DEPARTMENT
OF ECONOMICS

Essays in Applied Microeconomics

Nurfatima Jandarova

Thesis submitted for assessment with a view to
obtaining the degree of Doctor of Economics
of the European University Institute

Florence, 21 September 2021

European University Institute
Department of Economics

Essays in Applied Microeconomics

Nurfatima Jandarova

Thesis submitted for assessment with a view to
obtaining the degree of Doctor of Economics
of the European University Institute

Examining Board

Prof. Andrea Ichino, EUI, Supervisor
Prof. Giacomo Calzolari, EUI, Co-Supervisor
Prof. Stephen Machin, London School of Economics
Prof. Giulio Zanella, University of Bologna

Researcher declaration to accompany the submission of written work
Department Economics - Doctoral Programme

I Nurfatima Jandarova certify that I am the author of the work Essays in Applied Microeconomics I have presented for examination for the Ph.D. at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the Code of Ethics in Academic Research issued by the European University Institute (IUE 332/2/10 (CA 297).

The copyright of this work rests with its author. Quotation from it is permitted, provided that full acknowledgement is made. This work may not be reproduced without my prior written consent. This authorisation does not, to the best of my knowledge, infringe the rights of any third party.

I declare that this work consists of 44382 words.

Statement of inclusion of previous work:

I confirm that chapter 3 was jointly co-authored with Johanna Reuter and I contributed 50% of the work.
I confirm that chapter 4 was jointly co-authored with Prof. Aldo Rustichini and Prof. Michele Boldrin and I contributed 45% of the work.

Signature and date:

Nurfatima Jandarova
September 3, 2021

Abstract

This thesis consists of four essays in applied microeconomics. Chapter 1 studies the effects of parental job loss on various outcomes of children and provides new evidence on the heterogeneity of these effects along the cognitive ability distribution of children. I find that higher intelligence score protects children from the negative effects, but only in the long run. In the shorter term, instead of protecting, high intelligence exacerbates the cost of parental unemployment in terms of educational outcomes. This forces high-intelligence children with unemployed parents to start their careers at lower-paying jobs. Nevertheless, they can prove themselves via work performance and switch to better-paying jobs. I also provide suggestive evidence that their lifetime earnings could be higher had they continued their education.

Chapter 2, joint with Michele Boldrin and Aldo Rustichini, studies the relationship between fertility decisions and intelligence. We document that fertility may be negatively associated, at least in advanced societies, with higher intelligence. A possible explanation of the finding is provided in models describing the choice of individuals (in particular women) facing a trade-off between parenthood and career concerns. With positive complementarity between intelligence and effort in education and career advancement, higher intelligence individuals, particularly women, will sacrifice parenthood to education. Thus, current education and labor market policies may be imposing an uneven penalty on more talented women. We test and find support for the model in a large data set for the UK (Understanding Society), using several alternative measures of fertility. Our results provide a new interpretation of the well documented fact in demographic studies that education is negatively associated with fertility: it is not education as an outcome, but as an aspiration that reduces fertility.

Chapter 3 investigates the joint effect of local economic conditions on educational decisions and subsequent labour market outcomes using the instrumental variable approach. I find that adverse economic conditions at age 14 reduce educational attainment, except for the children aiming at university degrees. Second, exposure to a higher unemployment rate at age 14 permanently reduces real hourly wages over the life cycle. The IV estimator suggests that

a year of education lost due to initial economic conditions corresponds to about 8% lower wages at ages 26-30 and 6% lower wages at ages 41-45.

Chapter 4, joint with Johanna Reuter, attempts to differentiate the degree attainment in the UK by type of higher education institutions. Historically higher education in the UK has been shaped by a dual system: elite universities on the one hand and polytechnics and other higher education institutions on the other. Despite the formal equivalence of both degrees, the two institution types faced different financing, target populations, admission procedures and subjects taught. Nevertheless, in survey data they are often indistinguishable. We overcome this problem using a multiple imputation technique in the UKHLS and BHPS datasets. We examine the validity of inference based on imputed values using Monte Carlo simulations. We also verify that the imputed values are consistent with university graduation rates computed using the universe of undergraduate students in the UK.

To my family and Hanno

Acknowledgements

This dissertation would not have been possible without the support of many people.

First of all, I would like to express gratitude to my supervisors Andrea Ichino and Giacomo Calzolari for their guidance and support during this PhD journey. Besides providing feedback on my work and inspiring my research, I especially appreciate that they always found time for me and encouraged me when needed. I would also like to thank Sule Alan, Thomas Crossley, Michèle Belot, Alessandro Tondini, Marco Tabellini, Miguel Urquiola and all members of the EUI Microeconomics Working Group for helpful comments and lively discussions.

Special thanks to the external committee members, Stephen Machin and Giulio Zanella, for their invaluable comments on the thesis.

This thesis has also benefitted greatly from collaboration with Aldo Rustichini. I am looking forward to more fruitful collarborations throughout the post-doc and beyond.

The journey to this thesis has also brought me a lot of new friends that each made this time unforgettable. I cannot imagine my PhD experience without the help and support of Johanna Reuter, who is also my co-author, and Flavia Cavallini. Whether we were diving deep into econometric theory or sharing important moments in life, problems were half as difficult and successes were twice as gratifying. I hope that our future holds more interesting data, exciting projects and delightful moments to share.

While the first year of the program was challenging in many respects, it is also the most memorable because I met my partner, Hanno Kase. He has helped me countless times throughout these years: from summer paper to writing this thesis. Be it a model, data, making sense of codes or preparing for important presentations, I could always count on him. Together with him, the PhD journey was much easier and more joyful. We have shared numerous adventures, built a smart plant watering system and seen half the Europe. I hope that we continue to explore the world and grow both professionally and personally, together.

I am eternally grateful to my mom. From the very early ages, she encouraged me to learn and dream big. She did and continues doing everything, possible or impossible, to give me better chances in life and help achieve my goals. I hope that she lives long and healthy to continue seeing the dreams come true together with me. I am also deeply thankful to my sisters and brother, who are always there for me and believe in me.

I would also like to thank people that influenced my decision to pursue PhD program at the EUI: Madina Kurmangaliyeva, Stephen Machin and Sabit Khakimzhanov.

Finally, none of this would take place without the people of Italy, who have generously funded my PhD program.

Table of contents

Chapter 1

Does intelligence shield children from the effects of parental unemployment?

The topic of intergenerational effects of parental layoff on a variety of children's short- and long-term outcomes has recently received increased attention. The literature typically finds that having an unemployed parent has a negative impact on various educational and labour-market outcomes of children, especially from disadvantaged backgrounds[1]. In this paper, I examine whether higher levels of intelligence can protect children from these negative effects. I provide new evidence on the heterogeneity of the parental unemployment effects on children's outcomes along the distribution of intelligence score of children. Indeed, I find that children with higher ability score suffer less as a result of parental unemployment, but mostly in the long run.

To estimate how the effect of parental unemployment on children's outcomes differs across the ability distribution of children I use the UK Household Longitudinal Study (UKHLS) dataset. UKHLS is the largest panel survey in the UK with information about intelligence score and parents' employment status back when respondents were 14 years old. The estimation strategy relies on a difference-in-differences approach. I carefully examine and test the assumptions necessary for the causal interpretation of the results. The characteristics of the university admission system in the UK, in particular high dependence on prior academic achievements, make it possible for effects of negative shocks experienced during adolescence to go beyond school grades to university degree attainment. Given the positive returns to schooling and university degree[2], the effects of negative shocks at adolescence may propagate

1. Oreopoulos, Page, and Stevens (2008)
2. Gunderson and Oreopolous (2020) and Carneiro, Heckman, and Vytlacil (2011)

into the labour-market outcomes of children. The goal of this paper is to establish whether having high intelligence score helps children mitigate some of the negative effects.

I present two key findings. First, the cost of parental unemployment in terms of educational attainment of children is increasing in the intelligence score of children. Adolescents with unemployed parents experience a further decline in the probability of obtaining a tertiary degree by 1.5 percentage points for every 1 standard deviation rise in their intelligence score. In fact, high intelligence score exacerbates the losses due to parental unemployment. Though surprising, is consistent with dynamic complementarity theory of Cunha and Heckman (2010). In particular, the theory of dynamic complementarity of skills states that the productivity of human capital investments in adolescence depends on cognitive ability in childhood. A relevant implication of this theory is that loss of human capital investments has larger consequences at the higher end of the ability distribution. Consistent with the prediction that only poor households adjust human capital investments in response to income shocks (Mulligan 1997), I show that the above result is a reflection of losses incurred by smarter children from disadvantaged backgrounds.

My second result is that, despite the negative effect on educational achievement, higher intelligence improves the treatment effect of parental unemployment on labour-market outcomes later in life. A 1 standard deviation increase in the intelligence score improves the effect of parental unemployment on the probability of employment and earnings by 2.1 and 18.6 percentage points, respectively. Thus, in the longer term, high intelligence does indeed protect children from the effects of parental job loss during adolescence.

The two findings together suggest that despite initially aggravating the effect on educational attainment, high ability allows children to overcome these disadvantages with time. This result is consistent with employer learning theory[3], which extends a standard signalling model by allowing employers to learn about worker productivity by monitoring their work performance. According to this theory, the role of educational signal in the wage-setting process decreases as worker accumulates experience. A testable implication of this theory is that the effect of parental unemployment on initial labour-market outcomes of children should not vary by intelligence score. When individuals first enter the labour market, employers can only use their educational achievements to form a belief about worker productivity. Since children exposed to parental unemployment at the higher end of the ability distribution fail to obtain a university degree, they are initially unable to distinguish themselves from job candidates with lower ability. Therefore, their first-job characteristics should not depend on their ability. Consistent with this prediction, I find that occupational ranking of children' first

3. Farber and Gibbons (1996); Arcidiacono, Bayer, and Hizmo (2010); Altonji and Pierret (2001)

jobs does not vary with intelligence score. In addition, using a panel of predicted life-cycle wages I show that higher intelligence helps children with unemployed parents to overcome the initial setback at around age 30 onwards.

It is also worth noting that my results appear to contradict an established notion of positive returns to education in the labour market. This might raise a question of whether children with higher intelligence actually need a university education. The answer to this question depends on a careful examination of the benefits and costs of signalling via education vs on-the-job signalling[4], speed of employer learning and productivity dynamics[5]. While it is beyond the scope of my paper, the results could provide some evidence in favour of the educational signalling under two additional assumptions. First, parental unemployment has a negative treatment effect on children at the bottom of the ability distribution; and second, children whose parents were unemployed cannot do better than similar children whose parents stayed employed. Given these assumptions, the results in this paper imply that experience profile of high-ability children whose parents were unemployed is shifted downwards relative to their peers whose parents stayed employed, meaning that their lifetime earnings are also lower.

This paper contributes to a growing literature on the intergenerational effects of parental unemployment. This literature has examined the effect of parental job loss on a variety of educational, labour-market and non-cognitive outcomes of children (for a detailed summary see Table 1.B.2). Majority of the papers find large negative effects on educational outcomes[6], but small or zero effects on labour market outcomes of children[7]. Most of the papers that find zero results use data from the US or Scandinavian countries, which could be viewed as having more supportive and/or less history-dependent institutions. Lindemann and Gangl (2020) provide some evidence that differences in institutional settings matter for the magnitude of the intergenerational effects of parental unemployment. Therefore, the selectivity of the university programs in the UK could contribute to larger average treatment effects making it easier to detect heterogeneity in the effect of parental unemployment on children's outcomes. The literature has also explored heterogeneity of the intergenerational effects across a wide range of characteristics - family income/wealth, family composition, parental education, local labour market conditions, the magnitude of income shock, variation in tuition fees, gender and race - typically concluding that parental job loss imposes a higher cost on children from

4. Alós-Ferrer and Prat (2012)
5. Lange (2007); Kahn and Lange (2014)
6. Peter (2016); Brand and Thomas (2014); Pan and Ost (2014); Coelli (2011); Rege, Telle, and Votruba (2011); Stevens and Schaller (2011); Page, Stevens, and Lindo (2009); Bratberg, Nilsen, and Vaage (2008)
7. Mörk, Sjögren, and Svaleryd (2019); Hilger (2016); Page, Stevens, and Lindo (2009); Bratberg, Nilsen, and Vaage (2008)

disadvantaged backgrounds[8]. I contribute to this literature by examining the heterogeneity of the intergenerational effects of parental unemployment across the ability distribution of children.

The remainder of the paper is outlined as follows. Section 1.1 discusses the university admission system in the UK and how it can explain sizeable effects of parental unemployment on adolescents. I describe the dataset in more detail in section 1.2 and empirical strategy - in section 1.3. I review the main regression results and explore the mechanisms in section 1.4. In section 1.5 I discuss identifying assumption in detail and provide evidence supporting their validity. Finally, I conclude in section 1.6.

1.1 Institutional background

The university education in the UK has been for a very long time elitist and dominated by Oxford and Cambridge. While university sector has significantly expanded in the 1960s and 1990s, the universities in the UK, and more importantly, individual departments within universities, continue to be highly selective towards their applicants (Willetts 2017). The selectivity of university admission means that the applicants must demonstrate good knowledge of the subject they want to study before starting the university program.

Typically, the way students can demonstrate such knowledge is via GCE A-level grades. The A-level exams are subject-specific and students usually sit three or four of them. In principle, students are free to choose any combination of subjects; in reality, the choices are shaped by the entry requirements of the programs they wish to apply to. Students usually study the subjects in-depth for two years before taking the exam. However, the format has been changing over the years. Initially, the A-level exams could be taken at the end of the second year of studies. Between late 1980s and 2000s there was a shift towards so-called modular program, where subject exams are taken at the end of the module. In 1989 a new Advanced Supplementary (AS) exams were introduced, which were equivalent to half of an A-level exam. Under the modular approach, AS-level grades were counted towards an A-level grade. In 2015, the format was changed back to linear program where all A-level exams are taken at the end of the second year and AS-level grades no longer count towards A-level grades. Students usually sit A-level exams at around age 18.

8. Pan and Ost (2014); Coelli (2011); Rege, Telle, and Votruba (2011); Page, Stevens, and Lindo (2009); Oreopoulos, Page, and Stevens (2008)

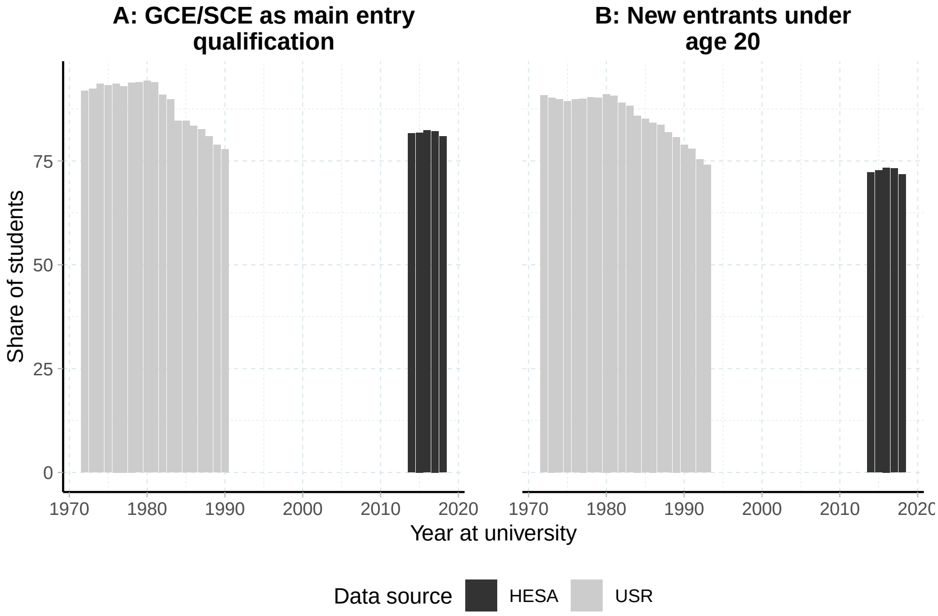
The admission to the programs that prepare for A-level exams often require good grades in GCSE (General Certificate of Secondary Education)[9] exams taken at the end of compulsory school, at age 16. Similarly to A-level exams, GCSEs are also subject-based examinations for which students study in the last two-three years of secondary school. Students usually sit at least 5 GCSE exams in subjects of their choice. Schools are encouraged to provide a pathway that leads to exams in 5 core subjects proposed by the UK government, including Math and English. Universities may also take into account GCSE grades when making admission decisions.

Scotland has its own system of school-leaving qualifications. For most of my analysis sample the relevant qualification is Scottish Certificate of Education (SCE) that was in place during 1962-1999. The SCE had two grades: Ordinary (later Standard) and Higher, which are broadly equivalent to GCSE and AS-levels in timing and importance. Ordinary Grades were typically taken at the age of 16, and Higher Grades - a year later. The main difference with GCSE and GCE is that Scottish qualifications aim at assessing broader knowledge; therefore, the exams were taken for a wider range of subjects. This and the equivalence of Higher Grades to AS level, not A level, exams is mainly a result of the differences in the higher education systems between Sctoland and England. Scottish university programs incorporated one year of foundation courses at the beginning, which allowed the admission criteria to be less specialised (Willetts 2017).

The data confirms that majority of university students enter via main route: passing three or more A-level exams in the specified subjects. Figure 1.1 demonstrates that about 80-90% of all first-time undergraduate students had GCE and/or SCE exam passes as the main entry qualification and were under age 20. Therefore, the suggested timeline of first passing GCSE exams at 16 and GCE A-level exams at 18 is relevant for most of the children considering a university education.

To sum up, the selectivity of the university programs makes the grades in entry qualifications a very important factor. This in turn, translates to selectivity of the places that prepare for A level exams and places a high importance on the qualifications obtained at the end of compulsory school. In addition, GCSE grades may also enter directly into the admission decisions. Both qualifications require an in-depth study of the test subjects in the preceding two or three years. Such selectivity and hierarchy also makes alternative routes of entering

9. Introduced in 1988, replacing the Certificate of Secondary Education (CSE) and more academically-targeted General Certificate of Education Ordinary Level (O level) qualifications, intended to unify the grading of the two. The reason for the unification was that CSE bunched together good and very good students, while O level - bad and very bad. Since they were two independent, separate qualifications, relatively better students at the tails of the distribution could not distinguish themselves.



Notes: The plots display the share of new entrants into university program by entry qualifications and age using two sources of data: Undergraduate Records of the Universities' Statistical Record (USR) and Higher Education Statistics Agency (HESA). USR contains detailed information on the population of undergraduate students in British universities funded by the Unversity Grants Committee (UGC) in the years 1972-1993. HESA publishes aggregated tables of student counts by personal characteristics and highest entry qualification since 1994.

Figure 1.1: Characteristics of new university entrants

university education more difficult. Therefore, if parental unemployment shock at the age of 14 introduces a disruption, the implications could go beyond a sinlge year of school grades all the way to the probability of university enrollment and beyond.

1.2 Data

For the analysis I am using University of Essex, Institute for Social and Economic Research (2020), also known as Understanding Society, a largest household panel study of 40K individuals in the UK starting from 2009. The study covers wide range of topics, including measures of cognitive ability and personality traits. The original study participants were sampled randomly from the UK population and their households were followed each year. The analysis relies on the data from wave 3, which in addition to original sample includes participants continuing from a preceding British Household Panel Survey (BHPS). The dataset includes cross-sectional weights that account for sampling probabilities, endogenous non-response and probability of 'surviving' until wave 3. The analysis sample of 20,988 includes individuals who were born in the UK after 1945, finished school between the minimum school leaving age and age of 20, were at least 24 years old at the time of interview, were living with a family at the age of 14, have non-missing degree information and have non-zero cross-sectional sampling weight (see table 1.B.1 for detailed observation counts surviving each filter).

1.2.1 Parental unemployment

Each respondent is asked about employment status of their father and mother at the time when the respondent was 14 years old. By default, I use fathers' unemployment indicator unless a respondent was raised by a single mother. The response rate to this question was fairly high in the analysis sample. Only 0.56% observations from single-mother households and 4.78% of the rest were missing information on parent's employment status.

1.2.2 Educational outcomes

To measure educational attainment of individuals I use indicator for staying in school past age 16 and indicator for having any tertiary degree. The latter includes degress obtained from both so-called old (pre-1992) and new (post-1992) universities. The new universities were created by the Further and Higher Education Act 1992, which allowed former polytechnics to obtain a university status. The distinction between the two types of universities may be important to understand the mechanisms. Based on the graduate activities data released by HESA for the academic year 2017/18, graduates from undergraduate programs in new universities are 4.9pp less likely to be full-time employed or engaged in further studies compared to 78.5% probability among graduates from old universities. Although UKHLS does not distinguish between the two types of universities, BHPS respondents were asked if the type of institution last attended was university or polytechnics among others[10]. I use the answers of BHPS respondents observed in wave 3 of UKHLS and multiply impute the variable to the rest of the sample (for more details see Chapter 4).

1.2.3 Labour market outcomes

For labour market outcomes I look at employment, occupational and earnings information. The employment indicator combines both employees and self-employed individuals. I do not remove self-employed or unemployed individuals from the sample because this decision could be an outcome of parental unemployment shock.

10. Ideally, I am interested in whether individuals have obtained their degrees from former polytechnics. The survey question only lists 'polytechnics' as a possible answer, not 'former polytechnics'. This is the reason the variable still overestimates share of graduates from pre-1992 universities among younger cohorts that attended universities after 1992.

I also use occupational ranking associated with first and current jobs. I rank the occupational codes according to median earnings of relevant population. The UKHLS releases occupational codes at the two-digit SOC levels, which I collapse to one-digit level to obtain major occupational groups. To account for possibly endogenous unemployment and non-employment, I either occupation codes of the last job in the preceding two years or set the median earnings close to zero. To rank current occupations I compute weighted median labour earnings among all individuals in the sample with same major occupational group and year of birth. To rank first occupations I merge the sample with the data on median earnings by major occupational groups among 18-21 year olds[11] in the year a respondent turned 20. I deflate the median earnings in current job using the recommended consumer price index[12] and the median earnings in first job - by retail price index[13].

Finaly, I use earnings information directly. The variable can take zero or negative values among unemployed and self-employed workers, respectively. This means that I cannot apply standard log transformation to the earnings. The popular alternative in such cases is an inverse hyperbolic sine (IHS) transformation defined as $\operatorname{arcsinh}(x) = \ln(x + \sqrt{x^2+1})$, which allows zero and negative values of transformed variable. However, this comes at an expense of interpretability: the regression coefficients cannot be directly interpreted as elasticities. Following Bellemare and Wichman (2020), I convert the estimated coefficients to percentage change units. Along with continuous measure of earnings, I also consider indicator for nominal earnings being above zero.

1.2.4 Cognitive ability

In wave 3 UKHLS administered cognitive ability questions among all respondents aged 16 and above. The five cognitive tests - word recall, serial 7 subtraction, number series, verbal fluency and numeric ability - were selected to be reliable, cover multiple domains of intelligence, and easy to administer (McFall 2013). I combine the counts of correct answers to each question into a single intelligence score using principal component analysis. To abstract from potential cohort-specific differences in intelligence scores (Flynn 1984), I performed principal component analysis separately in each cohort group, defined by five-year windows

11. For years 1972-1994 I computed the series using the General Household Survey; for years 1997-2019 the data is available from Office of National Statistics. Working with major occupational groups reduces the issue of comparability of occupational codes. It also reduces the missingness problem; for a few missing years I linearly interpolated the median earnings within a given occupational group. For individuals turning 20 in 1966-1971 I used earnings in 1972; for individuals turning 20 in 1995 and 1996, I used earnings in 1994 and 1997, respectively.

12. CPI excluding rent, maintenance and water charges (Fisher et al. 2019)

13. RPI series go further back in time than CPI.

of year of birth. The resulting score is normalized to have mean 0 and standard deviation equal to 1 in each cohort group.

The constructed intelligence score is positively correlated with all educational and labour market outcomes. Table 1.1 presents correlation measures between outcome variables and intelligence score. A 1 sd increase in intelligence score is associated with 9.2pp increase in degree attainment, 3.8pp increase in employment and 21.1% increase in earnings.

Table 1.1: Average outcomes by intelligence score

	Dependent variable:			
	Degree	Work	Current job	IHS earn
Const.	0.269***	0.741***	−0.886***	2.646***
	(0.003)	(0.003)	(0.052)	(0.012)
IQ	0.092***	0.038***	0.666***	0.209***
	(0.002)	(0.002)	(0.035)	(0.008)
Obs.	20,988	20,988	20,987	20,988

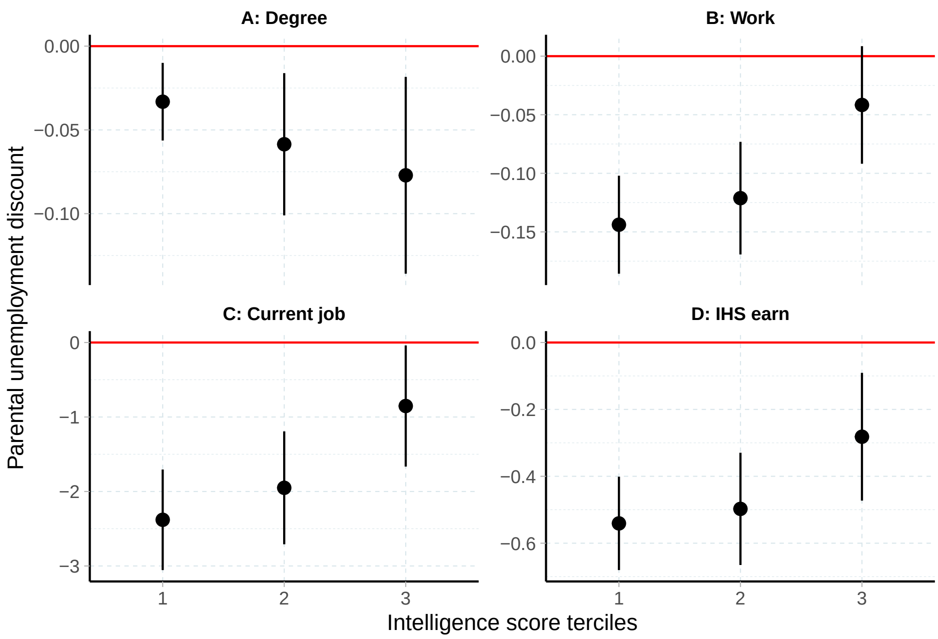
*p<0.1; **p<0.05; ***p<0.01

Notes: The table reports coefficients from weighted regressions of variables in columns on intelligence score. The IHS stands for inverse hyperbolic sine transformation; the corresponding coefficients could be converted to percentage change units (Bellemare and Wichman 2020). Standard errors are reported in parentheses.

1.2.5 Descriptive evidence

Before turning to the estimation strategy and results, I examine graphical evidence of the research question in the data. Figure 1.2 plots sample means of outcome variables by intelligence score and parental employment status. First, the figure suggests that there is a substantial discount associated with parental unemployment both in terms of educational attainment and labour market outcomes: the mean outcomes are lower among children whose parents were unemployed. Second, the magnitude of discount varies with intelligence score. The gap in degree attainment is increasing in intelligence score, whereas the gap in labour market outcomes is decreasing. Notably, the gap in occupational ranking of current jobs virtually disappears at high levels of intelligence score.

The graphical evidence suggests that cognitive ability is likely to play a protective role against negative family shocks experienced during adolescence, but only in terms of longer term outcomes measured here by labour market characteristics. Higher intelligence appears to be exacerbating the effects of the negative shock in the shorter term measured by degree attainment. The latter may be due to characteristics of the institutional setting discussed in section 1.1.



Notes: The figure plots parental unemployment discounts in terms of outcome variables in each panel by terciles of intelligence score. Parental unemployment discount is computed as the difference in sample mean among individuals with unemployed parents relative to those whose parents stayed employed.

Figure 1.2: Average outcomes by intelligence score and parental employment

1.3 Empirical strategy

Motivated by the results in Jacobson, LaLonde, and Sullivan (1993), most of the literature typically estimates the following regression equation

$$y_i = \alpha_0 + \alpha_1 UP_i + \alpha_2 \mathbf{X}_i + \alpha_3 \mathbf{P}_i + u_i \tag{1.1}$$

where y_i are outcome variables of individual i, UP_i is the indicator if a parent was unemployed when individual i was 14 years old; $\mathbf{X}_i$ is the vector of predetermined characteristics of individual i, and $\mathbf{P}_i$ is the vector of predetermined parental characteristics of individual i. The outcome variables are indicator for staying in school past age 16, tertiary degree indicator, work and self-employment indicators, IHS of real monthly earnings and categorical measures of nominal earnings being above zero or above cohort-specific median. The individual characteristics include fixed effects for gender, year of birth, country of birth, race, immigrant status (up to 3rd generation) and family composition at the age of 14. The parental characteristics include fixed effects for highest educational qualifications and country of birth of both mother and father.

To estimate the heterogeneous effects across ability distribution I modify equation (1.1) as follows

$$y_i = \beta_0 + \beta_1 UP_i + \beta_2 IQ_i + \beta_3 UP_i \times IQ_i + \beta_4 \mathbf{X}_i + \beta_5 \mathbf{P}_i + v_i \tag{1.2}$$

where IQ_i is the intelligence score of individual i. Here, β_2 estimates the difference in outcomes of children with unemployed parents relative to those whose parents stayed employed in the base group with $IQ = 0$. The coefficient of interest β_3 estimates the effect of parental unemployment for every unit increase in intelligence score relative to the base group.

The specification in equation (1.2) also allows me to interpret the coefficient of interest β_3 as a difference-in-differences estimator. A benefit of this strategy is that unlike the specification in equation (1.1) it allows parental unemployment indicator to be non-random. Allowing for selection bias is important in the current setting because the dataset can neither differentiate between different reasons for parental unemployment, nor provide a sufficient set of pre-displacement earnings information as suggested in Jacobson, LaLonde, and Sullivan (1993). So, even if β_2 cannot be causally interpreted, β_3 has a causal bearing under a modified parallel trends assumption and predeterminedness of intelligence score. The modified parallel trends assumption essentially requires selection bias to be constant across cognitive ability distribution.

To put it more formally, denote potential outcome of an individual if exposed to parental unemployment shock through y^1. Similarly, her potential outcome if her parents stayed employed - through y^0. The realised outcome is $y = y^0 \cdot (1 - UP) + y^1 \cdot UP$. For simplicity assume that IQ is a binary indicator equal to 1 if the intelligence score is above mean, and 0 otherwise. All expectations that follow include **X** and **P** in the conditioning set, which I omit for brevity. Modified parallelel trends assumption can be expressed as

$$\begin{aligned} \mathbb{E}(y^0|UP = 1, IQ = 1) - &\mathbb{E}(y^0|UP = 0, IQ = 1) = \\ &= \mathbb{E}(y^0|UP = 1, IQ = 0) - \mathbb{E}(y^0|UP = 0, IQ = 0) \end{aligned} \tag{1.3}$$

That is, if there is a selection bias present, the bias should be same in high- and low-ability groups conditional on a set of predetermined characteristics of children and parents.

The predeterminedness of intelligence score simply requires that IQ is not itself an outcome of parental unemployment. Denote a potential IQ score in case of treatment and no treatment as IQ^1 and IQ^0, respectively.

$$\mathbb{E}(y^0|UP,IQ^1) = \mathbb{E}(y^0|UP,IQ^0) = \mathbb{E}(y^0|UP,IQ) \tag{1.4}$$

Therefore, this assumption allows me to condition on the observed intelligence score IQ without being concerned about differences in intelligence score composition of children with and without unemployed parents. I formally test these assumptions in section 1.5.

Under this assumption, β_3 can be expressed as

$$\beta_3 = \mathbb{E}(y^1 - y^0|UP = 1, IQ = 1) - \mathbb{E}(y^1 - y^0|UP = 1, IQ = 0) \tag{1.5}$$

It describes treatment effect on treated in the high-ability group relative to the treatment on treated in the low-ability group.

1.4 Results and discussion

The main estimates of parental unemployment by intelligence score from equation (1.2) are reported in table 1.2. Panel A reports the effects on educational outcome variables. The results suggest that parental unemployment has larger negative effect on educational outcomes of children at the upper end of the ability distribution. Parental unemployment reduces the probability of degree attainment further by 1.5 pp for every 1 standard deviation increase in intelligence score. To get a sense of economic significance, compare this estimate to 9.2 pp increase in degree attainment associated with a 1 standard deviation increase in intelligence score (table 1.1). Thus, the estimated effect amounts to 16% of a marginal increase in educational attainment at the upper end of the ability distribution.

These results show that higher intelligence exacerbates even further the losses in terms of educational outcomes due to parental unemployment experienced at age 14. That is, instead of protecting children, higher intelligence makes them even more vulnerable to the negative shocks. Though surprising, the result is consistent with literature on human capital investments and skill formation. According to the theory of dynamic complementarity of skills (Cunha and Heckman 2010), human capital investments are more productive among children with already high levels of cognitive ability. A particular implication of this theory is that loss of human capital investments has larger cost for children with higher levels of intelligence score. The negative estimates in panel A of table 1.2 are in line with this prediction. To fully support the implication of the dynamic complementarity theory, I need to show that children at the higher

Table 1.2: Effect of parental unemployment on children's outcomes by intelligence score

	Regressors			
Depedent variable	Parent unemp	IQ	Parent unemp $\times$ IQ	Obs.
Panel A: Educational outcomes				
Post-16 school	-0.053***	0.085***	-0.021††	18,192
	(0.014)	(0.002)	(0.008)	18,192
Degree	-0.026*	0.086***	-0.015†	18,192
	(0.014)	(0.002)	(0.007)	18,192
Uni degree	-0.016	0.063***	-0.016††	18,192
	(0.014)	(0.004)	(0.007)	18,192
Panel B: Labour market outcomes				
Work	-0.056***	0.038***	0.021††	18,192
	(0.014)	(0.002)	(0.008)	18,192
IHS earnings	-0.252***	0.204***	0.041	18,192
	(0.050)	(0.008)	(0.029)	18,192
%Δ earnings	-22.483***	20.582***	4.327†	18,192
	(4.462)	(0.852)	(2.960)	18,192
Earn > 0	-0.055***	0.038***	0.022††	18,192
	(0.014)	(0.002)	(0.008)	18,192
Current job	-1.049***	0.619***	0.377††	18,192
	(0.226)	(0.037)	(0.138)	18,192

$^{\dagger}q < 0.1;^{\dagger\dagger}q < 0.05;^{\dagger\dagger\dagger}q < 0.01$
$^{*}p < 0.1;^{**}p < 0.05;^{***}p < 0.01$

Notes: The table reports coefficients from weighted regressions of variables in rows on parental unemployment indicator and intelligence score. All regressions control for respondents' (gender, year of birth, country of birth, race, migrant status, family composition at the age of 14) and parents' (highest educational qualifications and country of birth) characteristics. The row for university degree presents the coefficients combined using Rubin's rules using multiply imputed dependent variable. The IHS stands for inverse hyperbolic sine transformation; the corresponding coefficients are converted to percentage change units in the next row (Bellemare and Wichman 2020). Standard errors are reported in parentheses. The significance stars of interaction coefficients are assigned based on sharpened q-values (Benjamini, Krieger, and Yekutieli 2006; Anderson 2008)

Table 1.3: Effect of parental unemployment on degree attainment by parental qualifications

	Regressors			
Parent qualification	Parent unemp	IQ	Parent unemp $\times$ IQ	Obs.
No school	-0.004	0.164***	-0.157††	18,192
	(0.077)	(0.048)	(0.061)	18,192
Some school	0.214***	0.083***	-0.009	18,192
	(0.081)	(0.002)	(0.008)	18,192
Degree	0.191*	0.103***	-0.006	18,192
	(0.102)	(0.008)	(0.031)	18,192
Missing	0.288***	0.088***	-0.054††	18,192
	(0.088)	(0.005)	(0.014)	18,192

$^{\dagger}q < 0.1;^{\dagger\dagger}q < 0.05;^{\dagger\dagger\dagger}q < 0.01$
$^{*}p < 0.1;^{**}p < 0.05;^{***}p < 0.01$

Notes: The table reports coefficients from weighted regressions of degree indicator on parental unemployment indicator and intelligence score by parents' highest educational qualification groups. The regression controls for respondents' (gender, year of birth, country of birth, race, migrant status, family composition at the age of 14) and parents' (highest educational qualifications and country of birth) characteristics. Standard errors are reported in parentheses. The significance stars of interaction coefficients are assigned based on sharpened q-values (Benjamini, Krieger, and Yekutieli 2006; Anderson 2008)

end of the ability distribution do lose human capital investments. Unfortunately, the data does not allow me to verify this statement directly. Instead, I rely on the theory of intergenerational transmission of earnings (Becker and Tomes 1986; Mulligan 1997), which predicts that only poor households reduce human capital investments as a result of income shock. In figure 1.3, I show that most of additional negative effect on degree attainment at higher levels of intelligence score is concentrated among children with less educated parents[14].

Panel B of table 1.2 presents the estimates with labour market outcomes as dependent variables. These results suggest that cost of parental unemployment in terms of labour market outcomes is decreasing in intelligence score. Here, a 1 standard deviation increase in intelligence score improves the effect of parental unemployment on the probability of employment by 2.1 pp and earnings - by 4.3 pp. These results suggest that intelligence does protect individuals from the consequences of negative family shocks in the longer term. Even though the cost on educational outcomes is highest at the upper end of the ability distribution, the affected individuals with higher intelligence score are able to overcome the disadvantage later in the labour market.

14. Parental educational qualifications are self-reported by children and are missing for about a fifth of the sample. I treat missingness as a separate category in the estimations. For interpretation of the results, I assume missingness to be a signal of low educational attainment.

The result is consistent with employer learning theory, which extends a traditional signalling model by allowing employers to learn about worker productivity over time. In a traditional signalling model, workers can signal or reveal their ability only via education. That is, all signalling activity happens before entering the labour market. Wages are set according to the observed educational qualifications and do not change afterwards. Several papers have extended the traditional model by allowing employers to learn about worker productivity from their work performance (Farber and Gibbons 1996; Altonji and Pierret 2001; Arcidiacono, Bayer, and Hizmo 2010). When workers can send additional signals about their productivity after entering the labour market, the initial educational signal becomes less important in wage setting and returns to ability are increasing as work experience grows. Therefore, this theory offers an explanation for the positive results in terms of labour market outcomes: despite not being able to obtain a degree, high-ability workers can demonstrate their skills on the job and, thereby, mitigate the initial disadvantage.

The employer learning theory offers two testable implications. First, the effect of parental unemployment on early career outcomes should be flat with respect to intelligence score. Since high-ability children with unemployed parents fail to get a tertiary degree, initially they are not able to differentiate themselves from less able job candidates. Second, the rate at which higher intelligence score improves the effect of parental unemployment on labour market outcomes should increase with work experience. That is, the interaction coefficient β_3 should be larger among more experienced workers.

To test the first implication I estimate the effect of parental unemployment on occupational ranking of children's first job (table 1.4). Indeed, the effect of parental unemployment on median earnings in the first job does not vary with intelligence score. The point estimates are close to zero in magnitude and are statistically insignificant. Since these are not actual earnings in the first job, for better comparison I duplicate the results in terms of occupational ranking of the current job from table 1.2. Similar to the rest of the labour-market outcomes, a 1 standard deviation increase in intelligence score improves the effect of parental unemployment on occupational ranking of current job. Thus, high-ability individuals that did not get education due to parental unemployment are initially pooled together with low-ability workers and begin their careers at lower-paying jobs, confirming the implication of the signalling theory. By the time of the survey, high-ability workers whose parents were unemployed manage to switch to better jobs.

In order to test the second implication, I predict wages at ages 18, 20, 25, 30, 35, 40 and 45 for all individuals in wave 3 (for more information on prediction algorithm see appendix section 1.A). Then, I regress the predicted wages on age indicators interacted with parental

Table 1.4: Effect of parental unemployment on occupational ranking of children by intelligence score

	Regressors			
Depedent variable	Parent unemp	IQ	Parent unemp $\times$ IQ	Obs.
First job	-0.036***	0.018***	0.006	15,363
	(0.013)	(0.002)	(0.008)	15,363
Current job	-1.049***	0.619***	0.377††	18,192
	(0.226)	(0.037)	(0.138)	18,192

$^{\dagger}q < 0.1;^{\dagger\dagger}q < 0.05;^{\dagger\dagger\dagger}q < 0.01$
$^{*}p < 0.1;^{**}p < 0.05;^{***}p < 0.01$

Notes: The table reports coefficients from weighted regressions of occupation rankings on parental unemployment indicator and intelligence score. Both current and first occupations were aggregated to major occupational groups (one-digit SOC) prior to ranking. Current occupations were ranked according to log of real weighted median earnings of individuals born in the same year. First occupations were ranked according to log of real median earnings of 18-21 year olds in the year the respondent turned 20. The median occupational earnings of 18-21 year olds were computed using General Household Survey for years between 1972 and 1994 and downloaded from Office of National Statistics for years 1997 - 2019. All regressions control for respondents' (gender, year of birth, country of birth, race, migrant status, family composition at the age of 14) and parents' (highest educational qualifications and country of birth) characteristics. Standard errors are reported in parentheses. The significance stars of interaction coefficients are assigned based on sharpened q-values (Benjamini, Krieger, and Yekutieli 2006)

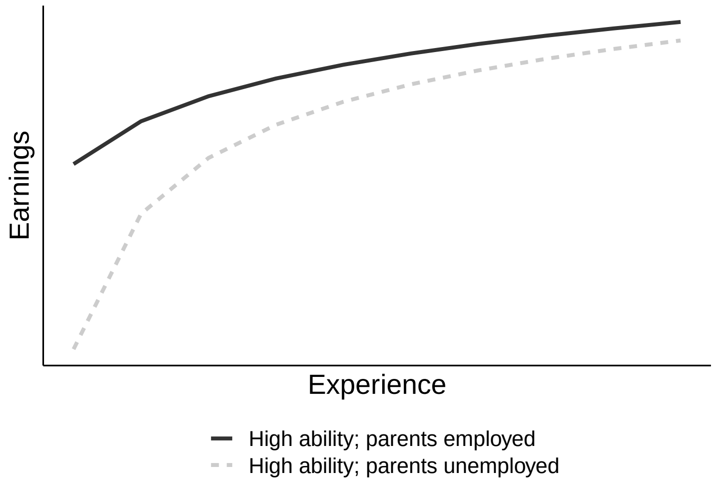
Table 1.5: Effect of parental unemployment on predicted wages by intelligence score and age

Age	Parent unemp	IQ	Parent unemp $\times$ IQ	Obs.
20	0.005***	0.000***	0.004***	140,371
	(0.001)	(0.000)	(0.000)	
25	0.004**	0.007***	0.007***	140,371
	(0.002)	(0.000)	(0.001)	
30	-0.001	0.011***	0.007***	140,371
	(0.003)	(0.000)	(0.002)	
35	0.010**	0.014***	0.013***	140,371
	(0.004)	(0.000)	(0.002)	
40	0.002	0.016***	0.013***	140,371
	(0.005)	(0.000)	(0.003)	
45	-0.001	0.018***	0.015***	140,371
	(0.006)	(0.000)	(0.003)	

$^{*}p < 0.1;^{**}p < 0.05;^{***}p < 0.01$

Notes: The table reports coefficients from fixed effect estimations of predicted real hourly wages on age indicators fully interacted with parental unemployment indicator and intelligence score. The coefficients are relative to age 18. Standard errors are clustered at the individual level and are reported in parentheses. Regressions are weighted by cross-sectional response weights from wave 3.

unemployment status and intelligence score using the fixed effects estimator. The results are presented in table 1.5. Consistent with the prediction, we can see that the effect of intelligence score on the treatment effect of parental unemployment is increasing with age.



Notes: The figure plots possible experience profile of earnings of high-ability children depending on the employment status of their parents. The curves are constructed under two additional assumptions that (i) parental unemployment has a negative treatment effect and (ii) experience profile of similar children whose parents stayed employed is an upper bound.

Figure 1.3: Implied experience profile of high-ability children

The estimation results show that children with higher cognitive ability skills are able to mitigate the effects of parental unemployment over time, despite being undereducated. On the other hand, more than 50% of individuals in the top quintile of intelligence score distribution have a tertiary degree. This could raise a question of whether high-ability individuals, in fact, need to get a university education. The answer to this question depends on a careful examination of benefits and costs of signalling via education vs on-the-job signalling (Alós-Ferrer and Prat 2012), speed of employer learning and productivity dynamics (Lange 2007; Kahn and Lange 2014). While it is beyond the scope of my paper, the results could provide some evidence in favour of the educational signalling under two additional assumptions. First, I need to assume that the effect of parental unemployment on children's labour-market outcomes at the bottom of the ability distribution is negative. Even though the identification strategy does not allow me to give causal interpretation to the main effect of parental unemployment, notice that all the estimates in the first column of table 1.2 are significantly negative. In addition, the literature has also found some negative average effects on the labour-market outcomes of children (table 1.B.2). Second, I need to assume that at any point along the experience profile of earnings, children whose parents were unemployed cannot do better than similar children whose parents were employed. In other words, the high-ability children could at best catch up with their peers whose parents stayed employed, but not overtake them. Under these two assumptions, the results in this paper imply that experience profile of high-ability children whose parents were unemployed is shifted downwards relative to the experience profile of their peers whose parents stayed employed, meaning that their lifetime earnings are also lower (figure 1.3).

1.5 Validity

1.5.1 Identifying assumption

The causal interpretation of the estimation results relies on the modified parallel trends assumption in equation (1.3), which essentially requires selection bias, if present, to be constant across ability distribution. I can formally test this assumption by estimating equation (1.2) using predetermined characteristics as dependent variables. If the modified parallel trends assumption holds true, then the interaction coefficients from these regressions should be zero. It is worth noting that in the current framework not every outcome realised prior to age 15 can be considered a predetermined characteristic. Since the estimation strategy allows for presence of selection bias, an ideal proxy for y_i^0 is a variable that may determine selection into treatment (parental unemployment) but cannot be influenced by either treatment, or selection. For example, consider a model where depending on the gender of a child parents decide how much they value their leisure and that parents of boys value their leisure twice as much as parents of girls. Then, for a given wage rate parents of boys will want to work less. In this situation, child's gender determines selection into treatment. Now consider school grades of children at any time before age 15. In this simple world, only financial investments into children matter for school performance. Since parents of boys in general work less, they also have less income, meaning they have less financial resources to invest into their child. This leads to boys' grades being lower than that of girls. So, despite observing school grades chronologically before treatment, they are, nevertheless, outcomes. Therefore, in this example, gender is a predetermined characteristic that describes selection, while school grades, even at earlier ages, are outcomes.

In Table 1.6, I present the regression results using predetermined variables in the UKHLS. Indeed, all the interaction coefficients, reported in column 3, are statistically insignificant and close to zero in magnitude. However, the set of predetermined variables available for the test in the UKHLS are rather limited: they are mostly related to ethnic background of parents and grandparents, which could be already captured by parents' country of birth indicator.

Table 1.6: Test of parallel trends assumption using predetermined characteristics

Variable	Regressors			Obs.	Mean variable
	Parent unemp	IQ	Parent unemp × IQ		
Father's father born UK	-0.004	0.002	0.000	20,060	0.75
	(0.008)	(0.001)	(0.004)		
Father's mother born UK	0.006	-0.001	0.000	20,060	0.76
	(0.007)	(0.001)	(0.004)		
Mother's father born UK	-0.015**	0.004***	-0.002	20,060	0.76
	(0.008)	(0.001)	(0.004)		
Mother's mother born UK	0.002	0.001	-0.001	20,060	0.77
	(0.006)	(0.001)	(0.004)		
White british father	0.014	0.002	-0.002	20,060	0.67
	(0.009)	(0.002)	(0.005)		
White british mother	0.010	0.001	-0.002	20,060	0.68
	(0.009)	(0.002)	(0.006)		
Has siblings	0.018*	0.000	-0.004	20,060	0.90
	(0.009)	(0.002)	(0.005)		

$^{\dagger}q < 0.1$;$^{\dagger\dagger}q < 0.05$;$^{\dagger\dagger\dagger}q < 0.01$
$^{*}p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$

Notes: The table shows the results from regressions of predetermined variables in UKHLS shown in the first column on parental unemployment and intelligence score. All regressions control for respondents' (gender, year of birth, country of birth, race, migrant status, family composition at the age of 14) and parents' (highest educational qualifications and country of birth) characteristics. Standard errors are reported in parentheses. The significance stars of interaction coefficients are assigned based on sharpened q-values (Benjamini, Krieger, and Yekutieli 2006)

Therefore, I repeat the test using another dataset: the 1970 British Cohort Study[15] (BCS70). It is an ongoing longitudinal survey following over 17,000 children[16] born in a week of 1970 in the Great Britian. I use early waves that took place when children were just born, 5, 10, and 16 years old. The advantage of cohort studies is that the respondents were given cognitive tests several times throughout life. Thus, I can combine the test results at ages 5, 10 and 16 into a single intelligence score at each age using the first principal component. In the estimations, I use intelligence score at age 10 as the main independent variable. The scores at ages 5 and 16 can be used as outcomes in the test, along with pre-determined characteristics at birth. The results are reported in Table 1.7. Again, all the interaction coefficients are statistically insignificant, both before and after multiple-inference adjustment. Moreover, the magnitudes of the estimates are small as evidenced by the ratio of coefficients to the means of the dependent variables in the sample. These results also support the identifying

15. Chamberlain, University of London, Institute of Education, Centre for Longitudinal Studies, and Chamberlain (2013); Butler et al. (2016); Butler, Bynner, and University of London, Institute of Education, Centre for Longitudinal Studies (2016); Bynner, University of London, Institute of Education, Centre for Longitudinal Studies, and Butler (2019)

16. Sample sizes vary across waves due to sample attrition and unit non-response. To account for this, I construct inverse-probability weights similar to (Mostafa and Wiggins 2014).

Table 1.7: Test of parallel trends assumption in BCS70

	Regressors				
Variable	Parent unemp	IQ	Parent unemp × IQ	Obs.	Mean of variable
At birth					
Birthweight, g	-40.227	60.880***	-7.167	4,890	3,282
	(28.969)	(10.369)	(26.475)		
Age of mother	0.641**	0.357***	0.315	4,894	26.21
	(0.267)	(0.084)	(0.265)		
Age of father	1.843***	0.435***	0.317	4,303	29.05
	(0.345)	(0.104)	(0.319)		
Height of mother, cm	-0.570*	0.369***	0.003	4,860	161
	(0.315)	(0.115)	(0.278)		
Mother married	0.003	0.001	-0.006	4,894	0.95
	(0.012)	(0.004)	(0.01)		
Age of mother at first birth	-0.522***	0.471***	0.026	4,874	21.68
	(0.176)	(0.063)	(0.178)		
At age 5					
Composite score (PC1)	-0.124*	0.263***	-0.040	2,076	-0.05
	(0.064)	(0.04)	(0.066)		
Age at test, days	-3.674**	-0.492	0.033	4,366	1,853
	(1.437)	(0.992)	(1.345)		
Reading score	-0.791***	1.447***	-0.801	2,157	3.09
	(0.272)	(0.183)	(0.299)		
English picture vocab. score	-0.230***	0.373***	0.037	4,453	-0.35
	(0.073)	(0.026)	(0.071)		
Copying designs score	-0.057	0.393***	0.072	4,453	-0.10
	(0.051)	(0.017)	(0.048)		
Draw-a-man score	-0.029	0.286***	0.094	4,453	-0.17
	(0.063)	(0.021)	(0.064)		
Complete-a-profile score	-0.419*	0.493***	-0.328	4,304	6.85
	(0.216)	(0.075)	(0.22)		
At age 10					
Has normal vision	-0.042**	0.004	0.009	4,648	0.86
	(0.019)	(0.006)	(0.019)		
At age 16					
Composite score (PC1)	-0.125*	0.585***	0.005	1,259	-0.07
	(0.073)	(0.027)	(0.075)		
Reading score	-1.865*	7.403***	0.778	1,336	53.55
	(1.004)	(0.369)	(1.01)		
Spelling score	1.169	14.626***	2.550	4,894	73.90
	(3.936)	(1.418)	(3.566)		
Vocabulary score	0.538	6.079***	-0.425	4,894	19.57
	(1.078)	(0.395)	(0.993)		
Math score	-1.112	6.102***	-0.519	1,591	36.10
	(0.901)	(0.298)	(0.864)		
Complete-matrix score	-0.203	0.579***	-0.038	1,368	8.81
	(0.132)	(0.052)	(0.149)		

$^{\dagger} q < 0.1$; $^{\dagger\dagger} q < 0.05$; $^{\dagger\dagger\dagger} q < 0.01$
$^{*} p < 0.1$; $^{**} p < 0.05$; $^{***} p < 0.01$

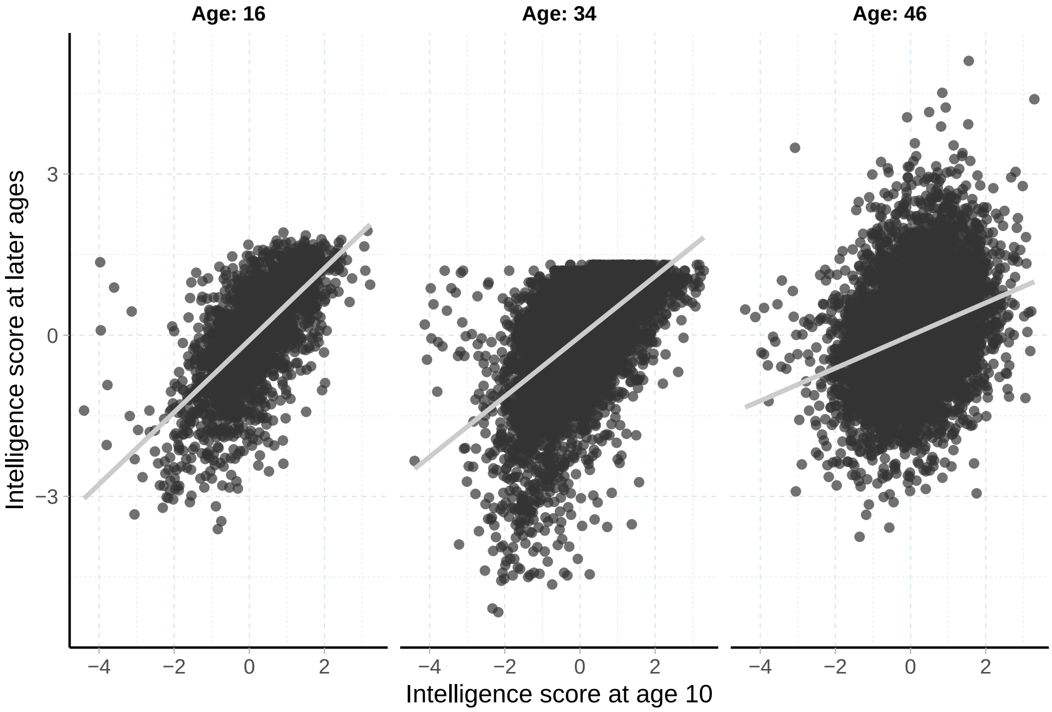
Notes: The table shows the results from regressions of predetermined variables shown in the first column on parental unemployment at age 16 and intelligence score at age 10 in the BCS70. All regressions control for respondents' (gender, country of birth) and parents' (country of birth and age left education) characteristics. Estimations are weighted with inverse probability of response (Mostafa and Wiggins 2014). Standard errors are reported in parentheses. The significance stars of interaction coefficients are assigned based on sharpened q-values (Benjamini, Krieger, and Yekutieli 2006) assumption of constant selection bias across intelligence scores; at least, based on observable pre-determined characteristics.

Another important observation from Table 1.7 is that interaction coefficients from regressions of intelligence scores measured at different ages are generally small and insignificant. That is, slope of the test scores at different ages with respect to test score at age 10 are the same between children with and without unemployed parents. These results provide support to the assumption that intelligence score is not itself an outcome of selection into treatment.

1.5.2 Stability of the relative intelligence score

The main results in Table 1.2 controls for the intelligence score measured at the time of the survey, decades after the exposure to treatment. There is evidence that large part of skill formation process is concentrated in certain periods of life (Cunha and Heckman 2010) with development of cognitive skills taking place by age 10 (Hopkins and Bracht 1975). Of course, the level of skills does not stay constant over time (Salthouse 2010). However, a crucial assumption for my analysis is the stability of the relative position of individuals along the cognitive ability distribution. That is, if a child scores at the top of the distribution at the age of 10, her score at later ages is also more likely to be at the higher end. Analysing population of Scottish cohorts born in 1921 and 1936 Deary (2014) estimates, conservatively, that about half of differences in intelligence score at age 70 can be traced back to relative standing in the distribution at age 11.

The UKHLS does not allow me to test this assumption as there is only a single set of cognitive ability test scores measured in wave 3. The BCS70, on the other hand, administered cognitive tests several times throughout life. For example, cognitive ability test scores are available at ages 5, 10, 16, 34, 42 and 46 in the BCS70. Using the tests at ages 10, 16, 34 and 46, I construct intelligence scores at these ages by extracting the first principal component. Figure 1.4 shows that intelligence scores at later ages are positively correlated with intelligence score at age 10. For example, a 1 standard deviation increase in intelligence score at age 10 is associated with 0.7 standard deviation increase in intelligence score at the age of 16. Figure 1.4 also shows that by age 50 the correlation coefficient reduces to 0.3. However, this is likely to be a lower bound due to sample attrition and differences in test composition. In order to measure intelligence correctly, the content of tests should be adapted to the age of test subjects. Tests appropriate to 10-year-old children might be too easy for 50-year-old individuals. Therefore, variations in the score across ages could reflect differences in test contents, even among tests measuring the same domain of cognitive ability. In addition to this,



Notes: The figure plots the scatterplot of standardized intelligence scores at ages 16 and 46/50 against the score at age 10/11. Intelligence scores are constructed as first PC and standardized to have mean 0 and standard deviation 1 within each dataset and age. The red line is the fitted linear regression line and the corresponding slope coefficient is printed at the upper left corner of each box.

Figure 1.4: Stability of intelligence score by ages

the cohort studies had different aims when testing children vs adults. For example, childhood tests were mostly examining the ability of children to solve new problems using their skills, while in adulthood they focused more on the ability of individuals to perform day-to-day tasks (Figure 1.B.1).

1.5.3 Recall bias

Another potential concern with the current parental unemployment measure is recall bias since the parental employment status is self-reported by children years or decades later. To assess the severity of the recall bias I plot the share of individuals reporting an unemployed parent against aggregate unemployment rates in the corresponding years in figure 1.5. I use two aggregate unemployment rates for comparison: one in the entire population of the UK and another - among British males at the ages 40-49, a superset of population of fathers of 14-year-old children. Reassuringly, for most of the sample the share of people with unemployed parent is comparable to both of the aggregate series. But, rather unexpectedly, the series diverge for the younger cohorts: average parental unemployment is much higher in the Understanding Society. It could be due to financial crisis of 2008-09 affecting retrospective perceptions of

Notes: The plot compares the average parental unemployment indicator in the UKHLS with aggregate unemployment rates in the UK. The shares in the UKHLS are weighted by individual cross-sectional weights. The two aggreage series are official unemployment rates from 1971 onwards and male unemployment rate in the age group 40-49 from 1983 onwards. The shaded areas correspond to recessions.

Figure 1.5: Parental unemployment and aggregate economy

these children. I test the sensitivity of the analysis results to the exclusion of cohorts that turned 14 in 1995 or later in table 1.8. The point estimates are largely similar to the baseline results, both qualitatively and quantitatively. The point estimates of the effects in terms of degree attainment are half as large as in the baseline results, but the effect in terms of staying on at school past age 16 are almost identical.

1.6 Conclusion

The intergenerational effects of parental layoff have recently received an increased attention, with many studies finding a significant cost imposed on various outcomes of children, especially pronounced among children from disadvantaged backgrounds. In this paper I provide new evidence on the heterogeneity of the parental unemployment effect on children's outcomes along the cognitive ability distribution. Using the UK survey data and difference-in-differences setting, I show that high ability does indeed help children overcome the effects of negative shocks, but only in terms of long-term outcomes. This conclusion is based on two key findings presented in this paper.

First, I show that high intelligence score exacerbates the effect of parental unemployment on educational attainment of children. This result is consistent with the dynamic complementarity theory (Cunha and Heckman 2010), which predicts that loss of human capital

Table 1.8: Effect of parental unemployment on children's outcomes by intelligence score: children born before 1981

Depedent variable	Regressors: Parent unemp	IQ	Parent unemp × IQ	Obs.
Panel A: Educational outcomes				
Post-16 school	-0.043***	0.138***	-0.034†	18,086
	(0.015)	(0.004)	(0.013)	18,086
Degree	-0.004	0.132***	-0.012	18,086
	(0.015)	(0.003)	(0.012)	18,086
Uni degree	-0.004	0.094***	-0.016	18,086
	(0.015)	(0.005)	(0.013)	18,086
Panel B: Labour market outcomes				
Work	-0.049***	0.054***	0.031†	18,086
	(0.014)	(0.004)	(0.014)	18,086
IHS earnings	-0.213***	0.297***	0.064	18,086
	(0.052)	(0.013)	(0.049)	18,086
%Δ earnings	-19.523***	30.222***	6.871	18,086
	(4.741)	(1.358)	(4.954)	18,086
Earn > 0	-0.044***	0.055***	0.029†	18,086
	(0.014)	(0.004)	(0.014)	18,086
Current job	-0.916***	0.892***	0.483	18,086
	(0.234)	(0.060)	(0.233)	18,086

$^{\dagger}q < 0.1;^{\dagger\dagger}q < 0.05;^{\dagger\dagger\dagger}q < 0.01$
$^{*}p < 0.1;^{**}p < 0.05;^{***}p < 0.01$

Notes: The table reports coefficients from weighted regressions of variables in rows on parental unemployment indicator and intelligence score in a restricted sample of children born before 1981. All regressions control for respondents' (gender, year of birth, country of birth, race, migrant status, family composition at the age of 14) and parents' (highest educational qualifications and country of birth) characteristics. The IHS stands for inverse hyperbolic sine transformation; the corresponding coefficients are converted to percentage change units in the next row (Bellemare and Wichman 2020). Standard errors are reported in parentheses. The significance stars of interaction coefficients are assigned based on sharpened q-values (Benjamini, Krieger, and Yekutieli 2006; Anderson 2008) investments affects high-ability children more. I show that most of the damaging effect of high intelligence is concentrated among children with less educated parents - they are more likely to have experienced loss in human capital investments due to parental unemployment.

Second, I show that despite the negative effect on educational outcomes, high levels of intelligence score demonstrate remedial properties in terms of labour-market outcomes. I also find that the effects in terms of occupational ranking of individuals' first jobs do not exhibit heterogeneity along the cognitive ability dimension. These results are consistent with the employer learning model. At the beginning of their career high-ability children that failed to get education due to parental unemployment are bundled together with low-ability children. However, as they accumulate work experience, high-ability workers can prove themselves on-the-job thereby improving their outcomes.

The two findings together show that high-ability children can overcome the effects of negative shocks, despite being undereducated. This might raise a question whether these children in fact need a university education. While it is beyond the scope of my paper, the results hint that obtaining a degree could increase their lifetime earnings. This conjecture is based on two assumptions: parental unemployment cannot have positive effect on children's outcomes and it has a strictly negative treatment effect on children at the bottom of the ability distribution.

Appendix 1.A Predicted panel of wages

Part of the analysis in this paper relies on estimations in the panel data with observations at various ages for each sample member. The UKHLS offers up to ten years worth of observations per individual. The time span, thus, is rather limited. But more importantly, these panel observations are at different segments of wage age profiles. To overcome these issues, I generate predicted wages at ages 18, 20, 25, 30, 35, 40 and 45 for each sample member.

First, I estimate the wage age profiles using methodology motivated by Lagakos et al. (2018). In particular, I estimate the following regression using fixed effects estimator

$$w_{iat} = \alpha + \gamma_a + \beta_a \mathbf{X}_i + \delta_t + \phi_t \mathbf{X}_i + \mu_i + \nu_{iat} \tag{1.6}$$

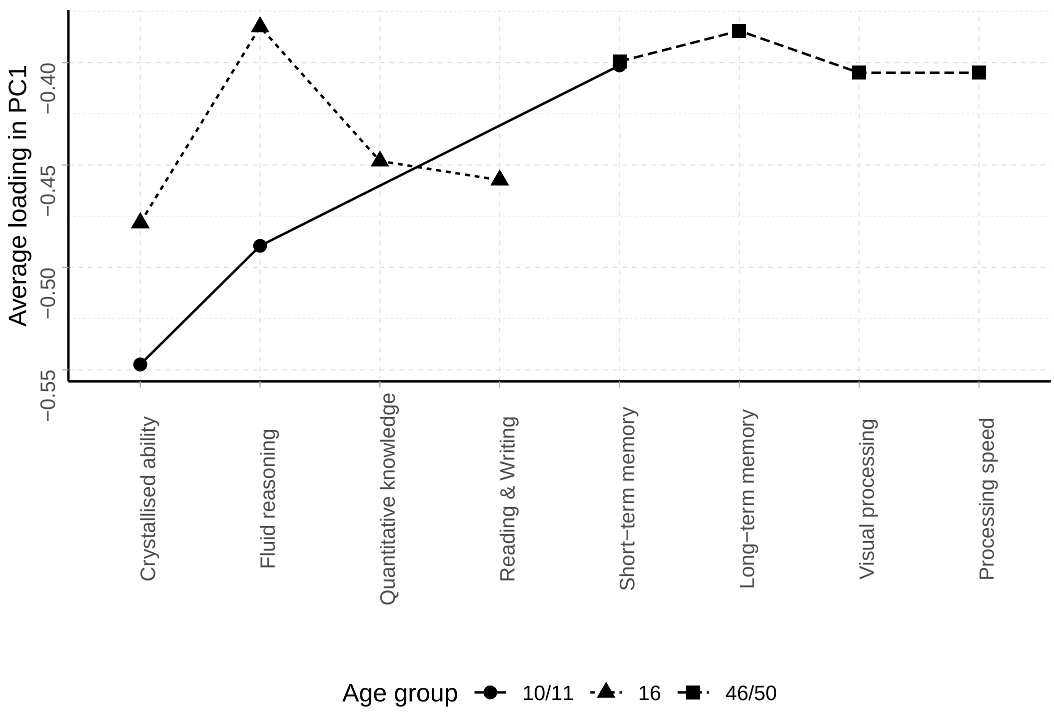
where w_{iat} is real hourly wage of person i at age a at time t, $\mathbf{X}_i$ are i's personal characteristics (gender, degree status and parental unemployment status when i was 14), and γ_a, δ_t and μ_i are age, year and individual fixed effects. The estimation allows for age- and year-specific effects of variables in $\mathbf{X}_i$.

Then, using the estimated coefficients, the predicted wages at age a are computed as follows

$$\hat{w}_{iat} = \hat{\alpha} + \hat{\gamma}_a + \hat{\beta}_a \mathbf{X}_i + \hat{\mu}_i \tag{1.7}$$

Finally, the data is reshaped such that each individual has six observations corresponding to ages 18, 20, 25, 30, 35, 40, and 45, instead of survey waves.

Appendix 1.B Supplementary Figures and Tables



Notes: The plot shows cognitive domains of tests administered at different ages. On the y-axis I plot simple average of test scores' loadings in PC1 in a given domain, dataset and age group.

Figure 1.B.1: Cognitive domains of tests by ages

Table 1.B.1: Sample filters

Filter	N
Initial sample	49,692
Attended and finished school	47,645
Born in UK	41,131
School-leaving-age law complier	38,134
Finished school by age 20	37,848
Born between 1950 and 1995	27,066
Lived with family at 14	26,628
Has degree information	26,454
Non-missing intelligence score	22,428
Non-zero sample weight	20,988

Notes: This table shows the counts of observations remaining after respective filters specified in column 1 applied to wave 3 of UKHLS.

Table 1.B.2: Literature summary

Paper	Identification strategy	Dataset	Result	Heterogeneity
Mörk, Sjögren, and Svaleryd (2019)	Propensity score matching	Swedish population-wide micro regster data 1987-2010	Childhood health, educational and early adult outcomes are not adversely affected by parental job loss	
Angelini, Bertoni, and Corazzini (2018)	Value-added models of personality	German Socio-Economic Panel Study (SOEP)	Parental unemployment makes offspring significantly more conscientious and - to smaller extent - less neurotic.	age at event, gender of child, gender of parent, parental educational, length of paternal unemployment
Hilger (2016)	Difference-in-differences	Federal tax returns 1996-2009	Layoffs only slightly redice college enrollment, college quality, and early career earnings.	family income, wealth, gender
Peter (2016)	Propensity score matching	German Socio-Economic Panel Study (SOEP)	Maternal job loss increases preschool chidlren's socio-behavioural problems and decreases adolescents' belief in self-determination.	
Pan and Ost (2014)	Conditional Independence Assumption	Panel Study of Income Dynamics (PSID)	Parental job loss decreases college enrollment by 10 pp.	parental education, home ownership, family income, magnitude of income shock, unemployment benefit generosity, tuition fees
Brand and Thomas (2014)	Propensity score matching	National Longitudinal Survey of Youth (NLSY) and National Longitudinal Survey's Child-Mother file (NLSCM)	Significant negative effect of job displacement among single mothers on children's educational attainment and social-psychological well-being in young adulthood. Effects are concentrated among older children and children whose mothers had a low likelihood of displacement.	age at event, propensity for displacement

Table 1.B.2: Literature summary *(continued)*

Paper	Identification strategy	Dataset	Result	Heterogeneity
Coelli (2011)	Conditional Independence Assumption	Canadian Survey of Labour and Income Dynamics (SLID)	Significant negative effect of parental job loss on any post-secondary education enrollment, lowering the probability by 10.5pp.	parental education, income, age at event, local unemployment rate, university tuition fees
Rege, Telle, and Votruba (2011)	Conditional Independence Assumption	Norwegian registry matched with student registry 2003-2007	Negative effect of paternal job loss on children's school performance, but non-significant positive effect from maternal job loss.	age at event, local economy, gender
Stevens and Schaller (2011)	Fixed effects	US Survey of Income and Program Participation (SIPP)	Parental job loss increases the probability of children's grade retention by 0.8 percentage points, or around 15%.	family income, parental education, family composition
Page, Stevens, and Lindo (2009)	Conditional Independence Assumption	Panel Study of Income Dynamics (PSID)	No evidence that firm closings have intergenerational effects on average, but found long-term costs on disadvantaged children.	family income, age at event
Oreopoulos, Page, and Stevens (2008)	Conditional Independence Assumption	Canadian Intergenerational Income Database (IID)	Children whose fathers were displaced have annual earnings about 9% lower. They are also more likely to receive unemployment insurance and social assistance. The estimates are driven by the experiences of children at the bottom of the income distribution.	pre-displacement income
Bratberg, Nilsen, and Vaage (2008)	Conditional Independence Assumption	Norwegian full population database of matched employer-employee data	No significant effects on earnings of children with fathers that experienced job loss.	pre-displacement income, father's education, industry

References

Alós-Ferrer, C., and J. Prat. 2012. "Job Market Signaling and Employer Learning" [in en]. *Journal of Economic Theory* 147 (5): 1787–1817.

Altonji, J. G., and C. R. Pierret. 2001. "Employer Learning and Statistical Discrimination" [in en]. *The Quarterly Journal of Economics* 116 (1): 313–350.

Anderson, M. L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–1495.

Angelini, V., M. Bertoni, and L. Corazzini. 2018. "Does Paternal Unemployment Affect Young Adult Offspring's Personality?" *Journal of Human Capital* 12 (3): 542–567.

Arcidiacono, P., P. Bayer, and A. Hizmo. 2010. "Beyond Signaling and Human Capital: Education and the Revelation of Ability" [in en]. *American Economic Journal: Applied Economics* 2 (4): 76–104.

Becker, G. S., and N. Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics* 4 (3): S1–S39.

Bellemare, M. F., and C. J. Wichman. 2020. "Elasticities and the Inverse Hyperbolic Sine Transformation" [in en]. *Oxford Bulletin of Economics and Statistics* 82 (1): 50–61.

Benjamini, Y., A. M. Krieger, and D. Yekutieli. 2006. "Adaptive Linear Step-up Procedures That Control the False Discovery Rate" [in en]. *Biometrika* 93 (3): 491–507.

Brand, J. E., and J. S. Thomas. 2014. "Job Displacement among Single Mothers: Effects on Children's Outcomes in Young Adulthood." *American Journal of Sociology* 119 (4): 955–1001.

Bratberg, E., Ø. A. Nilsen, and K. Vaage. 2008. "Job Losses and Child Outcomes" [in en]. *Labour Economics,* European Association of Labour Economists 19th Annual Conference / Firms and Employees, 15 (4): 591–603.

Butler, N., J. Bynner, and University of London, Institute of Education, Centre for Longitudinal Studies. 2016. *1970 British Cohort Study: Ten-Year Follow-Up, 1980,* SN3723.

Butler, N., A. Osborn, S. Dowling, and University of London, Institute of Education, Centre for Longitudinal Studies. 2016. *1970 British Cohort Study: Five-Year Follow-Up, 1975,* SN2699.

Bynner, J., University of London, Institute of Education, Centre for Longitudinal Studies, and N. Butler. 2019. *1970 British Cohort Study: Sixteen-Year Follow-Up, 1986,* SN3535.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil. 2011. "Estimating Marginal Returns to Education" [in en]. *American Economic Review* 101 (6): 2754–2781.

Chamberlain, G., University of London, Institute of Education, Centre for Longitudinal Studies, and R. Chamberlain. 2013. *1970 British Cohort Study: Birth and 22-Month Subsample, 1970-1972,* SN2666.

Coelli, M. B. 2011. "Parental Job Loss and the Education Enrollment of Youth" [in en]. *Labour Economics* 18 (1): 25–35.

Cunha, F., and J. J. Heckman. 2010. *Investing in Our Young People.* Working Paper, Working Paper Series 16201. National Bureau of Economic Research.

Deary, I. J. 2014. "The Stability of Intelligence From Childhood to Old Age" [in en]. *Current Directions in Psychological Science* 23 (4): 239–245.

Farber, H. S., and R. Gibbons. 1996. "Learning and Wage Dynamics" [in en]. *The Quarterly Journal of Economics* 111 (4): 1007–1047.

Fisher, P., L. Fumagalli, N. Buck, and S. Avram. 2019. "Understanding Society and Its Income Data."

Flynn, J. R. 1984. "The Mean IQ of Americans: Massive Gains 1932 to 1978." *Psychological Bulletin* (US) 95 (1): 29–51.

Gunderson, M., and P. Oreopolous. 2020. "Returns to Education in Developed Countries" [in en]. In *The Economics of Education,* Second, edited by S. Bradley and C. Green, 39–51. Academic Press.

Hilger, N. G. 2016. "Parental Job Loss and Children's Long-Term Outcomes: Evidence from 7 Million Fathers' Layoffs" [in en]. *American Economic Journal: Applied Economics* 8 (3): 247–283.

Hopkins, K. D., and G. H. Bracht. 1975. "Ten-Year Stability of Verbal and Nonverbal IQ Scores:" [in en]. *American Educational Research Journal* 12 (4): 469–477.

Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan. 1993. "Earnings Losses of Displaced Workers." *The American Economic Review* 83 (4): 685–709.

Kahn, L. B., and F. Lange. 2014. "Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures" [in en]. *The Review of Economic Studies* 81 (4): 1575–1613.

Lagakos, D., B. Moll, T. Porzio, N. Qian, and T. Schoellman. 2018. "Life Cycle Wage Growth across Countries." *Journal of Political Economy* 126 (2): 797–849.

Lange, F. 2007. "The Speed of Employer Learning." *Journal of Labor Economics* 25 (1): 1–35.

Lindemann, K., and M. Gangl. 2020. "Parental Unemployment and the Transition into Tertiary Education: Can Institutions Moderate the Adverse Effects?" [In en]. *Social Forces* 99 (2): 616–647.

McFall, S. 2013. *Understanding Society: UK Household Longitudinal Study: Cognitive Ability Measures* [in en]. User Guide. Colchester, Essex: Institute for Social and Economic Research, University of Essex.

Mörk, E., A. Sjögren, and H. Svaleryd. 2019. *Consequences of Parental Job Loss on the Family Environment and on Human Capital Formation: Evidence from Plant Closures* [in en]. SSRN Scholarly Paper ID 3445825. Rochester, NY: Social Science Research Network.

Mostafa, T., and R. D. Wiggins. 2014. *Handling Attrition and Non-Response in the 1970 British Cohort Study* [in en]. CLS Working Paper 2014/2. Institute of Education, University of London.

Mulligan, C. B. 1997. *Parental Priorities and Economic Inequality.* Chicago and London: The University of Chicago Press.

Oreopoulos, P., M. Page, and A. H. Stevens. 2008. "The Intergenerational Effects of Worker Displacement" [in en]. *Journal of Labor Economics* 26 (3): 455–483.

Page, M., A. H. Stevens, and J. Lindo. 2009. "Parental Income Shocks and Outcomes of Disadvantaged Youth in the United States" [in en]. In *The Problems of Disadvantaged Youth: An Economic Perspective,* 213–235. University of Chicago Press.

Pan, W., and B. Ost. 2014. "The Impact of Parental Layoff on Higher Education Investment" [in en]. *Economics of Education Review* 42:53–63.

Peter, F. 2016. "The Effect of Involuntary Maternal Job Loss on Children's Behaviour and Non-Cognitive Skills" [in en]. *Labour Economics* 42:43–63.

Rege, M., K. Telle, and M. Votruba. 2011. “Parental Job Loss and Children’s School Performance.” *The Review of Economic Studies* 78 (4): 1462–1489.

Salthouse, T. A. 2010. *Major Issues in Cognitive Aging.* Vol. 49. Oxford Psychology Series. New York: Oxford University Press.

Stevens, A. H., and J. Schaller. 2011. “Short-Run Effects of Parental Job Loss on Children’s Academic Achievement” [in en]. *Economics of Education Review* 20:289–299.

University of Essex, Institute for Social and Economic Research. 2020. *Understanding Society: Waves 1-9, 2009-2018 and Harmonised BHPS: Waves 1-18, 1991-2009,* SN6614.

Willetts, D. 2017. *A University Education* [in en]. Oxford University Press.

Chapter 2

Fertility Choice and Intelligence in Developed Countries

joint with Michele Boldrin and Aldo Rustichini

2.1 Introduction

Our Questions

Intelligence of individuals is known to have significant, and sometimes substantial, effect on life outcomes (Gottfredson (1997), Roberts et al. (2007), Strenze (2007), and Warne (2020)). In all cases considered in these reviews, the effect is positive on outcomes that are presumably individually desirable (such as educational attainment, income, general health and longevity), and negative on presumably undesirable outcomes (such as criminal record or unemployment).

We show that the effect of intelligence on fertility is an exception to this rule: higher intelligence is associated with lower fertility. The reason for this is discussed here. In summary, we propose that intelligence operates, in the case of all the variables indicated earlier, in the way in which a skill operates on a performance index, that is by improving the outcome directly. In the case of fertility, instead, intelligence operates instead by modifying the incentives of individuals, particularly women; these individuals are facing, in their decision to allocate limited time and effort available, the trade-off between career concerns and parenthood. A higher intelligence makes a more ambitious educational attainment and higher career target more realistic; these objectives might not be realistic with lower intelligence. The net outcome on fertility is postponement of parenthood, and over the entire fertile period a reduction of the number of children.

Understanding the reason for such penalty imposed on women with special cognitive talent is important, in particular in the design of public policies. Incentives operate within institutional arrangements, and are produced by the relevant public policies (in the case we are considering, labor market and education policies). To illustrate their effect we consider the relationship between education and parenthood. Demographic research, that we review below, has documented a negative association between fertility and education. But two very different pathways are possible. One pathway is the one usually considered in demographic research: the attitude to parenthood may be modified by education because new values and social norms are acquired with education. While this is an important factor, a different pathway is possible, and is the one we consider in this study: aspiration to education comes into conflict with parenthood. In this second case policies that make easier the pursue of education, in particular higher education, would make the uneven burden imposed on talented women lighter.

We show two important results. The first is that the negative effect we indicate operates only when there is a strict positive complementarity between effort provision and intelligence. The second is to test the general hypothesis in data and show that the effects are significant and large.

Advanced and Developing Economies

The conceptual structure we adopt here is suitable to examine the effects of intelligence on fertility in advanced societies, or more precisely societies in which birth control techniques are widely accessible and their use is the default norm, with little effect on their use induced by the level of education. In less advanced society this negative link may also be operating; but it is likely to work its way, at least in part, through a very different channel: intelligence allows (directly, or indirectly through acquired education) the proper and timely use of birth control devices and procedures. We do not examine these alternative pathways in this study, which is focused on the UK.

Over the recent years researchers have been examining the potential dysgenic effects on the human species of economic progress. [1]. This literature examines evidence in support of the general thesis that reduction of selective pressure may induce a shift in the direction of less fit populations, and in particular it can induce a decrease in the general level of cognitive

1. A controversial example is Lynn (1996)

skills. In our work here (differently from this literature) we examine a potential mechanism at work, within the tradition of analytical economics. We do not specifically attempt to estimate the long term effects of the distribution of intelligence in the population.

Intelligence and Education

A long tradition in the demographics literature explores another variable that has large effects on fertility, namely education. In this research, the general conclusion is that women with higher education are more likely to postpone the birth of the first child, or have fewer children over the fertile period (see Nitsche et al. 2018; Schofer and Meyer 2005; Gerster et al. 2007; Gustafsson 2001; Klesment et al. 2014; Kravdal and Rindfuss 2008; Miller 2011; Stange 2011; Wood, Neels, and Kil 2014). Some exceptions have also been found: see for example (Skirbekk 2008; McCrary and Royer 2011). Our research complements this line of investigation, clarifying some of the factors underlying the regularities that have been established there. As we mentioned already, we consider the hypothesis that intelligence impacts on education through the aspiration to education, in particular higher education, and more generally career concerns. The line of studies that we mentioned instead proceeds taking education that has already been acquired, and considers how education induces a different behavior through the change of norms, lifetime style and aspirations. We consider the hypothesis that an important part of the association between education and fertility found in those papers is induced not by the direct effect of acquired education on lifestyle and preferences, but instead by the aspiration to higher education and the benefits derived from it. The negative effects of fertility on career and earnings have been documented (see for example Lundborg, Plug, and Rasmussen 2017).

The correlation identified between educational attainment and fertility is probably therefore just a byproduct of the true association between education aspiration and fertility.

Multiperiod planning

The intuitive reason for a potential differential impact of intelligence on fertility can be provided already in a simple two-period model, in which a woman has to decide in the first period the number of desired children, and then, in the second, solves the time and effort allocation problem with this number taken as given. This analysis is presented in Rapallini et al. (2021), and tested on an experimental sample. The intuition derived in this simple case might be misleading when we compare it to the more realistic model of many periods, in which the woman plans childbearing over the entire horizon of her fertile years. There are two reasons. First, the differential effect of intelligence might be not a change of the total number of children at completed fertility, but only a shift of the age at which children are born. In the literature we mentioned, the effect of higher education is in fact identified mostly as a postponement of childbearing, and not necessarily as a reduction in number of children at completed fertility.

Second, a crucial feature of parenthood planning over the lifespan is the differential fecundity according to age, which is known to peak around age 24 and then decline. This characteristic is well known to women making the planning, so it is likely to enter into the decision process. It is obviously absent in the two-period model. Note that the first factor has a role independently of the fact that fecundity changes over time, although the latter fact complicates the analysis.

Some recent papers explore female labour supply and fertility choices jointly in the context of the dynamic life-cycle models. Gayle and Miller (2012) and Francesconi (2002) both study how women choose their labour supply and fertility jointly. Both of these papers find that women with a higher earnings profile have lower preferences for children. They do not take into account intelligence as part of the decision-making process.

Adda, Dustmann, and Stevens (2017), Keane and Wolpin (2010), and Sheran (2007) introduce some measure of ability into their dynamic life-cycle models. Adda, Dustmann, and Stevens (2017) allow women to choose consumption, occupation (abstract, routine or manual), mode of work (full-time or part-time) and whether to conceive or not. Women are ex-ante heterogeneous with respect to their characteristics: ability, taste for leisure, taste for children and potential infertility. These variables are latent. Ability is a variable affecting wages earned, not specifically the career or educational success, so although related to intelligence, it does not play the same role as intelligence in our theory. They find that the total fertility and occupational choices of women hardly depend on their ability: "Interestingly, we do not find much difference in terms of total fertility with respect to ability." (Adda, Dustmann, and Stevens 2017, p. 316-317). Instead, women with low taste for children choose steeper career paths, which "induces considerable costs through the sacrifice of fertility." (Adda, Dustmann, and Stevens 2017, p. 316)

Sheran (2007) also builds a dynamic life-cycle model, where women in each period choose labour force participation, school attendance, marriage status and contraception use. The period utility function of women also depend on their observed characteristics, including ability[2]. She finds that higher AFQT score has a strong positive effect on probability of

2. Ability is measured by the Armed Forces Qualification Test (AFQT) score. The test was administered to all the respondents in the sample.

working, attending school and using contraception, and a negative effect on probability of being married. Moreover, when returns to education increase, women stay in school longer, work more and spend less time being married. As a result, the average number of births per woman also goes down.

Our main contribution to this rich and important line of research is twofold. First, we examine how women's choices vary with their intelligence at different stages over the life cycle. And more importantly, we study how intelligence interacts with effort provision in order to produce the documented patterns in labour supply, education and fertility. In particular, we find that it is necessary for intelligence to be complementary with effort provision to explain the outcomes.

Organization of the Paper

The rest of the paper is organised as follows. In the next section 2.2 we develop a model of fertility and career choices. We study the implications with constant and age-dependent fecundity in sections 2.3 and 2.4, respectively. We present numerical analysis of the model with age-dependent fecundity in section 2.5 to illustrate the predictions of the model in the richer setup. In the rest of the paper we test the ideas and predictions of the model in the data. In section 2.6 we describe the data-set used in this paper and examine some preliminary descriptive patterns observed in the data against the general predictions of our model. Next, in sections 2.7 and 2.8, we describe the analysis of cross-sectional and longitudinal models, respectively, testing the theory developed in the earlier sections. In section 2.9 we estimate a cost of effort from observed variables and provide an additional test of the theory. We conclude in section 2.10.

2.2 A Model of fertility Choice

As we explained, the aim of the model we present is to investigate the way in which different intelligence affects the allocation of time and effort between educational attainment (or career advancement) and parenthood.

2.2.1 Setup

An important individual characteristics of each individual is summarized by a parameter in $\Theta = \mathbb{R}_+$, where $\theta \in \Theta$ describes intelligence. The state variable x of his or her planning problem is a vector:

$$x \equiv (n, \theta, h, c) \in \mathbb{N} \times \mathbf{\Theta} \times \mathbf{H} \times \mathbb{N} \equiv X \tag{2.1}$$

consisting of age in number of years (n), intelligence (θ), number of children (c) and human capital (h). We take $\mathbf{H} = \mathbb{R}_+$ with the natural order (larger is better).

The transition function on the first two coordinates (age and intelligence) is independent of the action of the individual. We take the intelligence to be constant over time (this is for simplicity, but it is an assumption that for the relevant age period, say 18 to 45, is approximately true). Time unit is the year, and thus age increases by one unit in every period.

The choice variable a is a pair

$$a \equiv (s, b) \in S \times B \equiv A \tag{2.2}$$

of attempt to have child's birth ($b \in \{0, 1\} \equiv B$), which models the fertility intention, and effort devoted to education and career ($s \in \mathbb{R}_+ \equiv S$).

Utility of leisure depends on the effort and number of children, and is described by a function $L : S \times C \to \mathbb{R}_+$. A wage function (which can be interpreted as utility from wages) gives earnings as function of human capital $W : H \to \mathbb{R}_+$. The variable W should be interpreted as describing the utility for the level of human capital that includes monetary and non-monetary factors, like personal satisfaction.

The space of children is naturally the integers. We set up some notation to deal with this fact. The utility benefit from children is described by a function $U : \mathbb{N} \to \mathbb{R}$. For a function (such as U) $f : \mathbb{N} \to \mathbb{R}$ we say that f is concave if for every $i, j, k \in \mathbb{N}$, $r \in [0, 1]$ such that

$$ri + (1 - r)j = k, \tag{2.3}$$

we have

$$U(k) \geq rU(i) + (1 - r)U(j)$$

(and we say it is strictly concave if the inequality is strict for all $i, j, k \in \mathbb{N}$, $r \in [0, 1]$).

Similarly for functions $g : X \times \mathbb{N} \to \mathbb{R}$, where X is a linear space we can define g to be concave (strictly concave) if we restrict the possible $i, j, k \in \mathbb{N}$, $r \in [0, 1]$ to satisfy 2.3. This is equivalent to:

$$\forall i, \Delta_{i+1} f \leq \Delta_i f$$

where $\Delta_i f \equiv f(i) - f(i-1)$, and strictly concave if a strict inequality holds for all i. Equivalently, the extension of the function f on $\mathbb{R}$ by linear interpolation is concave. We note that the sum of two concave functions is concave, and the sum of a strictly concave and a concave is strictly concave. We say the function f is eventually decreasing if $\Delta_i f < 0$ for all i larger than some finite value. Note that we endow the space of children with the reverse order $c \succeq d$ if and only if $d \geq c$.

We assume:

Assumption 2.2.1. *The per-period utility function depends only on the components (h,c) of the state, and s of the control. Also:*

1. *The per-period utility is separable, in the form:*

$$u(x,a) \equiv L(s,c) + W(h) + U(c) \tag{2.4}$$

2. *W is strictly increasing, differentiable, concave.*

3. *U is strictly concave, and for some integer c_U, $\Delta_{c_U} U < 0$*

Biological fertility is the probability of having a child in one year conditional on attempting ($b = 1$), and is described by a function $\phi : \mathbb{N} \to \{0,1\}$, from biological age to the probability. So at age n the probability of having a child with choice b is $b\phi(n)$. The utility from leisure (equal to 1, time available in one day, minus that spent in various activities) is modelled as follows:

Assumption 2.2.2. *Utility from leisure is described by a function L of the total weighted sum of effort allocated to education and child care:*

$$L(s,c) \equiv l(\gamma_H s + \gamma_R c), \tag{2.5}$$

with l decreasing and concave, and $\lim_{x\uparrow 1} l'(x) = -\infty$.

The total inter-temporal utility is the series discounted by $\delta \in (0,1)$.

2.3 Separable Utility

The analysis is considerably simplified, and a closed form solution is possible, when some restrictive assumption, in addition to (2.2.1) and (2.2.2) are made. The first one consists in ignoring the fact that fecundity changes with age:

Assumption 2.3.1. *The fertility function has a constant value ϕ, that is, for all n, $\phi(n) = \phi \in (0,1)$.*

From assumption 2.3.1 follows that the next period number of children is $c+1$ with probability $b\phi$ and c with the complementary probability. Note that this transition is independent of the choice of s.

The second assumption is a special form of the wage function:

Assumption 2.3.2. *There exists a function $w : S \times \Theta \to \mathbb{R}$ such that, if h' is the random variable denoting the next period human capital,*

$$E_{(h,s,\theta)}W(h') = W(h) + w(s,\theta) \tag{2.6}$$

Examples in which assumption 2.3.2 are satisfied are:

1. $W(h) = log(h)$ and the next period human capital is given by $h' = hY$, with the expectation of $\log(Y)$ depending on (s,x,θ);

2. $W(h)$ linear, $h' = h+Y$, with the expectation of Y depending on (s,x,θ).

Note that the transition to the next period human capital is independent of the choice of b.

Assumption 2.3.3. *The function w has strict positive complementarity, that is, it is strictly supermodular. (Topkis 1978, 1998)*

The separability condition 2.2.1 applies both to the running utility (which is the sum –see equation (2.4) – of three different components, utility from human capital, utility from children, and utility from leisure) and to the transition function. In fact, the number of children next period only depends on the decision variable $b \in \{0,1\}$. Also the human capital next period only depends on the current human capital, intelligence and effort. The number of children affects this transition only indirectly though the effect on the utility of leisure.

Proposition 2.3.4. *Given assumptions (2.2.1), (2.2.2), (2.3.1) and (2.3.2), the value function of the problem (i) can be written as function of (h,c,θ), (ii) it is the solution of the functional* *equation (2.8), and (iii) is separable in the form*

$$V(h,c,\theta) = \frac{1}{1-\delta}W(h) + B(c,\theta) \tag{2.7}$$

where the function B is the unique solution of the functional equation (2.9) below.

Proof. (i): By assumption (2.3.1) the age of the woman is in the current specification of the model is irrelevant, because in the original model the variable age of the mother is only entering into the transition probability for the variable children. So we write the value function as a function of (h,c,θ) only.

(ii): The Bellman functional equation of this problem is:

$$V(h,c,\theta) = \max_{(b,s)} \left(L(s,c) + W(h) + U(c) + \delta E_{(b,s,h,\theta)} V(h',c',\theta)\right). \tag{2.8}$$

(iii): Define the functional equation:

$$B(c,\theta) = u(c,\theta) + \delta \max\left(B(c,\theta), \phi B(c+1,\theta) + (1-\phi)B(c,\theta)\right) \tag{2.9}$$

where

$$u(c,\theta) \equiv U(c) + \max_s \left(L(s,c) + \frac{\delta}{1-\delta} w(s,\theta)\right) \tag{2.10}$$

The Bellman equation (2.8) has a unique solution, so to characterize the solution, it suffices to prove that a function of the form in equation (2.7) solves (2.8), where the function B is the solution of the problem defined in equation (2.9). This claim can be verified by direct substitution of (2.7) into (2.8), using assumption (2.3.2) and the fact that the transition on the human capital variable and the children variable are independent, as noted in the remarks after assumptions (2.3.1) and (2.3.2). □

We now examine the functional equation (2.9). This equation can be interpreted as describing the optimal plan for children, taking into account how θ is affecting the running utility $u(c,\theta)$ by changing the opportunity cost and benefit of larger s as described in (2.10). To characterize its solution, we let

$$c^*(\theta) = \min\{c' : c' \in \mathrm{argmax}_{d\in\mathbb{N}} u(d,\theta)\} \tag{2.11}$$

the optimal number of children for a woman of intelligence θ.

The characteristics of the optimal parenthood policy depend on the solution of the optimal effort problem defined by

$$D(c,\theta) = \max_s \left(L(s,c) + \frac{\delta}{1-\delta} w(s,\theta) \right) \tag{2.12}$$

Lemma 2.3.5. *In the optimal effort problem (2.12):*

1. *The function $c \to D(c,\theta)$ is concave for every θ.*
2. *The function $(c,\theta) \to D(c,\theta)$ is increasing in θ and decreasing in c;*
3. *The optimal effort policy $\theta \to \hat{s}(c,\theta)$ is increasing in θ;*
4. *The function $c \to u(c,\theta)$ is strictly concave, and eventually decreasing.*

Proof. Claim (1) follows by a standard argument adapted to our case of functions concave over the integers: take s_i, s_k, s_j optimal at i, j, k respectively, with an $r \in [0,1]$ such that $ri + (1-r)j = k$, and develop the chain of inequalities. Claim (2) follows because w is increasing in θ and L is decreasing in c. Claim (3) is a consequence of theorem 6.1 in Topkis (1978). Claim (4) follows because also for concave functions over the integers the sum of a concave and a strictly concave is strictly concave, and the sum of two eventually decreasing function is eventually decreasing. □

Proposition 2.3.6. *The equation (2.9) has solution $\hat{B}$ characterized as follows.*

1. *For all $c \geq c^*(\theta)$:*

$$\hat{B}(c,\theta) = \frac{1}{1-\delta} u(c,\theta)$$

2. *For all $c < c^*(\theta)$:*

$$\hat{B}(c,\theta) = (1-\delta(1-\phi))^{-1}(u(c,\theta) + \delta\phi\hat{B}(c+1,\theta))$$

The optimal policy is

$$\hat{b}(c,\theta) = 1 \text{ if } c < c^*(\theta), = 0 \text{ otherwise }. \tag{2.13}$$

Proof. One can check directly that the function satisfying the conditions (1) and (2) above satisfies the equation 2.9. In particular, condition (1) follows because the function $u(\cdot,\theta)$ is strictly concave and eventually decreasing by lemma 2.3.5. Condition (2) follows because the values of the function at all $c < c^*(\theta)$ are computed by backward induction. The final conclusion of policy follows for the feature of the value function $\hat{B}$ in conditions (1) and (2). $\square$

We can now draw our main conclusion:

Corollary 2.3.7. *The optimal parenthood policy $\theta \to c^*(\theta)$, defined in 2.11, is decreasing in θ.*

Proof. Recall that c^* is defined in 2.11 as the minimum in the set of optimal children, but by the fact that $u(\cdot,\theta)$ is strictly concave the set is a singleton, so we will refer to as the optimal solution. By theorem 4.3 of Topkis (1998) the function D defined in 2.12 is supermodular. Now note that for every $\theta \in \Theta$ the function $u(\cdot,\theta)$ is supermodular and satisfies strict increasing differences. The other conditions of theorem 6.1 in Topkis (1978) are satisfied and thus the conclusion that c^* is increasing in the reverse order (so decreasing in the natural one) follows. $\square$

For our purposes, the main result is the conclusion in corollary 2.3.7: intelligence has a negative effects on fertility. It is clear that this property follows because θ changes the tradeoff between effort and children in the maximization in equation (2.12).

2.3.1 A Simple Example

To illustrate the result we consider the specific functional form of the model where:

$$W(h) = \log(h), h' = h(1+\theta s), \tag{2.14}$$

U concave, with unique finite maximum $\hat{c}_U$.

In this case, an easy computation show that:

$$\hat{s}(c,\theta) = \left(\delta(1-\gamma_R c) - \frac{1-\delta}{\theta}\right)^+ \tag{2.15}$$

an increasing function of θ and decreasing function of c. The function u defined in equation (2.10) has the form: [3]

$$u(c,\theta) = U(c) + \frac{1}{1-\delta}\log(1+\theta(1-\gamma_R c)) - \log(\theta) + E(\delta) \tag{2.16}$$

Extending U in a piece-wise linear way and consider the first order conditions it is easy to see that the optimal child defined by $c^*(\theta)$ in this case is in fact decreasing in θ.

2.4 Varying Fecundity

In this section we examine the full model in which the tight restrictions on the functional form imposed in assumption (2.3.2). More important, we now consider in the model an essential feature of the real life problem of women, that fecundity depends on age (so we relax the assumption (2.3.1).

2.4.1 Multi-period Problem

With variables as defined, the individual has a flow utility in every period n of the lifetime (hence n is also the age) equal to:

$$L(s_n, c_n) + W(h_n) + B(c_n) \tag{2.17}$$

and conditional on the state (n, θ, h, c) chooses in every period the fertility intention b_n and the effort s_n; total utility is the expected discounted sum.

From our setup we derive that the transition on the state space leaves intelligence unchanged, increases age by one unit. The number of children can increase by at most one unit, depending on the choice of b and the age of the mother; the human capital increases by a random positive amount that depends on effort and intelligence. Thus the transition can be summarized by the transition on the components human capital and children of the state space.

We assume that the transition on children space and human capital space has a special form, namely it satisfies the following:

3. Here:

$$E(\delta) \equiv \frac{\delta\log(\delta) + (1-\delta)\log(1-\delta)}{1-\delta}$$

Assumption 2.4.1. *1. The probability on the new state is induced by a probability in the change (ξ, ζ) of human capital and children;*

2. *The two random variables ξ and ζ are independent, conditional on the state and choice of control;*

3. *The probability on ξ only depends on (θ, s), is represented by a probability distribution $f(\cdot; s, \theta)$ which is such that for any increasing function $w : \mathbf{H} \to \mathbb{R}$, for any θ, the function of s defined by the inner product:*

$$s \to (f(\cdot; s, \theta), w)$$

is strictly concave.

4. *the probability of ζ only depends on (n, b).*

A simple example of a function satisfying the conditions imposed on f is given by a function $\Pi : \Theta \times S \to [0, 1]$, where so that $\Pi(\theta, s)$ in the probability of transiting from h to $h+1$, if we assume

Assumption 2.4.2. Π *is given by:*

$$\Pi(\theta, s) \equiv \pi(\theta s) \tag{2.18}$$

where π is increasing in its argument and concave.

Optimal fertility intention is defined to be the optimal choice of the variable a as function of the state variables relevant for the choice, that is age, human capital and current number of children; thus it is a variable in $\{0, 1\}$.

Proposition 2.4.3. *Given assumptions 2.2.1, 2.2.2, 2.4.1,*

1. *The value function is the unique solution of:*

$$V(n, \theta, h, c) = \max_{a \in A} (u(x, a) + \delta \tag{2.19}$$

$$\int_{\mathbb{R}_+ \times \{0,1\}} V(n+1, \theta, h+\xi, c+\zeta) df(\xi; s, \theta) dg(\zeta; b, n))$$

2. *For every (n, θ, c) the function*

$$h \to V(n, \theta, h, c)$$

is strictly increasing.

We note that the value function equation in our case can be taken as the solution of the sequential choice (first the birth choice, then the effort choice):

$$V(n,\theta,h,c) = W(h) + B(c) + \max_{b\in\{0,1\}}\{ \tag{2.20}$$

$$\max_s \left(L(s,c) + \delta \int_{\mathbb{R}_+} V(n+1,\theta,h+\xi,c) df(\xi;s,\theta) \right),$$
$$\max_s \left(L(s,c) + \delta E_{b=1} \int_{\mathbb{R}_+} V(n+1,\theta,h+\xi,c+\cdot) df(\xi;s,\theta) \right) \}$$

We will denote $\hat{s}(x,b)$ the solution of the two sub-problems of effort choice in 2.20, conditional on a children birth choice b. The optimal pair at a state $x \equiv (n,\theta,h,c)$ is characterized by:

Proposition 2.4.4. $\hat{a}(x) = (\hat{s}(x),\hat{b}(x))$ *is the solution of:*

$$l'(\gamma_S\hat{s} + \gamma_R c) + \delta \int_{\mathbb{R}_+} E_{\hat{b}} V(n+1,\theta,h+\xi,c+\cdot) df_s(\xi;\hat{s},\theta) = 0 \tag{2.21}$$

$$\textit{if } l'(\gamma_R c) + \delta \int_{\mathbb{R}_+} E_{\hat{b}} V(n+1,\theta,h+\xi,c+\cdot) df_s(\xi;0,\theta) > 0 \tag{2.22}$$

and $\hat{s} = 0$ *otherwise; and*

$$l(\gamma_S\hat{s}(x,\hat{b}) + \gamma_R c) + \delta \int_{\mathbb{R}_+} E_{\hat{b}} V(n+1,\theta,h+\xi,c+\cdot) df(\xi;\hat{s},\theta) \geq \tag{2.23}$$

$$l(\gamma_S\hat{s}(x,b) + \gamma_R c) + \delta \int_{\mathbb{R}_+} E_b V(n+1,\theta,h+\xi,c+\cdot) df(\xi;\hat{s},\theta).$$

Proof. From proposition 2.4.3, the function giving the value of h for fixed (n,θ,c) is strictly increasing, hence by assumption 2.4.1, part 3, the function of s is strictly concave. The condition on the behavior of the function l at zero leisure implies that the optimal effort is either interior or zero. But $\hat{s} = 0$ if and only if the inequality 2.22 does not hold. The condition 2.23 is clear. □

2.5 Numerical Analysis

In this section we assume that human capital accumulation has a simple form, with b a positive real number, where we can choose for instance:

$$\pi(x) \equiv \min\{\max\{x^\alpha, 0\}, 1\} \tag{2.24}$$

for some $\alpha \in (0,1]$.

We also consider

$$l(x) = \eta \log(1-x). \tag{2.25}$$

2.5.1 Features of Optimal Choices and Parameters

General features of the optimal choice depend on a number of crucial parameters that must be identified from data. We list them here.

1. Biological fertility as function of age; that is the function ϕ introduced earlier. The computations below assume a peak at 24-25 years of age, with progressive decline reaching zero above the age of 40.

2. Utility from the number of children; that is the function B introduced earlier. The computations below assume a non-linear concave increasing function. The utility is independent from intelligence and human capital. This assumption is probably not entirely realistic, although the effect of deviation are easy to conjecture.

3. Wage profile as a function of human capital. Clearly, the steeper the function the higher is the incentive to make progress on human capital acquisition.

2.5.2 Value Function

Figure 2.5.1 displays the value function for different values of intelligence and human capital. Both variables increase the value. Due to the complementarity between effort and intelligence, the effect of intelligence is non-linear: at low values of human capital, when career paths are open, intelligence induces large increases in value. As we are going to see, this corresponds to a choice of not having children and preferring investment in career advancement.

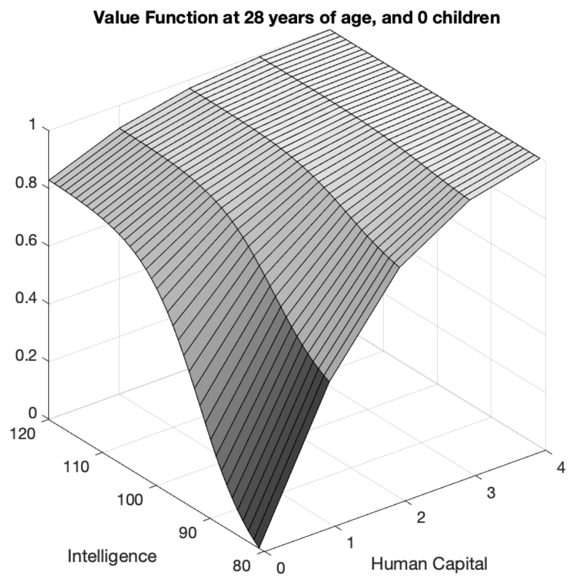


Figure 2.5.1: Value as Function of Intelligence and Human Capital. Intelligence is in normalized score. There are five levels of human capital: a higher number corresponds to higher level of human capital, so we may think of 1 as indicating school, and so on. Age as indicated, and currently no children.

2.5.3 Analysis of Optimal Fertility Intention

Figure 2.5.2 illustrates how fertility intention depends on age and intelligence.

For any given level of intelligence fertility intention has an inverse U-shaped, with a peak that depends on the level of intelligence. For any given age, fertility declines with intelligence.

Figure 2.5.3 illustrates the main result: it reports the difference between the value of fertility intention between high and low level of intelligence by age and number of children.

As the previous figure 2.5.2 makes clear, the lack of difference at high number of children between the high and low intelligence follows because in all cases the value is zero.

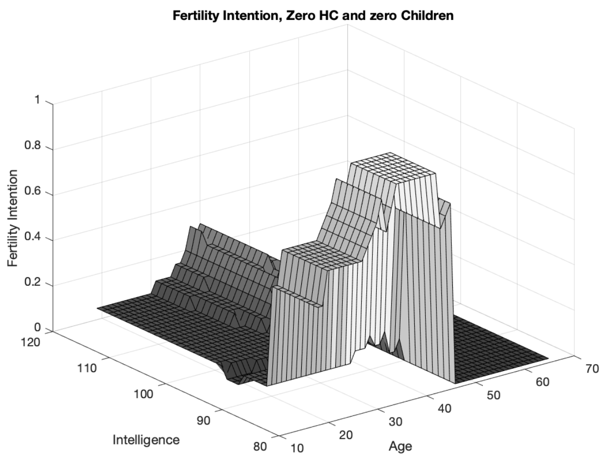


Figure 2.5.2: Fertility intention by age and intelligence. The vertical axis reports the fertility intention (equal to 1 if the intention is positive) by age and intelligence, at zero human capital and zero children.

2.5.4 Analysis of Optimal Effort

We complete the picture reporting the level of effort for each value of intelligence and human capital.

Effort is largest at lower level of human capitals (when career advancement is open), and high level of intelligence (that make these opportunities appealing).

2.6 Descriptive Evidence in the Data

In this section we analyse the evidence in the data and compare them to the general predictions from the model. We begin by describing the dataset used in this paper and key variables of interest.

2.6.1 Data

We use the UK Household Longitudinal Study (UKHLS), also known as Understanding Society. This is the largest household panel study in the UK, covering about 40,000 individuals in each wave since 2009. The participants were sampled from the UK population in 2009 and are followed almost every year. Starting from wave 2, the follow-up sample also includes

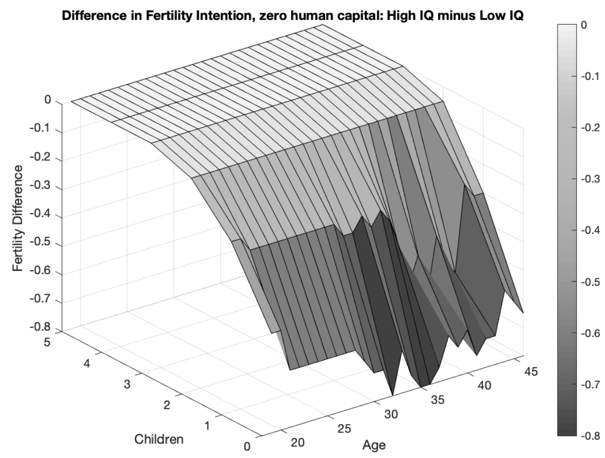


Figure 2.5.3: Difference in fertility intention by age and human capital. The vertical axis reports the difference in fertility intention (equal to 1 if the intention is positive) between high and low levels of intelligence. Intelligence is in normalized score.

former BHPS[4] respondents. The survey encompasses a wide range of topics, including education, employment, fertility and cognitive abilities. We examine in detail the variables that are most important for our analysis.

Education

The survey contains a variable reporting the highest qualification with six categories: tertiary degree, other higher degree, A-level or equivalent, GCSE or equivalent, other and no qualification. This variable is updated in every wave to take into account newly acquired qualifications, if applicable. We also note that individuals could earn a tertiary degree either from traditional universities or from polytechnics, which in turn gained university status in the early 1990s. The UKHLS variable on the highest qualification does not distinguish between the two types of higher education institutions the degree could be obtained from (Chapter 4). We convert the categorical variable on the highest qualification to a binary degree indicator D_i that takes value of 1 whenever individual i reports having a tertiary degree from any type of higher education institution in any wave. In the rest of the paper, whenever we refer to degree, we refer to the binary variable D_i indicating whether the respondents have ever earned a tertiary degree, regardless of the type of higher educational institution they have attended.

4. The British Household Panel Survey is a pre-decessor of the UKHLS. The BHPS ran from 1991 to 2008 covering about 10,000 individuals. In the final wave of the BHPS, the respondents were asked if they wished to continue as part of the UKHLS; about 80% did.

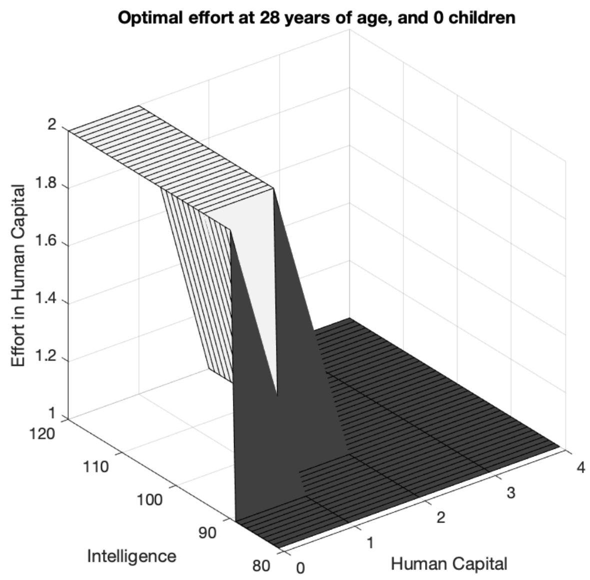


Figure 2.5.4: Effort level by Intelligence and Human Capital. The vertical axis reports the level of effort in human capital acquisition. Intelligence and human capital as in previous figures. Age in years and number of children as indicated.

Fertility

The UKHLS collects information on all household members and characterizes their relatedness. In particular, the dataset contains personal identifiers of mothers and fathers in case the parents live in the same household with children. In addition, the participants in wave 1 of the UKHLS and waves 11 and 12 of the BHPS were asked to provide information about children no longer living in the same household for any reason. Thus, we obtain children's year of birth for each parent-child pair. Using this information, we construct a variable cumulatively counting number of births c_{iat} at each age a of a given parent i at time t. For the cross-sectional analysis, we also use the total number of births ever observed in the dataset starting from age 14 as the proxy for the number of children at completed fertility. For more details see section 2.A.3.

Wages

In each wave the respondents are asked about their employment status, jobs and earnings. We use monthly labour earnings and usual hours worked in a month to construct hourly wages.

We then deflate the hourly wages using the CPI excluding rent, maintenance repairs and water charges, an index recommended by the UKHLS (Fisher et al. 2019).

Intelligence score

In wave 3 the participants were administered a set of five cognitive tests: word recall (immediate and delayed), serial 7 subtraction, number series, verbal fluency and numeric ability. The UKHLS then summarizes the results into counts of correct answers to each test. There are 40,889 individuals with non-missing test results out of the full sample of 49,692 respondents in wave 3. We then estimate the intelligence score using the maximum likelihood confirmatory factor analysis, adapting the model of Johnson and J.Bouchard (2005) to out variables. For more detailed information in the construction of the intelligence score see section 2.A.2.

2.6.2 Data Sample

The sample we work with in our analysis (called the working data set in the following) consists of wave 3 respondents with non-missing intelligence scores. We also restrict our sample to those born between 1950 and 1995 and with non-missing degree indicator. Furthermore, we select only those who have been observed in the survey at least once past the age of 22. This filter helps us remove individuals who have not yet completed their education phase and for whom the degree indicator is not accurate. The final sample consists of 29,360 individuals and 992,852 person-age observations.

We generate two working datasets: longitudinal and cross-sectional. The longitudinal dataset contains up to 49 observations for each person in our working sample. The observations correspond to ages starting from 16 to either 64, or age as of 2020, whichever is smaller. Using the information about all parent-child pairs described in section 2.6.1, we create a variable c_{iat} reporting the number of children of person i at age a at time t. For more details on the creation of this variable see section 2.A.3.

Using observations from all waves of the UKHLS, we also create a variable w_{iat} for the real hourly wage of individual i at age a at time t. But the UKHLS can only provide up to 10 direct wage observations per person. Therefore, we use predicted wages instead of actual wages. To generate predicted wages, we estimate wage age profiles controlling for gender, degree status and number of children. In our estimations we use the restriction dictated by the economic theory that wage age profile is flat towards the end of the career (Heckman, Lochner, and Taber 1998; Lagakos et al. 2018). For more detailed information, see section 2.A.4. Using the estimation results we generate the fitted values excluding time effects, denoted $\hat{w}_{ia}$.

Finally, we generate the cross-sectional dataset from the longitudinal by keeping only one observation per respondent. Before doing so, we construct a time-invariant variable corresponding to age at first birth. For each individual we take the lowest age (censored at 14 from below) at which we observe a non-zero cumulative count of births. We also construct a time-invariant variable with the predicted wage at age 45 (see section 2.A.4 for more details). Then, for each individual we keep the latest observation. At this point, the only time-varying variables we keep are cumulative count of births and age. Therefore, we obtain a cross-sectional dataset where every observation corresponds to a single individual with latest information on the number of children ever had during the entire fertile period starting from age 14 to the latest observed age, latest age when an individual was observed in the survey and the rest of their time-invariant characteristics.

2.7 Cross-sectional analysis

We begin by analysing the relationship between variables of interest and the intelligence score in the cross-sectional dataset. Table 2.7.1 reports the estimation results from simple linear regressions of the indicator of the variable degree, predicted wage at age 45 and number of children on intelligence score and gender.

As expected, intelligence score contributes positively to degree attainment and predicted wages. The effect on wages is slightly lower among women, but remains positive. For example, a one standard deviation increase in the intelligence score of women is associated with a 13.9 percentage point (pp) higher probability of earning a tertiary degree and 0.09 points higher real hourly wages at age 45. These figures correspond to 45% of the share of women with a degree and 14% of the average predicted real hourly wage of women at age 45. We can also see that higher intelligence score contributes negatively to the number of children. A one standard deviation increase in the intelligence score is associated with 0.03 less children for men and 0.16 less children for women, or 2% and 9% of the average number of children of men and women, respectively.

In the analysis of our conceptual framework we argued that the negative relationship between the intelligence score and fertility outcomes is stemming from the career-family trade-off. In particular, since intelligence is associated later in life with better educational and labour market outcomes, utility-maximizing individuals prefer to postpone fertility and give priority to education and career. In Table 2.7.2 we focus on the total number of children

Table 2.7.1: Intelligence and individual outcomes. The table shows relationship between individual outcomes in columns and intelligence scores.

	(1) Degree	(2) $\hat{w}45$	(3) # children ever
female	0.001	-0.145***	0.331***
	(0.005)	(0.004)	(0.013)
IQ	0.135***	0.116***	-0.026*
	(0.004)	(0.003)	(0.012)
female × IQ	0.003	-0.030***	-0.133***
	(0.005)	(0.004)	(0.016)
Constant	0.212***	0.688***	1.711***
	(0.008)	(0.007)	(0.027)
Mean (men)	0.307	0.748	1.337
Mean (women)	0.308	0.605	1.710
Observations	35107	31332	35107

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: $\hat{w}45$ is the predicted wage at age 45. All specifications control for birth cohort FE and parental qualifications. The estimations were weighted by the cross-sectional sampling weights of wave 3. Standard errors clustered at the primary sampling unit level are shown in the parentheses.

ever had (during the entire fertile period starting from age 14 up to the latest wave observed[5]) as the dependent variable and repeat estimations gradually adding controls for educational attainment and predicted wages. Since younger individuals are unlikely to have reached completed fertility, we repeat the estimations in the subsample of people observed at age 40 or older in columns 4-6.

The results are largely unchanged. Results in column 1 are identical to those in Table 2.7.1. A one standard deviation increase in the intelligence score of women is associated with 0.16 fewer children ever had. Going from column 1 to 2, we add controls for degree attainment interacted with both gender and the intelligence score. Now, a 1 standard deviation increase in intelligence score is associated with on average 0.10 less children among women. In column 3 we additionally control for predicted wages at age 45. This brings the average marginal effect of intelligence score further down to 0.04 less children among women. These results suggests that, indeed, part of the negative relationship between intelligence score and fertility outcomes is mediated via education and labour earnings.

The results in column 2 also show that degree status is negatively associated with the number of children: women with degree have on average 0.38 less children compared to

5. See section 2.A.3 for more details.

women without a degree. But the interaction term between degree and intelligence score has a positive coefficient. So, for example, marginal effect of intelligence score is 0.12 less children among women without a degree and only 0.03 less children among women with a degree. The positive contribution of the interaction between education and intelligence can be seen as a sign of complementarity. In the context of individuals optimising their time allocation, higher intelligence may allow them to achieve given results such as degree completion or getting a high-paying job with relatively less effort. This, in turn, leaves them with more time to devote to leisure and/or family formation. The positive coefficient would also be consistent with the quantity-quality trade-off in fertility choices, first described by Becker (1960). At higher levels of intelligence, the time spent with children will be more productive for their development. So, if quality of children enters utility function of parents, higher intelligence may have a positive effect on fertility choices.

Another measure of fertility we examine is the timing of birth of children. In particular, in Table 2.7.3 we study the age at first birth using Cox proportional hazards model. We again examine several specifications gradually adding controls for education and earnings. We also control for the total number of children ever had as a proxy for the desired level of completed fertility. We note that the estimation results show the coefficients in the hazard function. Therefore, negative coefficient means lower hazard and higher time to failure, which in this case is higher age at first birth. In the simple specification without controlling for education and earnings, we can see that higher intelligence score is associated with lower risk of birth; hence, higher age at first birth. Figure 2.7.1 plots the survival functions based on the results from column 1 in Table 2.7.3. The survival functions describe share of people that do not fail, that is do not have births, at a given age. We compute the survival functions for women with intelligence scores at mean (IQ = 0) and 1 standard deviation above the mean (IQ = 1) values. The plot shows that only 31% of women at the mean of intelligence score distribution have no children by age 30. Instead, a one standard deviation higher intelligence score brings the proportion of childless women by age 30 up to 39%. Since the regressions control for the level of completed fertility, these results suggest that women with higher intelligence score postpone their childbearing.

Once we add controls for degree status in column 2 and predicted wages in column 3, the coefficients in front of the intelligence score are lower in magnitude, both for men and women. Thus, we conclude that the negative relationship between intelligence and timing of birth is partially mediated through educational and labour market choices. But unlike the results in Table 2.7.2, intelligence score remains a strong predictor of age at first birth even after controlling for degree status and predicted wages. One potential explanation could be that at

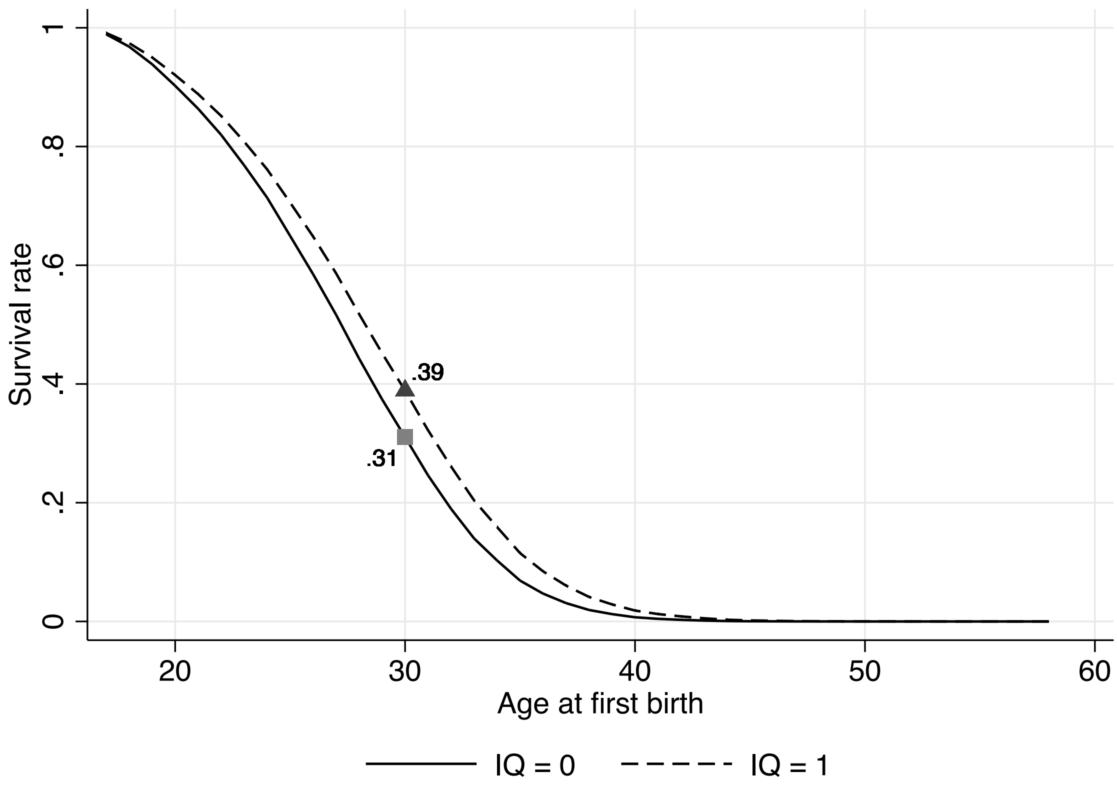
Table 2.7.2: Total number of children ever had. The table shows regression results from simple linear regressions of the total number of children ever had on intelligence score, degree status and predicted wages at age 45.

	All ages			Age 40+		
	(1)	(2)	(3)	(4)	(5)	(6)
female	0.331***	0.402***	0.650***	0.296***	0.357***	0.587***
	(0.013)	(0.017)	(0.052)	(0.016)	(0.020)	(0.056)
IQ	-0.026*	-0.018	-0.050**	-0.018	-0.015	-0.055*
	(0.012)	(0.015)	(0.017)	(0.015)	(0.019)	(0.022)
female × IQ	-0.133***	-0.109***	0.011	-0.129***	-0.105***	0.024
	(0.016)	(0.021)	(0.023)	(0.020)	(0.025)	(0.028)
Degree		-0.137***	-0.546***		-0.096**	-0.481***
		(0.026)	(0.102)		(0.034)	(0.112)
female × Degree		-0.254***	0.077		-0.252***	0.082
		(0.033)	(0.136)		(0.042)	(0.149)
Degree × IQ		0.031	0.033		0.027	0.032
		(0.028)	(0.029)		(0.037)	(0.039)
female × Degree × IQ		0.050	-0.034		0.052	-0.032
		(0.037)	(0.038)		(0.047)	(0.049)
$\hat{w}45$			0.049			0.033
			(0.056)			(0.061)
female × $\hat{w}45$			-0.685***			-0.631***
			(0.085)			(0.091)
Degree × $\hat{w}45$			0.393***			0.368**
			(0.109)			(0.118)
female × Degree × $\hat{w}45$			0.051			0.008
			(0.164)			(0.178)
Constant	1.711***	1.721***	1.745***	1.729***	1.736***	1.761***
	(0.027)	(0.028)	(0.049)	(0.028)	(0.029)	(0.051)
Observations	35107	35107	31332	26204	26204	23175

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The age is latest age when an individual was observed. $\hat{w}45$ is the predicted wage at age 45. All specifications control for birth cohort FE and parental qualifications. The estimations were weighted by the cross-sectional sampling weights of wave 3. Standard errors clustered at the primary sampling unit level are shown in the parentheses.



Notes: The survival function is generated for females at average level of total number of children ever had. The intelligence score is standardized to have mean 0 and standard deviation 1.

Figure 2.7.1: Birth hazard and age at first birth among women. The plot shows the survival function among women at mean and 1 standard deviation above mean intelligence scores using the results from column 1 in Table 2.7.3.

higher levels of intelligence, individuals might spend more time in education and go beyond a bachelor's degree. However, this does not seem to be the case. There is little difference in the distribution of intelligence scores of people with only a bachelor's degree compared to those who have master's or higher degree (see Figure 2.B.1). We estimate another specification in column 4, where we additionally control for the indicator of master's or higher degrees. The estimates of the effect of intelligence score change little between columns 3 and 4. In other words, education and labour earnings are unlikely to be the only channels through which intelligence affects fertility timing decisions. This could arise, if, for example, people with higher intelligence derive higher utility from their leisure at young ages without children.

The estimation results in Tables 2.7.2 and 2.7.3 assume that degree status and predicted wages at age 45 are uncorrelated with the error term. This assumption is clearly violated, if individuals decide on education, fertility and labour supply simultaneously, as we argue in this paper. Therefore, we estimate a system of four equations in Table 2.7.4. The equations describe fertility level and timing decisions, educational decisions and predicted wages at age 45. We acknowledge that these decisions are likely to be completely inter-related, that is, each decision affects another one. However, such system is hard to estimate; therefore, we drop predicted wages from fertility level estimation, and drop predicted wages and number of children from degree probability estimation.

The results are similar to those presented in Tables 2.7.2 and 2.7.3. People with higher intelligence score tend to have fewer children and later in life. These effects are especially pronounced among women. From the wage equation we can see that both degree and the intelligence score contribute positively to the predicted wages, but their interaction has no effect. This result again hints at the complementarity between intelligence score and labour supply decisions. We can also see that the predicted wages of women are on average lower and respond less to degree and the intelligence score. This finding is in agreement with gender differences in wage returns to skills found in Wu (2019).

2.8 Longitudinal Analysis

Our earlier analysis makes evident that timing is an important aspect of fertility decisions. In this section we use the longitudinal structure of our data, which allows us to easily analyse the age profile of fertility. Figure 2.8.1 plots average number of children across ages by gender and intelligence score quintiles of parents. The figure confirms the established result that women with higher intelligence scores tend to postpone childbearing and have fewer children

Table 2.7.3: Age at first birth. The table shows estimation results from Cox proportional hazards model in the full sample where time to failure is age at first birth.

	(1)	(2)	(3)	(4)
female	0.212***	0.278***	0.512***	0.513***
	(0.058)	(0.073)	(0.103)	(0.103)
IQ	-0.097***	-0.068***	-0.060***	-0.060***
	(0.012)	(0.015)	(0.017)	(0.017)
# children ever	0.313***	0.313***	0.323***	0.323***
	(0.026)	(0.032)	(0.040)	(0.040)
female × IQ	-0.116***	-0.085***	-0.070**	-0.070**
	(0.015)	(0.020)	(0.023)	(0.023)
female × # children ever	0.121***	0.119***	0.165***	0.165***
	(0.027)	(0.034)	(0.043)	(0.043)
Degree		-0.397***	-0.347**	-0.235
		(0.080)	(0.117)	(0.134)
female × Degree		-0.123	-0.096	-0.133
		(0.091)	(0.143)	(0.168)
Degree × IQ		0.049*	0.050	0.073*
		(0.025)	(0.027)	(0.036)
Degree × # children ever		0.019	0.020	-0.009
		(0.037)	(0.044)	(0.049)
female × Degree × IQ		-0.019	-0.026	-0.036
		(0.031)	(0.035)	(0.042)
female × Degree × # children ever		-0.026	-0.071	-0.055
		(0.042)	(0.050)	(0.057)
$\hat{w}45$			-0.057	-0.057
			(0.049)	(0.049)
female × $\hat{w}45$			-0.668***	-0.668***
			(0.078)	(0.078)
Degree × $\hat{w}45$			-0.036	-0.082
			(0.084)	(0.105)
female × Degree × $\hat{w}45$			0.326**	0.329*
			(0.123)	(0.150)
Masters or higher	No	No	No	Yes
Observations	24535	24535	21694	21694

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The negative coefficient means lower hazard and higher time to failure, i.e. higher age at first birth. $\hat{w}45$ is the predicted wage at age 45. All specifications control for birth cohort FE and parental qualifications. The estimations were weighted by the cross-sectional sampling weights of wave 3. Standard errors clustered at the primary sampling unit level are shown in the parentheses.

Table 2.7.4: System of equations: fertility, wage and degree probability. The table reports regression results from SEM estimations in the full sample.

	(1) # children	Age birth 1	$\hat{w}$	Degree
female	0.372***	0.469***	-0.113***	0.018
	(0.017)	(0.053)	(0.005)	(0.027)
Degree	-0.103***	-0.383***	0.298***	
	(0.024)	(0.097)	(0.006)	
female × Degree	-0.280***	0.030	-0.050***	
	(0.032)	(0.136)	(0.008)	
IQ	-0.075***	-0.051***	0.077***	0.842***
	(0.012)	(0.014)	(0.003)	(0.025)
female × IQ	-0.096***	-0.071***	-0.030***	0.003
	(0.017)	(0.019)	(0.003)	(0.032)
Degree × IQ	0.037	0.044	-0.015**	
	(0.026)	(0.028)	(0.005)	
female × Degree × IQ	0.078*	-0.031	0.007	
	(0.034)	(0.036)	(0.006)	
# children ever		0.295***	-0.008***	
		(0.009)	(0.002)	
$\hat{w}$45		0.005		
		(0.043)		
female × # children ever		0.211***	-0.008**	
		(0.013)	(0.002)	
female × $\hat{w}$45		-0.723***		
		(0.071)		
Degree × # children ever		0.031	0.018***	
		(0.021)	(0.003)	
Degree × $\hat{w}$45		-0.043		
		(0.089)		
female × Degree × # children ever		-0.114***	-0.003	
		(0.029)	(0.004)	
female × Degree × $\hat{w}$45		0.306*		
		(0.142)		
Constant	1.754***	-7.318***	0.618***	-1.496***
	(0.024)	(0.059)	(0.005)	(0.046)
Observations	35129			

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Equations for the total number of children ever had and predicted wages in columns 1 and 3 are estimated using simple linear regressions. Fertility timing equation in column 2 is fitted using Cox proportional hazards model. The degree equation in column 4 is estimated using logistic regression. $\hat{w}$45 is the predicted wage at age 45. All specifications control for birth cohort FE and parental qualifications.

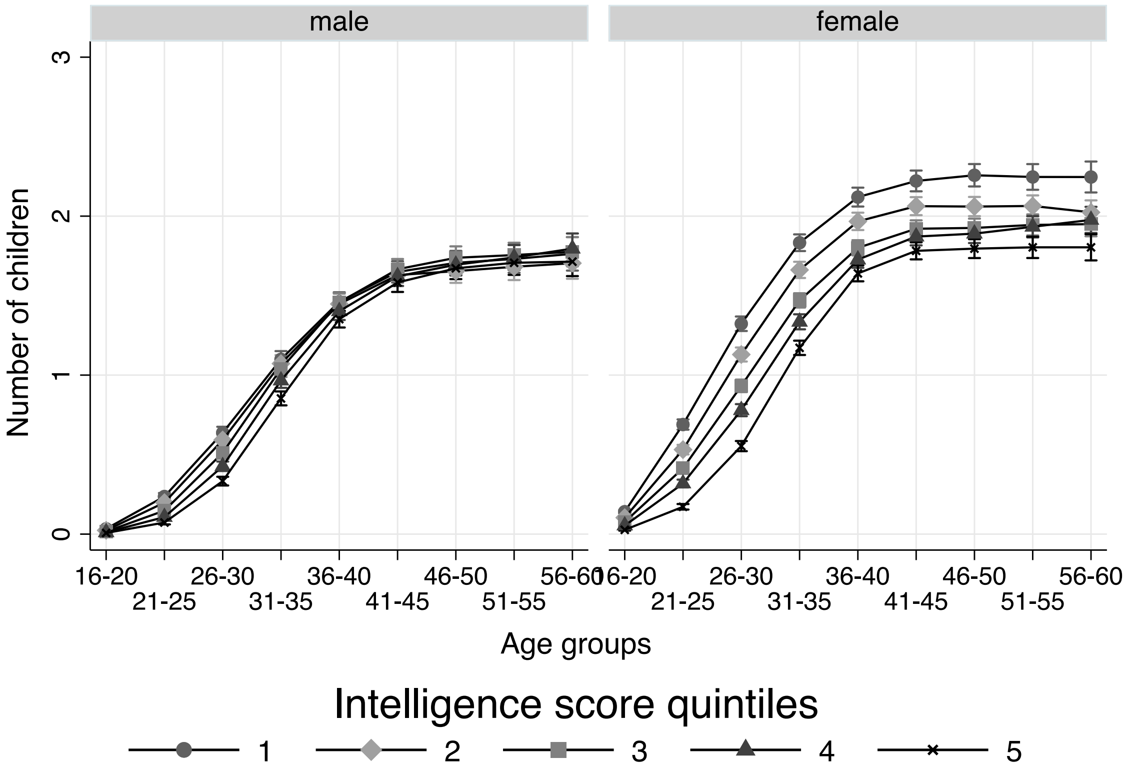


Figure 2.8.1: Fertility profile by age, gender and IQ scores. The plot shows average number of children across age groups by gender and IQ scores of parents. The IQ scores are binned into quintiles. The averages are weighted by the cross-sectional weights as at wave 3.

overall. There is almost a ten-year difference between women at the top and bottom quintile of the intelligence score with one child.

In fact, since reproductive capacity is strongly (negatively) correlated with age, the trade-offs between education, opportunities in the labour market and fertility change over time. To examine this effect we repeat our estimations in the longitudinal dataset. We fit a system of three equations: logistic regressions for probability of birth and probability of having a degree, and a linear regression for predicted wage estimated using fixed effect estimator. These equations allow for age-specific effects of intelligence score and degree indicator. Both intelligence score and degree indicator are constructed as time-invariant variables. This means that the degree indicator identifies individuals who have earned a degree at some point in time, but does not tell us when. The estimations also control for a lagged values of predicted wages and cumulative count of births, consistent with the model in section 2.4. Due to a time lag between conception and birth, these variables enter with a one-period lag.

Table 2.8.1 reports selected results from fertility equation from the system of equations. For brevity we omit the coefficients in front of the lagged wages and number of children. Panel A reports the estimates from regressions among men, panel B - among women. Similar to the estimations in the cross-sectional dataset, in panel C we repeat the estimations among women additionally controlling for the indicator of obtaining a degree at Master's or higher level. The base age group is 16-20 and the coefficients from other age groups should be interpreted relative to the base age group. For example, a one standard deviation higher intelligence score of women reduces their log odds ratio of birth probabilities by 0.196 points at ages 16-20 , but increases it by $0.383 - 0.196 = 0.187$ points at ages 41-45.

Examining first the results at ages 16-20, we can see that birth probabilities are decreasing with education and intelligence score. This is consistent with our hypothesis that fertility choices are significantly affected by the higher likelihood of success with higher intelligence. At higher levels of intelligence, individuals are more likely to invest their time in education to improve their earning potential. As a result, they are considerably less likely to start parenthood at early ages. But as people age, the negative effect of the intelligence score on birth probabilities diminishes and actually turns positive by ages 41-45. For example, a marginal effect of a one standard deviation increase in intelligence score on log odds ratio of birth is -0.196 at ages 16-20 and +0.186 at ages 41-45 among women with no degree. A similar figure for women with a degree is -0.473 at ages 16-20 and +0.166 at ages 41-45. These numbers also imply that the total effect of intelligence score over the entire fertile period is -0.01 among women without a degree and -0.31 among women with a degree. In other words, women with high intelligence that follow a steeper career path not only start

Table 2.8.1: Selected results from longitudinal SEM. The table reports regression results from fertility equation in SEM estimations in the longitudinal dataset. The dependent variable is probability of birth estimated using logistic regression.

		Relative to ages 16-20				
	Ages 16-20	21-25	26-30	31-35	36-40	41-45
Panel A: Men						
Constant	-4.353***	1.403***	1.966***	1.912***	1.608***	0.850***
	(0.070)	(0.076)	(0.074)	(0.075)	(0.083)	(0.107)
Degree	-1.416***	0.352*	1.165***	1.567***	1.777***	1.907***
	(0.161)	(0.173)	(0.165)	(0.165)	(0.167)	(0.177)
IQ	-0.213***	0.061	0.197***	0.202***	0.136*	0.180**
	(0.048)	(0.053)	(0.051)	(0.052)	(0.054)	(0.063)
Degree × IQ	-0.100	0.049	0.089	0.140	0.158	0.017
	(0.166)	(0.179)	(0.170)	(0.169)	(0.171)	(0.180)
Panel B: Women						
Constant	-2.737***	0.647***	0.867***	0.639***	-0.274***	-1.985***
	(0.048)	(0.053)	(0.054)	(0.058)	(0.075)	(0.163)
Degree	-1.609***	0.735***	1.416***	1.799***	1.938***	1.907***
	(0.080)	(0.088)	(0.085)	(0.086)	(0.094)	(0.158)
IQ	-0.196***	0.096**	0.202***	0.287***	0.262***	0.383***
	(0.025)	(0.030)	(0.030)	(0.032)	(0.040)	(0.083)
Degree × IQ	-0.239***	0.031	0.165*	0.218**	0.240**	0.282*
	(0.068)	(0.079)	(0.074)	(0.075)	(0.083)	(0.142)
Panel C: Women controlling for tertiary degree at Master's or higher level						
Constant	-2.740***	0.648***	0.866***	0.637***	-0.273***	-1.987***
	(0.048)	(0.053)	(0.054)	(0.058)	(0.075)	(0.163)
Degree	-1.497***	0.679***	1.346***	1.739***	1.852***	1.894***
	(0.090)	(0.100)	(0.096)	(0.097)	(0.106)	(0.174)
IQ	-0.196***	0.096**	0.202***	0.287***	0.262***	0.382***
	(0.025)	(0.030)	(0.030)	(0.032)	(0.040)	(0.083)
Degree × IQ	-0.277***	0.066	0.207*	0.251**	0.258**	0.257
	(0.079)	(0.091)	(0.086)	(0.087)	(0.096)	(0.162)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The education and wage equations are omitted from the table. Probability of birth is estimated using logistic regression and additionally controls for birth cohort fixed effects and lagged values of wage and number of children.

childbearing later, but have fewer children overall. These results are consistent with findings of Adda, Dustmann, and Stevens (2017). They too reach a conclusion that career choices of high-ability women comes at a cost of lower fertility.

We can also see that the interaction term between degree status and the intelligence score is large and negative at ages 16-20, close to zero at ages 31-35, and turns positive positive among women at ages 41-45. That is, higher intelligence among women with degree has even more negative effect on birth probabilities at early ages and even more positive effect at later ages compared to similarly intelligent women without a degree. These results confirm the evidence we have seen from the cross-sectional analysis.

The larger negative effect of intelligence on fertility choices at early ages of women pursuing degree qualifications could be related to overall duration of education. Although the evidence from cross-sectional analysis does not support this explanation, we test it again in the panel regressions. In column 3 of Table 2.8.1 we additionally control for indicator whether individuals have degrees at Master or higher levels. We again find that the estimates remain virtually the same compared to column 2, suggesting that the length of educational period is unlikely to explain the additional negative effect of intelligence score among women with degrees. An alternative explanation could be that higher education provides not only with training opportunities, but also with more appealing leisure opportunities, both of which may reduce time allocated to family management. For example, Gayle and Miller (2012) find that "education increases the value of leisure" (p.13).

The larger positive effect of intelligence on fertility choices at later ages of women with degree qualifications could also be related to the positive complementarity between intelligence and effort provision at work. Human capital that they have acquired early on in the career can help them enjoy higher-paying job at later ages with relatively less additional effort compared to a similarly-intelligent woman without a degree. Thus, their optimal time allocation decisions may substitute work time for more leisure or family time. Since ages 41-45 correspond to the end of the reproductive capacity for the majority of women, women are likely to choose to allocate this time more to childbearing than leisure. Thus, the results in Table 2.8.1 highlight the importance of age-specific reproductive capacity when studying fertility choices of women in a dynamic context.

2.9 Estimate of Cost of Effort

Both cross-sectional and longitudinal results suggest a complementarity between effort provision and intelligence in the decision-making process of an individual. Theory and numerical analysis of the model in section 2.2 show that such complementarity leads to highest effort being exerted by agents with high level of intelligence and low level of human capital. The model also predicts that individuals with high cost of effort will invest less into their human capital at the beginning of their career. In this section we test these hypotheses relying on an estimate of the effort cost for each individual in the sample.

We do not observe either effort provision or effort cost directly; but we can construct a proxy of the cost of effort using parental characteristics and personality traits scores. Among parental characteristics, we consider parents' educational qualifications and employment status when the respondents were 14 years old. The neuroticism score, one of the Big 5 personality traits, was collected by the UKHLS during wave 3 (2011-13), and is the most likely trait to have a negative effect on education attainment. Using factor analysis, we combine these variables into a proxy of the cost of effort that we denote $\hat{\eta}$. The higher is the $\hat{\eta}$ score, the higher is the cost of effort. We standardise the proxy cost of effort to have mean zero and unit standard deviation. For more details on the construction of the score see section 2.A.5.

Table 2.9.1: Cost of effort in fertility equation of women. The table reports logistic regression results from fertility equation from longitudinal SEM estimations among women. The dependent variable is birth probability.

		Relative to ages 16-20				
	Ages 16-20	21-25	26-30	31-35	36-40	41-45
Constant	-2.798***	0.700***	0.998***	0.818***	-0.166	-1.896***
	(0.057)	(0.064)	(0.065)	(0.069)	(0.087)	(0.186)
Degree	-1.552***	0.724***	1.355***	1.726***	1.834***	1.872***
	(0.094)	(0.103)	(0.100)	(0.101)	(0.109)	(0.179)
IQ	-0.170***	0.091*	0.167***	0.237***	0.189***	0.369***
	(0.031)	(0.037)	(0.036)	(0.039)	(0.047)	(0.098)
Degree × IQ	-0.212**	-0.015	0.150	0.203*	0.231*	0.217
	(0.080)	(0.091)	(0.086)	(0.087)	(0.095)	(0.160)
$\hat{\eta}$	0.100***	0.003	-0.087**	-0.117***	-0.110**	-0.107
	(0.029)	(0.033)	(0.032)	(0.032)	(0.037)	(0.066)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The education and wage equations are omitted from the table. Probability of birth is estimated using logistic regression and additionally controls for birth cohort fixed effects.

We repeat the SEM estimations from section 2.8 among women adding age-specific effect of $\hat{\eta}$ as an additional regressor. Table 2.9.1 reports the estimation results from fertility equation that estimates birth probabilities as a logistic function of regressors. Similar to the previous results, the base group is ages 16-20 and the rest of the coefficients should be interpreted relative to the base age group. For example, a one standard deviation increase in the intelligence score is associated with 0.238 higher log odds ratio of birth probabilities at ages 31-35 compared to the same intelligence effect at ages 16-20. Recall that we constructed the degree indicator as time invariant: that is, it identifies sample members who eventually earned a degree at some point in life.

The results suggest that women with higher cost of effort have higher probability of giving birth at young ages up to age 25. This is consistent with the theoretical predictions of the model. The effort is largest among high-intelligence individuals at the beginning of their career, when they are young. Young women with high cost of effort would optimally invest less into their human capital, making them more likely to begin childbearing earlier, as predicted by the model and observed in the data. We can also see that the cost of effort has virtually no effect on birth probabilities past age 30. That is, the trade-off between effort and fertility eases as women age. This could be related to decreasing fecundity and/or lower potential expected payoffs to human capital investments made at older ages[6].

2.10 Conclusions

Overview of Results

The model we presented provides a clear link between intelligence and fertility choice in advanced societies, where birth control techniques are widely available and accepted by prevailing social norms. The link is provided by the trade-off between alternative use of limited resources, the time and total effort available for child bearing and rearing on the one hand, and career concerns on the other. For biological and cultural reasons, the largest burden is likely to fall on the woman. The conceptual structure suggests that the higher intelligence should induce lower fertility, particularly among women.

Our data analysis supports these predictions, identifying a negative coefficient of intelligence on achieved fertility. The fact that the coefficient is larger for women is in agreement with the general idea in the model that the crucial pathway from intelligence to fertility is in the opportunity cost of effort in parenthood. Alternative explanations of the association could be (for instance) that individuals with different intelligence derive different utility from

6. Jayachandran and Lleras-Muney (2009) find that human capital investments increase with life expectancy "because a longer time horizon increases the value of investments that pay out over time." (p. 349) Sheran (2007) finds that "the gains from schooling are decreasing in age, most likely because the desire for additional years of schooling declines as the number of years women have to reap the benefits decreases." (p. 384)

parenthood. Our data indicate that such a difference in taste should be different in the two sexes. Although not implausible, Occam's razor cuts twice (a first time because there is no need to invoke a taste difference, and a second time because this hypothetical difference should also be stronger among women) in favor of the more parsimonious explanation we suggest.

Future Research

Further research on this topic seems necessary. Data sets large as the Understanding Society are available, for advanced countries (in particular USA, Germany, Australia) and the hypothesis will be tested in future research in these countries. The same fundamental mechanism is likely to operate differently in different economic, social cultural and institutional conditions.

Intelligence reducing fertility is not a natural necessity. Several factors may work in the opposite direction, for instance more intelligent individuals may have higher income, or may be better managers of their time. So effective policies may be possible. Our results point to an additional layer on top of the demographic crisis common to all the advanced society.

Appendix 2.A Data Used in the Analysis

Subsection 2.A.1 Understanding Society Data

The data we use in the analysis below are derived from the UK Household Longitudinal Study (UKHLS: https://www.understandingsociety.ac.uk/), also known as Understanding Society. It is a large household panel study for the UK. It includes about 40,000 individuals and started in 2009.

A previous longitudinal study, smaller in size, called British Household Panel Survey (BHPS), with about 10,000 participants, was conducted between 1991 and 2009. The BHPS could be thought of as the predecessor of the UKHLS. Questionnaires for the UKHLS were built on those of the BHPS, and and effort was made to insure the continuity between the two studies: 80 per cent of the participants in the last wave of the BHPS joined the UKHLS.

Subsection 2.A.2 Intelligence Score

In wave 3 (2011-13), the UKHLS administered give cognitive ability tests to the participants. The intelligence score was constructed using counts of correct answers from each test. [7] Specifically:

1. An index of *memory*, based on a word recall task: individuals are asked to recall as many as they can from a list of ten words. There were two parts, immediate (variable *cgwri_dv*: participants were asked to recall words immediately after presentation) and delayed (variable *cgwrd_dv*: participants were asked to recall words after the *serial 7 subtractions* test).

2. An index *serial 7 subtractions* (variable $cgs7cs\_dv$ for number of correct subtractions, as opposed to the number of correct answers). Individuals were asked to subtract 7 from the previous number, starting from 100, for five times in sequence.

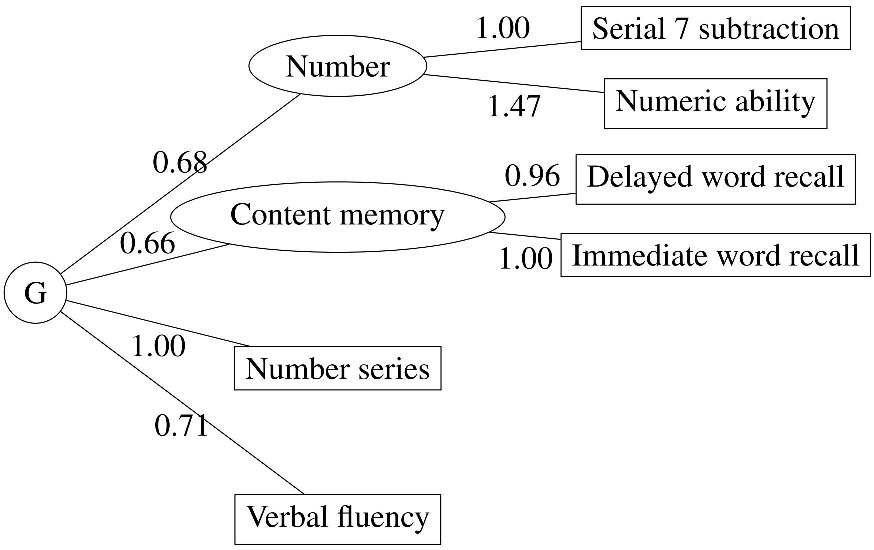
3. An index *number series* (variables $cgns1sc6\_dv$ and $cgns2sc6\_dv$). This task requires the respondent to look at a series of numbers. One number is missing from the series, and the respondent must provide the missing number in the series, after identifying the pattern in the series.

7. See a detailed description of the tasks in the User's Manual, pages 275-324. A user's guide specifically devoted to "concepts, questions and measures for the cognitive ability modules included in Wave 3 of the main and Innovation Panel surveys" is here. The complete set of instructions is reported here. The document illustrating some of the test already developed for the HRS are here.

4. An index of *verbal fluency* (variable $cgvfc\_dv$). Individuals were asked to name as many unique animals as they could in one minute.

5. An index of *numerical ability* (variable $cgna\_dv$). The variable is coded on a 0 to 5 scale, given by the number of items answered correctly. The tasks consisted of solving simple numerical problems based on everyday day life examples (such as computing the correct change after a purchase).

The correlation between the six scores is positive for all pairs (Table 2.A.1). We combine the scores into a single intelligence score using confirmatory factor analysis, following Johnson and J.Bouchard (2005). We first match the UKHLS cognitive tests to the tests used by Johnson and J.Bouchard (2005) based on the description of tasks. We also standardize the scores within each year of birth and gender cells to abstract from possible age-related differences in performances. We then adapt and simplified their preferred model and estimate the following model using maximum likelihood confirmatory factor analysis:



Notes: The round shapes mark latent variables and rectangular shapes mark observed test scores. The numbers on edges are the coefficients from the confirmatory factor analysis.

Figure 2.A.1: Structural model of intelligence score, adapted from Johnson and J.Bouchard (2005).

Figure 2.A.2 plots the densities of the predicted latent variable scores: general intelligence score (G), number score and content memory score. The distributions are negatively skewed.

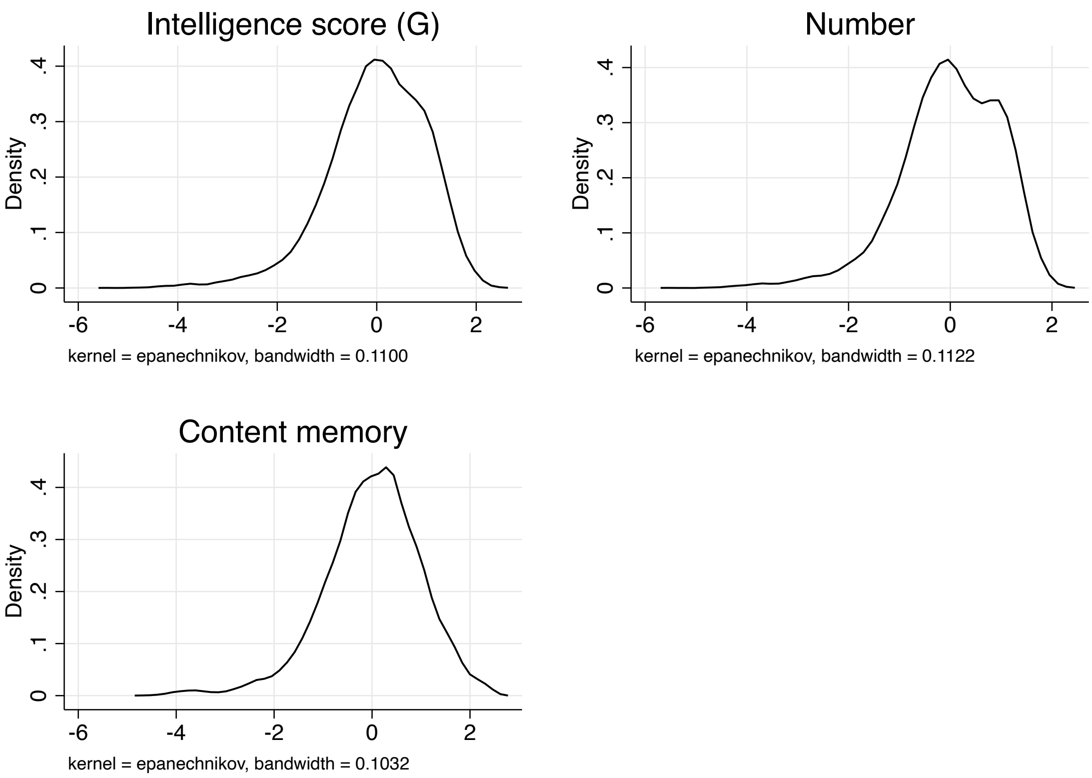
Table 2.A.1: Correlation among the cognitive test scores.

	IWR	DWR	S7S	NS	VF	NA
Immediate word recall (IWR)	1.0000					
Delayed word recall (DWR)	0.7248	1.0000				
	(0.0000)					
Serial 7 subtraction (S7S)	0.2247	0.1974	1.0000			
	(0.0000)	(0.0000)				
Number series (NS)	0.2846	0.2666	0.3321	1.0000		
	(0.0000)	(0.0000)	(0.0000)			
Verbal fluency (VF)	0.3916	0.3601	0.2343	0.3018	1.0000	
	(0.0000)	(0.0000)	(0.0000)	(0.0000)		
Numeric ability (NA)	0.3557	0.3232	0.4194	0.5000	0.3805	1.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	

Notes: IWR - immediate word recall, DWR - delayed word recall, S7S - serial 7 subtraction, NS - number series, VF - verbal fluency, NA - numeric ability. p-values in parenthesis.

Table 2.A.2: Cronbach's alpha for the six scores. ITS: item-test correlation, IRC: item-rest correlation; AITC: average inter-item covariance

Item	Obs	Sign	ITC	IRC	AITC	α
Immediate word recall	32,862	+	0.6233	0.4945	1.5280	0.4351
Delayed word recall	33,675	+	0.6391	0.4549	1.4490	0.4266
Serial 7 subtraction	32,630	+	0.3889	0.2827	2.0101	0.5085
Number series	31,687	+	0.5174	0.4112	1.8681	0.4899
Verbal fluency	33,272	+	0.8855	0.4242	0.7010	0.7110
Numeric ability	33,211	+	0.5345	0.4356	1.8156	0.4775
Test scale					1.5633	0.5272



Notes: The figure shows density plots of the predicted latent variable scores from model in Figure 2.A.1.

Figure 2.A.2: Distribution of the latent intelligence factors.

Subsection 2.A.3 Fertility

Constructing fertility panel dataset

The datasets from each wave of the UKHLS contain information on all household members, including newborns and children. In addition to that, the participants from wave 1 of the UKHLS and waves 11 and 12 of the BHPS were asked to provide information on all non-resident children. We use both of these sources to create a dataset of parent-child pairs, where using the year of birth of parents and children we infer the age at birth for each observation.

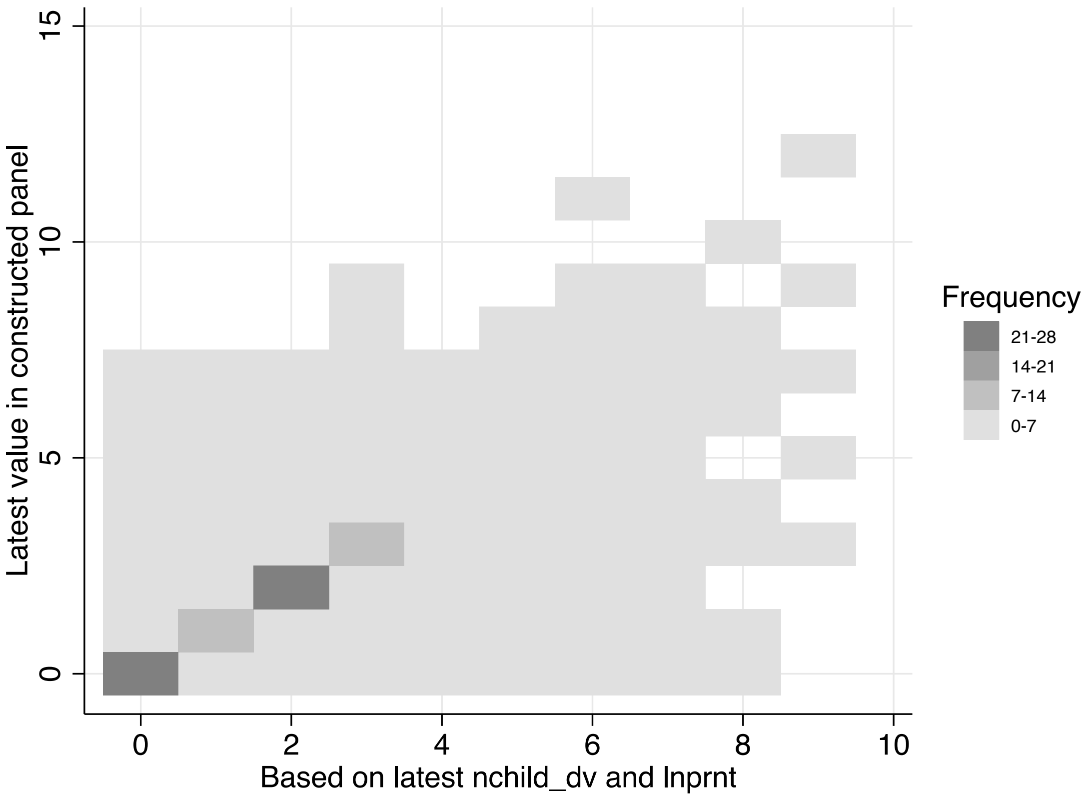
Next, we select all respondents from wave 3 of the UKHLS and rectangularize the dataset such that each individual has one observation per age between 14 and 64. That is, we create a rectangular panel with $N \times 50$ observations, where N is number of respondents from wave 3 of the UKHLS. We then merge it with the dataset of parent-child pairs based on parent identifiers and age. As a result of this merge, we create a variable $nbirth_{a}ge$ which describes number of births individuals had at a given age. Using this variable, we also compute cumulative number of births at each $nbirth_{c}um$, which we use as a proxy for the number of children at each age. It is only a proxy for the number of children for two reasons. First, cumulative number of births contains children that have subsequently died. Second, our sample filters at a level of parent-child pairs, which we describe below, may introduce measurement errors.

This procedure should ideally reflect information about number of children respondents ever had. In practice, our algorithm drops children with missing or inconsistent year of birth information. The missing year of birth issue is mostly prevalent among non-resident children, since parents could refuse to provide information or, in some cases, they report not knowing the year of birth. There are also cases of inconsistent information, where the implied age of parent at birth is either below 14 or above 50 for women and 70 for men. Since these issues mostly arise with non-resident children, we cannot use other variables and waves to correct these irregularities and are forced to drop these observations.

In order to assess the quality of the constructed fertility panel we compare the cumulative count of children with responses given in main questionnaires on number of biological children ever had (based on *nchild_dv* and *lnprnt*). In figure 2.A.3 we plot the bivariate distribution focusing on the latest information available for each person. We can see that most of the sample is concentrated on the diagonal. The correlation coefficient is 0.8.

Comparison of fertility measures by gender

Table 2.A.3 shows that women in the UKHLS are more likely to have children and on average have more children compared to men. Part of the gap can be explained by well-known age



Notes: The number of children ever had self-reported in the main questionnaire is on the x-axis. The number of children ever had implied from the constructed fertility panel is on the y-axis.

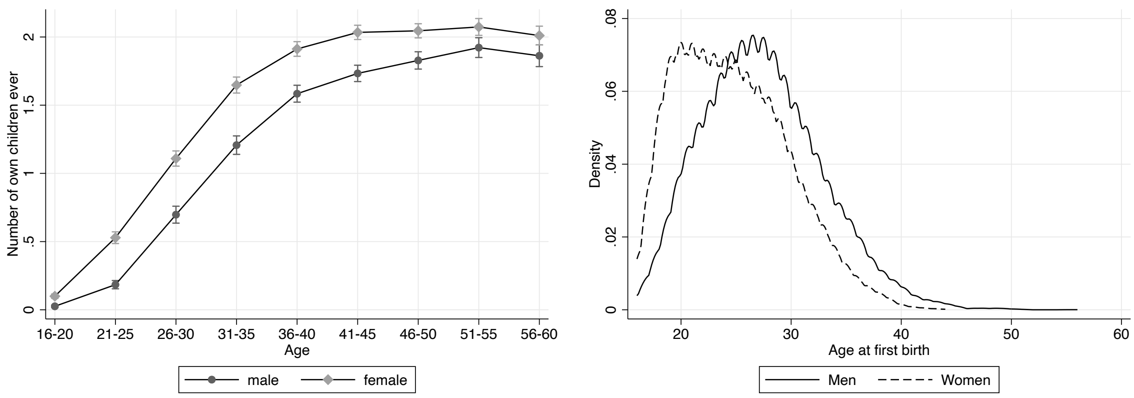
Figure 2.A.3: Bivariate distribution of reported and constructed number of children ever had.

differences: men typically have children later in life. For example, in the UKHLS men are typically two years older when they have their first-born child (figure 2.A.5b). This is similar to almost a three-year age difference between male or female respondents and their partners. Thus, men are likely to have fewer children compared to similarly-aged women. Since age distribution of male and female respondents in the UKHLS is similar, this can explain lower average male fertility in the sample.

Table 2.A.3: Fertility statistics by gender.

	Women			Men		
	N	mean	sd	N	mean	sd
Age at interview	19,722	37.50	12.42	15,180	37.11	12.45
Has any children	19,722	0.71	0.46	15,180	0.59	0.49
Number of own children ever	18,745	1.53	1.35	14,358	1.23	1.33
Number of own children in HH	19,722	0.71	1.01	15,180	0.57	0.96
Age at first birth	8,463	24.70	5.03	5,805	27.34	5.38
Partner characteristics						
Ever had a stable partner	19,722	0.85	0.36	15,176	0.78	0.41
Age of partner	9,791	43.23	11.31	8,380	39.56	10.49

Notes: The table is constructed using information as at wave 1 of the UKHLS. The sample includes only individuals between ages 16 and 60 (born from 1950 onwards). The summary statistics are weighted by individual sampling weights.

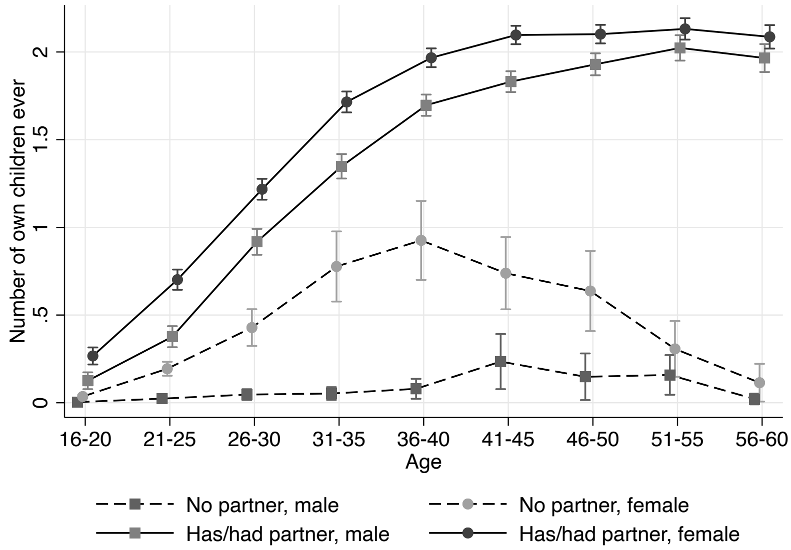


(a) Number of own children ever **(b) Age at first birth**

Notes: The plot shows average number of children across individuals of different ages and distribution of age at first birth by gender. The figures are constructed using information as at wave 1 of the UKHLS. The sample includes only individuals between ages 16 and 60 (born from 1950 onwards). The means are weighted by individual sampling weights.

Figure 2.A.4: Distribution of fertility statistics by gender.

Another possible explanation is measurement error. For example, some men may not know about existence of their children. This situation is more likely to arise if the parents were not in a stable relationship at the time of conception. We construct a stable partnership indicator as having either been married or cohabited with a partner. Figure 2.A.6 shows that indeed the gender gap in number of children is highest among people that report not having had a stable partnership. Moreover, there is no catch-up-later pattern between men and women with no stable relationships: men at all ages report having no children.



Notes: The plot shows average number of children across individuals of different ages by gender and partnership status. The figure is constructed using information as at wave 1 of the UKHLS. The sample includes only individuals between ages 16 and 60 (born from 1950 onwards). The statistics are weighted by individual sampling weights. Stable partnership indicator is constructed as having either been married or cohabited.

Figure 2.A.6: Age-specific gap in number of children between men and women by type of partnership.

Subsection 2.A.4 Predicted earnings

We use the UKHLS dataset to construct also a panel of earnings at each age for each respondent. The latest release of the UKHLS data set includes ten waves, and therefore we have at most ten wage observations for each worker. More importantly, these ten observations cover different sections along the age profile of wages depending on the year of birth of the respondents.

We address this issue by using predicted wages according to estimated wage age profiles. For wage age profiles, we estimate the following equation

$$w_{iat} = \alpha_t + \eta_t \mathbf{X}_i + \gamma_a + \beta_a \mathbf{X}_i + \lambda c_{iat} + \kappa c_{iat} \mathbf{X}_i + \mu_i + \nu_{iat} \tag{2.26}$$

where w_{iat} are real hourly wages of person i at age a observed in year t, $\mathbf{X}_i$ are individual characteristics (gender and highest qualification) and c_{iat} is the number of children of person i at age a at time t. The regression controls for time fixed effect α_t, age fixed effect γ_a and allow variables in $\mathbf{X}_i$ to have age-specific β_a and time-specific η_t effects on wage.

Following the preferred specification in Lagakos et al. (2018), we also impose age restrictions as in Heckman, Lochner, and Taber (1998). The restriction is that wage profiles are flat between ages 51 and 60. In terms of parameters in equation 2.26 this restriction can be written as $\delta_a = \beta_a = 0, \forall a \in [51, 60]$.

We then calculate predicted wages as

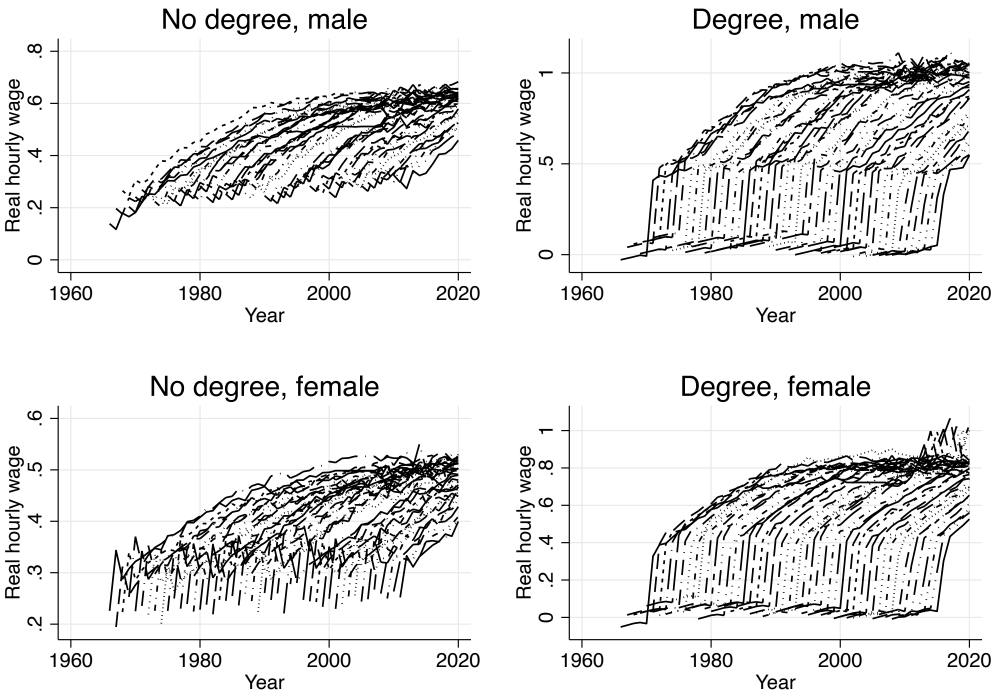
$$\hat{w}_{iat} = \hat{\gamma}_a + \hat{\beta}_a \mathbf{X}_i + \hat{\lambda} c_{iat} + \hat{\kappa} c_{iat} \mathbf{X}_i + \hat{\mu}_i \tag{2.27}$$

Notice that equation 2.26 flexibly controls for year effects without imposing functional assumptions on time trend. The drawback is that we cannot extrapolate the year effects to out-of-sample observations. Therefore, we remove year effects $\hat{\alpha}_t + \hat{\eta}_t \mathbf{X}_i$ from predicted wages.

Figure 2.A.7 plots average $\hat{w}_{it}$ within each birth cohort (i.e., all people born in the same year) by gender and degree status. We can see that the predicted profiles are consistent with several well-known observations. First, female wages are on average lower than male wages. Second, wages of workers with degree are higher compared to those without a degree. And third, wage-age profiles are steeper among degree-holders compared to workers without a degree.

In addition to predicting entire wage profiles, we also generate a prediction at age 45, $\hat{w}_{i45}$, for cross-sectional analysis. Denote $\bar{c}_{i45}$ average number of children at age 45 conditional on $\mathbf{X}_i$. In other words, we compute average number of children of 45-year-olds in each gender-degree cell. Depending on which cell individual i is in, the variable $\bar{c}_{i45}$ takes the corresponding value. Then, $\hat{w}_{i45}$ can be written as

$$\hat{w}_{i45} = \hat{\gamma}_{45} + \hat{\beta}_{45} \mathbf{X}_i + \hat{\lambda} \bar{c}_{i45} + \hat{\kappa} \bar{c}_{i45} \mathbf{X}_i + \hat{\mu}_i \tag{2.28}$$



Notes: Each line corresponds to average predicted real hourly wage within a single birth year cohort between 1950 and 1995 in a given gender-degree cell.

Figure 2.A.7: Predicted wage profiles

Subsection 2.A.5 Cost of effort

To construct the proxy for cost of effort we apply factor analysis to parental characteristics and neuroticism score. The UKHLS contains information about the highest educational qualifications of each of respondents' parents. It is a categorical variable with the following values: no education, some school with no qualifications, school with qualifications, post-school qualifications, degree, or other. The UKHLS also asked participants to think back to when they were 14 years old and report whether their parents were working, unemployed, deceased or absent at the time. The neuroticism score is part of the Big 5 personality traits questionnaire administered during wave 3 of the UKHLS (2011-13).

Table 2.A.4 reports correlation measures between individual outcomes and components of the cost of effort. Higher values of parental outcomes generally have a positive effect on individual outcomes of children. Only mother's employment status seems uncorrelations with future degree attainment of her children. The neuroticism score has a relatively smaller but negative impact on individual outcomes.

Next, we apply exploratory factor analysis to the components of the cost of effort. Table 2.A.5 shows the rotated factor loadings of the first five factors. In general, more favourable conditions such as having employed and university-educated parents have positive contributions. Factor 2 suggests that individuals whose mother only has school qualifications have higher score than the ones whose mother continued her education, but not to the university level. Factor 3 has the opposite picture of school versus post-school qualifications of fathers. Individuals whose father continued education beyond school, even if not all the way to the universities, have higher score. These results could reflect the relative value of the time parents spent in the labour market versus time spent with children. Factor 4 picks up the neuroticism score.

We denote the difference of between factor 4 and the sum of factors 1-3 and 5 as the proxy of the cost of effort and call it $\hat{\eta}$. Table 2.A.6 shows the relationship between individual outcomes and $\hat{\eta}$. As expected, individual outcomes and $\hat{\eta}$ display negative correlation.

Appendix 2.B Supplementary Tables and Figures

Table 2.A.4: Components of the cost of effort and individual outcomes. The table shows the estimation results from the regressions of the variables in columns on the components of the cost of effort.

	Degree	Predicted wage at 45
Neuroticism	-0.008**	-0.023***
	(0.002)	(0.002)
Mother: degree	0.219***	0.129***
	(0.017)	(0.010)
Father: degree	0.289***	0.149***
	(0.015)	(0.009)
Mother: post-school quals	0.129***	0.088***
	(0.011)	(0.008)
Father: post-school quals	0.073***	0.044***
	(0.010)	(0.007)
Mother: school quals	0.078***	0.068***
	(0.009)	(0.006)
Father: school quals	0.041***	0.030***
	(0.010)	(0.007)
Mother was working	-0.009	0.018***
	(0.007)	(0.005)
Father was working	0.064***	0.046***
	(0.010)	(0.008)
Constant	0.141***	0.625***
	(0.014)	(0.010)
R^2	0.094	0.082
Observations	27674	24794

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.A.5: Rotated factor loadings. Factor loadings after promax rotation. The table only reports loadings higher than 0.2 in absolute value.

	Factors					
Variable	1	2	3	4	5	Uniqueness
Neuroticism				0.2600		0.9892
Mother: degree	0.6145					0.5857
Father: degree	0.7000					0.5312
Mother: post-school quals		-0.5843				0.5601
Father: post-school quals			0.5831			0.5542
Mother: school quals		0.5941				0.5633
Father: school quals			-0.5841			0.5744
Mother was working					0.3905	0.8765
Father was working					0.3499	0.9251

Table 2.A.6: Individual outcomes and cost of effort. The table shows estimation results from regressions of the variables in columns on η.

	Degree	Predicted wage at 45
$\hat{\eta}$	-0.101***	-0.056***
	(0.004)	(0.002)
Constant	0.303***	0.690***
	(0.004)	(0.003)
R^2	0.048	0.036
Observations	27674	24794

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

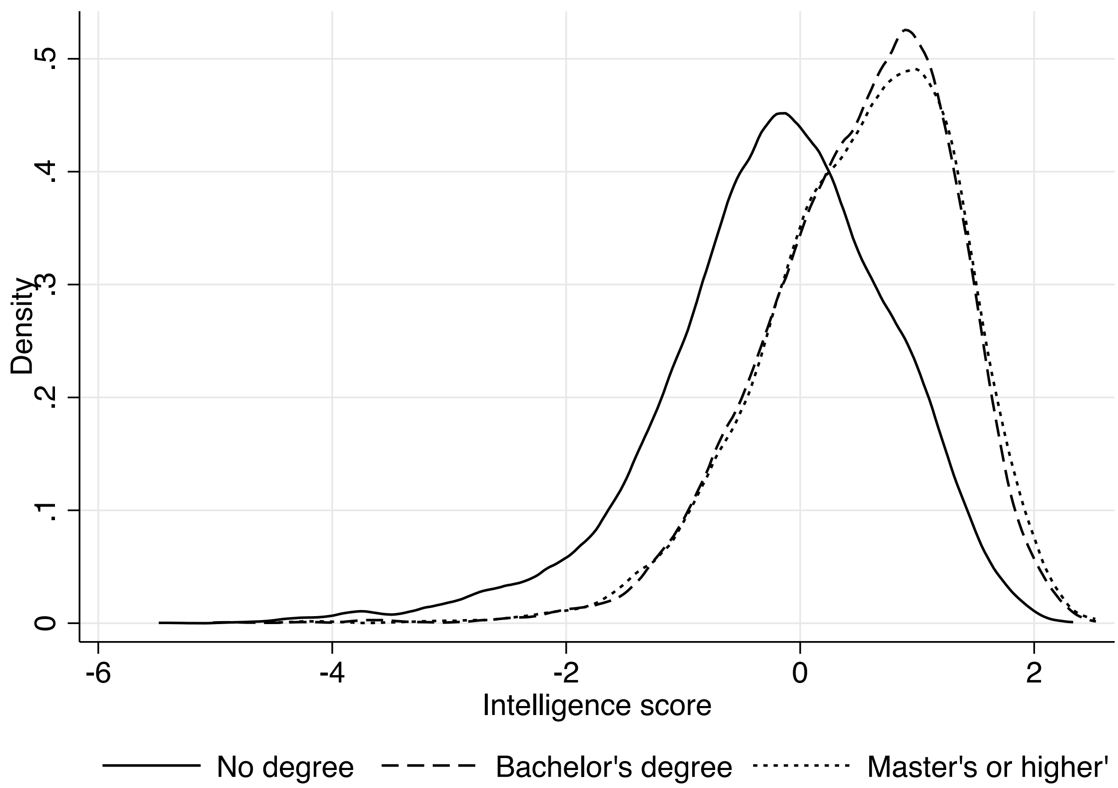


Figure 2.B.1: Distribution of intelligence score by degree status.

Table 2.B.1: Age at first birth: 40+ subsample.

	(1)	(2)	(3)	(4)
female	0.257***	0.301***	0.506***	0.506***
	(0.062)	(0.077)	(0.107)	(0.107)
IQ	-0.092***	-0.063***	-0.060**	-0.060**
	(0.013)	(0.016)	(0.018)	(0.018)
# children ever	0.307***	0.305***	0.314***	0.314***
	(0.027)	(0.033)	(0.040)	(0.040)
female × IQ	-0.121***	-0.094***	-0.083***	-0.083***
	(0.016)	(0.021)	(0.024)	(0.024)
female × # children ever	0.110***	0.110**	0.160***	0.160***
	(0.028)	(0.035)	(0.044)	(0.044)
Degree		-0.400***	-0.371**	-0.287*
		(0.085)	(0.121)	(0.137)
female × Degree		-0.086	-0.034	-0.020
		(0.097)	(0.147)	(0.173)
Degree × IQ		0.052*	0.055	0.088*
		(0.026)	(0.028)	(0.037)
Degree × # children ever		0.022	0.024	0.004
		(0.039)	(0.045)	(0.051)
female × Degree × IQ		-0.017	-0.018	-0.036
		(0.034)	(0.037)	(0.045)
female × Degree × # children ever		-0.029	-0.077	-0.077
		(0.044)	(0.052)	(0.060)
$\hat{w}45$			-0.040	-0.040
			(0.051)	(0.051)
female × $\hat{w}45$			-0.610***	-0.611***
			(0.080)	(0.081)
Degree × $\hat{w}45$			-0.024	-0.075
			(0.086)	(0.107)
female × Degree × $\hat{w}45$			0.281*	0.288
			(0.126)	(0.152)
Masters or higher	No	No	No	Yes
Observations	21045	21045	18624	18624

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table shows estimation results from cross-sectional Cox proportional hazards model where time to failure is age at first birth. The estimation sample is restricted to people ever observed at age 40 or older. The negative coefficient means lower hazard and higher time to failure, i.e. higher age at first birth. $\hat{w}45$ is the predicted wage at age 45. All specifications control for birth cohort FE and parental qualifications. The estimations were weighted by the cross-sectional sampling weights of wave 3. Standard errors clustered at the primary sampling unit level are shown in the parentheses.

Table 2.B.2: System of equations: fertility, wage and degree probability: 40+ subsample.

	(1) # children	Age birth 1	$\hat{w}$	Degree
female	0.327***	0.462***	-0.093***	-0.099**
	(0.020)	(0.055)	(0.007)	(0.032)
Degree	-0.058	-0.371***	0.315***	
	(0.030)	(0.101)	(0.009)	
female × Degree	-0.285***	0.071	-0.062***	
	(0.041)	(0.142)	(0.012)	
IQ	-0.088***	-0.048**	0.089***	0.831***
	(0.015)	(0.015)	(0.003)	(0.029)
female × IQ	-0.092***	-0.078***	-0.037***	0.062
	(0.020)	(0.020)	(0.004)	(0.038)
Degree × IQ	0.034	0.041	-0.011	
	(0.032)	(0.029)	(0.006)	
female × Degree × IQ	0.083	-0.023	0.005	
	(0.043)	(0.039)	(0.008)	
# children ever		0.288***	-0.004*	
		(0.010)	(0.002)	
$\hat{w}45$		0.012		
		(0.044)		
female × # children ever		0.207***	-0.013***	
		(0.014)	(0.003)	
female × $\hat{w}45$		-0.667***		
		(0.073)		
Degree × # children ever		0.029	0.010**	
		(0.022)	(0.004)	
Degree × $\hat{w}45$		-0.034		
		(0.092)		
female × Degree × # children ever		-0.120***	0.001	
		(0.031)	(0.005)	
female × Degree × $\hat{w}45$		0.273		
		(0.146)		
Constant	1.772***	-7.263***	0.605***	-1.439***
	(0.026)	(0.062)	(0.006)	(0.048)
Observations	26221			

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table reports regression results from cross-sectional SEM estimations in the subsample of people ever observed at age 40 or older. Equations for the total number of children ever had and predicted wages in columns 1 and 3 are estimated using simple linear regressions. Fertility timing equation in column 2 is fitted using Cox proportional hazards model. The degree equation in column 4 is estimated using logistic regression. $\hat{w}45$ is the predicted wage at age 45. All specifications control for birth cohort FE and parental qualifications.

References

Adda, J., C. Dustmann, and K. Stevens. 2017. "The Career Costs of Children." *Journal of Political Economy* 125 (2): 293–337.

Becker, G. S. 1960. "An Economic Analysis of Fertility" [in en]. In *Demographic and Economic Change in Developed Countries,* edited by Universities–National Bureau Committee for Economic Research, 209–240. New York: Columbia University Press.

Fisher, P., L. Fumagalli, N. Buck, and S. Avram. 2019. "Understanding Society and Its Income Data."

Francesconi, M. 2002. "A Joint Dynamic Model of Fertility and Work of Married Women." *Journal of Labor Economics* 20 (2): 336–380.

Gayle, G.-L., and R. A. Miller. 2012. *Life-Cycle Fertility and Human Capital Accumulation.* Technical report.

Gerster, M., N. Keiding, L. B. Knudsen, and K. Strandberg-Larsen. 2007. "Education and second birth rates in Denmark 1981-1994." *Demographic research* 17:181–210.

Gottfredson, L. S. 1997. "Why g matters: The complexity of everyday life." Special Issue Intelligence and Social Policy, *Intelligence* 24 (1): 79–132.

Gustafsson, S. 2001. "Optimal age at motherhood. Theoretical and empirical considerations on postponement of maternity in Europe." *Journal of population economics* 14 (2): 225–247.

Heckman, J. J., L. Lochner, and C. Taber. 1998. "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics* 1 (1): 1–58.

Jayachandran, S., and A. Lleras-Muney. 2009. "Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines." *Quarterly Journal of Economics* 124 (1): 349–397.

Johnson, W., and T. J.Bouchard. 2005. "The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized." *Intelligence* 33 (4): 393–416.

Keane, M. P., and K. I. Wolpin. 2010. “The Role of Labor and Marriage Markets, Preference Heterogeneity, and the Welfare System in the Life Cycle Decisions of Black, Hispanic, and White Women.” *International Economic Review* 51 (3): 851–892.

Klesment, M., A. Puur, L. Rahnu, and L. Sakkeus. 2014. “Varying association between education and second births in Europe: Comparative analysis based on the EU-SILC data.” *Demographic Research* 31:813–860.

Kravdal, Ø., and R. R. Rindfuss. 2008. “Changing relationships between education and fertility: A study of women and men born 1940 to 1964.” *American sociological review* 73 (5): 854–873.

Lagakos, D., B. Moll, T. Porzio, N. Qian, and T. Schoellman. 2018. “Life Cycle Wage Growth across Countries.” *Journal of Political Economy* 126 (2): 797–849.

Lundborg, P., E. Plug, and A. W. Rasmussen. 2017. “Can women have children and a career? IV evidence from IVF treatments.” *American Economic Review* 107 (6): 1611–37.

Lynn, R. 1996. *Dysgenics: Genetic Deterioration in Modern Populations.* Praeger Publishers.

McCrary, J., and H. Royer. 2011. “The effect of female education on fertility and infant health: evidence from school entry policies using exact date of birth.” *American economic review* 101 (1): 158–95.

Miller, A. R. 2011. “The effects of motherhood timing on career path.” *Journal of population economics* 24 (3): 1071–1100.

Nitsche, N., A. Matysiak, J. Van Bavel, and D. Vignoli. 2018. “Partners’ educational pairings and fertility across Europe.” *Demography* 55 (4): 1195–1232.

Rapallini, C., A. Minello, D. Vignoli, and A. Rustichini. 2021. *Education Aspirations, Career Concerns and Fertility.* Technical report. University of Florence.

Roberts, B. W., N. R. Kuncel, R. Shiner, A. Caspi, and L. R. Goldberg. 2007. “The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes.” *Perspectives on Psychological Science,* 313–345.

Schofer, E., and J. W. Meyer. 2005. “The worldwide expansion of higher education in the twentieth century.” *American sociological review* 70 (6): 898–920.

Sheran, M. 2007. "The career and family choices of women: A dynamic analysis of labor force participation, schooling, marriage, and fertility decisions." *Review of Economic Dynamics* 10 (3): 367–399.

Skirbekk, V. 2008. "Fertility trends by social status." *Demographic research* 18:145–180.

Stange, K. 2011. "A longitudinal analysis of the relationship between fertility timing and schooling." *Demography* 48 (3): 931–956.

Strenze, T. 2007. "Intelligence and socioeconomic success: A meta-analytic review of longitudinal research." *Intelligence* 35 (5): 401–426.

Topkis, D. M. 1978. "Minimizing a Submodular Function on a Lattice." *Oper. Res.* 26 (2): 305–321.

———. 1998. *Supermodularity and Complementarity.* Princeton University Press.

Warne, R. T. 2020. *In the Know: Debunking 35 Myths about Human Intelligence.* Cambridge University Press.

Wood, J., K. Neels, and T. Kil. 2014. "The educational gradient of childlessness and cohort parity progression in 14 low fertility countries." *Demographic research* 31:1365–1416.

Wu, A. H. 2019. *Life-cycle Returns to Social and Math Skills: The Roles of Gender, Sorting and Employer Learning.* Working Papepr. Harvard University.

Chapter 3

From bad to worse: long-term effects of recession in adolescence

There is increasing evidence in the literature that entering the labour market at the time of recession has adverse and lasting effects on a wide range of outcomes from earnings and employment to higher mortality and lower fertility (Schwandt and von Wachter 2019; Oreopoulos, von Wachter, and Heisz 2012; Kahn 2010; Schwandt and von Wachter 2020). The literature has found that such "scarring" effects of recessions are especially pronounced among the less educated workers. However, the existing research typically treats educational attainment of labour market entrants as a pre-determined characteristic.

In this paper, I study the effect of local economic conditions both on educational decisions and subsequent labour market outcomes using the largest panel survey data in the UK. The estimation strategy relies on the instrumental variables approach. I use the information on individuals' school location, county of birth and year of birth to match them with the prevailing unemployment rate in that area at the time when respondents were 14 years old. At this time, individuals were subject to compulsory schooling laws, ruling out the possibility of endogenous timing of schooling. I also provide evidence that after conditioning on year and county of birth fixed effects, as well as county-specific linear time trends, the local unemployment rate is uncorrelated with other pre-determined characteristics such as gender, race and parental characteristics.

I find that higher local unemployment rate at the age of 14 reduces educational attainment of children. This effect is entirely driven by non-degree qualifications and degrees earned from former polytechnics that became universities in 1992, while degree attainment from traditional universities are not affected by initial economic conditions. In particular, a 1 pp increase in unemployment rate reduces years of education by 1 months, probability of obtaining non-degree qualifications by 1 pp and probability of obtaining a degree qualification from either pre- or post-1992 universities - by 1pp. I also find that this effect is more pronounced among children from more disadvantaged backgrounds, but it cannot be fully explained by parental job loss. Previous research by Taylor and Rampino (2014) suggests that the effects could be mediated through beliefs about the value of education. They show that children of less educated parents adjust downwards the perceived importance of post-compulsory education as a result of higher unemployment rates.

Second, I find that children exposed to higher unemployment rates at age 14 have permanently lower wages over the ife cycle. The initial hike in wages together with first-stage results suggest that these children substitute education to entering the labour market sooner. Thus, growing up in a recession can have an additional negative effect on lifetime earnings of workers by reducing their educational attainment. In particular, a year of education lost due to higher initial local unemployment rate translates to 8% lower hourly wages at ages 26-30 and 6% lower hourly wages at ages 41-45.

This paper contributes to the growing literature on persistent effects of initial labour market conditions, summarized by von Wachter (2020). For example, Schwandt and von Wachter (2019) show that entering the labour market during recession leads to about 10% reduction in annual earnings at the time of entry. The effect then gradually fades away in the next 10-15 years. They also find larger effect sizes for high-school graduates compared to college graduates. Adverse effects of entering the labour market during recession are not limited to earnings and employment outcomes. Schwandt and von Wachter (2020) report higher mortality and lower fertility. The main contribution of this paper is that it studies the effect of local economic conditions at early ages jointly on education and subsequent labour-market outcomes.

The results in this paper also contribute to the discussion on the effect of recessions on educational decisions of children. A large body of literature has studied how children and families adjust their educational decisions as a result of recessions, and the Great Recession of 2008-09 in particular. Using the US data, Terry Long (2014) and Dellas and Sakellaris (2003) find that college attendance increases during recession, despite reductions in family income and increases in tuition fees. Using the children of UKHLS and BHPS respondents, Taylor and Rampino (2014) find that on average higher unemployment rate leads to more positive attitudes towards and aspirations for post-compulsory education. They also show that this effect only holds among children with highly-educated parents and/or parents who themselves have strong positive view of education. On the other hand, children of less-educated parents reduce the value placed on post-compulsory education at times of high unemployment, controlling for family income.

The rest of the paper is organised as follows. In the next section I describe the dataset used in the paper, followed by the description of the empirical strategy in section 3.2. I discuss the results in section 3.3, examine assumptions necessary for identification in section 3.4 and conclude in section 3.5.

3.1 Data

I use the UK Household Longitudinal Study[1] (UKHLS), also known as Understanding Society. This is the largest household panel study in the UK covering about 40,000 individuals since 2009.

I observe local unemployment counts starting from 1981; therefore, I restrict my analysis sample to people who turned 14 in 1981 or later. I also restrict the sample to people born no later than 1995 to remove individuals who were still in education. Furthermore, I remove individuals who have either prematurely dropped out of school or stayed in school past the age 20. I also remove 1,033 individuals with missing degree information.

In order to match individuals to the local labour market conditions at the time of their teenage years, I rely on two variables. First, unique school codes where respondents have finished school[2], which I can match to postcodes using school registries. The school codes, however, have been only asked from a subset of people born since 1981. Out of 78,218 people who have responded at least once in UKHLS waves 1-10, 23,342 are born since 1981 and 14,114 have non-missing school codes. Therefore, whenever school codes are unavailable, I use the county of birth information[3], which has been asked from all respondents.

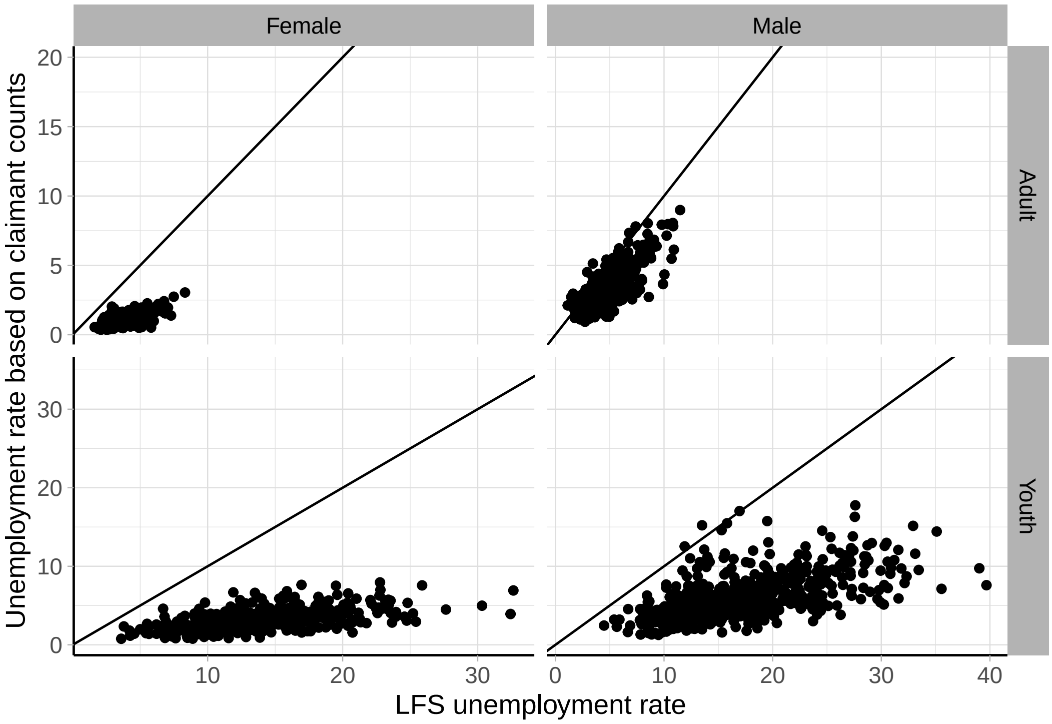
For the local labour market conditions, I use the information published by the Office for National Statistics. In particular, I use aggregated tables of stock of unemployment benefit claimants broken down by gender, age groups and UK counties. These series are available from 1981 to 2014 using the county boundaries prevailing from 1974 to 1996. In order to get unemployment rate measure, I divide the stock of claimants to the working-age population of same gender and age group resident in the same county. However, this measure is only an approximation to the true unemployment rate. I classify the unemployment rates into young

1. University of Essex, Institute for Social and Economic Research (2020)
2. University of Essex, Institute for Social and Economic Research (2021b)
3. University of Essex, Institute for Social and Economic Research (2021a)

(ages 16-24) and adult (ages 24-64). Figure 3.1.1 plots my measure of local unemployment rate against the official aggregate unemployment rate in the UK. It is evident from the plot that the constructed measure underestimates the unemployment rate among female and young populations. Figure 3.1.2 compares my unemployment rate measure with the official county unemployment rates computed based on the Labour Force Survey (LFS). The LFS data at the county level is only available in 1994-98 and 2003-11. This time frame makes the LFS impractical for the main research question, but presents me with the benchmark to compare my measure of local unemployment rates in the corresponding years. Most of the observations land significantly below the the 45 degree line, confirming that my measure is significantly underestimating the local unemployment rates among female and young populations. Moreover, the relationship between the two measures also has nearly flat slope, suggesting that my measure also underestimates variability of local unemployment rates. Only a measure of adult male unemployment rates performs well when compared with official rates both in terms of level and variability. Therefore, in the rest of the paper, I match all individuals from the UKHLS with adult male unemployment rate prevailing in their county at the time when they were 14.

Notes: The figure plots the unemployment rate measure based on unemployment benefit claimants and resident population at the county level (circles) against the aggregate unemployment rate in the UK (solid line).

Figure 3.1.1: Unemployment rate measure vs. aggregate unemployment rate

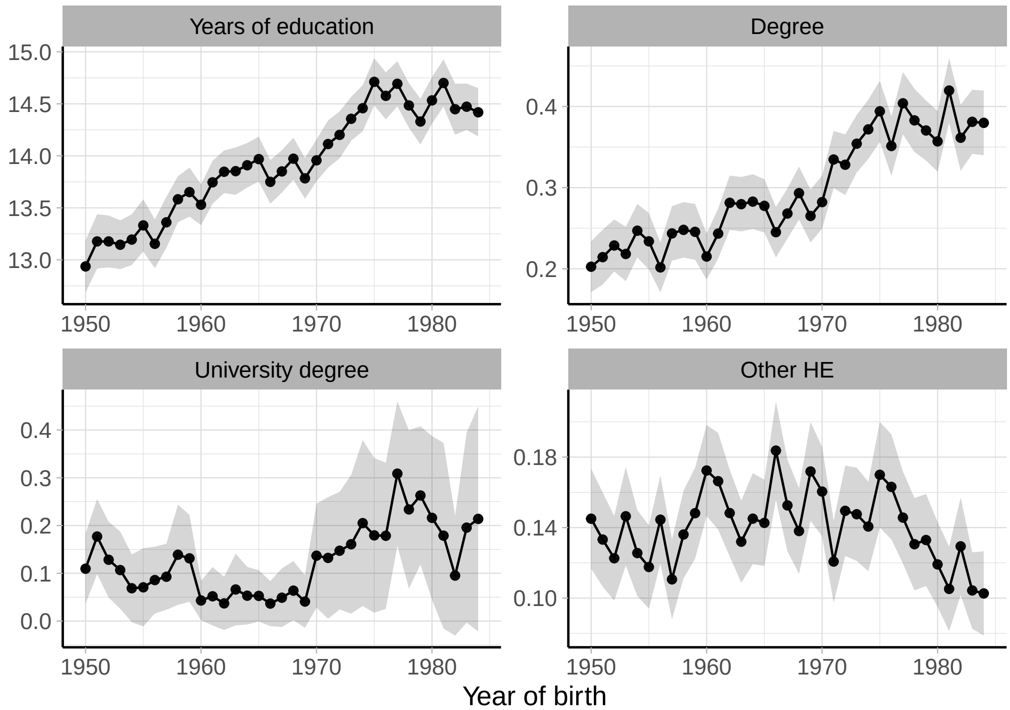


Notes: The figure compares the unemployment measure based on unemployment benefit claimants and resident population against the official county unemployment rates from the Labour Force Survey. The LFS unemployment rates in 1994-1998 are not disaggregated by gender. The solid line indicates the 45 degree line.

Figure 3.1.2: Local unemployment rates based on claimant counts vs. Labour Force Survey

3.1.1 Educational outcomes

The main focus of this paper is to study the effect of local economic conditions on earnings mediated through educational decisions. I construct years of education from reported ages when people left school and further education and from their highest qualification information. I also use several measures of higher education: degree attainment, university degree attainment (imputed, see Chapter 4) and other higher education. The degree attainment variable takes value of 1 if an individual has reported having a degree qualification. However, this variables combines degrees obtained from traditional universities, such as Oxford or Camrbidge, with the degrees obtained from former polytechnics (Chapter 4). The information to distinguish between the two types of institutions is not readily available in the UKHLS, but is present in the predecessor survey, the BHPS. Using multiple imputation technique, I can separate the degrees earned from traditional universities from the set of all degree-holders in the UKHLS. The indicator for other higher education predominantly includes people with higher education diploma[4], teaching or nursing qualifications.



Notes: The figure plots average educational outcomes of cohorts. The averages are weighted by the cross-sectional sampling weights. For degree and non-degree qualifications, the plot shows average qualification attainment rate across cohorts.

Figure 3.1.3: Degree attainment by gender and year of birth

4. Higher education diploma is typically awarded after two years of full-time study at a degree program.

3.1.2 Earnings

The main variable of interest is labour earnings of individuals. The identification relies on variation across counties and birth cohorts. To be able to compare individuals from different birth cohorts, I need to use their earnings information at a similar age. Although, the UKHLS is a panel dataset, it only has started in 2009[5] and a direct approach would only be possible for a very small subset.

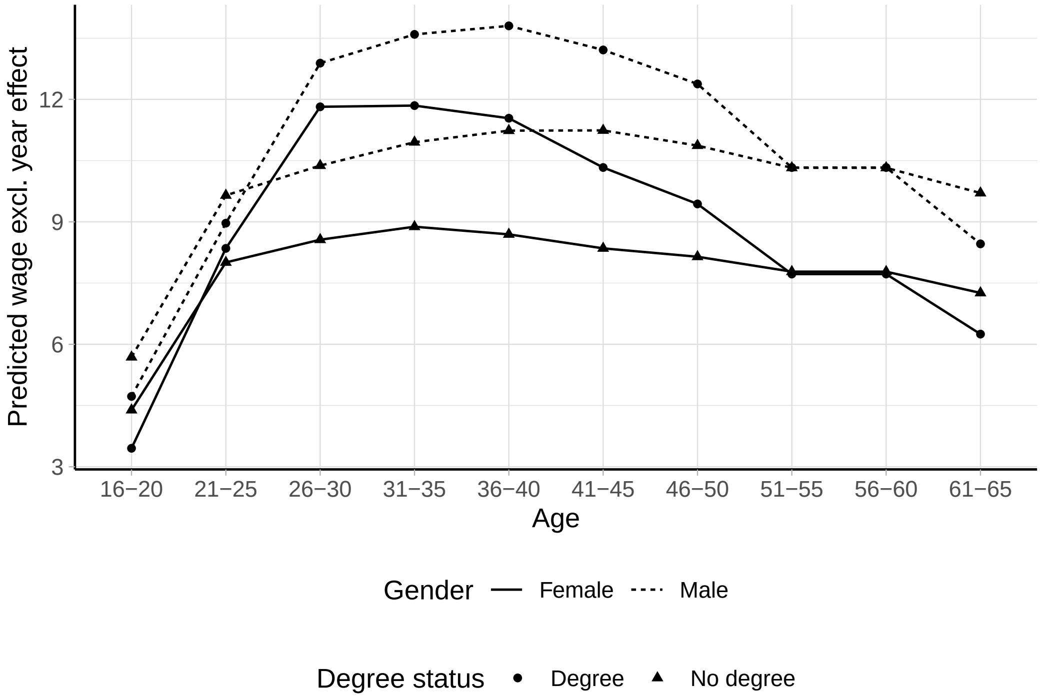
Therefore, I use predicted wages given the estimated wage age profiles. I estimate the profiles following the logic similar to Lagakos et al. (2017). In particular, I assume no wage growth at the ages of 51-60. Hence, all wage growth in this age group is attributed to time and cohort effects. I begin by calculating the hourly wage rate using the information on monthly labour earnings and usual hours worked in a week. I also deflate the hourly wage rate using the CPI excluding rents, maintenance repairs and water charges. Then, I set the outlier observations to missing where outliers are identified as observations with wages below 1st or above 99th percentiles in each birth cohort, gender and degree status cells. Finally, I estimate the wage profiles by regressing the real hourly wages on year and age indicators fully interacted with gender and degree attainment variables using the fixed effect estimator. The estimated profiles are shown in Figure 3.1.4. Reassuringly, the estimated profiles demonstrate higher lifetimes earnings and steeper profiles among degree-holders. The profiles of younger cohorts are also lower compared to that of older cohorts, which could be explained by increased participation rate in university education and larger supply of university-educated labour force.

3.2 Estimation strategy

The main objective of this paper is to study the effect of local economic conditions on educational decisions and subsequent labour market outcomes. The selectivity of university admission and format of school exams in the UK means that children start thinking about their educational goals before age 16. Therefore, I use prevailing local unemployment rate at the time individual was age 14 as the measure of initial local economic conditions. I use counties of school/birth as a measure of local labour markets.

I use instrumental variable approach described by equations (3.1) and (3.2).

5. The BHPS has been around since 1991, but the dataset is much smaller.



Notes: The figure plots the estimated age profiles for real hourly wages net of year effects. The identifying restriction is that age profile should be flat at ages 51-60. The estimation sample includes wage observations from UKHLS between 1st and 99th percentile for each birth cohort, gender and degree attainment cell.

Figure 3.1.4: Wage profiles by gender and degree attainment

$$y_{i,t} = \alpha + \beta_a e_i + \gamma_a + \theta_i + \varepsilon_{i,t} \tag{3.1}$$

$$e_i = \phi_0 + \phi_1 u_{i,g,0} + \zeta_i + v_i \tag{3.2}$$

where $y_{i,t}$ is real hourly wage of individual i at time t, e_i is the educational attainment of individual i, and $u_{i,g,0}$ is the unemployment rate in i's county g at the time she was 14 years old. The regression equation allows the effect of education to change throughout individual's career. Thus, ϕ_1 captures the effect of initial local unemployment rate on the educational decisions and β_a - the resulting changes in the age profiles of real hourly wages. For the discussion of identifying assumptions see section 3.4.

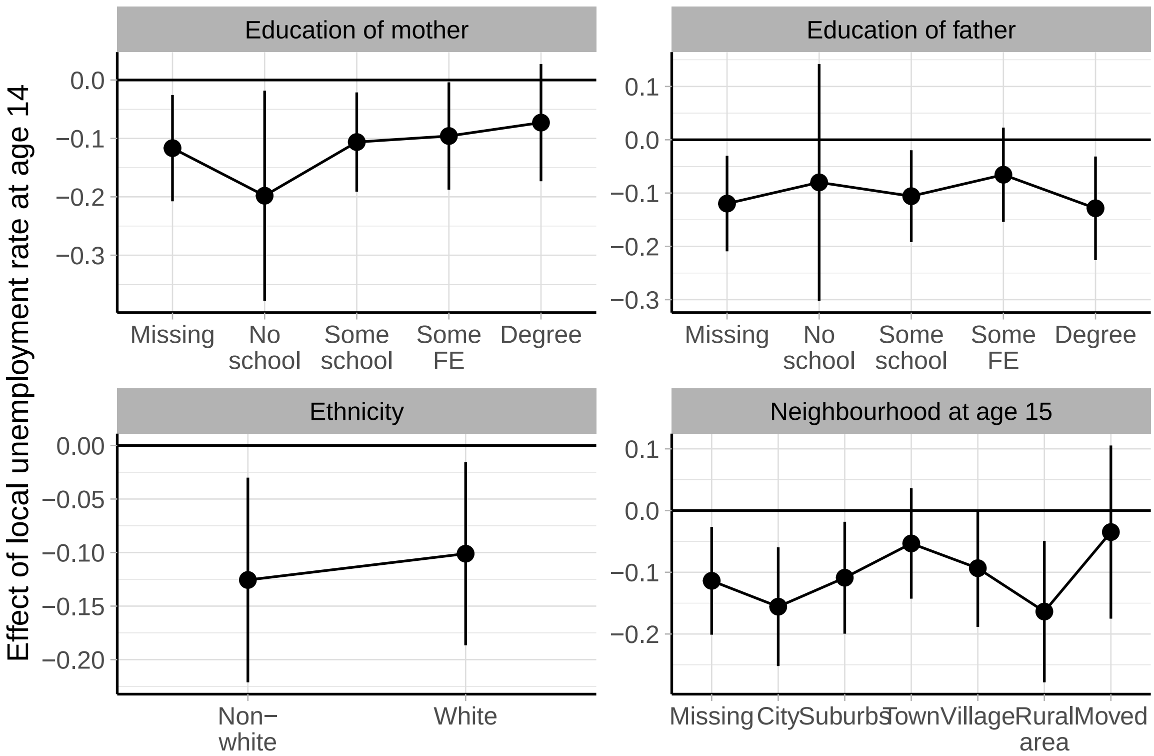
3.3 Results and discussion

3.3.1 First stage

The first-stage estimation results show that higher local unemployment rate at the age of 14 reduces educational outcomes of children with lower educational intentions (Table 3.3.1). A 1 pp increase in local unemployment rate at the age of 14 reduces years of education by 1.2 months. The probability of obtaining a non-degree qualification reduces by 1pp and the probability of having a degree - by 1.4pp, entirely driven by the degrees earned from former polytechnics. The effects are similar across gender.

The first-stage estimates are similar to Tumino and Taylor (2015), who find that educational decisions of children from less affluent families respond more to prevailing economic conditions. In particular, they also find that higher adult unemployment rate reduces post-compulsory education of students in rented and social housing. I perform a similar exercise by interacting the local unemployment rate with various family characteristics in Figure 3.3.1. Typically, the effect of local unemployment rate on years of education is more negative among children with less educated parents and children that used to live in inner city or rural areas at the age of 15. Somewhat surprisingly the effect is also more negative among children whose fathers have a degree.

A stronger effect on children from more disadvantaged households may suggest that income effect is dominating. If their families are more likely to lose income during recessions, the children may put a higher value to entering the labour market sooner. The dataset also allows me to test this effect directly via parental unemployment indicators. Every



Notes: The plots show the marginal effect of 1 pp increase in local unemployment rate on years of education by categories of parental education, respondents' ethnicity and neighbourhood at age 15. All regressions control for year of birth, county, country of birth, race, gender, parental qualifications fixed effects and county-specific linear trends. The standard errors are clustered at the primary sampling unit. The whiskers correspond to 95% confidence intervals.

Figure 3.3.1: The effect of local unemployment rate on years of education by family characteristics

Table 3.3.1: First-stage estimates

	Years of edu	Degree	Traditional uni degree	Other HE
Panel A: Average effect				
L. unemp at 14	-0.102**	-0.014*	-0.001	-0.010**
	(0.044)	(0.007)	(0.010)	(0.005)
Panel B: Interacted by gender				
L. unemp at 14	-0.103**	-0.014*	-0.001	-0.009*
	(0.044)	(0.007)	(0.011)	(0.005)
Female	0.008	0.013	0.020	0.041**
	(0.134)	(0.022)	(0.039)	(0.017)
L. unemp at 14 $\times$ Female	0.003	0.001	0.000	-0.003
	(0.015)	(0.002)	(0.003)	(0.002)
Obs.	15,147	15,136	10,009	15,136
F statistic	5.407	3.544	0.004	4.273

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: The table shows the first-stage estimation results. The dependent variables are in columns. All regressions control for year of birth, county, country of birth, race, gender, parental qualifications fixed effects and county-specific linear trends. The standard errors are clustered at the primary sampling unit and reported in parentheses. The estimations are weighted by the cross-sectional sampling weights.

respondent was asked if their parents were unemployed at the time the respondent was 14 years old. Although, the parental unemployment indicator roughly corresponds to aggregate unemployment rate over the longer horizon, the two appear uncorrelated in the estimation sample (Figure 3.C.1). Therefore, estimations in Table 3.3.2 use aggregate unemployment rate series available from 1971, in place of local unemployment rates. In the first column, I verify that variations in aggregate unemployment rate produce similar first-stage estimates. The magnitude of the coefficient is smaller compared to the results in Table 3.3.1, but qualitatively are similar. A 1 pp increase in aggregate unemployment rate at age 14 reduces years of education by 0.4months. I also verify that aggregate unemployment rate is a strong predictor of parental unemployment indicator in column 2. A 1 pp higher aggregate unemployment rate increases the probability of having an unemployed father by 0pp. However, parental unemployment accounts for little, if any, of the estimated treatment effect as suggested by the results in column 3: the point estimate of the effect of aggregate unemployment rate on the years of education changes little.

An alternative explanation is that the recessions alter beliefs, akin to findings of Giuliano and Spilimbergo (2014). For example, Taylor and Rampino (2014) study children of UKHLS and BHPS respondents and find that children of less-educated parents place lower value to post-compulsory and university education.

Table 3.3.2: First-stage estimates: mediation via parental unemployment

	Years of edu	Father unemp at 14	Years of edu
Agg. unemp at 14	-0.030*	0.004***	-0.027
	(0.017)	(0.001)	(0.017)
Father unemp at 14			-0.110
			(0.341)
Agg. unemp at 14 × father unemp at 14			-0.028
			(0.039)
Obs.	21,572	21,572	21,572

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

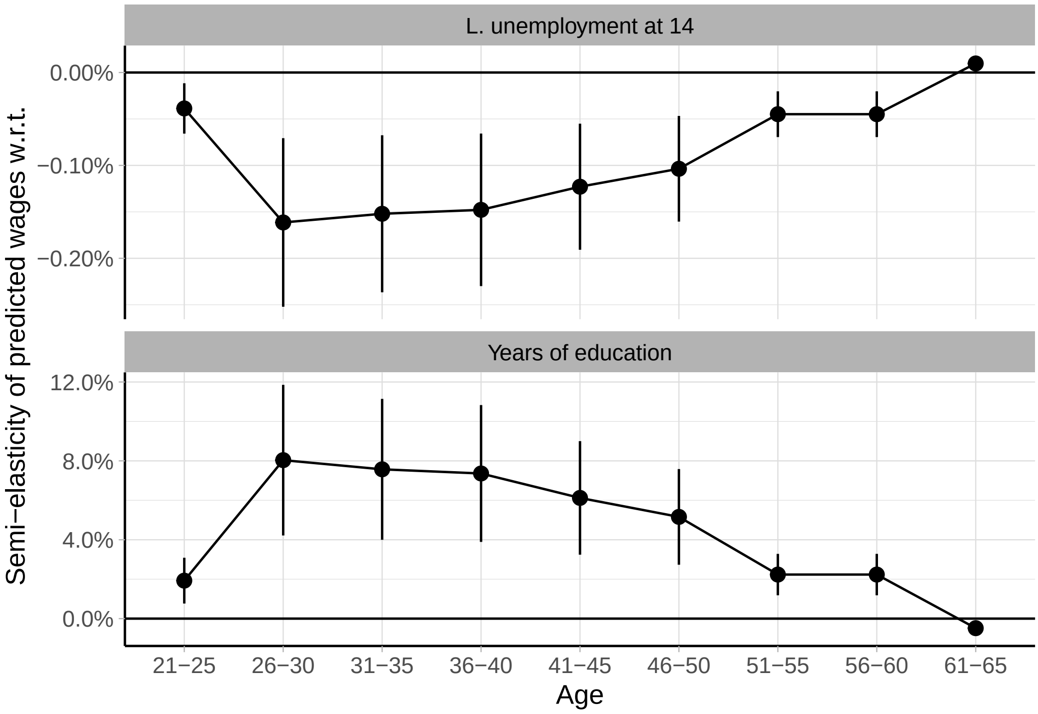
Notes: The table shows the first-stage estimation results using aggregate unemployment rate. The dependent variables are in columns. All regressions control for quiennial birth cohort, county, country of birth, race, gender, parental qualifications fixed effects and count-specific linear trends. The standard errors are clustered at the primary sampling unit and reported in parentheses. The estimations are weighted by the cross-sectional sampling weights.

In summary, previous research has shown that recessions have a lasting negative effect on labour market entrants, especially among the less educated ones. My results in Table 3.3.1 suggest that recessions can have an additional negative effect on labour market outcomes through lowering educational attainment of children.

3.3.2 Second stage

Next, I study the effect of local unemployment rate on labour market outcomes using the predicted hourly wage profiles. Figure 3.3.2 plots the reduced-form and IV estimates by age. First, exposure to higher unemployment rate at the age of 14 has somewhat positive effect on wages up to age 20. Together with the first-stage estimates, these results suggest that people substitute education to labour market entrance. However, higher initial unemployment rate reduces wages past age 40 until retirement. For example, a 1 pp increase in local unemployment rate at age 14 reduces real hourly wages at age 45 by 0.23%.

The second stage results suggest that a year of education lost due to higher unemployment rate at the age of 14 has a considerable effect on real hourly wages throughout the career. For example, a 1 year of education lost translates to 9.93% lower wages at age 45. These are comparable to the wage returns to education, previously estimated in the literature (Gunderson and Oreopolous 2020).



Notes: The top and bottom panels plot reduced-form and second-stages estimates, respectively. The fixed-effects estimators are transformed into semi-elasticities, that is, percentage change in real hourly wages with respect to a unit change in the independent variable. The whiskers correspond to 95% confidence interval. The standard errors are clustered at the individual level.

Figure 3.3.2: Reduced-form and second-stage estiamtes

Table 3.4.1: Instrument balancing test

	Female	White	Mother degree	Father degree	Mother born UK	Father born UK
L. unemp at 14	-0.003	0.004	0.000	0.002	0.010	0.006
	(0.007)	(0.004)	(0.006)	(0.006)	(0.005)	(0.005)
	[0.900]	[0.725]	[0.944]	[0.900]	[0.236]	[0.725]
Obs.	15,157	15,147	12,203	11,852	14,994	14,935

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: The table shows the effect of local unemployment rate at age 14 on pre-determined characteristics. The dependent variables are in columns. All regressions control for year of birth and county indicators and county-specific linear trends. The standard errors are clustered at the primary sampling unit and reported in parentheses. The Benjamini-Hochberg adjusted p-values are reported in square brackets. The estimations are weighted by the cross-sectional sampling weights.

3.4 Validity

The causal interpretation of β_a in equation (3.1) depends on validity of instrumental variable approach. To begin with, the instrument should be as good as random. This is partially ensured by the timing of the shock. Here, the unemployment rate is measured at the time people were still in compulsory schooling. This rules out the possibility of endogenously timing the exposure to the shock. However, the assumption may be violated if households self-select into local areas in a way that correlates with local unemployment rates at age 14. To account for these I control for county fixed effects and county-specific linear trends. I also verify that local unemployment rate at age 14 is uncorrelated with pre-determined family and children characteristics after conditioning on year of birth, county and county-specific linear time trends in Table 3.4.1.

The relevance of the instrument is estimated from the first-stage regression in equation (3.2) and presented in Table 3.3.1. Since all the variables are time-invariant, I run estimations using the cross-section of individuals, approximating ζ_i by a vector of personal characteristics including indicators of year of birth, county and country of birth, gender, race and parents' highest qualification levels and county-specific linear trends. The F-statistic of the local unemployment rate at 14 for years of education is 5.4, which approximately translates to a size of 5%-level Wald test above a 25% threshold (Stock and Yogo 2005).

Finally, exclusion restriction requires that unemployment rates at the age of 14 can only affect individuals through their educational decisions. This assumption is harder to justify. Failure of exclusion restriction would mean that $u_{i,g,0}$ should also be included in (3.1), in which case β_a is unidentified. This can arise if experiencing recession at the age of 14 affects children's beliefs that affect both their willingness to continue education and their behaviour in the labour market. For example, Giuliano and Spilimbergo (2014) show that exposure to recessions at young age increases preferences for redistribution and beliefs that luck is more important than effort. Despite this, the estimates shown in Figure 3.3.2 imply that a year of education lost due to initial conditions reduces wages by 2-8% between ages 21 and 60. These estimates are similar to the rate of private return to a year of education typically found in literature (Psacharopoulos and Patrinos 2018). Such similarity may suggest that a possible failure of the exclusion restriction has a small effect on bias in the estimator of β_a.

3.5 Conclusion

There has been an increasing evidence of "scarring" effects of entering the labour market during a recession, especially among less-educated workers. However, the existing literature has treated educational attainment as a pre-determined characteristic. In this paper I study the effects of local unemployment rate at age 14 on educational decisions and subsequent labour market outcomes using instrumental variable approach. This framework rules out the possibility of endogenously timing the exposure to the shock. In addition, the institutional setting makes age 14 crucial for educational decisions.

From the first-stage estimates I find that higher initial unemployment rate reduces years of education, primarily among students aiming at non-degree qualifications and degrees from post-1992 universities (former polytechnics). This result is in contrast with previous findings of increased college attendance during recessions in the US (Terry Long 2014). I also find that this effect is more pronounced among children from more disadvantaged backgrounds, but it cannot be fully explained by parental job loss. Previous research by Taylor and Rampino (2014) suggests that these effects could be mediated through a reduction in perceived value of education among children with less-educated parents.

Next, I show that exposure to higher unemployment rate at the age of 14 has minor positive effect on wages up to age 20. This suggests that children are substituting education to entering the labour market sooner. However, wages of these children when they are 40 years or older are permanently lower compared to their peers growing up in more favourable conditions. The IV estimator suggests that a year of education lost due to higher initial local unemployment rate translates to 10% lower hourly wages at age 45.

Appendix 3.A Unemployment rate measures

I calculate the baseline measure of unemployment rate as the ratio of unemployment benefit claimants to the population in the respective age group (3.3).

$$u_{gk} = \frac{\text{claimants}_{gk}}{\text{population}_{gk}} \tag{3.3}$$

where g stands for county and $k \in \{\text{adult}, \text{youth}\}$. Adult group is defined as aged between 24 and 64 and youth - between 16 and 24.

The stock of unemployment benefit recipients is available from 1978 (with gaps) to 2014 at the pre-1996 county level. The population estimates are available from 1981 to 2017 at the pre-2009 and pre-2015 county levels. Fortunately, most of these county definitions can be translated easily to pre-1996 county definitions (see section 3.B).

This statistic is different from the actual unemployment rate in two respects. First, not all unemployed workers may be receiving unemployment benefits. Second, population is a superset of labour force. In fact, the consistent underestimation of the true unemployment rate among female and youth groups in Figure 3.1.1 is possibly due to their lower benefit take-up rates.

Appendix 3.B Local area boundaries

I used school codes and county of birth to assign individuals to local areas. Using school registries from the Department of Education, I merge school codes with school postcodes. Then, using postcode directory from the ONS I merge postcodes with longitude and latitude coordinates. Given coordinates of county boundaries, school could then be assigned to a county. Out of 4,555 unique school codes in the UKHLS, 4,304 are assigned to a county.

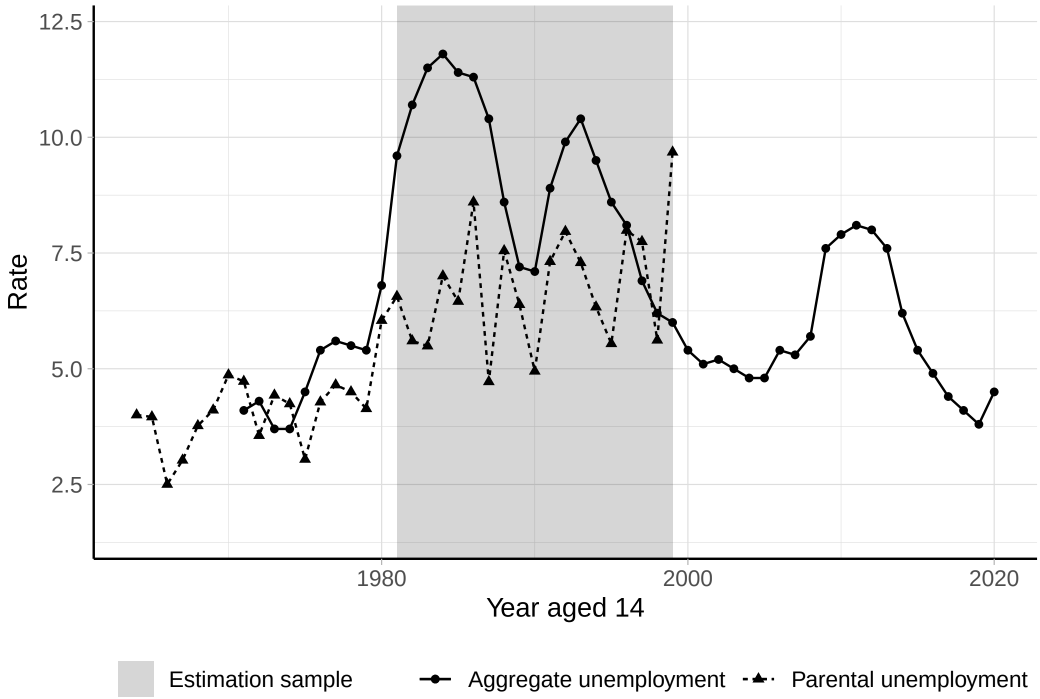
Unfortunately, the variable with county of birth has been coded with several county definitions. In my sample of interest, these are mainly county classifications that existed before 1974 (pre-1974) and during 1974-1996 (pre-1996). To be able to merge with unemployment statistic later, I use pre-1996 classifications as the baseline. Unfortunately, the old and new names are not always easy to separate and there is no clear distinction by year of birth either. Therefore, I examine closely the boundary changes to determine if a unique mapping is possible.

The *Local Government Act 1972* (1974) had abolished historical counties and introduced a two-tier system of counties and districts in England and Wales. Similar changes have been introduced by the *Local Government (Scotland) Act 1973* (1975), which abolished historical subdivisions and introduced two-tier structure of regions and districts. These reorganisations also introduced significant changes to county borders. I create a lookup table between counties using the area overlap between pre-1974 and pre-1996 counties. I classify a match to be unique if more than 90% of pre-1974 county is contained within a given pre-1996 county, and approximate if area overlap is between 75% and 90%. This results in a unique match for 72, approximate match for 18 and no match for 6[6] of the 96 pre-1974 counties. To maximise the sample size, I use all individuals whose county information can be mapped to pre-1996 definitions with some level of confidence. But I also check the robustness of the results when restricting the sample to individuals with unique match.

The population estimates at the county level are available with counties coded using pre-2009 classification. These were introduced in 1996 as a result of *Local Government Act 1992*. Fortunately, these boundary reorganisations were less complicated and a unique match was possible in the vast majority of the cases. Cardiff and the Vale of Glamorgan could only be matched approximately to South Glamorgan. Simlarly, Slough could be matched approximately to Berkshire. Only Caerphilly and Conwy in Wales could not be matched.

6. These are Berkshire, Lancashire, Lincolnshire (parts of Lindsey), Glamorgan, Midlothian and West Riding of Yorkshire.

Appendix 3.C Supplementary tables and figures



Notes: The plot shows the share of respondents reporting unemployed parent at age 14 by year of birth against the aggregate unemployment rate at the time. The average parental unemployment rate is weighted by the cross-sectional sampling weight and standard errors of means are clustered at the primary sampling unit level. The whiskers correspond to 95% confidence intervals.

Figure 3.C.1: Comparison of the parental unemployment with aggregate unemployment

References

An Act to Make New Provision, by Giving Effect to Proposals in Cm. 1599 (The Citizen's Charter) Relating to Publicity and Competition, for Securing Economy, Efficiency and Effectiveness in the Manner in Which Local Authorities Carry on Certain Activities; and to Make New Provision in Relation to Local Government in England for Effecting Structural, Boundary and Electoral Changes. [in eng].

An Act to Make Provision with Respect to Local Government and the Functions of Local Authorities in England and Wales; to Amend Part II of the Transport Act 1968; to Confer Rights of Appeal in Respect of Decisions Relating to Licences under the Home Counties (Music and Dancing) Licensing Act 1926; to Make Further Provision with Respect to Magistrates' Courts Committees; to Abolish Certain Inferior Courts of Record; and for Connected Purposes. [in eng]. 1974.

An Act to Make Provision with Respect to Local Government and the Functions of Local Authorities in Scotland; to Amend Part II of the Transport Act 1968; and for Connected Purposes. [in eng]. 1975.

Dellas, H., and P. Sakellaris. 2003. "On the Cyclicality of Schooling: Theory and Evidence." *Oxford Economic Papers* 55 (1): 148–172.

Giuliano, P., and A. Spilimbergo. 2014. "Growing up in a Recession." *The Review of Economic Studies* 81 (2): 787–817.

Gunderson, M., and P. Oreopolous. 2020. "Returns to Education in Developed Countries" [in en]. In *The Economics of Education,* Second, edited by S. Bradley and C. Green, 39–51. Academic Press.

Kahn, L. B. 2010. "The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy." *Labour Economics* 17 (2): 303–316.

Lagakos, D., B. Moll, T. Porzio, N. Qian, and T. Schoellman. 2017. "Life Cycle Wage Growth across Countries." *Journal of Political Economy* 126 (2): 797–849.

Oreopoulos, P., T. M. von Wachter, and A. Heisz. 2012. "The Short- and Long-Term Career Effects of Graduating in a Recession." *American Economic Journal: Applied Economics* 4 (1): 1–29.

Psacharopoulos, G., and H. A. Patrinos. 2018. *Returns to Investment in Education: A Decennial Review of the Global Literature* [in en]. Policy Research Working Papers 8402. World Bank Group.

Schwandt, H., and T. M. von Wachter. 2019. "Unlucky Cohorts: Estimating the Long-Term Effects of Entering the Labor Market in a Recession in Large Cross-Sectional Data Sets." *Journal of Labor Economics* 37 (S1): S161–S198.

———. 2020. "Socioeconomic Decline and Death: Midlife Impacts of Graduating in a Recession." Cambridge, MA.

Stock, J. H., and M. Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression" [in en]. In *Identification and Inference for Econometric Models,* edited by D. W. K. Andrews, 80–108. New York: Cambridge University Press.

Taylor, M., and T. Rampino. 2014. "Educational Aspirations and Attitudes over the Business Cycle" [in en]. *Economica* 81 (324): 649–673.

Terry Long, B. 2014. "The Financial Crisis and College Enrollment: How Have Students and Their Families Respond?" [In en]. In *How the Financial Crisis and Great Recession Affected Higher Education,* edited by J. R. Brown and C. M. Hoxby, 209–233. University of Chicago Press.

Tumino, A., and M. Taylor. 2015. *The Impact of Local Labour Market Conditions on School Leaving Decisions.* Working Paper 2015-14. Institute for Social and Economic Research.

University of Essex, Institute for Social and Economic Research. 2020. *Understanding Society: Waves 1-9, 2009-2018 and Harmonised BHPS: Waves 1-18, 1991-2009,* SN6614.

———. 2021a. *Understanding Society: Waves 1-10, 2009-2019 and Harmonised BHPS: Waves 1-18, 1991-2009: Special Licence Access,* SN6931.

———. 2021b. *Understanding Society: Waves 1-10, 2009-2019: Special Licence Access, School Codes,* SN7182.

von Wachter, T. 2020. "The Persistent Effects of Initial Labor Market Conditions for Young Adults and Their Sources" [in en]. *Journal of Economic Perspectives* 34 (4): 168–194.

Chapter 4

Multiple Imputation of University Degree Attainment

joint with Johanna L. Reuter

The second half of the 20th century has seen a massive expansion of university education throughout the world. In the UK, the participation rate in higher education rose from 4.1% in 1960[1] to about 20% in 1990[2]. But not all higher education is made equal. From 1965 to 1992, students in the UK could earn their degrees either from traditional universities or from public sector colleges led by polytechnic institutions. In 1992 the vast majority of the polytechnics were converted to universities. Formally, degrees from polytechnics were of the same standard as university degrees. Nevertheless, the institutions faced different target populations, admission procedures, subjects taught, organization and financing schemes. These differences, together with the elite image of the traditional universities, contributed to a public perception of polytechnics degrees as inferior to that of universities (Willetts 2017; Pratt 1997).

This perceived inferiority hints at something that has been established in the literature: the type of higher education institution can act as a signal of education quality. Thus, types of higher education institutions could serve as signals of education quality. College quality is an important determinant of educational decisions and returns to education. Brewer, Eide, and Ehrenberg (1999), Black and Smith (2004) and Black and Smith (2006) report that attending a higher quality institution is associated with sizeable private wage returns. Dillon and Smith (2017) find that students and their families prefer high-quality universities, even when it is not

1. *The Robbins Report* (1963)
2. *The Dearing Report* (1997)

the best match given the ability of the student. It is clear that the type of higher education institution can be considered of importance.

However, common survey datasets often offer limited information about the types of institutions from which individuals earned their degrees. For example, the UK Household Longitudinal Study (UKHLS), the largest panel study in the UK, up to 2019-21[3] only asked a small subset of participants for details about the higher education institutions they attended. Furthermore, additional restrictions may apply before one is granted access to such information, as is the case for the UKHLS.

In this paper, we try to overcome the issue of missing higher education institution types by using a multiple imputation technique. We rely on the institution type information available in the British Household Panel Study (BHPS), a smaller panel study carried out from 1991 to 2008, as well as the close relationship between the two panel studies. In particular, the BHPS specifically asked its participants to indicate the type of institution last attended, distinguishing between universities and polytechnics. The survey designs are highly comparable between the two studies. Thus, we can transform the lack of institution type into a missing data problem in a combined dataset of the BHPS and the UKHLS. In addition, many of the former BHPS respondents are now part of the UKHLS, presenting us with a second strategy of using the BHPS subsample within the UKHLS for imputation.

To properly reflect the uncertainty about imputed values we use a multiple imputation technique (Rubin 1977). By imputing multiple values for each missing observation, we can perform our analysis of interest multiple times and combine the estimated parameters. The combined estimators then reflect both sampling and imputation uncertainty. For them to deliver valid inference, two crucial assumptions must hold. First, the probability of missing data cannot depend on the missing value. This is the so-called missing at random assumption. We provide evidence from covariate balance tests to support this assumption. Second, the imputation model must be proper. To check this assumption we use the simulation-based evaluation method proposed by Brand et al. (2003).

When constructing the imputation model, another important consideration we take into account is the agreement between imputation and analysis models (Schafer 1997). The agreement means that the imputation model should be consistent with the model that the researchers want to estimate given the research question. In this paper, we adopt the following research question from our companion paper Dimbou et al. (n.d.): how did the expansion of the higher education in the UK change the composition of students in terms of their

3. All adult participants were asked for such details in wave 11 carried out in 2019-21. The dataset from wave 11 will be released at the end of 2022.

intelligence scores? Differentiating between traditional universities and former polytechnics is crucial for such an analysis. Since the two institution types targeted different types of applicants, it is reasonable to expect that they also faced distinct trends in the composition of the student body. In the context of the current paper, we are interested in the relationship between the probability of getting a degree from a traditional university and the intelligence score of the student. We also would like to examine how this relationship changed over time. Therefore, the imputation model should ideally control for time trends and intelligence scores. In practice, the intelligence score variable is only available for a subset of the UKHLS panel. Therefore, we test three versions of the imputation model that differ in estimation samples and the inclusion of intelligence score variable.

We find that the imputation models with and without intelligence scores perform similarly across all dimensions. In the simulation-based evaluation, the two models produce combined estimators with similar bias and efficiency statistics for the marginal effect of the intelligence score on the average university degree attainment. We also show that the combined estimators of the average university degree attainment across cohorts is, in general, similar to the benchmark graduation rates computed using the USR and the HESA dataset. This similarity could allow us to use a simpler imputation model without the intelligence score in our companion paper.

The rest of the paper is structured as follows. In the next section we discuss the institutional differences between universities and polytechnics. We describe and compare the BHPS and the UKHLS datasets in section 4.2. In section 4.3, we provide a brief introduction to multiple imputation, examine crucial assumptions, construct imputation models and describe the simulation-based evaluation method. Finally, we discuss the results in section 4.4 and conclude in section 4.5.

4.1 Institutional background

The higher education system in the UK was characterised by a binary divide until 1992: tertiary education was provided both by independent universities and public colleges[4]. This division was motivated by the desire of the government to adapt training according to "national economic needs for specific skills" (Willetts 2017, p.52). These public sector colleges were

4. Most of the public colleges have been present in the UK from at least 1870 onwards alongside the traditional universities, such as Oxford and Cambridge (Gillard 1998). However, many of them were merged in the 1960s to create new polytechnics that were distinctively defined by the government (Pratt 1997).

funded and organised by the local education authorities and provided vocational and other training necessary to meet local demand for skills.

The role of these public sector colleges was an important one. They included teacher training colleges, nursing colleges and polytechnics, where much of the higher education in technical and scientific subjects took place (Gillard 1998). Already in 1983, the first polytechnics were founded in order to "promote the industrial skill, general knowledge, health and well-being of young men and women belonging to the poorer classes" (Lawson and Silver, cited in Gillard 1998, p.83). However, over time the view on polytechnics and their role in higher education shifted away from being just for the poor, but rather institutions which provided a technical and scientific higher education.

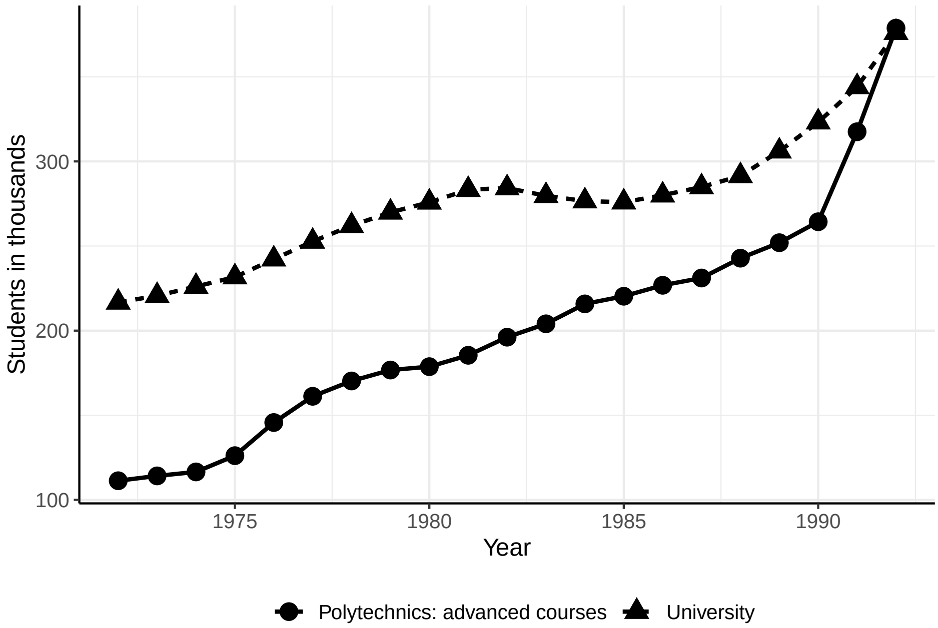
The early 1960s saw two policy changes important for higher education decisions of young individuals: the end of military conscription in 1960[5] and introduction of the centralised applications via the Universities Central Council on Admissions (UCCA) in 1961. As a result, in 1961 the Prime Minister Harold MacMillan announced creation of a committee headed by Lord Robbins "to review the pattern of ... higher education ... and advise ... the government on what principles its long-term development should be based" (*The Robbins Report* 1963, p.1). The report of this committee, also known as the Robbins Report, suggested the unification of the higher education system. Nevertheless, the government followed the idea of the Education Secretary at the time, Anthony Crosland, to adhere to the policy of a binary divide between universities and public sector colleges (Gillard 1998). Within the public sector colleges, polytechnics were the main instrument through which the binary policy was implemented. In 1966 the Government published a White paper detailing creation of 28[6] polytechnic institutions. These new polytechnics were formed by merging over 50 existing colleges (Pratt 1997).

Only in 1992 did the 'binary divide' come to an end with the Further and Higher Education Act, which allowed polytechnics to obtain university status. A majority of the institutions used this option immediately: the number of universities and university students almost doubled over the next two years. The abolishment of the binary system gave rise to what are known as old universities (universities that existed before 1992) and new universities (former polytechnics that became universities after 1992).

In order to better understand the differences and similarities between the different types of higher education, it is important to shed further light on their histories and students. The initial idea of the binary divide was that polytechnics would constitute a parallel form of

5. Call-ups for military conscription ended on 31 December 1960.
6. Later the number was increased to 30.

higher education with a special "commitment to non-degree students and to part-time courses" (Pratt 1997). Thus, polytechnics allowed the government to cater higher education to a wide range of students. The number of students in advanced courses, including degree courses, in polytechnics had been steadily rising over the years, catching up with the number of university students in 1992 (Figure 4.1.1). Unlike universities, the polytechnics could not award their own degrees. Therefore, most of the degree courses offered by the polytechnics were validated by the Council for National Academic Awards (CNAA). The CNAA was established in 1964 with the aim of granting awards to non-university students who had completed courses of study comparable in standards to university (Pratt 1997).



Notes: Figure 3.2 in Pratt (1997).

Figure 4.1.1: Number of students in polytechnics and universities

Despite polytechnics shifting the focus from non-advanced to advanced courses, they were more oriented towards undergraduate and part-time students compared to universities, based on the data reported in Pratt (1997). Although the overall share of part-time students in polytechnics fell from over 70% in 1965 to about 30% in 1988, most of the decline was driven by reduction in non-degree courses. The share of part-time students on advanced courses fell modestly from 40% in 1972 to 30% in 1992. At the same time, part-time students in universities rose from 9% in 1972 to 15% in 1992. In terms of level of study, the share of undergraduate students in degree courses in polytechnics remained at about 87%.

Polytechnics encompassed students from a wider range of backgrounds. For example, non-white students accounted for 14% of students in degree courses in polytechnics in 1991, compared to 8% in universities. Students aged 21 and over constituted about 50% of all full-time and more than 85% of part-time students in polytechnics. The focus on a wider population also led polytechnics to have broader admission criteria. While 70% of university students were admitted based on 3 or more GCE A-level passes in 1990, the corresponding share among polytechnics entrants was only 34%.

These figures suggest that the polytechnics did indeed provide access to higher education to a larger group of people. But it came at the expense of quality perception: polytechnics were not viewed as a parallel form of higher education, rather they were seen as secondary to universities (Willetts 2017; Pratt 1997).

All of these factors suggest that distinguishing the degree-holders by types of institutions is important, especially given the context laid out by our companion paper Dimbou et al. (n.d.). As explained above, in this companion paper, we are interested in changes in the composition of students during the massive expansion of higher education in the UK from 1960 to 1990. We are especially interested in how individuals of different abilities sort into different types of higher education institutions (HEIs). In this regard, it is important to separate HEI types for the following reasons. First, universities and polytechnics taught different subjects and had different objectives. Second, differences in public perception of the two types of HEIs could also translate to different returns to higher education. These concerns could imply different sorting patterns of individuals in types of higher education institutions based on their abilities. These ability-sorting patterns could have persisted even past 1992.

4.2 Data

We use the UK Household Longitudinal Study[7] (UKHLS), also known as the Understanding Society. This is the largest household panel study in the UK of about 40,000 individuals that started in 2009. The UKHLS is tightly related to a previous smaller longitudinal study, the British Household Panel Survey (BHPS), that was carried out in 18 waves between 1991 and 2009 and covered around 10,000 individuals. The UKHLS questionnaires were built upon those of the BHPS, ensuring continuity of many variables between the two studies. In addition to this, respondents in the last wave of the BHPS were asked if they were willing to join the UKHLS and about 80% of them agreed and were followed within the UKHLS.

The UKHLS covers a wide range of topics and is one of the most popular data sources for research. In the context of our project, however, the UKHLS offers very limited information

7. University of Essex, Institute for Social and Economic Research (2020)

about the type of institutions respondents received their higher education qualifications from. This limitation comes from the fact that only individuals who were full-time students at the time of the interview were asked about the type of institution attended[8]. We seek to exploit the close relationship between the BHPS and the UKHLS to overcome this limitation. The main advantage of the BHPS for our purposes is that it attempts to distinguish between different types of higher education institutions (HEIs), possibly reflecting the importance of reorganisation in the university sector at the time. In particular, respondents were asked to categorise their further education institution last attended among the following: nursing school, college of further education, other training establishment, polytechnic, university or other. Thus, we can use the information on the institution type in the BHPS to impute the variable in the UKHLS.

Our working sample consists of respondents born in the UK between 1950 and 1984[9] with non-missing information on the highest qualification obtained (from any institution type) and non-zero response weights. We exclude respondents who were still in education. We also exclude respondents from the minority boost samples. This results in 6,800 observations in wave 18 of the BHPS and 20,771 observations in wave 3 of the UKHLS, of which 5,117 are former BHPS subjects.

Our main variable of interest is the highest qualification achieved. This is a categorical variable with the following values: no qualification, GCSE or equivalent, A-level or equivalent, other higher degree[10], degree and other qualifications. We construct a binary degree attainment measure equal to one if the respondent has a degree (from any type of HEI). We define a university degree attainment variable as equal to one if a person both has a degree and attended a university as indicated by the type of HEI last attended. The latter variable is only defined for the BHPS subsample and is missing for all the UKHLS respondents.

We also need measurement of cognitive abilitiy of respondents since we adopt the analysis framework of our companion paper Dimbou et al. (n.d.). Therefore, we focus on wave 3 of the UKHLS as it is the only wave in which the participants were administered cognitive ability tests. The tests were composed of five parts: word recall, serial 7 subtraction, number series, verbal fluency and numeric ability. We combine the counts of correct answers to each of

8. In wave 5, respondents who completed their studies since 1995 were asked to give the name of institutions they attended. In the currently undergoing wave 11, every respondent is asked about their higher education institution, irrespective of completion status and date of completion. The data from wave 11 is to be released at the end of 2021.

9. That is, we exclude people born before or during the war. We also exclude cohorts born after 1985 as members of these cohorts had not yet completed their education. We verified that the share of respondents in UKHLS wave 3 having a degree as their highest qualification (from any institution type) starts declining past 1985.

10. Diploma in HE, teaching qualification and nursing/other medical qualification.

these tests into a single intelligence score using principal component analysis and extracting the first component[11]. In order to abstract from age-related differences in the test scores, we standardize both the correct answer counts and the resulting intelligence score within each decennial year of birth group.

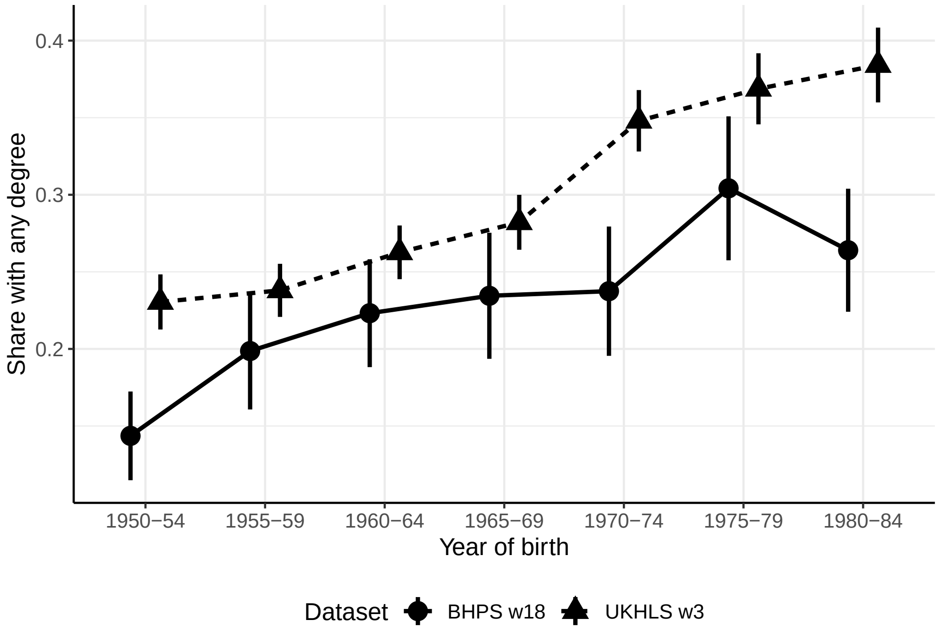
Furthermore, we use two external data sources on university graduates to establish a benchmark for university degree attainment over time. The Universities Statistical Records (USR) provides detailed annual information about all universities funded by the University Grants Committee during the period from 1972 to 1993. Specifically, we use undergraduate records with student-level microdata. From 1994 onwards this data was released by the Higher Education Statistics Agency (HESA) in the form of aggregated tables. In particular, we use data on the number of first-degree graduates at each university. We combine the two sources to create a time-series of number of graduates from old universities, a measure closely related to university degree attainment we are interested in. Note that the USR only covers the old universities, so we count all first-degree British graduates in the dataset. In the HESA tables, we exclude student counts in new universities (i.e., former polytechnics). To compute graduation rates we divide the series by the population of 21-year-olds at the time of graduation. In the rest of the paper, we call this series the graduation rate from old universities.

4.2.1 Degree attainment in the BHPS and the UKHLS

Before moving onto the discussion of our multiple imputation strategy, we examine degree attainment measures in the UKHLS and the BHPS. In Figure 4.2.1 we plot the average degree attainment by year of birth in wave 18 of the BHPS and wave 3 of the UKHLS. We can see that for most of the sample the two series are very close to each other with an exception of cohorts born after the mid-1970s. These differences hint that the BHPS might no longer reflect a representative sample of individuals born from 1975 onwards. This could be due to the fact that the BHPS was designed to reflect the population of Great Britain in 1991, whereas the UKHLS reflects the UK population as of 2009[12].

11. It explains 28% of the variation in the data.

12. The BHPS is a representative sample of the adult population of Great Britain in 1991, Scotland and Wales in 1999 and Northern Ireland in 2001. Ideally, adulthood outcomes of children in the BHPS would be similar to outcomes of adults in the UKHLS born in the respective years. In practice, life events such as migration, institutionalization or death could make the two samples different. For example, it could be that pursuing university education and the subsequent career paths are more likely to involve migration, implying higher chances of these children dropping out of the sample coverage. Therefore, among the BHPS children born after 1975 we see fewer degree-holders than in the corresponding adult population in the UKHLS.

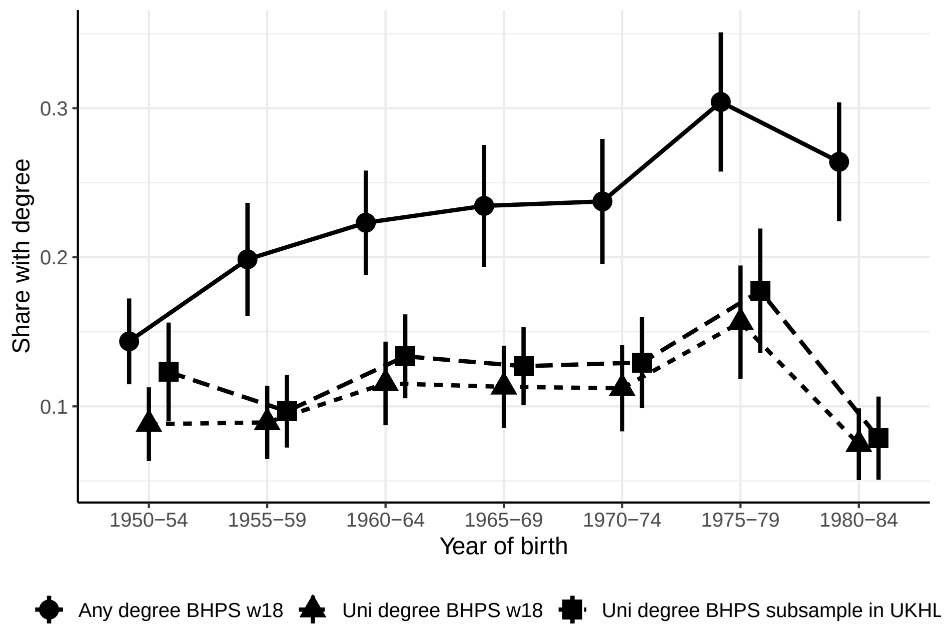


Notes: The figure compares average degree attainment from any type of HEI between the two samples: wave 3 of the UKHLS and wave 18 of the BHPS. The averages were weighted using the respective cross-sectional response weights. The whiskers corresponds to 95% confidence interval due to sampling uncertainty.

Figure 4.2.1: Degree attainment in UKHLS and BHPS

In Figure 4.2.2 we compare the degree attainment rates by types of HEI in the two samples. Recall that the UKHLS also follows former BHPS subjects from wave 2 onwards. Thus, we can recover institution types of former BHPS subjects in the UKHLS using their previous responses. Therefore, the figure essentially compares university degree attainment in the full BHPS sample and in the BHPS subsample of the UKHLS. We can see that the two are very close to each other in magnitude and dynamics. This is reassuring as it suggests that the decision to continue from the BHPS to the UKHLS is unlikely to be related to the university degree type.

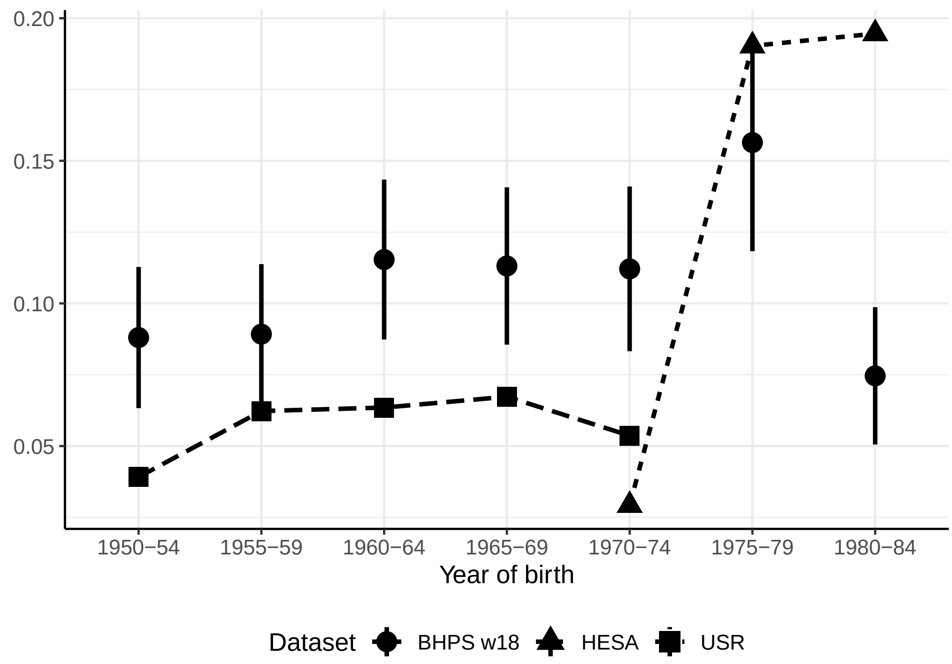
We also note that among cohorts born in 1960s and 1970s about half of degrees were obtained from universities. But almost all degrees of younger cohorts were obtained from universities, at least in wave 18 of the BHPS. Some of this difference could be related to lower representativeness of BHPS for cohorts born after 1975. But it could also be linked to the questionnaire design. The available options for the HEI type only differentiate universities from polytechnics, but not necessarily from former polytechnics. Therefore, our measure of university degree attainment based on the BHPS data may also contain the graduates who earned their degrees from former polytechnics. By the time people born from the mid-1970s onwards turned 20, 30 polytechnic institutions had already obtained university status.



Notes: The figure compares shares of people with degree by type of HEI among BHPS respondents. The degree attainment is calculated using the full sample from wave 18 of the BHPS. The university degree attainment is computed both in the full sample from wave 18 of the BHPS and in the BHPS subsample from wave 3 of the UKHLS. The averages are weighted using cross-sectional response weights from respective waves. The whiskers correspond to 95% confidence intervals due to sampling uncertainty.

Figure 4.2.2: Degree attainment by HEI type

To have a better understanding of the quality of the university degree attainment measure in the BHPS we further compare it with the benchmark graduation rates from old universities in Figure 4.2.3. First, we note that the USR ends in 1993 and the HESA starts in 1994. Thus, both datasets do not include the full cohorts of graduates born in 1970-74, which explains the sharp discontinuity in the graduation rates of this birth cohort. Second, we can see that university degree attainment measure in the BHPS is considerably higher than the benchmark graduation rates from old universities, for most of the birth cohorts. One possible explanation is that the benchmark graduation rate from old universities is computed with a larger denominator. We use total population counts of 21-year-olds in the graduation year, which also includes people not born in the UK. Another explanation could be related to the questionnaire design issue, mentioned earlier. Since the survey does not necessarily differentiate universities from *former polytechnics*, respondents may have categorised their institutions as universities based on the status at the time of interview, not at the time of graduation.



Notes: The figure compares university degree attainment in wave 18 of the BHPS with the benchmark graduation rates from old universities computed using data from the USR and HESA. The graduation rate is computed as the ratio of number of British first-degree graduates from old universities to the population of 21-year-olds at the time of graduation. The averages in the BHPS wave 18 are weighted using cross-sectional response weights. The whiskers corresponds to 95% confidence interval due to sampling uncertainty.

Figure 4.2.3: University degree attainment in the BHPS against the benchmark

4.3 Multiple imputation

We use multiple imputation as a way to deal with imputation uncertainty. Almost any imputation carries a level of uncertainty because the true value is unknown. Imputing a single value for each missing data point and giving it the same importance during the estimation as one gives an observed value fails to account for this uncertainty. In his seminal work, Rubin (1977) presented a multiple imputation technique where imputing multiple values for each missing data allows to incorporate the imputation uncertainty into the main analysis of interest.

The general concept of working with multiply imputed datasets is simple as we explain using Figure 4.3.1, adapted from van Buuren (2018). First, each missing value in the original dataset is assigned multiple imputed values. This could be thought of as creating multiple datasets that only differ in the values assigned to originally missing cells. All these datasets are identical to the original dataset in terms of originally observed values. These datasets are referred to as completed datasets. In each of the completed datasets we run the analysis of interest using standard techniques and obtain corresponding estimates and standard errors. That is, in each completed dataset we can "forget" that some data points were imputed and use standard regression analysis. Finally, the estimators from each completed dataset are combined together into a single estimator using “Rubin’s rules”. This way the combined estimator accounts for both sampling and imputation uncertainty.

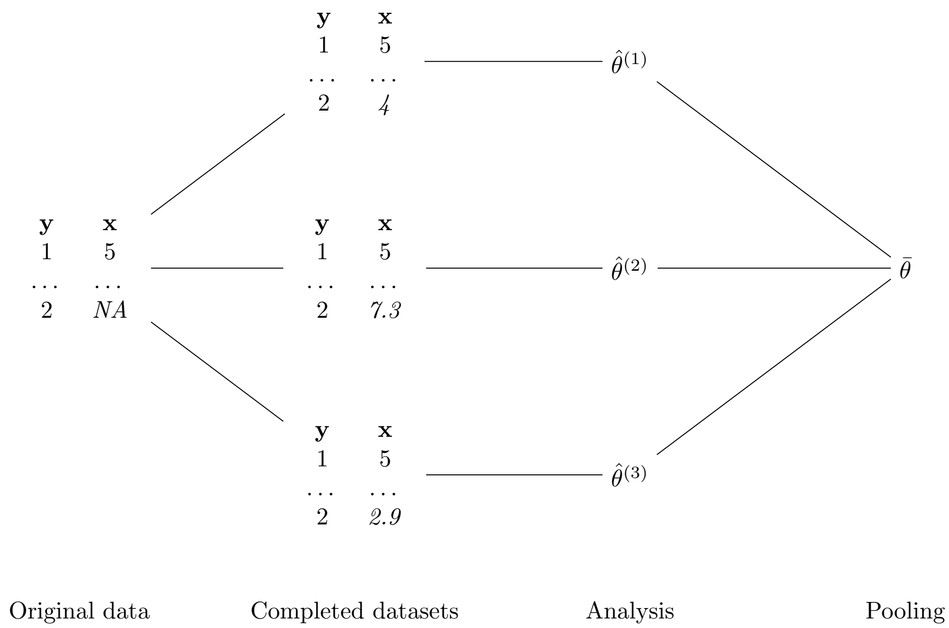


Figure 4.3.1: General concept of working with multiply imputed data

For a formal definition of the combined estimator, denote the estimator from a completed dataset $m \in \{1, \dots, M\}$ by $\hat{\theta}^{(m)}$ and its variance-covariance matrix by $U^{(m)}$. Then,

$$\bar{\theta} = \frac{1}{M} \sum_{m=1}^{M} \hat{\theta}^{(m)} \tag{4.1}$$

$$T = \bar{U} + \left(1 + \frac{1}{M}\right) B \tag{4.2}$$

$$\bar{U} = \frac{1}{M} \sum_{m=1}^{M} U^{(m)} \tag{4.3}$$

$$B = \frac{1}{M-1} \sum_{m=1}^{M} \left(\hat{\theta}^{(m)} - \bar{\beta}\right)^2 \tag{4.4}$$

The combined estimator $\bar{\theta}$ is a simple average of the estimators from M completed datasets. The combined or total variance T of the estimator $\bar{\theta}$ consists of two parts: the within- $\bar{U}$ and between-imputation variance B in equations (4.3) and (4.4), respectively.

Before proceeding to a formal overview of multiple imputation and the necessary assumptions, let us introduce some notation. Suppose we have a sample of size n where a variable Y is missing for some observations. We can construct a corresponding indicator variable $R = \mathbb{1}\{Y \text{ is observed}\}$. It is also convenient to denote the vector with only observed values as Y_{obs} and similarly the vector with missing values Y_{mis}. Note Y_{mis} is a latent variable that contains true values of Y for observations with missing data, but we as researchers do not observe it. Assume the other variables in the sample do not have missing values and denote them by $\mathbf{X}$.

4.3.1 Missing data mechanism

The aim of imputation is to fill in the Y_{mis} values. In most cases, we do not know the true values of Y_{mis} with certainty, but we can characterise their distribution conditional on observed information. In other words, the imputation process could be thought of as drawing observations from the conditional distribution characterised by $\Pr(Y_{mis}|Y_{obs}, X, R)$. In case of multiple imputation, we draw observations from this distribution multiple times. The conditional probability $\Pr(Y_{mis}|Y_{obs}, X, R)$ is the imputation model. How we define the imputation model depends on the processes that generate the missing data.

The missing data mechanisms can be categorized in three ways: missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). MCAR, as suggested by its name, assumes that the response indicator R is determined completely randomly and depends neither on observed nor missing values. Formally, this can be written as $\Pr(R = 1|Y, \mathbf{X}) = \Pr(R = 1)$. In the context of our paper Y is an indicator for having a university degree. If survey designers only asked a subsample about their HEI type and that subsample was determined randomly, then the missing data would be MCAR.

MAR relaxes the MCAR assumption by allowing the missingness to depend on observed values. That is, $\Pr(R = 1|Y, \mathbf{X}) = \Pr(R = 1|Y_{obs}, \mathbf{X})$. Expanding on the previous example, suppose older people could not correctly remember their institution type and therefore did not provide an answer. In such a case, the missing data would be MAR conditional on age.

Finally, under MNAR whether or not a value is missing may not only depend on observed variables, but also on the value itself. So, $\Pr(R = 1|Y_{mis}, Y_{obs}, X)$ does not simplify. In our example, if specifically people who obtained a degree from former polytechnics did not disclose this information, the missing data would be MNAR.

The missing data mechanism is closely related to the concept of ignorable nonresponse. Nonresponse can be called ignorable if the MAR assumption is satisfied and the parameters of the response model are distinct from parameters of the data-generating model in a sense that knowing one does not provide information about the other (Rubin 1987; Schafer 1997). Ignorable nonresponse allows us to build the imputation model using only observed data without explicitly modelling the missing data mechanism. That is, we can draw imputations from $\Pr(Y|Y_{obs},X,R=1) = \Pr(Y_{mis}|Y_{obs},X,R=0)$. If nonresponse is non-ignorable, we would have to model the missing data mechanism explicitly because $\Pr(Y|Y_{obs},X,R=1) \neq \Pr(Y_{mis}|Y_{obs},X,R=0)$.

4.3.2 Imputation model

Notice that the example we give to describe MCAR is very similar to the case of the UKHLS and the BHPS. People in the BHPS were asked about their institution type, people in the UKHLS were not. If the BHPS sample is not systematically different from the UKHLS one, i.e. they are drawn randomly from the same population, then indeed we have a case of MCAR. We examine this claim in Table 4.3.1 by comparing socio-economic and family characteristics of two samples: wave 18 of the BHPS and wave 3 of the UKHLS. In wave 18 of the BHPS we consider the respondents with valid information on university degree attainment, i.e., those who have non-missing highest qualification and HEI type information. Panel A of Table 4.3.1 compares the BHPS sample with the entire wave 3 of the UKHLS including former BHPS respondents; panel B - with wave 3 of the UKHLS excluding former BHPS respondents. From Table 4.3.1 we conclude that the samples are very different from one another. This suggests that the UK population has changed significantly in the 20 years between the BHPS and the UKHLS. Individuals in the UKHLS are more likely to have continued past compulsory schooling and have some tertiary degree, have slightly lower earnings (although this comparison may be confounded by the financial crisis), and have higher educated parents.

Thus, we need to construct an imputation model that is likely to satisfy the MAR assumption. In order to do so, we need to include covariates that help explain both the university degree attainment and the missing data mechanism. We already know from Table 4.3.1 that degree attainment, earnings, and parental educational qualifications are highly correlated with the missing data mechanism. Furthermore, the imputation of university degree attainment only makes sense if an individual has a degree. Therefore, we can focus on building an imputation model among degree-holders, setting the university degree attainment variable to zero for everyone else.

Another important consideration in developing our imputation model is its congruence with the analysis model (Schafer 1997). Failure to include the terms of interest means that the imputation model restricts their coefficients to zero. This, in turn, results in attenuated coefficients of interest in the analysis stage. As mentioned earlier, we adopt the analysis context from our companion paper Dimbou et al. (n.d.). In particular, we would like to know

Table 4.3.1: Testing MCAR assumption between BHPS wave 18 and UKHLS wave 3

	UKHLS w3			BHPS w18			Diff			
	Mean	S.D.	N	Mean	S.D.	N	Coef	S.E.	FWER p-val	% mean BHPS
Panel A: wave 18 of the BHPS vs wave 3 of the UKHLS										
Individual characteristics										
Female	0.525	0.499	20,771	0.523	0.500	6,563	0.002	0.005	1.0	0.4
White british	0.891	0.311	20,771	0.900	0.300	6,563	-0.009	0.007	1.0	-1.0
Born in England	0.747	0.434	20,402	0.776	0.417	6,094	-0.029**	0.008	0.0	-3.7
Age in 2008	41.545	9.817	20,771	42.680	9.330	6,563	-1.134***	0.180	0.0	-2.7
Post-compulsory edu	0.587	0.492	20,771	0.491	0.500	6,563	0.096***	0.010	0.0	19.5
Any degree	0.298	0.458	20,771	0.193	0.395	6,563	0.105***	0.009	0.0	54.4
Ever married	0.793	0.405	20,771	0.809	0.393	6,563	-0.016	0.007	0.3	-2.0
Any children	0.785	0.411	20,771	0.789	0.408	6,563	-0.003	0.007	1.0	-0.4
Working	0.751	0.433	20,771	0.800	0.400	6,563	-0.050***	0.007	0.0	-6.2
Real monthly earnings	17.947	17.732	20,771	20.036	19.481	6,563	-2.089***	0.366	0.0	-10.4
Family characteristics at age 14										
Father has degree	0.117	0.321	16,993	0.093	0.290	5,380	0.024***	0.006	0.0	26.1
Mother has degree	0.075	0.264	17,555	0.055	0.229	5,535	0.020***	0.005	0.0	35.9
Father employed	0.886	0.317	20,403	0.925	0.263	6,158	-0.039***	0.005	0.0	-4.2
Mother employed	0.657	0.475	20,509	0.632	0.482	6,276	0.025*	0.009	0.1	4.0
Father born in Eng	0.689	0.463	20,425	0.717	0.451	6,523	-0.028**	0.009	0.0	-3.9
Mother born in Eng	0.696	0.460	20,491	0.721	0.449	6,545	-0.025*	0.010	0.1	-3.5
Panel B: wave 18 of the BHPS vs wave 3 of the UKHLS excl. former BHPS respondents										
Individual characteristics										
Female	0.526	0.499	15,654	0.523	0.500	6,563	0.003	0.006	1.0	0.5
White british	0.888	0.315	15,654	0.900	0.300	6,563	-0.012	0.008	1.0	-1.3
Born in England	0.775	0.417	15,651	0.776	0.417	6,094	-0.001	0.010	1.0	-0.1
Age in 2008	41.339	9.840	15,654	42.680	9.330	6,563	-1.341***	0.205	0.0	-3.1
Post-compulsory edu	0.595	0.491	15,654	0.491	0.500	6,563	0.105***	0.011	0.0	21.3
Any degree	0.311	0.463	15,654	0.193	0.395	6,563	0.118***	0.010	0.0	61.2
Ever married	0.790	0.407	15,654	0.809	0.393	6,563	-0.019	0.008	0.2	-2.3
Any children	0.783	0.412	15,654	0.789	0.408	6,563	-0.006	0.008	1.0	-0.8
Working	0.744	0.436	15,654	0.800	0.400	6,563	-0.056***	0.008	0.0	-7.0
Real monthly earnings	17.840	17.921	15,654	20.036	19.481	6,563	-2.196***	0.405	0.0	-11.0
Family characteristics at age 14										
Father has degree	0.121	0.327	12,857	0.093	0.290	5,380	0.029***	0.007	0.0	31.0
Mother has degree	0.078	0.268	13,311	0.055	0.229	5,535	0.023***	0.006	0.0	40.9
Father employed	0.881	0.324	15,597	0.925	0.263	6,158	-0.045***	0.005	0.0	-4.8
Mother employed	0.663	0.473	15,609	0.632	0.482	6,276	0.031**	0.010	0.0	4.9
Father born in Eng	0.712	0.453	15,587	0.717	0.451	6,523	-0.005	0.011	1.0	-0.7
Mother born in Eng	0.720	0.449	15,638	0.721	0.449	6,545	-0.001	0.011	1.0	-0.1

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The Table compares average characteristics of individuals in wave 18 of the BHPS with those in wave 3 of the UKHLS. The BHPS sample is restricted to individuals with valid university degree attainment variable, i.e., those who have non-missing highest qualification and HEI type information. Panel A uses the entire wave 3 of the UKHLS for comparison and Panel B is restricted to wave 3 of the UKHLS excluding former BHPS respondents. The last four columns report the statistics of the difference in means between the BHPS and the UKHLS samples. The standard errors of the difference are clustered at the sampling strata-wave level. FWER p-values are computed as in Holm (1979) to adjust for multiple inferences and are used to assign significance stars to the coefficient of difference. The last column reports the size of the coefficient of difference relative to the mean in the BHPS sample in %. The estimates are weighted using the cross-sectional response weights in the respective waves.

how the expansion of higher education in the UK changed the composition of the university students in terms of their intelligence scores. In other words, how did the relationship between intelligence scores and university degree attainment change over time. Thus, we are interested in estimating the following equation:

$$\frac{\Pr(U_i = 1)}{1 - \Pr(U_i = 1)} = \exp\left(\alpha + \gamma_y + \delta \mathbf{X}_i + \beta_y \mathbf{X}_i\right) \tag{4.5}$$

where U_i is the university degree attainment variable, $\mathbf{X}_i$ contains gender, intelligence score and their interaction and γ_y are birth cohort fixed effects. Here, δ describes the relationship between intelligence score and university degree attainment by gender in the base birth cohort group, and β_y shows how this relationship has changed across birth cohorts relative to the base group. We acknowledge that the simple specification in equation (4.5) may suffer from an omitted variable bias. For example, it does not control for the educational qualifications of parents, a variable that could explain both the intelligence score of respondents and their university degree attainment status. For the purposes of this paper, we abstract from this issue and treat the intelligence score as if it were randomly distributed in the population. That is, we assume that parameters δ and β_y are true causal parameters describing the effect of intelligence and gender on university degree attainment probabilities. Nonetheless, the implication remains: the analysis and imputation models should be consistent with each other. If parental education enters the analysis model, it should also be part of the imputation model.

Thus, our imputation model should ideally contain all the regressors from equation (4.5), including their interaction terms. Our imputation model can be written as follows

$$\Pr(U_i = 1) = \begin{cases} f\left(\zeta + \eta_y + \rho \ddot{\mathbf{Z}}_i + \lambda_y \tilde{\mathbf{Z}}_i\right) & \text{if } d_i = 1 \\ 0 & \text{if } d_i = 0 \end{cases} \tag{4.6}$$

where d_i is the degree attainment variable. That is, the probability of having a university degree is a function of a constant ζ, birth cohort fixed effects η_y and personal characteristics in $\ddot{\mathbf{Z}}_i$ and $\tilde{\mathbf{Z}}_i$, the latter of which is allowed to have an effect specific to each birth cohort. Given that we specify a non-trivial imputation model only among the degree-holders, the estimation sample consists of 1,540 observations. Therefore, we differentiate between the characteristics that enter the model linearly $\ddot{\mathbf{Z}}_i$ and those that are interacted with birth cohort indicators $\tilde{\mathbf{Z}}_i$. We note that $\mathbf{X}_i \in \tilde{\mathbf{Z}}_i \in \ddot{\mathbf{Z}}_i$. So, in addition to gender, intelligence score and their

Table 4.3.2: Imputation model versions

	Equation	Estimation sample
Model 1	without intelligence score	BHPS wave 18 and UKHLS wave 3
Model 2	without intelligence score	UKLHS wave 3
Model 3	with intelligence score	UKHLS wave 3

interaction term, the set of regressors in $\tilde{\mathbf{Z}}_i$ includes country of birth, race, an indicator for whether an individual has ever cohabited, an indicator for whether an individual has ever been married, an indicator for whether an individual has any children, the number of children, the second-order polynomial of real earnings, and the employment status at the time of interview. Besides $\tilde{\mathbf{Z}}_i$, the set of linear regressors $\ddot{\mathbf{Z}}_i$ includes

- *Individual characteristics*: years of education, age when left further education, residence in England at the time of interview, indicator whether current residence is in the country of birth, car ownership, indicator for having a second job, major occupational group of main job, major occupational group of first job
- *Parental characteristics*: countries of birth of father and mother, highest educational qualifications of father and mother, employment statuses of father and mother when the respondent was 14 years old
- *Design variables*: survey design weight, survey response weight.

Unfortunately, cognitive ability tests were only administered in wave 3 of the UKHLS. On the one hand, failing to add intelligence score to the imputation model will attenuate the correlation between university degree attainment and intelligence (Schafer 1997) unless it is not captured by the rest of the predictors. On the other hand, using former BHPS respondents in wave 3 of the UKHLS for estimating the imputation model presents an additional concern. Not only were they sampled from a different UK population, but they have also self-selected to continue into the UKHLS. This could violate MAR assumption, if the decision to continue from the BHPS is correlated with the HEI type even after controlling for a set of observed characteristics. Due to the potential benefits and drawbacks of both including and excluding the intelligence score terms, we consider three different versions of the imputation model presented in Table 4.3.2. These versions differ based on the estimation samples and the inclusion of intelligence score terms.

Next, we test if our model specifications violate the MAR assumption in Table 4.3.3. Determining if missing data is MAR or MNAR is essentially impossible. However, we can test if the missingness indicator is correlated with other variables even after conditioning on all the variables included in the model. If so, then the imputation model clearly violates the MAR assumption and the given variable should be included. According to the results in Table 4.3.3, the strongest signals of failure of the MAR assumption are observed in Model 1. Interest in politics and neighbourhood characteristics are both strongly correlated with the missing data indicator. Models 2 and 3 display fewer, if any, statistically significant violations. But even models 2 and 3, for some variables the estimated differences between missing and non-missing subsamples are large in magnitude: they constitute more than 10% of the sample mean in the UKHLS. We add these variables to $\ddot{\mathbf{Z}}_i$.

4.3.3 Evaluation

Denote by F_i the indicator which takes the value of 1 if individual i is female and denote by I_i her intelligence score. Then, $\mathbf{X}_i = (F_i, I_i, F_iI_i)$ and the analysis model in equation (4.5) can be rewritten as

$$\frac{\Pr(U_i = 1)}{1 - \Pr(U_i = 1)} = \exp\left(\alpha + \gamma_y + (\delta^F + \beta_y^F)F_i + (\delta^I + \beta_y^I)I_i + (\delta^{FI} + \beta_y^{FI})F_iI_i\right)$$

Here, γ_y are cohort fixed effects, $\delta = (\delta^F, \delta^I, \delta^{FI})$ describe the effects of gender, intelligence score and their interaction term on the log odds ratio of the probability of having a university degree in the base birth cohort group. Given our sample restrictions and our definition of birth cohort groups, the base cohort are people born in 1950-54. Then, the parameters $\beta_y = (\beta_y^F, \beta_y^I, \beta_y^{FI})$ capture how the effects of gender, intelligence score and their interaction term change across birth cohorts relative to the base group. We collectively denote these parameters by $\theta = (\gamma_y, \delta, \beta_y)$. The goal in this subsection is to study the properties of the combined estimator $\bar{\theta}$ from the multiply imputed data.

In the absence of missing data and under random sampling we could obtain the consistent estimator $\hat{\theta}$ and rely on its asymptotic distribution to draw an inference. The question here is whether we can draw a valid inference using the combined estimator $\bar{\theta}$. Rubin (1987, chapter 4) studies the properties of the combined estimator from the random-response randomization-based perspective. In short, he outlines two sufficient conditions for the randomization-validity of the combined estimator $\bar{\theta}$. First, the complete-case estimator $\hat{\theta}$ should be randomization-valid for θ. That is, in the absence of missing data, our estimator $\hat{\theta}$ should be consistent for the parameter of interest θ. Moreover, the 95% confidence interval around $\hat{\theta}$ should contain

Table 4.3.3: Testing MAR assumption

	Model 1				Model 2				Model 3			
Dependent variable	$1-R$	N obs	N miss	% mean	$1-R$	N obs	N miss	% mean	$1-R$	N obs	N miss	% mean
Individual characteristics												
Has mobile	0.001	1,304	5,049	0.1	0.000	1,015	5,049	0.0	0.000	967	4,877	0.0
	(0.009)				(0.009)				(0.010)			
	[1.000]				[1.000]				[1.000]			
Supports a polit party	-0.006	1,304	5,049	-1.6	-0.005	1,015	5,049	-1.3	-0.024	967	4,877	-6.7
	(0.023)				(0.020)				(0.021)			
	[1.000]				[1.000]				[1.000]			
Responsible for child under 16	-0.019	1,304	5,049	-8.0	-0.013	1,015	5,049	-5.4	-0.009	967	4,877	-3.7
	(0.011)				(0.011)				(0.011)			
	[0.595]				[1.000]				[1.000]			
Likely to move	-0.045	1,304	5,049	-7.3	-0.043	1,015	5,049	-6.9	-0.030	967	4,877	-4.9
	(0.023)				(0.021)				(0.021)			
	[0.352]				[0.431]				[1.000]			
Interested in politics	0.116***	1,304	5,049	17.7	0.052	1,015	5,049	8.0	0.031	967	4,877	4.7
	(0.021)				(0.020)				(0.021)			
	[0.000]				[0.126]				[1.000]			
Received interest on savings	0.033	1,259	4,848	7.4	0.007	967	4,848	1.6	0.004	928	4,685	0.9
	(0.024)				(0.025)				(0.024)			
	[0.819]				[1.000]				[1.000]			
Interest on savings missing	-0.003	1,304	5,049	-7.4	0.006	1,015	5,049	14.6	0.015	967	4,877	37.5
	(0.009)				(0.008)				(0.008)			
	[1.000]				[1.000]				[0.728]			
Good fin situation (subjective)	-0.009	1,304	5,049	-1.3	-0.001	1,015	5,049	-0.2	-0.011	967	4,877	-1.5
	(0.022)				(0.021)				(0.021)			
	[1.000]				[1.000]				[1.000]			
Current residence: access to services												
Good shopping services	0.058	1,304	5,049	9.4	-0.067*	1,015	5,049	-10.9	-0.049	967	4,877	-7.9
	(0.026)				(0.025)				(0.026)			
	[0.201]				[0.093]				[0.728]			
Good public transp services	0.143***	1,304	5,049	27.5	-0.015	1,015	5,049	-2.9	-0.027	967	4,877	-5.2
	(0.025)				(0.024)				(0.025)			
	[0.000]				[1.000]				[1.000]			
Good medical services	0.094***	1,304	5,049	12.7	0.037	1,015	5,049	5.0	0.043	967	4,877	5.8
	(0.025)				(0.022)				(0.023)			
	[0.001]				[0.827]				[0.728]			
Good leisure services	0.086***	1,304	5,049	17.4	0.004	1,015	5,049	0.8	0.003	967	4,877	0.6
	(0.023)				(0.024)				(0.024)			
	[0.001]				[1.000]				[1.000]			
Likes neighbourhood	0.021	1,304	5,049	2.2	0.003	1,015	5,049	0.3	0.006	967	4,877	0.6
	(0.010)				(0.008)				(0.008)			
	[0.334]				[1.000]				[1.000]			

${}^{*}p < 0.1$; ${}^{**}p < 0.05$; ${}^{***}p < 0.01$

Notes: The Table reports the results from linear regressions of dependent variables in the first column on the missingness indicator conditional on all the terms in the corresponding imputation model. In particular, models 1 and 2 do not control for intelligence score terms, while model 3 does. The estimation sample of model 1 includes wave 18 of the BHPS and wave 3 of the UKHLS; that of models 2 and 3 is restricted to wave 3 of the UKHLS only. The estimations are weighted using the respective cross-sectional weights and clustered at the level of sampling strata and wave. Regular standard errors are reported in parentheses and FWER adjusted p-values - in square brackets. The significance stars are assigned based on the FWER adjusted p-values. The Table also reports the size of the estimated coefficients relative to the sample mean in wave 3 of the UKHLS in %.

the true parameter θ in 95 out of 100 samples from the population. Second, the imputation model should be *proper* meaning that $\bar{\theta}$ should be randomization-valid for $\hat{\theta}$.

It is easy to see that the first assumption is unrelated to missing data. Even if there were no missing data, we would need to satisfy this assumption to be able to draw inference about θ based on the estimator $\hat{\theta}$. When it comes to the second assumption, analytical studies of the properties of the combined estimator $\bar{\theta}$ are usually very difficult (Schafer 1997). Therefore, researchers have developed simulation-based methods. We adopt the algorithm proposed by Brand et al. (2003) to evaluate bias and efficiency of the combined estimator $\bar{\theta}$ in each imputation model and then compare them across models. The idea is the following:

1. Estimate equation (4.5) in the subsample with both nonmissing university degree attainment information and nonmissing intelligence score. We obtain estimates $\tilde{\theta}$, which we now treat as true parameters. Use the estimated $\tilde{\theta}$ fill in the missing values of the university degree attainment variable U. Denote the simulated variable by $\widetilde{U}$, which has no missing values.

2. Specify a missing data mechanism. We select one of the following response models:

   a. MCAR: $\Pr(R_i = 1) = \alpha, \forall i$ where α is a constant.

   b. MAR: $\Pr(R_i = 1|\ddot{\mathbf{Z}}_i, \tilde{\mathbf{Z}}_i)$ consistent with the imputation models specified in the previous subsection.

   c. MNAR: $\Pr(R_i = 1|\widetilde{U}_i) = \alpha + 0.2\left(\widetilde{U}_i - \Pr(\widetilde{U}_i = 1)\right)$ so that individuals with a university degree are 20% more likely to have a nonmissing value.

3. Generate L incomplete datasets under the chosen missing data mechanism. This is done by drawing response indicators $R^{(l)}, \forall l \in \{1, \ldots, L\}$ according to the probabilities specified in the previous step and setting $\widetilde{U}_i^{(l)}$ to missing whenever $R_i^{(l)} = 0$. Brand et al. (2003) recommend L in the range between 200 and 1000. We set $L = 500$.

4. Within each of the L simulated incomplete datasets, we multiply impute the university degree attainment variable and estimate $\bar{\theta}^{(l,v)}, \bar{U}^{(l,v)}, B^{(l,v)}, \quad \forall l \in \{1, \ldots, L\}$ and $\forall v \in \{\text{Model 1}, \text{Model 2}, \text{Model 3}\}$.

5. Given the combined estimators, compute the bias and efficiency statistics:

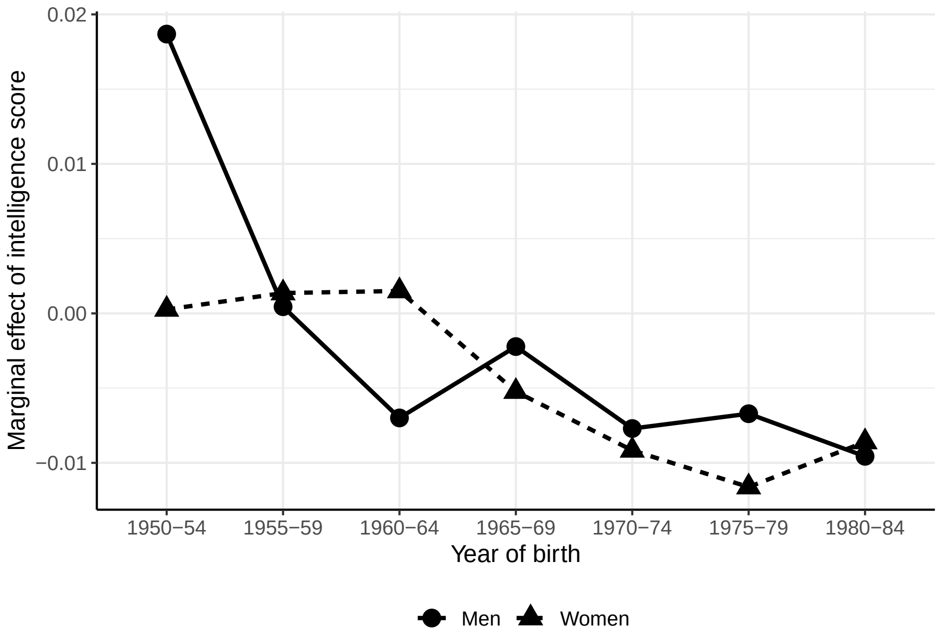
   a. Raw bias: $\frac{1}{L}\sum_{l=1}^{L} \bar{\theta}^{(l,v)} - \tilde{\theta}$

   b. Coverage rate: $\frac{1}{L}\sum_{l=1}^{L} \mathbb{1}\{\tilde{\theta} \in 95\% \text{ CI in dataset } l\}$

c. Average width: $\frac{1}{L}\sum_{l=1}^{L}$ width of 95% CI in dataset l

6. Repeat steps 3 - 5 for all missing data mechanisms described in step 2.

Let's first examine the "true" parameters from step 1. Figure 4.3.2 plots the marginal effect of one standard deviation increase in intelligence score on the probability of having a university degree by gender across birth cohorts. The marginal effects are computed using the estimates of equation (4.5) in the subsample with no missing data. As mentioned earlier, we treat these estimates as true for the purposes of this paper. They are likely to change when the analysis model includes possible confounders as controls, such as parental education. Nevertheless, we observe a declining trend. A one standard deviation increase in the intelligence score of individuals born in 1950-54 had a positive effect on university degree attainment probabilities for men. But by the time people born in 1975-79 were of getting their higher education, an increase of one standard deviation in the intelligence score meant a 0.7-1.2pp decrease in the probability of university degree attainment.



Notes: The figure plots the marginal effects of intelligence score on the university degree attainment probabilities by gender and birth cohorts. The marginal effects were computed using the estimates of equation (4.5) in the BHPS subsample of the UKHLS. The estimation sample only includes individuals with non-missing university degree variable and non-missing intelligence score. The marginal effects are computed at mean intelligence score.

Figure 4.3.2: "True" parameters used to initiate the evaluation algorithm

We now turn to examining the results of our simulation-based evaluation. We simulated 500 datasets where some observations were randomly set to missing according to the MAR response model. In each of the 500 simulated datasets, the missing values were multiply imputed using the three versions of our imputation model. Thus, we get three combined estimators from each simulated dataset. Table 4.3.4 reports the bias and efficiency statistics of the combined estimators for the marginal effect of intelligence score on the university degree attainment across imputation models. It is evident from the table that the imputation model 1 has the worst performance across all indicators compared to models 2 and 3. The estimators from model 1 have the largest in magnitude bias statistic and lowest coverage rate. For example, looking at the estimators resulting from imputation model 1, the 95% confidence interval of the marginal effect of intelligence score on university degree attainment for men born in 1965-69 contains the true value of the marginal effect only 20.4% of the time. We can also see that the imputation models 2 and 3 display very similar results, both in terms of bias and in terms of efficiency. This result is somewhat surprising as it would suggest that the imputation model does not need to control for the intelligence score. One possible explanation for such similarity between the results from imputation models 2 and 3 could be that conditioning on the intelligence score is to a large extent redundant after controlling for the rest of the variables in the imputation model.

Thus, given the results in Table 4.3.4, we impute the missing observations in the university degree variable in the UKHLS using imputation models 2 and 3. That is, we use the UKHLS respondents in wave 3, where only the former BHPS subjects have a non-missing university degree information.

4.4 Results and discussion

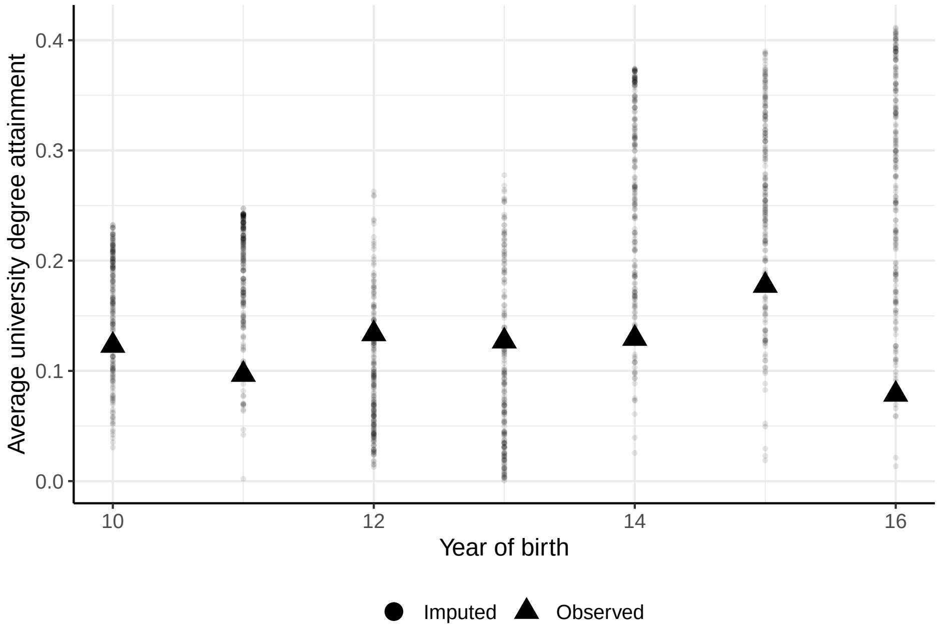
We first examine if the imputed values are plausible. Figure 4.4.1 plots the average university degree attainment over time among observed and imputed subsamples. Each dot in the plot corresponds to the average university degree attainment among imputed observations in each of the completed datasets. The triangles correspond to the average university degree attainment among the BHPS subsample of the UKHLS. Overall, the distribution of the imputed values are in plausible ranges of average university degree attainment. Furthermore, the imputed values seem to follow the general trend of rising university education.

In Figure 4.4.2 we directly compare average university degree attainment over time with the graduation rates computed using the USR and the HESA datasets. Reassuringly, the estimates of the average university degree attainment produced by the two imputation models are, in general, close to the benchmark graduation rates in the USR and the HESA datasets. It is notable that despite the fact that average university degree attainment is considerably underestimated in the observed subsample born in 1980-84 (see Figure 4.2.3), the multiple

Table 4.3.4: Evaluation results under MAR mechanism

	Model 1		Model 2		Model 3	
Birth cohort	Men	Women	Men	Women	Men	Women
Raw bias, pp						
1950-54	6.5	3.9	5.1	3.5	5.3	3.7
1955-59	2.1	2.5	1.5	2.0	1.4	2.1
1960-64	4.9	6.3	2.6	5.3	2.5	5.5
1965-69	5.5	4.6	3.8	4.0	3.9	4.0
1970-74	3.6	3.7	2.5	3.2	2.9	4.0
1975-79	8.7	5.2	6.3	3.6	5.5	3.0
1980-84	4.1	2.3	2.5	1.6	2.8	1.7
Coverage rate, %						
1950-54	13.8	35.4	46.8	75.2	49.0	70.2
1955-59	97.8	95.4	99.8	99.8	100.0	99.8
1960-64	33.2	0.6	95.0	16.6	95.2	12.6
1965-69	20.4	7.4	84.4	37.4	81.8	35.6
1970-74	91.2	78.0	99.6	75.6	95.4	53.6
1975-79	0.0	24.8	12.2	74.4	21.4	85.4
1980-84	92.8	100.0	98.6	99.8	95.6	99.6
Average width, pp						
1950-54	10.8	7.4	10.1	7.7	10.5	8.0
1955-59	7.2	6.9	6.5	6.9	6.4	7.0
1960-64	8.9	9.2	7.2	9.1	7.2	9.3
1965-69	9.7	7.6	8.8	7.5	9.0	7.6
1970-74	9.2	8.4	8.5	7.6	9.1	8.2
1975-79	11.7	9.3	10.6	8.3	9.8	8.1
1980-84	10.9	8.4	9.1	7.7	9.6	8.0

Notes: The table reports the bias and efficiency statistics for the marginal effect of intelligence score on university degree attainment probability computed as described in step 5 of the evaluation algorithm. The table uses the combined estimators from 500 simulated datsets with missing data. The misisng data were generated under the MAR mechanism.

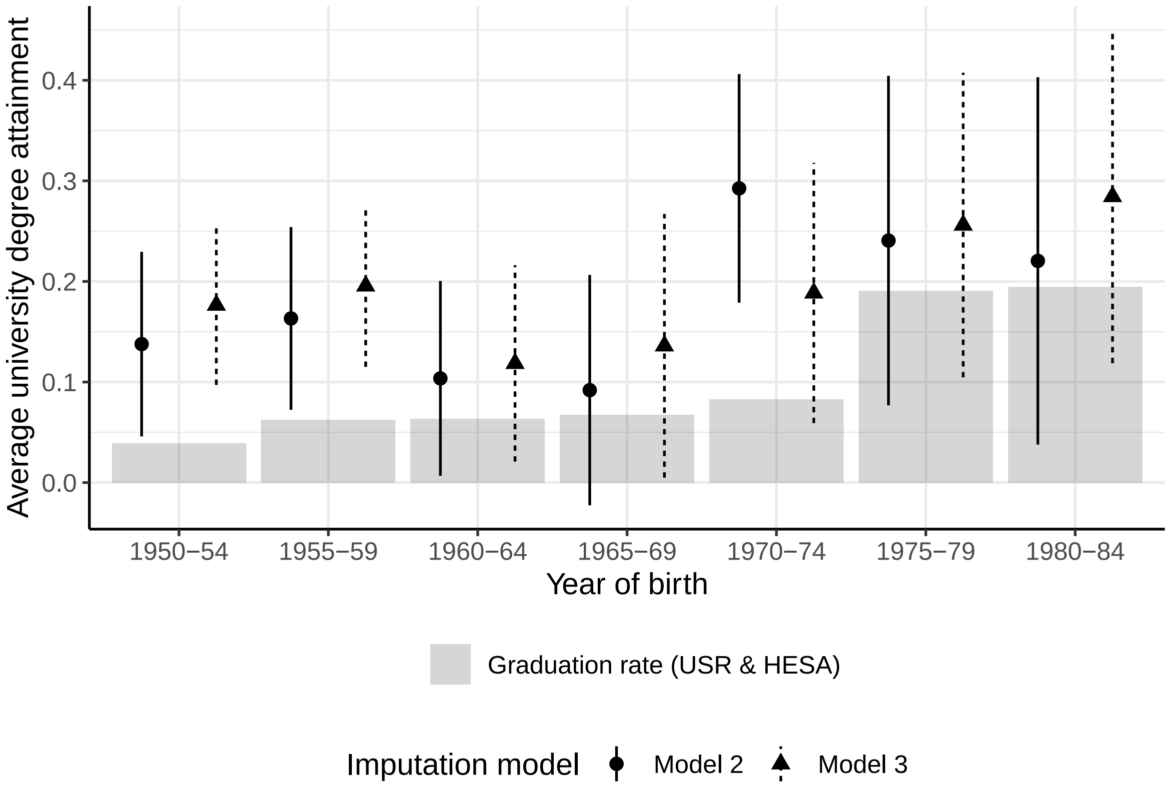


Notes: The figure plots the average university degree attainment over time in observed and imputed subsamples. The averages are calculated using wave 3 of the UKHLS and weighted using the corresponding cross-sectional weights. The triangles correspond to the average university degree attainment of the BHPS subsample, i.e. observed. The dots correspond to average university degree attainment among individuals with originally missing data. Their averages are computed within each imputed set out of total M sets.

Figure 4.4.1: Average university degree attainment over time

imputation yields estimates comparable to the benchmark graduation rates. However, multiple imputation did not perform well in the older subsample. The combined estimators from both imputation models 2 and 3 considerably overestimate the university degree attainment rates for this cohort. So, for example, the combined estimators for average university degree attainment among individuals born in 1950-54 are at 14-18%, whereas the graduation rate for that cohort is at 4%, according to the USR dataset. For the cohorts born after 1960, the imputation model 3 produces completed datasets that are most consistent with the benchmark graduation rates obtained from the USR and the HESA datasets.

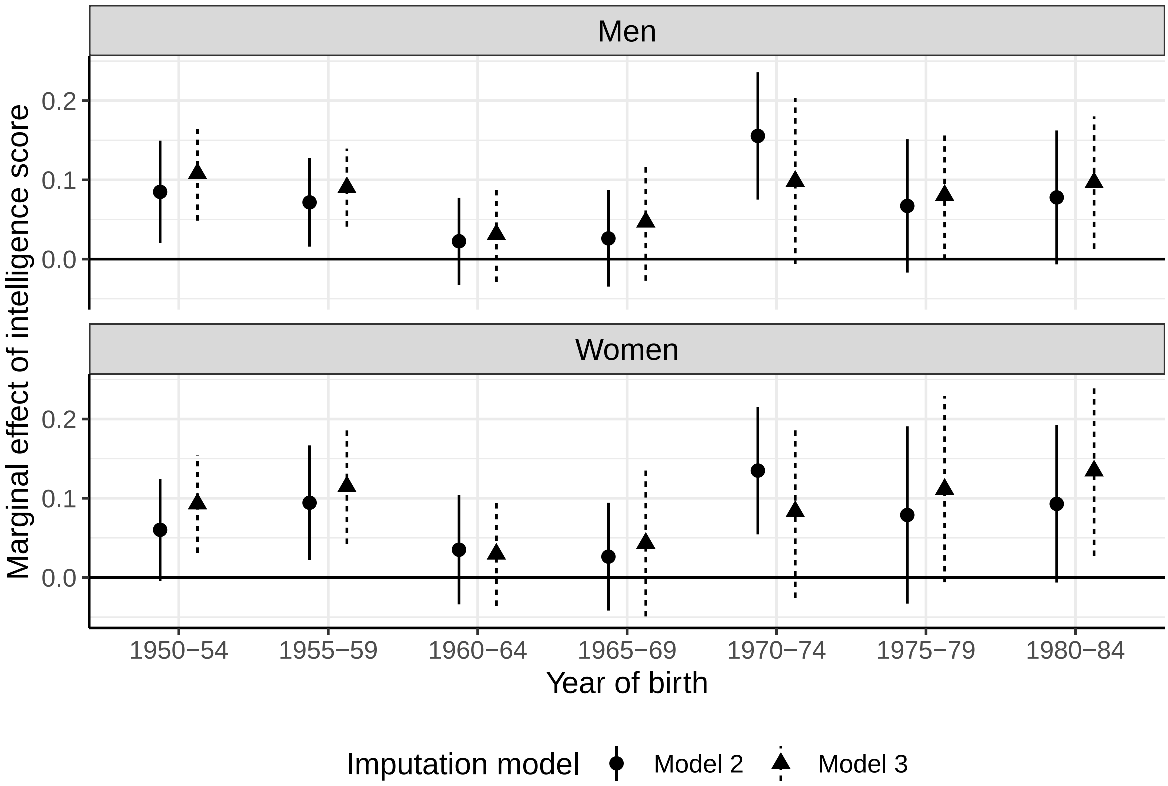
Figure 4.4.3 plots the marginal effect of the intelligence score on the probability of obtaining a university degree estimated from the analysis model in equation (4.5). Similar to the conclusion from the simulation-based evaluation, the imputation models 2 and 3 produce similar estimates of the marginal effect. The results suggest that in early 1970s, that is when people born in 1950-54 were attending HE institutions, a one standard deviation higher intelligence score raised the probability of having a university degree by 6-11pp. However, within 10 years the marginal effect of the intelligence score went down to 2-4pp. These estimates could suggest that the expansion of higher education in the UK led to a lower threshold in terms of intelligence score for admission and completion of university education. But starting from 1990s, the intelligence scores again become a significant predictor of



Notes: The figure plots the average university degree attainment by gender over time. The estimations are done in each completed dataset separately. The estimation sample is restricted to UKHLS wave 3 subjects. The estimations are weighted using cross-sectional survey weights. The whiskers show the 95% confidence interval based on standard errors clustered at the sampling strata level.

Figure 4.4.2: Average university degree attainment by gender

university degree attainment. The point estimates of the marginal effect of the intelligence score are at about 7-16pp for cohorts born from 1970 onwards. We also note that the confidence intervals of the estimates for cohorts born after 1970 are wider, compared to older cohorts. This could be due to the fact that the educational attainment of these cohorts were underestimated in the BHPS, leading to a smaller sample size available during the estimation of the imputation model.



Notes: The figure plots the estimates of the marginal effect of the intelligence score evaluated at the mean in logit model in equation (4.5). The estimations are done in each completed dataset separately. The estimation sample is restricted to UKHLS wave 3 subjects. The estimations are weighted using cross-sectional survey weights. The whiskers show the 95% confidence interval based on standard errors clustered at the sampling strata level.

Figure 4.4.3: Marginal effect of intelligence score at the mean

4.5 Conclusion

The second half of the 20th century has seen a massive expansion of higher education throughout the world, also in the UK. The share of individuals with a higher education degree (HE) has been steadily rising. But not all degrees are made equal. From 1965 to 1992, students in the UK could earn their degrees either from traditional universities or from public sector colleges, led by polytechnics. Despite formal equality of the degrees earned from either type of institutions, these institutions faced different target populations, admission procedures, subjects taught, organization and financing schemes. These differences, together with the elite image of the traditional universities, contributed to a public perception of polytechnics degrees as inferior to that of universities (Willetts 2017; Pratt 1997).

This perceived inferiority hints at something that has been established in the literature: the type of higher education institution can act as a signal of education quality. Therefore, differentiating between types of HE institutions is an important consideration in the analysis of the higher education sector in the UK.

However, common survey datasets often offer limited information about the types of institutions from which individuals earned their degrees. For example, the UK Household Longitudinal Study (UKHLS), the largest panel study in the UK, until wave 11 only asked a small subset of participants for details about the HE institutions they attended. Furthermore, additional restrictions may apply before one is granted access to such information, as is the case for the UKHLS.

In this paper, we try to overcome the issue of missing HE institution types by using a multiple imputation technique. We use the two British panel surveys, BHPS (1991-2008) with about 10,000 individuals in each wave and UKHLS (2009-present) with about 40,000 individuals in each wave. The BHPS specifically asked its participants about the type of higher education institution they last attended, distinguishing between universities and polytechnics. Moreover, 80% of the respondents from the last wave of the BHPS continue as part of the UKHLS. We use the close relationship between the BHPS and the UKHLS and transform the lack of institution types in the UKHLS into a missing data problem. To properly reflect the uncertainty about imputed values we use a multiple imputation technique (Rubin 1977).

We build our imputation models taking into account assumptions about the missing data mechanisms and the agreement between the imputation and analysis models. In this paper, we adopt the analysis model from our companion paper Dimbou et al. (n.d.), which studies how the expansion of higher education sector in the UK changed the composition of students in terms of their intelligence scores. Thus, the agreement between imputation and analysis models requires us to include the intelligence score into the imputation model. However, the intelligence score is only available for a subset of BHPS respondents that continued to the UKHLS. Therefore, we differentiate between three versions of our imputation model differing in the BHPS sample used for the estimation and the inclusion of the intelligence score. For the imputed datasets to produce estimators with properties necessary to draw valid inferences, the imputation model has to be proper. To check whether our imputation models are proper, we use a simulation-based evaluation method.

We find that the imputation models with and without intelligence scores perform similarly across all dimensions. In the simulation-based evaluation, the two models produce combined estimators with similar bias and efficiency statistics for the marginal effect of intelligence score on average university degree attainment. We also show that the combined estimators of the average university degree attainment across cohorts are, in general, similar to the benchmark graduation rates computed using the datasets on the universe of undergraduate students. This similarity could allow us to use a simpler imputation model without the intelligence score in our companion paper.

References

Black, D. A., and J. A. Smith. 2004. "How Robust Is the Evidence on the Effects of College Quality? Evidence from Matching" [in en]. *Journal of Econometrics,* Higher Education (Annals Issue), 121 (1): 99–124.

———. 2006. "Estimating the Returns to College Quality with Multiple Proxies for Quality." *Journal of Labor Economics* 24 (3): 701–728.

Brand, J. P., S. Buuren, K. Groothuis-Oudshoorn, and E. S. Gelsema. 2003. "A Toolkit in SAS for the Evaluation of Multiple Imputation Methods" [in en]. *Statistica Neerlandica* 57 (1): 36–45.

Brewer, D. J., E. R. Eide, and R. G. Ehrenberg. 1999. "Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings." *The Journal of Human Resources* 34 (1): 104–123.

Dillon, E. W., and J. A. Smith. 2017. "Determinants of the Match between Student Ability and College Quality." *Journal of Labor Economics* 35 (1): 45–66.

Dimbou, L. Y., A. Ichino, N. Jandarova, J. L. Reuter, A. Rustichini, P. M. Visscher, and G. Zanella. n.d. *Intelligence and Tertiary Education.* Technical report.

Gillard, D. 1998. *Education in England: A History.*

Higher Education in the Learning Society: Main Report. 1997. Report. London: The National Committee of Inquiry into Higher Education.

Higher Education: Report of the Committee Appointed by the Prime Minister under the Chairmanship of Lord Robbins. 1963. Report. London: Committee on Higher Education.

Pratt, J. 1997. *The Polytechnic Experiment: 1965-1992.* Buckingham, UK; Bristol, PA, USA: Society for Research into Higher Education & Open University Press.

Rubin, D. B. 1977. “Formalizing Subjective Notions about the Effect of Nonrespondents in Sample Surveys.” *Journal of the American Statistical Association* 72 (359): 538–543.

———. 1987. *Multiple Imputation for Nonresponse in Surveys.* New York, Chichester, Brisbane, Toronto, Singapore: John Wiley & Sons.

Schafer, J. L. 1997. *Analysis of Incomplete Multivariate Data* [in en]. CRC Press.

University of Essex, Institute for Social and Economic Research. 2020. *Understanding Society: Waves 1-9, 2009-2018 and Harmonised BHPS: Waves 1-18, 1991-2009,* SN6614.

van Buuren, S. 2018. *Flexible Imputation of Missing Data.* Second. Boca Raton, FL: CRC/Chapman & Hall.

Willetts, D. 2017. *A University Education* [in en]. Oxford University Press.