the role of compulsory education on decisions to pursue additional education by conditioning age-16 cognitive ability on preschool cognitive ability, where the latter is assumed to be largely determined by initial human capital endowments and family background. Thirdly, unlike previous studies that mostly focus on education in early adulthood, our research considers educational decisions which are recorded up to midlife (age 46), allowing individuals to consider and complete qualifications at a time convenient to them. For example, an individual may want to enrol in university studies at the first opportunity but may fail to do so due to insufficient financial support. Instead, they may first seek employment and at some later date go on to complete an undergraduate education. This longer option period allows for greater expression of individual educational preferences while helping to mitigate the effects of financial constraints. Fourthly, we modify the dynamic sequential decision model developed by Heckman et al. (2018) to estimate the selection effect of childhood abilities incorporating a measurement model to estimate latent abilities. To simulate the developmental of early cognitive abilities, we use a linear value-added plus lagged inputs model of ability formation which was originally proposed by Todd and Wolpin (2007) and amended by Dickerson and Popli (2016). The full empirical formulation of the model is estimated using a structural equation modelling (SEM) approach. Inverse probability

[1]In 1972, the UK government raised the school leaving age to 16. Now under the *Education and Skills Act 2008*, children in England can leave school on the last Friday in June if they turn 16 at the end of the summer holidays. But before they turn 18, they must choose between three options: continuing in full-time education, starting an apprenticeship or traineeship, or spending 20 hours or more a week working or volunteering, while in part-time education or training. In the rest of the UK, the school-leaving age remains at 16. During our target study period, the minimum school leaving age for the sample group was 16.

weights are applied to account for selective attrition and non-response. This set-up offers a flexible framework for linking educational decisions to skills formation in the context of the cohort data we employ. It is also useful in informing further related research, for example, in exploring how early cognitive ability interacts with educational attainment in determining midlife returns for health and income.

We find that preschool cognition (age 5) and non-cognitive ability substantively influence the development of post-compulsory school cognition (age 16). Both non-cognitive and cognitive abilities (preschool and post-compulsory school) have positive effects on sequential choices to pursue post-compulsory schooling, undergraduate, and post-graduate education. However, preschool cognitive ability is not statistically significant at conventional levels. Conditional on age 5 cognitive ability the significant influence of age 16 cognition on the decision to continue in formal education affirms the importance of compulsory education. The magnitude of impact of the latter is greater in decisions to pursue post-compulsory and postgraduate education than for undergraduate education. This is not surprising, given the reality that most students who enter post-compulsory schooling do so with plans to progress to undergraduate education. The importance of non-cognitive ability on educational choices increases with the level of education. We find differenes by gender. For females educational decisions are more strongly influenced by preschool cognitive ability (and age 16 cognitive ability to a lesser extent) and non-cognitive ability. The former does not appear relevant for males where age 16 congnitive ability dominates, and non-cognitive ability appears important for postgraduate education only.

2 Methods

We begin by introducing a dynamic sequential educational decision model to investigate the effect of early abilities on sequential educational decisions. This model is an extension of Heckman et al. (2018) to fit the educational system in Britain. We then present the empirical estimation framework, which is made up of a measurement model and a structural model. In addition to the dynamic choice model, the structural model also includes a value-added plus lagged inputs model to reflect the early cognitive development in childhood drawing on the approaches of Dickerson and Popli (2016) and Todd and Wolpin (2007). Following the findings of Heckman et al. (2018), we hypothesise that selection bias in the decision equation is the result of a combination of cognitive and non-cognitive abilities, both of which are latent variables. The purpose of the measurement model is to estimate these latent abilities based on a set of relevant measurements. The full empirical model is estimated jointly using a Generalised Structural Equation Modelling (GSEM) approach, whereby all equations from both the measurement and structural models are simultaneously estimated, so that standard errors are estimated appropriately.

2.1 Sequential educational decision model

Heckman et al. (2018) present a multistage sequential model of educational choices with transitions and decision nodes and applied to the US setting[2]. We adjust this to fit the British context and investigate the selection effects of early cognitive abilities on sequential educational decisions over time. This is shown in Figure 1. The minimum school leaving age was 16 for our target population, which indicates that children are allowed to make their educational decisions freely at age 16, which is the starting point of our dynamic sequential decision model. Our dynamic decision model starts with the decision whether or not to proceed to post-compulsory secondary schooling (A Level or equivalent qualifications) after completing compulsory schooling. This is followed by a decision to enter university education, and thereafter whether to enrol in postgraduate education.

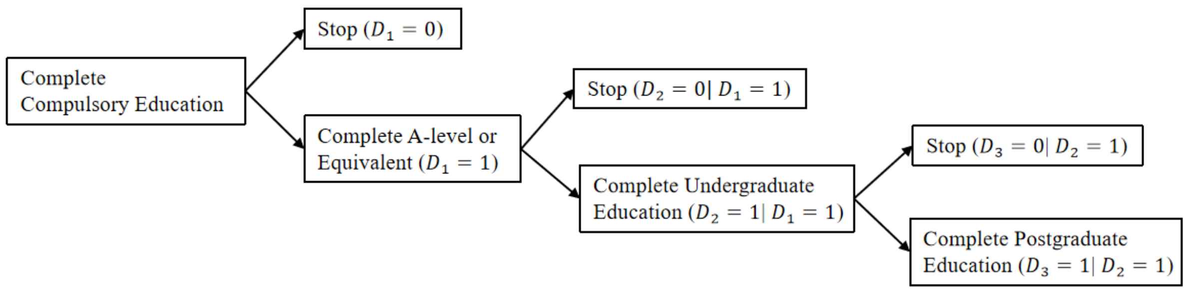

Figure 1: A decision tree for educational decisions applicable to the UK education system.

Assuming that the sequence of decisions is irreversible, we use the nodes $j \in \{1, ..., \overline{p} - 1\}$ to describe different educational stages. At the same time, $P = \{1, ..., \overline{p}\}$ is the set of stopping states and $\overline{p}$ is the highest educational attainment. For each node, the agent has two possible options: remain in the node j or progress to the next node $j + 1$. An indicator D_j is used to denote the agent's educational choice at node j: $\mathrm{D}_j = 0$ means the agent stops at the node j; $\mathrm{D}_j = 1$ indicates that the agent does not stop and proceeds to the next node $j+1$ (continues to the next level of education)[3]. Thus, $\mathrm{D}_p = 0$ indicates that agents stop their education at the state $p \in P$.

We assume this decision process depends on the agent's potential net utility I_j at node j of going on to the next node. People are assumed to continue education when its related net utility is not less than zero:

$$\mathrm{D}_j = \begin{cases} 1, & \mathrm{I}_j \geq 0 \\ 0, & \mathrm{I}_j < 0 \end{cases} \qquad for \quad \mathrm{Q}_j = 1, j \in J, J = \{1, ..., \overline{p} - 1\}$$

[2]This sequential decision model is also analysed in Cunha and Heckman (2007), and Heckman and Navarro (2007).

[3]Heckman et al. (2018) use the opposite notation in that $\mathrm{D}_j = 0$ if a person at node j transits to the next node, and $\mathrm{D}_j = 1$ if a person stops at node j.

where Q_j indicates whether an agent reaches decision node j. $Q_j = 0$ if the agent never progresses to node j, while $Q_j = 1$ indicates the person reaches node j and makes a related educational decision Q_j. $Q_j = 1$ also implies that agents have provided positive responses to all decision nodes prior to j. By conditioning on $Q_j = 1$, we ensure that we pay attention to agents who are eligible to make the transition.

Individuals make their educational decisions depending on the perceived gains (utility). We assume that selection into schooling can be fully accounted for by using observed characteristics and unobserved abilities. Conditional on $Q_j = 1$, the unobserved and continuous utility I_j is approximated by a model:

$$I_j = \phi_j \left(X^D, \theta, \eta \right), j \in \{1, ..., \overline{p} - 1\} \tag{1}$$

where X^D is a vector of observed exogenous variables that determine the transition decisions of the agent at different nodes, and θ is a vector of unobserved abilities which are latent cognitive and non-cognitive abilities. η is an idiosyncratic error term and assumed to be normally distributed with mean zero.

Empirically, we cannot verify the agent's utility of every option but only observe their educational choice at each stage (proxied by an individual's observed highest educational attainment). Accordingly, for the specific binary educational decision at node j, utility I_j is assumed to be linearly identified by the abilities (θ) and a vector of exogenous controls (X_j^D)[4]. Thus, we have:

$$I_j = \beta_j \theta + \pi_j X_j^D + \eta_j \tag{2}$$

where β_j indicates the selection effect of early abilities on the educational decision D_j. The error term η is assumed to be independent of factors θ and X^D. It is also assumed to be uncorrelated with the error terms (e and ϵ) in the measurement and ability models described in Sections 2.2.1 and 2.2.2. Conditional on X^D, η is assumed to be independent across individuals and transitions ($\eta_k \perp \eta_{k'}, k \neq k',\ and\ k, k' \in \{1, ..., \overline{p} - 1\}$). This sequential educational decision model is estimated by probit regression.

Endogeneity caused by unobserved factors may be a potential concern. Our model framework and assumptions strictly adhere to those specified by Heckman et al. (2018). A fundamental assumption of this framework is that the endogeneity of socioeconomic factors influencing educational decisions primarily stems from unobserved cognitive and non-cognitive abilities. After controlling for these two factors, the model is assumed to be free from endogeneity concerns. Similarly, by constructing measurement models to

[4]Linearity is assumed for ease of interpretation, but it is not necessary. For example, using data from National Longitudinal Survey of Youth in America, Ganzach (2000) estimates the influence of cognitive ability in a non-linear formulation and finds that cognitive ability and mother's education has an offsetting relationship on educational expectation and educational attainment, while cognitive ability has a synergistic relationship with educational expectation in determining educational attainment.

estimate cognitive and non-cognitive abilities, our educational decision equation is also assumed to be free from endogeneity issues.

2.2 Empirical analysis

2.2.1 Measurement model

Following Cunha and Heckman (2008), the set of cognitive and non-cognitive abilities $\theta \in \left\{\theta^C, \theta^{NC}\right\}$ are assumed to be latent, which means they cannot be observed and measured directly by researchers. Accordingly, we have several test scores, each of which contains partial information about the relevant latent ability. We can think of each measure as measuring the relevant ability with "measurement error". The purpose of the measurement model is to extract information about these latent abilities from each test score and to predict their "true" values.

For each θ_t, we have m related measures available. Let $\mathrm{M}_{m,t}$ denotes the m^{th} measurement result (e.g. cognitive test score) for θ_t at time t. Since each measure contains partial information about θ_t, $M_t = (\mathrm{M}_{1,t}, ..., \mathrm{M}_{m,t})$ is systematically defined by:

$$M_t = \phi\left(\theta_t, e_t\right)$$

where $e_t = (\mathrm{e}_{1,t}, ..., \mathrm{e}_{m,t})$ is a vector of measurement errors[5]. We assume a linear form for the measurement equations:

$$\mathrm{M}_{m,t} = \alpha_{m,t}\theta_t + \mathrm{e}_{m,t} \tag{3}$$

where α is a vector of factor loading that captures the association between the observed measure and the unobserved ability, which presents the part of the information about the latent variables contained in the measurements. To deal with the different scaling, we standardise all measurement scores and normalise the factor loading of one of the measures for each factor in each period to unity[6]. We assume the measurement errors to be normally distributed with mean zero, be independent across measurement equations and over time ($\mathrm{e}^z_{l,t} \perp \mathrm{e}^{z'}_{l',t}$, for $l \neq l'$, $t \in \{1, ..., T\}$, $l, l' \in \{1, ..., m\}$ and $z, z' \in \left\{\theta^C_t, \theta^{NC}_t\right\}$), and be independent of θ.

After estimating the measurement models, we can obtain the estimated cognitive abil-

[5] We assume that cognitive ability influences only cognitive measures and non-cognitive ability influences only non-cognitive measures. There is no cross complementarity. This assumption differs from Heckman et al. (2018). For example, Heckman et al. (2018)'s cognitive measurement model assumes that non-cognitive abilities affect cognitive variables in addition to including some control variables in the function. The main reason they do this is that one of their cognitive indicators is 9th grade GPA. They argue that academic success, while largely determined by cognitive ability, also depends on socio-emotional characteristics. However, our cognitive indicators are scores on cognitive tests, which are less likely to be influenced by other factors. We adopt the same specification as Conti et al. (2010).

[6] Switching the normalisation to the loading on other measures has no substantive effect on the results.

ities ($\widehat{\theta}^{C}$) and estimated non-cognitive ability ($\widehat{\theta}^{NC}$). We then substitute these estimates into the subsequent structural model.

2.2.2 Early cognitive development model

The early cognitive development model aims to capture changes in cognition across periods in childhood. Let us assume that the stock of cognitive ability at time t (θ_t^C) is a function of the past cognitive ability stock (θ_{t-1}^C), some exogenous factors (X_t^C) and an error term ϵ_t:

$$\theta_t^C = f\left(\theta_{t-1}^C, \mathrm{X}_t^C, \epsilon_t\right)$$

where $t \in \{1, ..., T\}$ represents different time periods, and $t = 0$ indicates the time of birth. We assume the development of cognitive ability over time has a linear formation[7]:

$$\theta_t^C = \gamma_t \theta_{t-1}^C + \lambda_t \mathrm{X}_t^C + \epsilon_t$$

where γ_t is a vector of time-varying parameters to be estimated which denotes the time effect of cognitive development, and ϵ_t is an error term that is normally distributed with zero mean and is assumed to be independent across individuals and over time. Conditional on X^C, ϵ_t is assumed to be independent of the lagged cognitive ability θ_{t-1}^C[8] and measurement errors e.

For period $t = 0$, as we do not have specific measures to identify the initial cognitive ability, θ_0^C, we assume that this initial cognitive ability is proxied by a linear combination of initial circumstance at birth, X_0^C. Hence, we assume:

$$\theta_0^C = \gamma_0 \mathrm{X}_0^C + \epsilon_0$$

In the empirical model, we consider two periods for cognitive abilities: preschool cognitive ability measured at age 5 (θ_5^C) and post-compulsory school cognitive ability measured at age 16 (θ_{16}^C). Thus, we estimate the following equations:

$$\begin{aligned}
\theta_5^C &= \gamma_5 \theta_0^C + \lambda_5 \mathrm{X}_5^C + \epsilon_5 \\
&= \gamma_5 \left(\gamma_0 \mathrm{X}_0^C + \epsilon_0\right) + \lambda_5 \mathrm{X}_5^C + \epsilon_5 \\
&= \gamma_5 \gamma_0 \mathrm{X}_0^C + \lambda_5 \mathrm{X}_5^C + (\gamma_5 \epsilon_0 + \epsilon_5) \qquad (4) \\
\theta_{16}^C &= \gamma_{16} \theta_5^C + \lambda_{16}^{NC} \theta_{10}^{NC} + \lambda_{16} \mathrm{X}_{16}^C + \epsilon_{16} \qquad (5)
\end{aligned}$$

[7]This linear formation is known as the value-added specification of cognitive production function.

[8]We identify the cognitive model and assumptions following Dickerson and Popli (2016), except that we exclude a latent parental investment variable from the equations, as this is not relevant to our research interests.

where $\gamma_5\gamma_0$ represents the effect of birth conditions on preschool cognitive ability and γ_{16} represents the time effect of preschool cognitive ability on post-compulsory school cognitive ability. In line with our research focus, we treat non-cognitive abilities separately from the set of control variables (X^C).

2.2.3 SEM framework

Figure 2 is a simplified path diagram, presenting the structural and measurement model estimated by the SEM approach[9]. The unobserved variables (e.g. latent abilities and error terms) are drawn in ellipses and the observable variables are in rectangles. Single-headed arrows represent unidirectional causal connections between two variables.

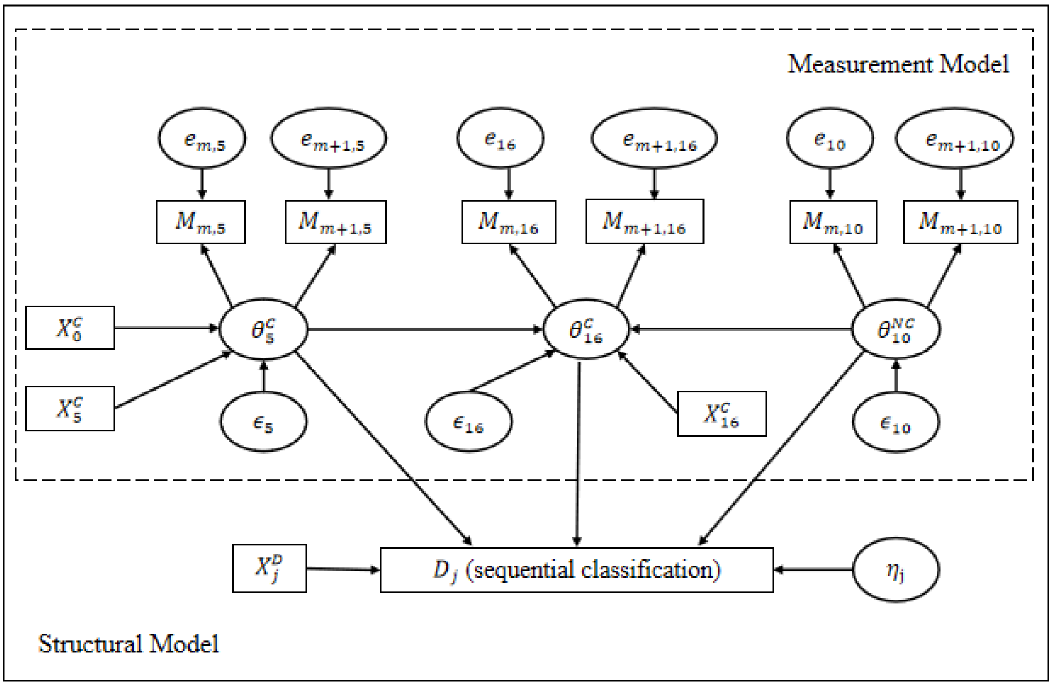

Figure 2: Structural equation modelling framework

The dashed rectangle shows the measurement models of preschool cognition θ_5^C and post-compulsory school cognition θ_{16}^C (given by Equation (3)). For the measurement model of latent cognitive ability at age t, to simplify, we present only two cognitive measures $M_{s,t}$ and $M_{s+1,t}$. Our measurement model has a reflective format which presumes that latent factors affect observed indicators, and not vice versa (Hoyle, 2012), which is estimated via the Confirmatory Factor Analysis (CFA) in SEM.

The structural model (given by Equation ((2), (4), and (5)) is illustrated in the large rectangle. Latent preschool cognition (θ_5^C) is influenced by birth conditions (X_0^C) and exogenous covariates (X_5^C), while post-compulsory school cognition (θ_{16}^C) is affected by past

[9]Since the key variables in the model are binary, we specify a generalised lineal model SEM to better fit the data generating process. Estimation is performed by using the GSEM command in Stata version MP18.

cognitive ability (θ_5^C), non-cognitive ability (θ_{10}^{NC}) and other exogenous covariates (X_{16}^C). Each educational decision (D_j) is determined by both early cognitive abilities (θ_5^C,θ_{16}^C) and non-cognitive ability (θ_{10}^{NC}) together with exogenous covariates (X^D). These equations are estimated by maximum likelihood. In the empirical application, we adopt a one-step estimation approach where the measurement and structural models are estimated simultaneously.

3 Data

Our individual-level data come from the BCS70, which is a longitudinal and multipurpose study, following the lives of around 17,000 people who were born in one week of 1970 in Scotland, England and Wales. There are currently 11 waves in the BCS70, covering interviews with respondents from birth to age 51[10]. Each wave collected detailed information on health, educational and social development, and economic circumstances among other factors. Most importantly, the BCS70 tracks and measures individual cognitive ability from early childhood to later adulthood, which allows researchers to discover the development pattern of human cognition and study the effect of cognitive ability on other aspects of life. This fits well with our research question.

3.1 Sample attrition

The sample for this paper is drawn from the nine main waves of BCS70, as detailed in Table 1. Due to the extended time span, natural attrition and item non-response occurs, for example, due to relocation and changes in contact information, or failure to respond to a question. After merging data across waves, we define two samples. First, a "baseline sample", which includes only the main variables of interest — namely, all ability test scores related to early cognitive and non-cognitive abilities together with education in adulthood — and has a sample size of 3,799. Our study requires participants to have taken at least one cognitive test at ages 5 and 16. This leads to substantial reductions in the sample due to the fact that significantly fewer individuals took cognitive tests at age 16 compared to age 5. In addition we include information on educational decisions made after completing compulsory education. To minimise information loss, we incorporated education data from the waves corresponding to ages 30 through 46, including intermediate years. For participants who did not report education information at age 46, we used their earlier (post-30) education data for interpolation. We excluded the 26-year-old wave because most individuals who pursued postgraduate education had not yet completed their degree at that point. We assume that educational changes are

[10]People were interviewed at the ages of birth, 5, 10, 16, 26, 30, 34, 38, 42, 46, and 51. The latest wave for the 51-year-old interview was not available at the time of this study.

relatively minimal after age 30.

Secondly, we define a "estimation" sample, which additionally includes control variables, and has a sample size of 2,363. Due to missing data, including control variables significantly reduces the available sample for estimation. We compare the distributions of the control variables in our "estimation" sample with those in the original data, using t-tests. Results are summarised in Table A.7. With the exception of mother's age at birth, the t-tests show that the distributions of the control variables differ across the estimation sample and original data. Taken together, it is possbile that the loss of sample may bias our estimates. In all subsequent analyses, we apply inverse probability weighting (IPW) to address sample attrition and provide estimates that can be interpreted as representative of the original survey population. The details of IPW estimation are provided in Appendix A.1.

Table 1: Number of observations changed when merging data

Data	Achieved sample
Birth Sweep	17196
Age 5 Sweep	13049
Age 10 Sweep	12608
Age 16 Sweep	6044
Age 30 Sweep	11226
Age 34 Sweep	9665
Age 38 Sweep	8874
Age 42 Sweep	9841
Age 46 Sweep	8581
No. obs in baseline sample	3799
No. obs in selected sample	2363

Note: The baseline sample includes only the main variables of interest: all ability test scores related to early cognitive and non-cognitive abilities together with education in adulthood. The reduction in sample size at this stage is primarily due to natural attrition in the data and non-response on relevant variables. The education decision variables are generated from the last five sweeps. The estimation sample includes non-missing observations in baseline sample and covariates. The decrease in sample size at this stage is due to non-response on covariates. Source: the BCS70 wave 1, 2, 3, 4, 6, 7, 8, 9 and 10.

3.2 Measure educational decisions

We consider three sequential educational decisions that people face after compulsory secondary education, which are listed in Table 2. Only after completing a given level of education may an individual decide whether or not to move on to the next educational decision. Accordingly, the number of observations decreases as educational level increases. Each educational decision is a dummy indicator, with 1 representing a move to the next level and 0 otherwise. Educational decisions are inferred based on the highest educational qualification they obtain up to midlife (30-46 years old). This contains four groups: compulsory schooling, post-compulsory education, undergraduate education, and postgraduate education[11]. This categorical variable is obtained by combining two variables from the BCS70 — the individual's highest National Vocational Qualification (NVQ) level from an academic qualification and the highest NVQ level from a vocational qualification - and transforming NVQ levels into educational achievements[12]. The first and second NVQ levels are classed as compulsory secondary education. The third level of the NVQ is equivalent to post-compulsory education. Additionally, the fourth level of the NVQ equals undergraduate education, while the fifth level of the NVQ is analogous to postgraduate education.

Table 2 shows that in the sample, around 63% of people continue their education after completing compulsory secondary schooling. This is 8 percentage points higher than the general population in the wave 10 sweep. Following post compulsory education, 76% of these individuals chose to undertake and finish their undergraduate degrees (compared to 75% of the Wave 10 sweep), while only about 22% of these individuals continued to complete the postgraduate education (compared with 20% of the Wave 10 sweep). The proportions of educational decisions in the sample are close to the proportions observed in the raw BSC70 sample, with the exception for the proportion of those who completed post-compulsory schooling, which is higher in the sample.

3.3 Cognitive ability measures

Preschool cognitive ability and post-compulsory school cognitive ability are the two early cognitive abilities that we focus on in this paper. Both are latent variables and cannot be observed directly by researchers. Instead, multiple age-specific cognitive ability tests designed by psychologists are conducted at the relevant age sweeps of the BCS70. We construct a measurement model to measure early cognitive abilities separately using these cognitive tests. Table A.8 presents descriptive statistics for all related cognitive test scores. We compute the total score for each cognitive ability test and use the standardised score in the measurement model (Moulton et al., 2020).

Preschool cognitive ability is assessed at age 5 via five tests. The copying design test requires the child to make two copies of eight shapes to show their visuo-spatial abilities.

[11] Our measure of educational decision is obtained by extrapolating back from the individual's highest educational achievement. A positive answer to an educational decision implies that the person has chosen to pursue a certain level of education and has obtained the appropriate degree certificate. For example, if a person chooses to go to university but for some reason drops out and does not receive a diploma, in our case, the answer to this educational decision is by default negative. In addition, we miss those who are undertaking some stage of education at the time of the interview but have not yet obtained a certificate as this information is not available.

[12] The transformation follows the guidance from Centre for Longitudinal Studies (2011).

Table 2: Descriptive statistics of sequential educational decisions

	N	Min	Max
D1: Whether to complete post-compulsory schooling	2363	0	1
No	874		
Yes	1489		
D2: Whether to finish undergraduate education, after post-compulsory schooling	1489	0	1
No	363		
Yes	1126		
D3: Whether to complete postgraduate education, after undergraduate education	1126	0	1
No	875		
Yes	251		

Source: the BCS70 wave 10.

The English picture vocabulary test asks a child to pick one from four photographs that match a particular word (a total of 56 sets). This test aims to examine the child's verbal ability. The human figure drawing test measures general perceptual ability, in which a child is asked to draw a picture of a man or a woman. To examine children's spatial development, the complete profile test requires children to fill in the features of a profiled human face, such as a nose, eyes, and so on. Last, the Schonell reading test is used to evaluate children's "reading age" by asking them to read 50 words.

Post-compulsory school cognitive ability is also examined by five cognitive tests. For the spelling test, cohort members must distinguish whether 100 words are spelled correctly. For the vocabulary test individuals are asked to select a word from a multiple-word choice list (75 words in total), that shares the same meaning as the term that is presented. The five-subscales of the condensed Edinburgh Reading Test measure a teenager's verbal (reading) skills from vocabulary, grammar, sequencing, comprehension, and retention. An arithmetic test includes 60 multiple-choice questions on topics such as probability, arithmetic, and other subjects. The final test poses 11 out of the 28 matrix questions from the matrices section of the British Ability Scales (BAS) test, which assesses the non-verbal pattern reasoning ability of teenagers. Each pattern matrix is missing a piece in the bottom right corner. Children need to identify the missing piece and select it from five square tiles provided.

3.4 Non-cognitive ability measures

The BCS70 measures children's non-cognitive ability at age 10. Table A.9 presents six selected measurements following Conti et al. (2010): the locus of control (caraloc) scale, the perseverance scale, the cooperativeness scale, the persistence scale, the attentiveness scale and the completeness scale. The locus of control scale, which has 20 questions, measures children's perceived achievement control. Five distractor questions are omitted and accordingly we obtain a raw score that ranges from 0 to 15 where high scores show greater self-esteem and internalisation[13]. The perseverance scale is based on the question "How much perseverance does the child show in the face of difficult tasks?", while the cooperativeness scale is derived from the question "How cooperative is the child with his peers?". Next, the persistence scale originates from the question "Does the child show perseverance and persist with difficult or routine work?". The attentiveness scale derives from the question "Does the child pay attention to what is being explained in class?". Last, the completeness scale is measured by the question "Does the child complete tasks which are started?". The raw scores of the latter five scales range from 1 to 47.

3.5 Socioeconomic background

We control for an individual's initial birth conditions and family circumstances. Table 3 shows the set of control variables used in each estimated equation, while Table 4 displays corresponding descriptive statistics.

Birth weight is measured in kilograms, and together with mother's age at delivery, accounts for initial endowment and early disadvantage that the child might face. Gender is a dummy indicator where 1 equals male.

Table 3: Selection of control variables in each structural equation

	Preschool cognitive ability	Post-compulsory school cognitive ability	Educational decisions
Mother's age at birth	$\surd$		
Birth weight	$\surd$		
Gender	$\surd$	$\surd$	$\surd$
Mother's education (age 5)	$\surd$	$\surd$	$\surd$
Number of siblings (age 5)	$\surd$	$\surd$	$\surd$
Family income (age 16)		$\surd$	$\surd$

Note: Preschool cognitive ability refers to Eqn (4); Post-compulsory school cognitive ability refers to Eqn (5); Educational decisions refers to Eqn (2).

We select three factors to proxy early family circumstance: number of siblings, mother's education, and family income. We count the number of siblings of cohort members and construct the following three categories: none, one, and two or more siblings. Respondents report the types of educational qualifications that their mother holds. Due to a number of missing values across parents, we construct a derived variable representing mother's highest educational qualification, or father's if missing. This variable contains

[13] Questions 4, 7, 11, 15, and 19 are deleted. Each "No" response counts as one point, except for question 10 where the "Yes" response earns one point. Conti et al. (2010) only deleted 4 questions and received raw scores ranging from 0 to 16. This transformation follows the guidance of the UCL website (University College London, 2017)

three categories: no qualification, below A level, A level and above. The BCS70 offers a derived variable for income which groups weekly household income into 11 categories, ranging from less than £50 to more than £500. We recode these into three groups: low, medium, and high income, according to classification guidance published by the government[14]. The guidance suggests that low income includes households with incomes less than 60% of the national median, and that high income consists of households with incomes in the top 10% of the national distribution. This categorisaton leads us to selecting income up to £150 per week as defining low income households, and income greater than £350 per week as high income households. The remainder of the sample are placed in the medium income group. For estimation, dummy indicators are used to reflect these groups.

Table 4: Sample sizes of control variables

Variable	N	Min	Max
Initial birth conditions			
Mother's age at birth	2363	15	46
Birth weight in kilograms	2363	1.16	6.46
Gender	2363	0	1
female	1384		
male	979		
Early family circumstances			
Number of siblings at age 5	2363	1	3
none	258		
one sibling	1275		
two or more siblings	830		
Mother's education at age 5	2363	1	3
no qualification	1080		
below A level	898		
A level and above	385		
Family income at age 16	2363	1	3
low-income group	692		
medium-income group	1289		
high-income group	382		

Source: the BCS70 wave 1, 2, and 4.

4 Results

We start by introducing the association between the estimated abilities and educational choices (Sections 4.1), then move to the results from the SEM (Section 4.2 and 4.3). There are three latent variables that need to be estimated: preschool cognitive ability

[14]The classification of the low-income group is follows a government website (House of Commons Library, 2024).

at age five, post-compulsory school cognitive ability at age 16, and non-cognitive ability measured at age 10. See Appendix A.2 for descriptive statistics for all the measurement variables.

4.1 Distribution of estimated abilities

Figure 3 presents distribution graphs for the three estimated abilities. The bottom row displays three latent abilities estimated using the full sample (n = 2,363). The top row shows the corresponding ability distributions estimated from the original waves of data (age 5 sweep, n= 13,049; age 16 sweep, n= 6,044; age 10 sweep, n=12,608). As illustrated, the shape of the distribution curve for the sample closely resembles the overall shape of the survey distribution. Predicted preschool cognitive ability follows a normal distribution, which remains robust under tests for skewness and kurtosis. In contrast, the distributions of the predicted post-compulsory school cognitive ability and non-cognitive ability are negatively skewed. Changes in the distribution of cognitive abilities over time may be attributable to family background and compulsory education.

Figure 4 shows density plots of early abilities seperately by the three educational decisions. The density curves for positive respondents who select into the next level of education are shifted to the right of respective curves for the negative respondents who do not transition to the next level of education. This holds for all plots with the possible exception for the graph in the upper right corner. This indicates that the overall level of early ability rises with the level of education and that groups with higher ability are more likely to respond positively to each educational decision. This trend is more evident in post-compulsory school cognitive ability and early non-cognitive ability. These patterns suggest a positive correlation between early abilities and educational decisions.

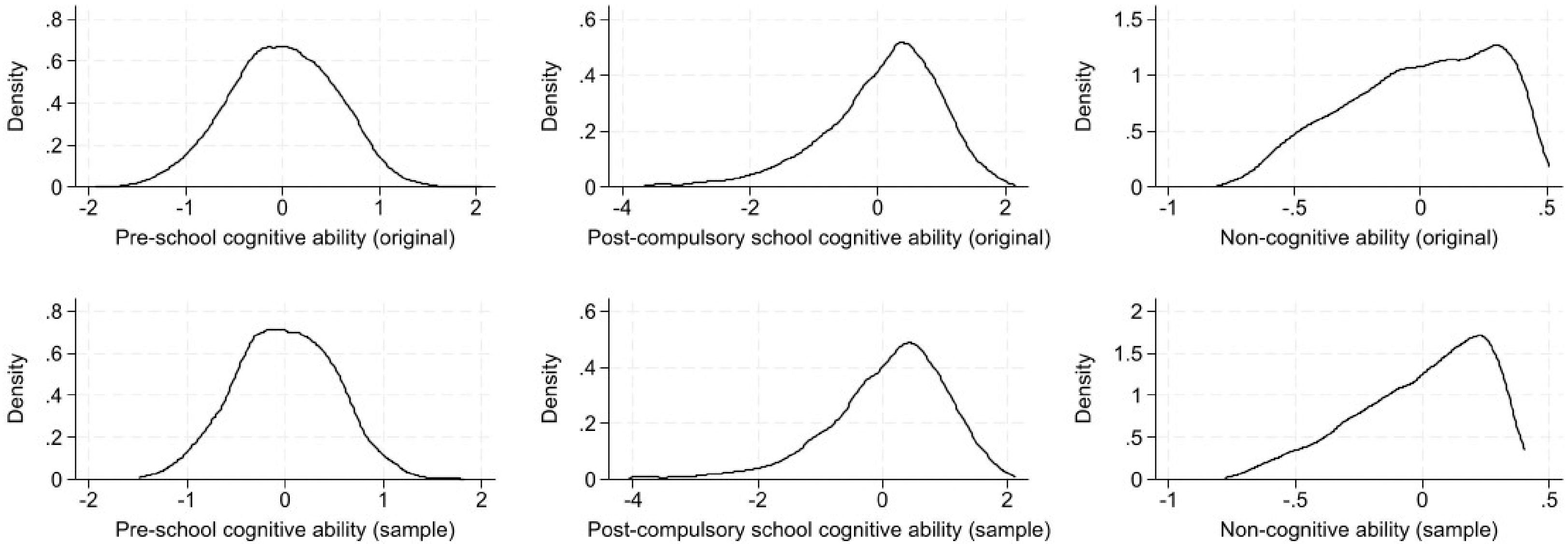

Figure 3: Kernel density estimates for three early abilities predicted from measurement model, for the survey data (the top) and the selected sample (the bottom). Notes: There were 13,049 individuals attended at least one cognitive test at the age 5 wave, while there were 6,044 individuals at the age 16 wave. There were 12,608 individuals at the age 16 wave. There were 2,363 individuals in the sample. We performed t-tests to compare the means across the three groups, and found no statistically significant differences.

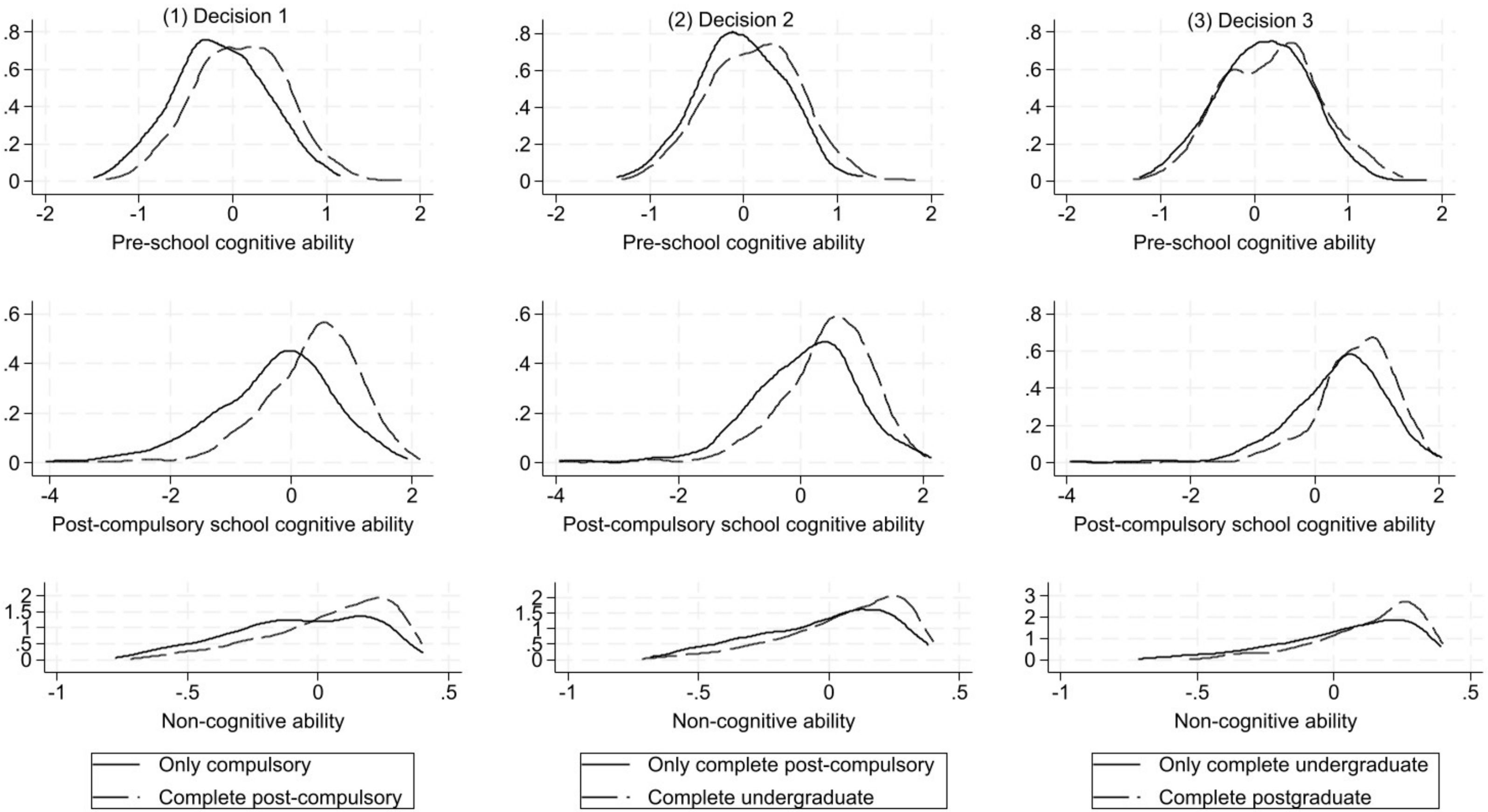

Figure 4: Kernel density estimates for predicted early cognitive abilities, for the first educational decision (the left), the second educational decision (the middle) and the third educational decision (the right). Notes: The dashed line illustrates density for having a positive response for a given decision ($D_j = 1$), while the solid line presents density for having a negative response ($D_j = 0$). All group-wise t-tests indicate statistically significant differences between group means.

4.2 Effect of cognitive development

Table 5 shows the effect of cognitive development. The first column displays regression results for preschool cognitive ability (Equation (4)), while the second column presents regression results of post-compulsory school cognitive ability (Equation (5))[15]. We find that preschool cognitive ability is positively associated with the two birth conditions: birth weight and mother's age at delivery. Controlling for early family background, a one kilogram increase in birth weight is associated with 0.208 standard deviation increase in preschool cognitive ability. This is consistent with Hawkes and Joshi (2012) and Nakamuro et al. (2013), since children with low birth weight or a younger mother are most likely to come from a disadvantaged background. The number of children in the household is likely to influence the allocation of parental resources, affecting cognitive development (Azmitia and Hesser, 1993; Dai and Heckman, 2013). We find that a higher number of siblings (two or more) within the family is associated with lower preschool cognitive ability. This reflects a greater resource dilution in larger families (Azmitia and Hesser, 1993), and a reduction in parental attention and investment plays an important role in early cognitive development.

Preschool cognitive ability and early non-cognitive ability significantly (at the 1% level) predict post-compulsory school cognitive ability. A one standard deviation increase in preschool cognitive ability leads to a 0.543 standard deviations improvement in post-compulsory school cognitive ability, while a one extra standard deviations of non-cognitive ability increases post-compulsory school cognitive ability by 1.365 standard deviations, holding other factors constant. It appears that preschool cognitive ability and non-cognitive ability play a dominant role in the development of post-school cognitive ability. This is intuitive that children who have greater patience and self-control are more likely to gain more knowledge and engage in cognitive abilities in their early home and preschool education - and thus score better on cognitive tests. Also, children with higher preschool cognition gain en early advantages in early cognitive development.

Khanam and Nghiem (2016) discuss the association between family income and cognitive development. From Table 5 , it is also apparent that children from middle- and high-income groups have an advantage in early cognitive development over children from low-income groups. Mother's education has a sustained positive influence on cognitive development (González et al., 2020; Schady, 2011) - the higher the level of mother's education, the greater the improvement in the child's cognitive development. For example, the average post-compulsory school cognitive ability of children whose parents have an A-level or above (primary school) qualification is 0.328 (0.146) standard deviations greater than comparable children whose parents have no educational qualifications. In general, girls are likely to have a higher level of preschool cognitive ability than boys, with the

[15]Equations are estimated using maximum likelihood (ML).

gender gap appearing to widen over time.

Table 5: Results of cognitive development models

	Cognitive ability at age 5		Cognitive ability at age 16	
	Coef.	Std. Err.	Coef.	Std. Err.
Preschool cognitive ability (age 5)			0.543***	(0.063)
Non-cognitive ability (age 10)			1.365***	(0.126)
Mother's age at delivery	0.021***	(0.004)		
Birth weight	0.208***	(0.035)		
Gender (baseline = female)	-0.058	(0.038)	-0.088**	(0.044)
Number of siblings at age 5 (baseline = no sibling)				
one sibling	-0.014	(0.054)	-0.092	(0.068)
two or more siblings	-0.323***	(0.057)	-0.223***	(0.074)
Mother's education at age 5 (baseline = no qualification)				
below A level	0.277***	(0.039)	0.146***	(0.051)
A level and above	0.467***	(0.055)	0.328***	(0.070)
Father interaction term	0.054	(0.044)	-0.081	(0.051)
Family income at age 16 (baseline = low-income)				
middle-income			0.158***	(0.055)
high-income			0.319***	(0.072)

Note: $^{**}p \leq 0.05$, $^{***}p \leq 0.01$. The table presents the estimated structural parameters from the cognitive development model, corresponding to Equations 4 and 5. Since the measurement and structural models are estimated simultaneously, all standard errors are estimated appropriately. For binary and categorical independent variables, the reference categories are indicated. All reported statistics are weighted using IPW. For respondents missing mother's education data, we use the father's highest education level as a substitute. To test whether father's education has a different effect, we add an interaction term which interacts the father indicator with mother's education.

4.3 Selection effect of early abilities

Table 6 displays results from the sequential educational decision model. The first row lists the selection effects of preschool cognitive ability on the sequential educational decisions, while the second row presents the selection effects of post-compulsory school cognitive ability. After controlling for other factors, we find no significant effect of preschool cognitive ability on educational choices. In contrast, cognitive ability measured after compulsory education shows a positive and significant impact on all educational decisions. Post-compulsory school cognitive ability exhibits the strongest influence on the completion of post-compulsory education, followed by its impact on postgraduate education, and lastly, on undergraduate education. For instance, a one standard deviation increase in post-compulsory school cognitive ability is associated with a 37.9 percentage point increase in the probability of completing post-compulsory education, while the probability of completing postgraduate education increases by 20.8 percentage points. Relatively speaking, preschool cognitive ability reflects more an individual's innate cognition, whereas post-compulsory school cognitive ability is influenced by compulsory education, building upon preschool cognition. The impact of preschool cognition on educational decisions is partly direct and selective, but also occurs indirectly by affecting post-compulsory school cognitive ability, which in turn influences educational decisions. In other words, post-compulsory school cognitive ability partially mediates the effect of preschool cognitive ability on educational outcomes. This relationship is captured in our structural model. Regarding the magnitude and significance of estimated coefficients, the influence of post-compulsory school cognitive ability on educational decisions appears to be more consistent and robust compared to preschool cognitive ability. This suggests that post-compulsory school cognitive ability may be a more effective predictor of educational outcomes than preschool cognitive ability.

From the third row of Table 6, we find a strong positive association of non-cognitive ability with all educational decisions. For an undergraduate, a one standard deviation increase in non-cognitive ability at age 10 raises the probability of completing a postgraduate education by 80.9 percentage points, compared with 28.9 percentage points for post-compulsory school cognitive ability. Although a direct comparison of the coefficients is not strictly meaningful due to differences in measurement, the results still highlight the significance of non-cognitive abilities in education. Our findings are consistent with Heckman et al. (2018), that is, individuals with high ability are more willing to pursue higher education beyond compulsory education than those with relatively low ability. They argue that the high-ability group has a higher probability to continue higher education not only because they are able to, but also because of other potential benefits such as sorting on gains.

Consistent with Hegna and Smette (2017), our result show a positive impact of mother's education on the educational decisions of children. Higher mother's education is associated with a greater probability of their child completing further education after compulsory education,a phenomenon referred to as the intergenerational transmission of educational attainment. In particular, children of parents with an A-level qualification or higher are about 47.6 percentage points more likely to choose to pursue a postgraduate education after being an undergraduate than children of less-educated parents. Ganzach (2000) suggests that more educated parents are better equipped to facilitate children's learning by creating a better social and physical environment. The importance of family background in shaping educational outcomes is well established in the literature (see e.g. De Graaf and Huinink, 1992; White, 1982; Wilson, 2001).

The number of children in the household is likely to influence the allocation of parental resources, affecting educational decisions (Jensen and McHale, 2015; Karwath et al., 2014). We find that individuals with more than two siblings have a 25.1 percent points lower probability of completing post-compulsory education and, as a result, lower opportunity to continue subsequent education. This may be attributed to the fact that parents in larger families face greater financial burdens and may encourage their children to enter the workforce earlier to ease the family's budget constraint. Given this, it is unsurprising that families with higher incomes are more willing to allow their children to continue with post-compulsory education[16].

For those who have completed post-compulsory education, we find that family income has no statistically direct impact on whether they complete tertiary education. This finding differs from that of Taubman (1989). One possible explanation is that, at the time, students were largely exempt from university tuition fees[17], and the cost of living for university students was typically covered by a combination of family support, government grants[18] and student loans. Therefore, it is unlikely that family finances played a significant role in influencing educational decisions.

[16] There are two primary reasons why family income has an impact on educational decisions. The first is the willingness to pursue education. Higher-income parents have greater access to educational resources that enhance their children's academic performance and college readiness (Pfeffer, 2018; Looker, 1997), thereby shaping their children's attitudes towards education and expectations of future attainment. A second factor is the cost of tuition and living expenses that can deter students from pursuing higher education. Lunn and Kornrich (2018) found that even in times of economic uncertainty, families with higher incomes continued to prioritise investment in education. Accordingly, the ability to receive direct financial support from family while at university is crucial for students, particularly those facing academic challenges (Pfeffer, 2018). Using the same dataset, Bratti (2007), controlling for potential endogeneity, found that parental income had a strong effect on the likelihood of a child leaving school at age 16. This suggests that lower family income may lead to earlier exits from the education system, thereby reducing access to higher education.

[17] British university students typically complete their undergraduate education between the ages of 21 and 23. For the individuals in our sample, this would generally have occurred between 1991 and 1993, assuming they pursued undergraduate studies. Notably, university tuition fees were low and paid by the local government in the UK until 1998.

[18] The amount of the living grant is related to family income. Students from high-income families may receive a smaller grant, or even none at all, while students from low-income families may receive a larger grant to cover living expenses.

Meanwhile, gender differences in educational preference are not observed in all education stages. One possible explanation is that the UK government's higher education reforms in the 1990s, along with the growing demand for high-skilled labour, significantly increased female participation in higher education, gradually equalising it with that of men. Favaro and Sciulli (2022) also mention that UK performs relatively well in terms of gender equality among OECD countries. Thus, although insignificant, while there was a gender gap in post-compulsory education decisions made before 1990, this gap became less pronounced in decisions related to subsequent education.

Table 6: Results of sequential educational decision models

	Post-compulsory schooling D1		Undergraduate education D2		Postgraduate education D3	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Preschool cognitive ability (age 5)	0.127	(0.083)	0.170	(0.109)	-0.068	(0.124)
Post-compulsory school cognitive ability (age 16)	0.379***	(0.059)	0.165**	(0.072)	0.208**	(0.093)
Non-cognitive ability (age 10)	0.289*	(0.148)	0.376**	(0.189)	0.809***	(0.242)
Gender (baseline = female)	0.087	(0.063)	-0.085	(0.079)	0.003	(0.093)
Number of siblings (baseline = no sibling)						
one sibling	-0.090	(0.101)	0.054	(0.127)	-0.071	(0.155)
two or more siblings	-0.251**	(0.107)	0.041	(0.136)	-0.139	(0.171)
Mother's education (baseline = no qualification)						
below A level	0.270***	(0.068)	0.161	(0.090)	0.126	(0.114)
A level and above	0.414***	(0.105)	0.418***	(0.129)	0.476***	(0.138)
Father interaction term	0.019	(0.071)	-0.299***	(0.080)	-0.137	(0.115)
Family income (baseline = low-income)						
middle-income	0.040	(0.073)	-0.004	(0.097)	-0.122	(0.123)
high-income	0.182*	(0.106)	0.104	(0.134)	0.004	(0.148)

Note: $^*p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$. The table presents the estimated structural parameters from the dynamic educational decision model, corresponding to Equation 2. Since the measurement and structural models are estimated simultaneously, all standard errors are estimated appropriately. For binary and categorical independent variables, the reference categories are indicated. All reported statistics are weighted using IPW. AIC: 459420, BIC: 460008.3. For respondents missing mother's education data, we use the father's highest education level as a substitute. To test whether father's education has a different effect, we add an interaction term which interacts the father indicator with mother's education.

We further explore the influence of gender by exploring heterogeneity in the effect of abilities on educational choices[19]. Table 7 presents the results from stratifying the sample by gender. Holding other factors constant, preschool cognitive ability has a significant and positive effect on the completion of post-compulsory and undergraduate education for females, but no such effect is found for males. Males with higher cognitive ability at age 16 are more likely to pursue undergraduate education, while this trend is not present among females. In contrast, females with higher cognitive ability at age 16 are more inclined to pursue postgraduate education, a pattern not observed among males. Consistent with the overall findings, early non-cognitive skills positively influence educational choices for females. However, no significant effects are found on males' decisions regarding post-compulsory and undergraduate education.

Overall, female educational choices appear to be more influenced by abilities than those for males. One possible reason is that, Adolescents internalise gender role expectations shaped by cultural messages from their social environment—including peers, parents, media, and schools—into their gender ideology (Van der Vleuten et al., 2016). For instance, when it comes to university education, females may be more influenced by other factors - such as family expectations and gender role norms - which weakens the impact of cognitive ability at age 16 on their educational decisions. Although the proportion of women participating in higher education has increased in recent years, men and women continue to differ in their choices of academic fields and occupations (Barone, 2011; Gerber and Cheung, 2008). The persistence of these differences leads to divergent educational pathways and affects subsequent educational opportunities and labour market prospects (Van der Vleuten et al., 2016). Regarding postgraduate education, for example, the gender difference in the returns to education may play a role. High-ability females may be more motivated to pursur further education as a way to signal their value, whereas capable males may find it easier to secure desirable employment opportunities upon completing their undergraduate studies.

[19]The estimations are adjusted using IPW defining by gender. The distributions of propensity scores are provided in the Appendix.

Table 7: Heterogeneity by gender

	(1) main			(2) female			(3) male		
	D1	D2	D3	D1	D2	D3	D1	D2	D3
Preschool cognition (age 5)	0.127	0.170	-0.068	0.297**	0.453**	0.033	0.029	0.017	-0.177
	(0.083)	(0.109)	(0.124)	(0.130)	(0.180)	(0.195)	(0.119)	(0.151)	(0.182)
Post-compulsory school cognition (age 16)	0.379***	0.165**	0.208**	0.284***	0.073	0.245*	0.439***	0.222**	0.121
	(0.059)	(0.072)	(0.093)	(0.076)	(0.097)	(0.144)	(0.088)	(0.108)	(0.127)
Non-cognitive ability (age 10)	0.289*	0.376**	0.809***	0.543***	0.498**	0.594**	0.062	0.222	0.995***
	(0.148)	(0.189)	(0.242)	(0.182)	(0.239)	(0.292)	(0.199)	(0.247)	(0.335)
Covariates		√			√			√	
Cognitive model		√			√			√	
Measurement model		√			√			√	
N		2363			1384			979	
AIC		459420			218306.3			237984.7	
BIC		460008.3			218813.8			238458.7	

Note: $*p \leq 0.1$, $**p \leq 0.05$, $***p \leq 0.01$. The table presents the estimated structural parameters from three different specifications of the educational decision model, corresponding to Equation (4) in the paper. D1, D2 and D3 represent the three educational decisions separately: post-compulsory schooling, undergraduate education, and postgraduate education. Model (1), which uses the full sample of both males and females, represents our main results. Model (2) includes only female participants, while Model (3) includes only male participants. Comparing the coefficients across the three models allows us to examine gender heterogeneity in the effects of abilities on educational choices. The measurement and structural models are estimated simultaneously, all standard errors are estimated appropriately. For binary and categorical independent variables, the reference categories are indicated. All reported statistics are weighted using IPW. For Models (2) and (3), the IPW is constructed separately by gender.

5 Conclusions

Educational choice, though a key predictor of later occupations, primarily reflects individual preferences and marks the first major career-related decision made before entering the labour market or starting a family (Oguzoglu and Ozbeklik, 2016; Humlum et al., 2019). The purpose of this paper is to investigate the selection effect of early cognitive and non-cognitive abilities on three sequential educational decisions made after completing compulsory education in Britain, using data from the BCS70. These selection effects refer to the fact that individuals with different abilities may have different preferences regarding educational decisions.

Consistent with the existing literature (see e.g. Frederick, 2005; Heckman et al., 2018), our findings show a positive selection effect of early cognitive abilities on individual educational decisions. The direct influence of post-compulsory school cognitive ability persists in all three educational decisions, while the effect of preschool cognitive ability is insignificant, holding other factors constant. While the importance of early cognition is widely recognised, the periods during child development during which cognitive ability has greatest influence have not, as yet, been investigated. Our findings help to fill this research gap by confirming that, after controlling for preschool cognitive ability, the selection effect of post-compulsory school cognitive ability on educational decisions remains substantive and significant. This may be because post-compulsory education cognitive ability is temporally closer to the time when individuals make decisions about additional education. As cognitive abilities dynamically change, individuals are able to update their capabilities and as a result, are more likely to base decisions on their current cognitive state rather than on a past one.

Moreover, our findings show the importance of non-cognitive ability for educational decision-making, alongside cognitive ability. This is consistent with the literature. For instance, Heckman et al. (2006) found that improvements in non-cognitive abilities such as self-control and self-esteem, significantly increase the likelihood of completing a four-year college degree, even when controlling for cognitive abilities. They emphasise that non-cognitive skills, including perseverance and social skills, are crucial to success in the labour market and are comparable to cognitive abilities in predicting educational attainment and employment outcomes. Furthermore, we find the effect of non-cognitive ability increases with the level of education. Kautz et al. (2014) state that non-cognitive skills tend to be malleable during adolescence, indicating that they can be enhanced through targeted interventions. In addition to early abilities, our findings also confirm the importance of the family environment in shaping educational choices.

A great deal of literature has emphasised the importance of early investments, especially for infancy and early childhood (see e.g. Cunha and Heckman, 2008; Ozawa et al., 2022). Feinstein and Bynner (2004) reports that changes in mid-childhood strongly impact early adult outcomes - even more so than the effects of cognitive development before age five. Our finding from the early cognitive development model is that preschool cognition and non-cognitive skills together with family background govern the development of post-compulsory school cognition. Accordingly, post-compulsory school cognitive ability may act as a mediator for the influence of preschool cognitive ability on educational decisions. Our findings suggest that children's cognitive development during middle and later childhood also warrants greater attention.

The literature suggests that, in contrast to cognitive abilities, non-cognitive skills exert a more pronounced influence on women's educational choices (Lavy and Schlosser, 2011; Buchmann and DiPrete, 2006; Heckman et al., 2006). Furthermore, women are generally more cautious in selecting academic pathways and are more likely to engage in continuous learning compared to men (Buchmann and DiPrete, 2006). Although the overall results suggest that gender differences in educational preferences are not pronounced, our heterogeneity analysis reveals that the influence of early abilities on educational choices varies by gender. Specifically, females' educational decisions are more strongly associated with both preschool cognitive ability and non-cognitive ability at age 10, while males are more influenced by post-compulsory school cognitive ability. Notably, early non-cognitive ability plays a particularly important role in males' decisions to pursue postgraduate education.

The range of educational decisions people make after compulsory education is linked to subsequent educational attainment and later adult outcomes. Understanding the determinants of the educational decision-making process can support policymakers in designing more effective policies to influence the take-up of additional educational opportunities. Our findings also underscore the need for active intervention in early cognitive development and highlight the significance of compulsory education policy.

Acknowledgements

The Economic and Social Research Council funds the Centre for Longitudinal Studies (CLS) Resource Centre (ES/W013142/1) which provides core support for the CLS cohort studies. While the CLS Resource Centre makes these data available, CLS does not bear any responsibility for the analysis or interpretation of these data by researchers(University College London, 2024). The research data are distributed by the UK Data Service. The CLS cohorts are only possible due to the commitment and enthusiasm of their participants, their time and contribution is gratefully acknowledged.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We are grateful for comments from Gurleen Popli, Jack Britton and participants at the Health, Econometrics, and Data Group (HEDG) Seminar. For the purpose of open access, a Creative Commons Attribution (CC BY)

licence is applied to any Author Accepted Manuscript version arising from this submission.

Appendix

A.1 IPW estimation

The data we use span a relatively long period, from birth to age 46. Inevitably, each wave of the survey suffers from some form of attrition. Some of these dropouts may directly or indirectly influence educational decisions, for instance, through early death due to health problems or parental separation. In addition, we exclude observations with item non-response for any of the variables used in the education models. As a result, our final analytical sample is only about one-fifteenth the size of the original sample. Compared to the full sample, the long-term "survivors" tend to have higher average levels of education and better ability test scores. Failing to account for non-response may therefore lead to biased estimates of the relationship between educational choices and cognitive ability. To address this issue, we apply Inverse Probability Weighting (IPW) to our model estimation (Robins et al., 1995; Fitzgerald et al., 1998; Wooldridge, 2002).

Following Jones et al. (2013) [see chapter 10, page 277], our IPW approach includes an additional set of observed variables, denoted as z. This relies on selection on observables and implies that non-response is ignorable conditional on z (Fitzgerald et al., 1998; Wooldridge, 2002). The selection-on-observables assumption requires that z includes variables that both predict non-response and are correlated with the outcome of interest, but are excluded from the main education model by design.

The original sample includes individuals who provided complete interviews and usable data. We separate the observations into two groups: response and non-response. Non-response is defined as failure to provide valid observations for the education models. To compute the IPW estimates, we estimate a binary response model, conditional on a vector of characteristics z measured in the early waves. In practice, z includes all covariates used in the GSEM equations, except for family income at age 16, which was excluded due to large missing values. In addition, z includes the number of previous pregnancies of the respondent's mother (a continuous variable ranging from 0 to 17) and region indicators (a categorical variable with 11 categories) from the initial wave.

Relating to the number of previous pregnancies, if the parents already had other children, the share of family resources available to the newborn—such as parental attention and financial support—might be reduced. This could influence the child's cognitive development and subsequent educational decisions. If we are estimating a structural model, then from the perspective that the number of siblings influences the allocation of family resources—and subsequently affects cognitive development and educational outcomes — the number of siblings at age 5 seems more appropriate than the number of pregnancies the mother had by the time of birth. Hence, the latter is reserved for estimating IPWs only. From the probit results (see Table A.1), we find that conditional on number of siblings, the coefficient of number of previous pregnancies remains significant. This may reflect that drop-out may be driven less by family resource allocation and more by maternal health conditions. After controlling for maternal age at childbirth, a higher number of pregnancies by the same age likely indicates a heavier physical burden, which might increase the probability of drop-out. As for region, it is mainly included to account for potential regional disparities in access to educational resources. This variable also appears in the list of controls in the Heckman et al. (2018), although we later removed it from the GSEM due to its lack of significance.

Figure A.1 presents the distribution of the predicted propensity scores for the responders and non-responders groups. The distribution curve is relatively smooth, with good overlap across groups and without any extreme values. This suggests that our weights are relatively reasonable.

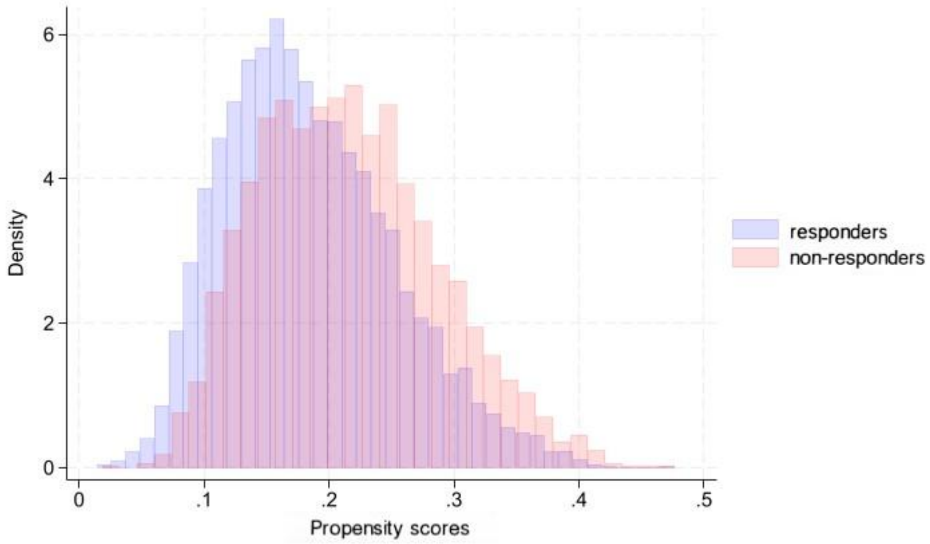

Figure A.1: IPW: distribution of propensity score for responders and non-responders groups

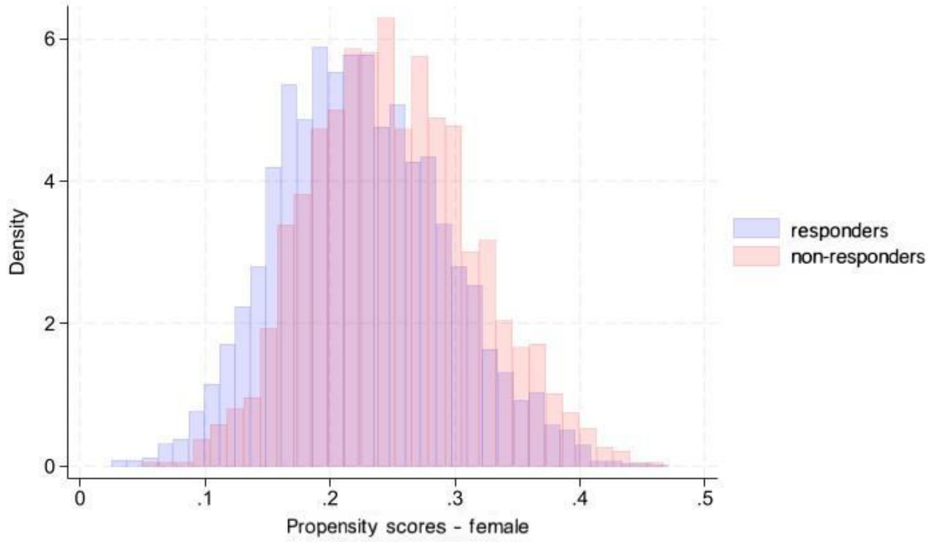

Figure A.2: histogram graph of propensity score - ipwfemale

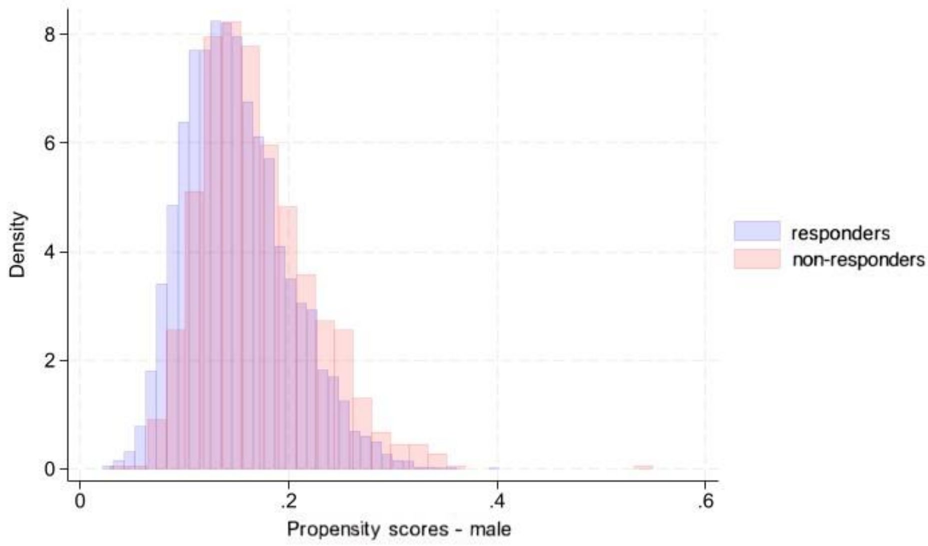

Figure A.3: histogram graph of propensity score - ipwmale

A.2 Descriptive analysis of abilities

There are three latent variables that need to be estimated: preschool cognitive ability at age five, post-compulsory school cognitive ability at age 16, and non-cognitive ability measured at age 10. The BCS70 offers multiple measurements for each latent variable, which we have introduced in Sections 3.3 and 3.5. Tables A.1 and A.2 shows the pairwise correlations between cognitive test scores, and Table A.3 displays the correlation between non-cognitive measurements[20]. We find a positive correlation between all the measures. On average the correlations between measures for post-compulsory school cognitive ability are larger than those for age 5 cognitive and non-cognitive ability. The correlation between the shortened Edinburgh Reading Test and the vocabulary test especially is 0.757. This reveals that the measurements at age 16 have some overlap, but still contain idiosyncratic information on cognitive ability.

Table A.1: The correlation between cognitive test scores at age five

	cd5	cp5	epvt5	hfd5	srt5
Copying designs test (cd5)	1				
Complete a profile test (cp5)	0.153	1			
English picture vocabulary test (epvt5)	0.201	0.112	1		
Human figure drawing test (hfd5)	0.270	0.224	0.110	1	
Shortened Edinburgh reading test (srt5)	0.213	0.050	0.085	0.123	1

Source: the BCS70 wave 2.

Table A.2: The correlation between cognitive test scores at age sixteen

	srt16	m16	vt16	at16	st16
Shortened Edinburgh reading test (srt16)	1				
BAS - matrices test (m 16)	0.504	1			
Vocabulary test (vt16)	0.757	0.387	1		
Arithmetic test (at16)	0.679	0.503	0.635	1	
Spelling test (st16)	0.559	0.369	0.567	0.543	1

Source: the BCS70 wave 4.

We estimate latent variables using the measurement model[21]. Table A.4 presents the estimated factor loading which indicate the impact of the latent variable on the related measure. Results reveals that all measures are loaded positively and significantly

[20] A preliminary factor analysis has been performed on these measurements, while the Velicer (1976) minimum average partial correlation criterion is suggested to retain one component.

[21] In our sample, we only require participants to complete at least one cognitive ability test. Hence, some observations may contain missing values in some cognitive tests. The SEM approach by default applies an equation-wise deletion approach for models with continuous latent variables, that allows estimating all observations even with missing values, while the traditional SEM approach requires no missing values in the sample.

Table A.3: The correlation between non-cognitive measures at the age ten

	Loc	Pes	Cop	Com	Att	Pet
Locus of control scale (Loc)	1					
Perseverance scale (Pes)	0.251	1				
Cooperativeness scale (Cop)	0.136	0.393	1			
Completeness scale (Com)	0.136	0.581	0.293	1		
Attentiveness scale (Att)	0.221	0.624	0.351	0.593	1	
Persistence scale (Pet)	0.212	0.713	0.326	0.564	0.617	1

Source: the BCS70 wave 3.

at the 1% significance level. The factor loading of the copying designs test, shortened Edinburgh Reading Test and locus of control scale is constrained to equal one (anchoring). The magnitudes of the standardised loading of most remaining measures are above 0.50, which indicates that each measure is a significant indicator of its underlying construct.

To assess how much variability there is in cognitive performance between ages 5 and 16, we group cognitive ability in each period by quartile so that individuals are each classified into one of the four quartile groups. By cross-tabulating the quartile groups across the two periods, we obtain a quartile transition matrix to which we assign labels to each group based on cognitive ranking changes (See Table A.5) (Feinstein and Bynner, 2004). Table A.6 displays summary statistics (proportions) across socio-economic statius for the five groups. We find that the low-low group has the lowest average early mother's education and family income, while the high-high group correlates with socio-economic advantage. Early family background is clearly associated with cognitive development. Comparing Groups 1 (low-low) and 2 (escapers), we find that it is children with more educated parents and higher family income who are able to overcome early cognitive developmental disadvantage and catch-up in adolescence. By comparing Group 4 (fallers) with the other groups (Groups 3, 5), we find that children with less-educated parents and slightly lower family incomes appear more likely to fall behind in adolescent cognitive development, even if they gain an advantage in early cognitive development. Since the number of escapers (Group 2) is close to twice the number of fallers (Group 4), the distribution of cognitive ability changes from normal to skewed in Figure 3.

Table A.4: Predicted latent abilities: Loading of measurement models

	Cognitive ability at age 5		Cognitive ability at age 16		Non-cognitive ability	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Copying designs test	1	constrained				
Complete a profile test	0.430***	(0.047)				
English picture vocabulary test	0.738***	(0.077)				
Human figure drawing test	0.761***	(0.049)				
Shortened Edinburgh reading test	0.520***	(0.080)				
Shortened Edinburgh reading test			1	constrained		
BAS - matrices test			0.632***	(0.046)		
Arithmetic test			0.854***	(0.034)		
Spelling test			0.627***	(0.041)		
Vocabulary test			0.808***	(0.030)		
Locus of control scale					1	constrained
Perseverance scale					3.092***	(0.267)
Cooperativeness scale					1.568***	(0.162)
Completeness scale					2.622***	(0.238)
Attentiveness scale					2.789***	(0.242)
Persistence scale					2.930***	(0.255)

Note: $^{***}p \leq 0.01$. The table reports the estimated parameters from three ability measurement models. All test scores were standardised prior to estimation. As the measurement and structural models are estimated simultaneously, no additional adjustment to the standard errors is required.

Table A.5: Definition of groups indicating change or persistence in childhood

<table>
<tr><td></td><td colspan="4">Quartile at age 16</td></tr>
<tr><td>Quartile at age 5</td><td>1 (Low)</td><td>2</td><td>3</td><td>4 (High)</td></tr>
<tr><td>1 (Low)</td><td>Group 1 (Low-Low)</td><td colspan="3">Group 2 (Escapers)</td></tr>
<tr><td>2</td><td colspan="4">Group 3 (Median)</td></tr>
<tr><td>3</td><td colspan="4">Group 3 (Median)</td></tr>
<tr><td>4 (High)</td><td colspan="2">Group 4 (Fallers)</td><td colspan="2">Group 5 (High-High)</td></tr>
</table>

Note: This table references the design of Feinstein and Bynner (2004).

Table A.6: The number of observations in each group (5 and 16)

	Gender	Parental education (5)	Number of siblings (5)	Family income (16)	N	%
Group 1 (Low-Low)	0.46	1.30	2.41	1.51	259	10.96%
Group 2 (Escapers)	0.42	1.59	2.33	1.80	332	14.05%
Group 3 (Median)	0.40	1.71	2.22	1.88	1182	50.02%
Group 4 (Fallers)	0.43	1.65	2.26	1.85	169	7.15%
Group 5 (High-High)	0.40	2.05	2.14	2.11	421	17.82%

A.3 Supplementary tables

Table A.7: Comparing the distributions of control variables in sample and in original dataset

Variable	Sample	Original wave	T statistic	p value
Initial birth conditions				
Mother's age at birth				
mean	26.04	25.97	0.60	0.55
N	2363	13135		
Birth weight				
mean	3.34	3.27	6.00	0.00
N	2363	13135		
Gender				
mean	0.41	0.52	-10.43	0.00
female	1384 (59%)	6327 (48%)		
male	979 (41%)	6808 (52%)		
N	2363	13135		
Early family circumstances				
Number of siblings at age 5				
mean	2.24	2.31	-5.21	0.00
none	258 (11%)	1352 (10%)		
one sibling	1275 (54%)	6378 (49%)		
two or more siblings	830 (35%)	5405 (41%)		
N	2363	13135		
Mother's education at age 5				
mean	1.71	1.57	9.04	0.00
no qualification	1080 (46%)	7090 (55%)		
below A level	898 (38%)	4248 (33%)		
A level and above	385 (16%)	1535 (12%)		
N	2363	12873		
Family income at age 16				
mean	1.87	1.78	6.53	0.00
low-income group	692 (29%)	2639 (37%)		
medium-income group	1289 (55%)	3523 (49%)		
high-income group	382 (16%)	1023 (14%)		
N	2363	7185		

Source: the BCS70 wave 1, 2, and 4. "Original wave" refers to the raw BCS70 data with non-response cases removed. T tests showed that the mean of controls in sample group was significantly different from the mean of controls in population group (except for mother's age at birth).

Table A.8: Descriptive statistics of cognitive test scores

	N	Mean	S.d.	Min	Max
Cognitive ability tests at age 5					
Copying designs test	2362	5.07	1.93	0	8
Complete a profile test	2292	7.16	3.93	0	16
English picture vocabulary test	1796	34.76	8.72	6	51
Human figure drawing test	2341	10.77	3.03	1	21
Shortened Edinburgh reading test	1175	4.08	6.21	0	49
Cognitive ability tests at age 16					
Shortened Edinburgh reading test	1097	56.58	12.37	14	75
BAS matrices test	1130	9.00	1.62	1	11
Arithmetic test	1361	38.18	11.37	0	60
Spelling test	2230	164.81	26.42	0	198
Vocabulary test	2212	43.92	12.43	0	75

Source: the BCS70 wave 2 and 4.

Table A.9: Descriptive statistics of non-cognitive test scores

	N	Mean	S.d.	Min	Max
Locus of control scale	2363	7.62	2.92	0	15
Perseverance scale	2307	30.93	10.78	1	47
Cooperativeness scale	2331	32.73	8.68	1	47
Completeness scale	2331	35.40	12.58	1	47
Attentiveness scale	2332	34.43	12.17	1	47
Persistence scale	2340	30.87	13.03	1	47

Source: the BCS70 wave 3.

Table A.10: Female: Predicted latent abilities: Loading of measurement models

	Cognitive ability at age 5		Cognitive ability at age 16		Non-cognitive ability	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Copying designs test	1	constrained				
Complete a profile test	0.363***	(0.060)				
English picture vocabulary test	0.929***	(0.118)				
Human figure drawing test	0.781***	(0.058)				
Shortened Edinburgh reading test	0.625***	(0.106)				
Shortened Edinburgh reading test			1	constrained		
BAS - matrices test			0.607***	(0.062)		
Arithmetic test			0.803***	(0.039)		
Spelling test			0.539***	(0.046)		
Vocabulary test			0.863***	(0.035)		
Locus of control scale					1	constrained
Perseverance scale					2.817***	(0.298)
Cooperativeness scale					1.476***	(0.189)
Completeness scale					1.938***	(0.233)
Attentiveness scale					2.383***	(0.265)
Persistence scale					2.485***	(0.285)

Note: ***$p \leq 0.01$. The table reports the estimated parameters from three ability measurement models. All test scores were standardised prior to estimation. As the measurement and structural models are estimated simultaneously, no additional adjustment to the standard errors is required.

Table A.11: Female: Results of cognitive development models

	Cognitive ability at age 5		Cognitive ability at age 16	
	Coef.	Std. Err.	Coef.	Std. Err.
Preschool cognitive ability (age 5)			0.723***	(0.101)
Non-cognitive ability (age 10)			1.100***	(0.137)
Mother's age at delivery	0.024***	(0.004)		
Birth weight	0.163***	(0.044)		
Number of siblings at age 5 (baseline = no sibling)				
one sibling	-0.045	(0.065)	0.091	(0.091)
two or more siblings	-0.328***	(0.067)	-0.023	(0.100)
Mother's education at age 5 (baseline = no qualification)				
below A level	0.222***	(0.047)	0.004	(0.062)
A level and above	0.500***	(0.065)	0.101	(0.089)
Father interaction term	-0.004	(0,042)	-0.061	(0.050)
Family income at age 16 (baseline = low-income)				
middle-income			0.077	(0.065)
high-income			0.293***	(0.083)

Note: $**p \leq 0.05$, $***p \leq 0.01$. The table presents the estimated structural parameters from the cognitive development model for *female participants*, corresponding to Equations 4 and 5. Since the measurement and structural models are estimated simultaneously, all standard errors are estimated appropriately. For binary and categorical independent variables, the reference categories are indicated. All reported statistics are weighted using IPW defined by *female*. For respondents missing mother's education data, we use the father's highest education level as a substitute. To test whether father's education has a different effect, we add an interaction term which interacts the father indicator with mother's education.

Table A.12: Female: Results of sequential educational decision models

	Post-compulsory schooling D1		Undergraduate education D2		Postgraduate education D3	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Preschool cognitive ability (age 5)	0.297**	(0.130)	0.453**	(0.180)	0.033	(0.195)
Post-compulsory school cognitive ability (age 16)	0.283***	(0.076)	0.073	(0.097)	0.245*	(0.144)
Non-cognitive ability (age 10)	0.543***	(0.182)	0.498**	(0.239)	0.594**	(0.292)
Number of siblings (baseline = no sibling)						
one sibling	-0.018	(0.127)	0.112	(0.160)	0.325*	(0.197)
two or more siblings	-0.116	(0.135)	0.204	(0.171)	0.149	(0.215)
Mother's education (baseline = no qualification)						
below A level	0.320***	(0.088)	0.038	(0.120)	0.029	(0.142)
A level and above	0.461***	(0.137)	0.325*	(0.172)	0.551***	(0.178)
Father interaction term	-0.086	(0.089)	-0.296***	(0.110)	-0.088	(0.156)
Family income (baseline = low-income)						
middle-income	-0.043	(0.091)	0.018	(0.126)	0.006	(0.153)
high-income	0.289**	(0.137)	0.105	(0.169)	0.006	(0.193)

Note: *$p \leq 0.1$, **$p \leq 0.05$, ***$p \leq 0.01$. The table presents the estimated structural parameters from the dynamic educational decision model for *female participants*, corresponding to Equation 2. Since the measurement and structural models are estimated simultaneously, all standard errors are estimated appropriately. For binary and categorical independent variables, the reference categories are indicated. All reported statistics are weighted using IPW defined by *female*. N:1384; AIC: 218306.3; BIC: 218813.8. For respondents missing mother's education data, we use the father's highest education level as a substitute. To test whether father's education has a different effect, we add an interaction term which interacts the father indicator with mother's education.

Table A.13: Male: Predicted latent abilities: Loading of measurement models

	Cognitive ability at age 5		Cognitive ability at age 16		Non-cognitive ability	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Copying designs test	1	constrained				
Complete a profile test	0.507***	(0.081)				
English picture vocabulary test	0.677***	(0.108)				
Human figure drawing test	0.765***	(0.081)				
Shortened Edinburgh reading test	0.454***	(0.134)				
Shortened Edinburgh reading test			1	constrained		
BAS - matrices test			0.622***	(0.061)		
Arithmetic test			0.873***	(0.048)		
Spelling test			0.657***	(0.060)		
Vocabulary test			0.770***	(0.042)		
Locus of control scale					1	constrained
Perseverance scale					2.880***	(0.324)
Cooperativeness scale					1.428**	(0.200)
Completeness scale					2.708***	(0.307)
Attentiveness scale					2.702***	(0.300)
Persistence scale					2.805***	(0.308)

Note: $^{***}p \leq 0.01$. The table reports the estimated parameters from three ability measurement models. All test scores were standardised prior to estimation. As the measurement and structural models are estimated simultaneously, no additional adjustment to the standard errors is required.

Table A.14: Male: Results of cognitive development models

	Cognitive ability at age 5		Cognitive ability at age 16	
	Coef.	Std. Err.	Coef.	Std. Err.
Preschool cognitive ability (age 5)			0.4575***	(0.089)
Non-cognitive ability (age 10)			1.384***	(0.166)
Mother's age at delivery	0.022***	(0.006)		
Birth weight	0.228***	(0.052)		
Number of siblings at age 5 (baseline = no sibling)				
one sibling	0.051	(0.087)	-0.270***	(0.103)
two or more siblings	-0.290***	(0.094)	-0.415***	(0.111)
Mother's education at age 5 (baseline = no qualification)				
below A level	0.296***	(0.062)	0.262***	(0.080)
A level and above	0.414***	(0.085)	0.496***	(0.104)
Father interaction term	0.096	(0.074)	-0.089	(0.085)
Family income at age 16 (baseline = low-income)				
middle-income			0.220**	(0.089)
high-income			0.353***	(0.115)

Note: $^{**}p \leq 0.05$, $^{***}p \leq 0.01$. The table presents the estimated structural parameters from the cognitive development model for *male participants*, corresponding to Equations 4 and 5. Since the measurement and structural models are estimated simultaneously, all standard errors are estimated appropriately. For binary and categorical independent variables, the reference categories are indicated. All reported statistics are weighted using IPW defined by *male*. For respondents missing mother's education data, we use the father's highest education level as a substitute. To test whether father's education has a different effect, we add an interaction term which interacts the father indicator with mother's education.

Table A.15: Male: Results of sequential educational decision models

	Post-compulsory schooling D1		Undergraduate education D2		Postgraduate education D3	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Preschool cognitive ability (age 5)	0.029	(0.119)	0.017	(0.151)	-0.177	(0.182)
Post-compulsory school cognitive ability (age 16)	0.439***	(0.088)	0.222**	(0.108)	0.121	(0.127)
Non-cognitive ability (age 10)	0.062	(0.199)	0.222	(0.247)	0.995***	(0.335)
Number of siblings (baseline = no sibling)						
one sibling	-0.137	(0.165)	0.025	(0.201)	-0.459*	(0.237)
two or more siblings	-0.349**	(0.172)	-0.069	(0.212)	-0.4071	(0.259)
Mother's education (baseline = no qualification)						
below A level	0.201*	(0.107)	0.257*	(0.137)	0.276	(0.185)
A level and above	0.353**	(0.157)	0.455**	(0.190)	0.454**	(0.224)
Father interaction term	0.121	(0.118)	-0.302**	(0.119)	-0.167	(0.179)
Family income (baseline = low-income)						
middle-income	0.097	(0.113)	-0.023	(0.149)	-0.251	(0.198)
high-income	0.064	(0.159)	0.123	(0.210)	0.092	(0.237)

Note: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$. The table presents the estimated structural parameters from the dynamic educational decision model for *male participants*, corresponding to Equation 2. Since the measurement and structural models are estimated simultaneously, all standard errors are estimated appropriately. For binary and categorical independent variables, the reference categories are indicated.
All reported statistics are weighted using IPW defined by *male*. N:979; AIC: 237984.7; BIC: 238458.7. For respondents missing mother's education data, we use the father's highest education level as a substitute. To test whether father's education has a different effect, we add an interaction term which interacts the father indicator with mother's education.

References

Almlund, M., Duckworth, A. L., Heckman, J., and Kautz, T. (2011). Personality psychology and economics. In *Handbook of the Economics of Education*, volume 4, pages 1–181. Elsevier.

Azmitia, M. and Hesser, J. (1993). Why siblings are important agents of cognitive development: A comparison of siblings and peers. *Child Development*, 64(2):430–444.

Barone, C. (2011). Some things never change: Gender segregation in higher education across eight nations and three decades. *Sociology of education*, 84(2):157–176.

Bratti, M. (2007). Parents' income and children's school drop-out at 16 in england and wales: evidence from the 1970 british cohort study. *Review of Economics of the Household*, 5:15–40.

Brunello, G. and Schlotter, M. (2011). Non-cognitive skills and personality traits: Labour market relevance and their development in education & training systems. IZA Discussion Paper 5743, IZA Institute of Labor Economics.

Buchmann, C. and DiPrete, T. A. (2006). The growing female advantage in college completion: The role of family background and academic achievement. *American Sociological Review*, 71(4):515–541.

Centre for Longitudinal Studies (2011). Highest educational qualification in ncds and bcs70. `https://cls.ucl.ac.uk/wp-content/uploads/2018/06/Deriving-highest-qualification-in-NCDS-and-BCS70.pdf`. Accessed: 2024-11-14.

Coneus, K. and Laucht, M. (2014). The effect of early noncognitive skills on social outcomes in adolescence. *Education Economics*, 22(2):112–140.

Conti, G., Heckman, J., and Urzua, S. (2010). The education-health gradient. *American Economic Review*, 100(2):234–238.

Cunha, F. and Heckman, J. J. (2007). Identifying and estimating the distributions of ex post and ex ante returns to schooling. *Labour Economics*, 14(6):870–893.

Cunha, F. and Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, 43(4):738–782.

Dai, X. and Heckman, J. J. (2013). Older siblings' contributions to young child's cognitive skills. *Economic Modelling*, 35:235–248.

De Graaf, P. M. and Huinink, J. J. (1992). Trends in measured and unmeasured effects of family background on educational attainment and occupational status in the federal republic of germany. *Social Science Research*, 21(1):84–112.

Dickerson, A. and Popli, G. K. (2016). Persistent poverty and children's cognitive development: evidence from the uk millennium cohort study. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 179(2):535–558.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2018). On the relationship between cognitive ability and risk preference. *Journal of Economic Perspectives*, 32(2):115–134.

Ermisch, J. and Francesconi, M. (2001). Family matters: Impacts of family background on educational attainments. *Economica*, 68(270):137–156.

Favaro, D. and Sciulli, D. (2022). Female long-term labour market outcomes: the role of early-life abilities and education. *Applied Economics*, 54(3):340–353.

Feinstein, L. and Bynner, J. (2004). The importance of cognitive development in middle childhood for adulthood socioeconomic status, mental health, and problem behavior. *Child Development*, 75(5):1329–1339.

Fitzgerald, J., Gottschalk, P., and Moffitt, R. A. (1998). An analysis of sample attrition in panel data: The michigan panel study of income dynamics. NBER Working Paper 0220, National Bureau of Economic Research.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4):25–42.

Ganzach, Y. (2000). Parents' education, cognitive ability, educational expectations and educational attainment: Interactive effects. *British Journal of Educational Psychology*, 70(3):419–441.

Gerber, T. P. and Cheung, S. Y. (2008). Horizontal stratification in postsecondary education: Forms, explanations, and implications. *Annual Review of Sociology*, 34(1):299–318.

Glewwe, P., Huang, Q., and Park, A. (2017). Cognitive skills, noncognitive skills, and school-to-work transitions in rural china. *Journal of Economic Behavior & Organization*, 134:141–164.

Glick, P. and Sahn, D. E. (2010). Early academic performance, grade repetition, and school attainment in senegal: A panel data analysis. *The World Bank Economic Review*, 24(1):93–120.

González, L., Cortés-Sancho, R., Murcia, M., Ballester, F., Rebagliato, M., and Rodríguez-Bernal, C. L. (2020). The role of parental social class, education and unemployment on child cognitive development. *Gaceta Sanitaria*, 34:51–60.

Harris, D. (1940). Factors affecting college grades: a review of the literature, 1930-1937. *Psychological Bulletin*, 37(3):125.

Hawkes, D. and Joshi, H. (2012). Age at motherhood and child development: Evidence from the uk millennium cohort. *National Institute Economic Review*, 222:R52–R66.

Heckman, J. and Vytlacil, E. (2001). Identifying the role of cognitive ability in explaining the level of and change in the return to schooling. *Review of Economics and Statistics*, 83(1):1–12.

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(S1):S197–S246.

Heckman, J. J. and Navarro, S. (2007). Dynamic discrete choice and dynamic treatment effects. *Journal of Econometrics*, 136(2):341–396.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.

Hegna, K. and Smette, I. (2017). Parental influence in educational decisions: Young people's perspectives. *British Journal of Sociology of Education*, 38(8):1111–1124.

House of Commons Library (2024). Research briefing: Sn07096. `https://commonslibrary.parliament.uk/research-briefings/sn07096/`. Accessed: 2024-11-14.

Hoyle, R. H. (2012). *Handbook of structural equation modeling.* Guilford press.

Humlum, M. K., Nandrup, A. B., and Smith, N. (2019). Closing or reproducing the gender gap? parental transmission, social norms and education choice. *Journal of Population Economics*, 32(2):455–500.

Jensen, A. C. and McHale, S. M. (2015). What makes siblings different? the development of sibling differences in academic achievement and interests. *Journal of Family Psychology*, 29(3):469.

Jones, A. M., Rice, N., d'Uva, T. B., and Balia, S. (2013). *Applied health economics.* Routledge.

Karwath, C., Relikowski, I., and Schmitt, M. (2014). Sibling structure and educational achievement: how do the number of siblings, birth order, and birth spacing affect children's vocabulary competences? *Zeitschrift für Familienforschung*, 26(3):372–396.

Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., and Borghans, L. (2014). Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success. NBER Working Paper 20749, National Bureau of Economic Research.

Khanam, R. and Nghiem, S. (2016). Family income and child cognitive and noncognitive development in australia: Does money matter? *Demography*, 53(3):597–621.

Lavy, V. and Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 3(2):1–33.

Looker, E. D. (1997). In search of credentials: factors affecting young adults' participation in postsecondary education. *Canadian Journal of Higher Education*, 27(2/3):1–36.

Lunn, A. and Kornrich, S. (2018). Family investments in education during periods of economic uncertainty: Evidence from the great recession. *Sociological Perspectives*, 61(1):145–163.

Michalos, A. C. (2014). *Encyclopedia of quality of life and well-being research*, volume 171. Springer Netherlands Dordrecht.

Moulton, V., McElroy, E., Richards, M., Fitzsimons, E., Northstone, K., Conti, G., Ploubidis, G. B., Sullivan, A., and O'Neill, D. (2020). A guide to the cognitive measures in five british birth cohort studies. https://discovery.ucl.ac.uk/id/eprint/10112765/1/250820-Guide-to-cognitive-measures-in-five-British-birth-cohorts.pdf.

Nakamuro, M., Uzuki, Y., and Inui, T. (2013). The effects of birth weight: Does fetal origin really matter for long-run outcomes? *Economics Letters*, 121(1):53–58.

Oguzoglu, U. and Ozbeklik, S. (2016). Like father, like daughter (unless there is a son): Sibling sex composition and women's stem major choice in college. Technical report, IZA Discussion Papers.

Ozawa, S., Laing, S. K., Higgins, C. R., Yemeke, T. T., Park, C. C., Carlson, R., Ko, Y. E., Guterman, L. B., and Omer, S. B. (2022). Educational and economic returns to cognitive ability in low-and middle-income countries: A systematic review. *World Development*, 149:105668.

Pfeffer, F. T. (2018). Growing wealth gaps in education. *Demography*, 55:1033–1068.

Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association*, 90(429):106–121.

Ryberg, R. (2018). Positive self-concept predicts youth staying in school longer in india. *Advances in Life Course Research*, 37:1–14.

Schady, N. (2011). Parents' education, mothers' vocabulary, and cognitive development in early childhood: Longitudinal evidence from ecuador. *American Journal of Public Health*, 101(12):2299–2307.

Taubman, P. (1989). Role of parental income in educational attainment. *The American Economic Review*, 79(2):57–61.

Todd, P. E. and Wolpin, K. I. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital*, 1(1):91–136.

University College London (2017). Bcs70 16-year data- work and caraloc. `https://cls.ucl.ac.uk/wp-content/uploads/2017/07/CARALOC_LAWSEQ.pdf`. Accessed: 2024-11-14.

University College London, UCL Social Research Institute, C. f. L. S. (2024). 1970 british cohort study. [data series]. 11th release. *UK Data Service. SN: 200001, DOI: http://doi.org/10.5255/UKDA-Series-200001.*

Van der Vleuten, M., Jaspers, E., Maas, I., and van der Lippe, T. (2016). Boys' and girls' educational choices in secondary education. the role of gender ideology. *Educational Studies*, 42(2):181–200.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41:321–327.

Weisbrod, B. A. (1962). Education and investment in human capital. *Journal of Political Economy*, 70(5, Part 2):106–123.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91(3):461.

Wilson, K. (2001). The determinants of educational attainment: Modeling and estimating the human capital model and education production functions. *Southern Economic Journal*, 67(3):518–551.

Wooldridge, J. M. (2002). Inverse probability weighted m-estimators for sample selection, attrition, and stratification. *Portuguese Economic Journal*, 1(2):117–139.