Dilnoza Adykhamjanovna Muslimova

Complementarities in Human Capital Production

The Importance of Gene-Environment Interactions

Complementarities in Human Capital Production:

The Importance of Gene-Environment Interactions

Complementarities in Human Capital Production:

The Importance of Gene-Environment Interactions

Complementariteiten in de Productie van Menselijk Kapitaal:

Het Belang van Interacties tussen Genen en Omgeving

Thesis

To obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. A.L. Bredenoord

and in accordance with the decision of the Doctorate Board.

The public defence shall be held on
Friday 2 June 2023 at 13.00 hrs

by

Dilnoza Adykhamjanovna Muslimova
born in Shymkent city, Kazakhstan.

Erasmus University Rotterdam

Doctoral Committee:

Promotors:	prof.dr. J.L.W. van Kippersluis prof. S. von Hinke
Other members:	prof.dr. O.R. Marie prof.dr. T. Buser prof.dr. E. Del Bono
Co-promotor:	dr. S.F.W. Meddens

Tinbergen Institute, Erasmus School of Economics dissertation series number: 814
ISBN 978-90-5892-665-4

Design: PanArt, www.panart.nl
Print: OBT bv, www.obt.eu
Cover Artwork: OpenAI & Dilnoza Muslimova

This publication (cover and interior) is printed on FSC® paper Magno Satin MC.

TO FAMILY

Table of Contents

Chapter 1
Introduction

1.1 Nature and nurture

Human capital outcomes (e.g., education, skills, health) are thought to be shaped by a complex interplay between nature (one's genes) and nurture (one's environment). For example, the weight reducing effects of interventions targeting diet or exercising might depend on an individual's genetic propensity for diet or exercise. This dependency of the effect of the environment on the effect of genes and *vice versa* is known in the scientific literature as gene-environment interaction. Gene-environment interaction research investigates how genetic differences between individuals modify their responses to environmental interventions, or *vice versa*, how genetic effects are modified by environmental conditions. In turn, this sheds light on how different environmental changes and policy interventions affect different subgroups of population, more precisely, groups of individuals with higher genetic predisposition as compared to those with lower genetic predisposition to a trait of interest (Biroli, Galama, Von Hinke, et al., 2022).

Historically, there have been numerous attempts to quantify gene-environment interplay. The field has progressed from twin and adoption studies to having opportunities to use molecular genetic data, which nowadays are integrated in a growing number of social science data sets, to name a few, HRS, ALSPAC, and biobanks, for example, UK Biobank, Lifelines, MoBA. Increasing availability of genetic data for social scientists over the last decade offers promising opportunities to deepen our understanding of the fundamental causes of inequalities, especially, in human capital formation (Biroli, Galama, von Hinke, et al., 2022).

In this thesis, I show through three distinct applications how gene–environment interplay (G×E) studies, where genetic endowments (variation therein), measured by polygenic scores (PGS, also known as polygenic indices, PGIs), can contribute to testing of economic theories as well as understanding heterogeneities in treatment effects in the formation of human capital. My applications focus on a particular home environment in childhood (Chapter 2) and a policy environment in early adulthood (Chapter 3). In Chapter 4, I address methodological considerations researchers need

to be aware of when using polygenic scores in gene-environment interplay research and personalized interventions.

Further in the current chapter, I introduce the reader to the relevant areas of scientific literature. Moreover, I introduce key concepts and methods from the genoeconomics literature that are used throughout the subsequent chapters of this thesis by first providing a brief description of the human genome and then elaborating on the Genome-Wide Association Studies which make use of the human genome data. Afterwards, I introduce the polygenic scores and discuss gene-environment interplay framework and important nuances in this field of research. I conclude this chapter by discussing the contributions of the thesis and current status of the chapters.

1.2 Human capital

The origins of economists' interest in human capital stem from the evidence that physical capital and labour could not fully explain income growth observed in the United States in the previous century (Becker, 1964). Since then, human capital was recognized as one of the important drivers of economic development (Mincer, 1984). As compared to other forms of capital, human capital is embedded within people and encompasses traits that make us more productive, e.g., one's educational attainment, skills or health (Becker, 1964; Mincer, 1984). Human capital is a deciding factor in productivity at the micro and macro level (Acemoglu & Autor, 2012; Barro, 2001; Romer, 1990). Moreover, human capital and investments therein are an important input in skill production to match the market demands driven by technological advancements and growth (Goldin & Katz, 2018). A dramatic increase in returns to schooling among the baby boom cohort in the US in the previous century (Lemieux, 2006) is sometimes interpreted as a race between human capital and technology, where investments in human capital do not fully keep up with technological advances (Goldin & Katz, 2018). It is fair to say, therefore, that understanding how human capital is produced is a core building block for economists trying to understand individual well-being as well as societies' economic development.

This thesis investigates the role of nature and nurture in the production of human capital. Given the nature of the available datasets, I define human capital as a combination of skills accumulated by the end of adulthood following Cunha & Heckman (2008) and Cunha et al. (2010), and approximate it by years of education. I use years of education and educational attainment interchangeably throughout the chapters, since I convert the highest degree obtained by the respondents (by the time they reach adulthood) to years of education. This conversion also facilitates consistency with the approach of genome-wide association studies for educational attainment discussed in the next section.

Formally, building on Becker & Tomes (1986, eq. 4), I specify an amended version of human capital production function of the form:

$$H_t = f(x_{t-1}, s_{t-1};\ G, E_{t-1}), \quad (1)$$

where H stands for human capital at time t, x for private investments, s for public investments, G refers to genetic and cultural endowments, and E denotes the broader environment (e.g., prevailing education policies, societal norms, etc.). In this thesis, I will focus on private investments, coming from either parents (Chapter 2) or the individual herself (Chapter 3). To explain intergenerational persistence of earnings and assets, a critical assumption is made: the existence of complementarities between endowments and investments (Becker & Tomes, 1986). That is,

$$\frac{\partial^2 H_t}{\partial x_{t-1} \partial G_t} > 0, \quad (2)$$

This complementarity assumption also serves as a basis of the concept of dynamic complementarity, where skills produced at an earlier stage raise the productivity of future investments (Cunha & Heckman, 2007). Galama & Van Kippersluis (2022) distinguish between two aspects of dynamic complementarity relevant to this thesis: 1) *investments are more productive when the stock of skills is higher* (Ben-Porath, 1967; Heckman, 1976; Rosen, 1976), 2) *individuals optimally choose to invest more when their stock of skills is higher and when the environment is conducive to investments*. The first aspect is interpreting complementarity as a property of the production function and would suggest a positive interaction between the genetic endowments for human capital and environments conducive to investments in human capital. The second feature is informative of the endogenous response to investment-enhancing environments given the educational endowments of individuals. So, the second feature would imply that endowments are associated with investments in human capital when environments are conducive to such investments.

Chapters 2 and 3 of this thesis consider these two aspects of complementarity and two sources of changes in private investments, parental investment changes driven by birth order of children and changes in individual investments driven by a policy change, specifically, legalization of the contraceptive pill. The differences in the nature of these investments and the resulting interactions with genetic endowments are discussed in Section 1.6 of this chapter.

1.3 Human capital and genetics

Since 2013, a series of genome-wide association studies discovered specific genetic variants robustly associated with educational attainment (Lee et al., 2018; Okbay et al., 2016, 2022; Rietveld et al., 2013a), confirming the findings of earlier twin studies that human capital outcomes have some genetic basis, that is they are partially heritable (Branigan, Mccallum, & Freese, 2013). However, we also learned that genes only explain a relatively small proportion of variation in human capital outcomes, up

to 16% in the most recent genome-wide association study of educational attainment (Okbay et al., 2022). Hence, the role of environments and interaction between environments and genetic variation should not be overlooked (Turkheimer, 2000, Rutter, 2006, Heckman, 2007). Recent gene-environment interplay studies on human capital formation use unique opportunities of exogenous variation in environmental exposures, for example, quasi-experiments due to policy changes such as Raising of School Leaving Age in the UK (Barcellos, Carvalho, & Turley, 2018, 2021) or unexpected peak in pollution (von Hinke & Sorensen, 2022). They also use exogenous variation in genetic endowments among siblings in the UK Biobank (Muslimova, van Kippersluis, Rietveld, von Hinke, & Meddens, 2021) or after controlling for parental genes in MoBa (Cheesman et al., 2022). The above studies show how environments modify the effect of genes in important ways: they might cushion against low genetic endowments thereby reducing inequalities (Cheesman et al., 2022), or they might enhance the underlying genetic advantages (Muslimova et al., 2021) and widen the existing gaps in education.[1]

1.4 Concepts and methods in social-science genetics

Human genome

A human genome consists of 23 pairs of chromosomes with the 23rd pair determining the biological sex of a person. One of each pair of chromosomes an individual inherits from their father, and the other one from the mother. A chromosome is composed of two intertwined strands of deoxyribonucleic acid (DNA), each made up of a sequence of four possible nucleotide molecules: adenine (A) and thymine (T), cytosine (C) and guanine (G). Adenine (A) on one strand is always paired with thymine (T) on the other strand. Cytosine (C) is always paired with guanine (G). The human genome consists of approximately 3 billion such base pairs and stretches of base pairs, which in turn form approximately 20,000 genes coding for proteins (Ezkurdia et al., 2014).

Two unrelated individuals have approximately 99.6% of their DNA in common, with most genetic differences across humans (<1%) stemming from single nucleotide polymorphisms (SNP) (Auton et al., 2015). A SNP is a locus (location) on the DNA at which two different nucleotides, alleles for that SNP, can be observed in the population. An individual's genotype is coded as 0, 1, or 2, depending on the count of *effect* alleles (Auton et al., 2015), where the *effect* allele is the allele which the effect estimate refers to (Wootton & Sallis, 2020).

[1] Dias Pereira et al. (2022) provide a comprehensive review of the gene-environment interplay literature in economics and social sciences.

Genome Wide Association Studies

In my research, I use Genome-Wide Association Studies (GWASs) to identify the strength of association between SNPs and a particular trait of interest by regressing an outcome (also known as *trait*) of interest residualised with respect to year and month of birth, genotyping batch number, gender, and first 40 principal components of genetic relatedness matrix (to account for population stratification) on each SNP in a hypothesis-free approach. These associations, as a part of GWAS summary statistics, are then used as weights for the construction of polygenic scores in the independent holdout samples. GWASs can be conducted using different software, I report software specifications in each chapter.

GWASs are quite demanding to the data. Firstly, individual genes typically exhibit small effect sizes, therefore, well-powered samples are required to measure them. Secondly, to achieve such large samples, researchers use meta-analysis of multiple datasets (also known as *cohorts*), however, even cohorts of the same ancestry (e.g. European) may come from different institutional environments and generations and this might affect the GWAS weights. Subtle population differences can take place even within the same countries (see for example Abdellaoui et al. 2013), hence, it is important to control for population stratification, which is typically done using the principal components of genetic relatedness matrix. Tam et al. (2019) provide further overview of current limitations and benefits of GWAS studies.

Polygenic scores and their interpretation

Since individual SNPs usually explain a very small portion (<0.02%) of the variance in behavioural outcomes (Chabris, Lee, Cesarini, Benjamin, & Laibson, 2015; Visscher et al., 2017), it is a common practice to combine multiple SNPs into a polygenic score - PGS (Dudbridge, 2013), constructed as a sum of SNPs weighted by their GWAS effect size as follows:

$$PGS_i = \sum_{j=1}^{J} \beta_j x_{ij}, \tag{3}$$

where PGS_i is the value for the polygenic score for individual i, β_j is the regression coefficient of SNP j ($j = 1, \dots, J$) from a GWAS in an independent training sample, and x_{ij} is the genotype of individual i for SNP j (coded as 0, 1 or 2, indicating the number of "effect" alleles). To facilitate the interpretation, PGSs are usually standardized with mean 0 and standard deviation 1 in the holdout (prediction) sample.

The polygenic score measures the genetic predisposition towards an outcome of interest within the environmental context and demographic characteristics of the discovery GWAS sample (Domingue, Trejo, Armstrong-Carter, & Tucker-Drob, 2020). It is therefore preferable to select discovery and analysis samples from the same environmental context, especially when analyzing gene-environment interactions. At the same time, the discovery sample should be independent of the analysis sample to

avoid overfitting, meaning no related or overlapping individuals between the samples (Dudbridge, 2013).

Polygenic scores have become appealing to use in social science research given their several advantages over other genetic measures. First, the underlying SNPs are fixed at conception, they do not change over time and are random within families (Kong et al., 2018). Second, polygenic scores can capture the influence of a whole spectrum of genes, which is especially important for complex traits. Unlike mendelian or monogenic traits, complex traits are individual outcomes which are usually influenced by numerous genes. Educational attainment and subjective well-being are examples of complex traits. Thirdly, polygenic scores can be more flexibly used within a gene-environment interplay framework as compared to twin or candidate gene studies. While twin studies do not directly measure environment and only infer about it after accounting for common environment and genetic variation from the zygosity of twins, candidate gene studies are known for being underpowered and they generally do not replicate (Duncan & Keller, 2011; Duncan, Pollastri, & Smoller, 2014).

Despite the above advantages, there are important nuances to using polygenic scores. First, since GWASs are based on finite samples, the resulting polygenic scores are prone to measurement error (Benjamin, Cesarini, Laibson, & Turley, 2020). Fortunately, depending on the nature of the study, measurement error can be addressed using at least one of the two approaches (Van Kippersluis et al., 2022): (1) correction using external SNP-based heritability of the trait developed in the PGS repository (Becker et al., 2021); (2) the Obviously Related Instrumental Variables approach (Van Kippersluis et al., 2022), where two PGSs for the same trait based on two independent GWAS samples are used as instruments for each other. I show applications of the latter method in Chapters 2 and 3. Second, given the differences in genetics of individuals with different ancestries, the training sample for GWAS and a holdout sample need to be from the same population of individuals with homogenous ancestry. Lack of genetic correlation between the GWAS discovery sample and holdout sample might impair the predictive power of the polygenic scores (Muslimova et al., 2022). Because of this, most of the current studies are based on European ancestry individuals, and our knowledge on polygenic scores is therefore heavily biased towards European-ancestry individuals.[2] Third, the causal effect of a PGS on an outcome is usually confounded by genetic nurture, i.e., parental genes being responsible both for the genotype of their children and the environment in which they are raised (Kong et al., 2018). So, even though availability of parental genotypes would remedy this issue and estimate the direct genetic (causal) effect of a PGS on an outcome of interest, samples that include genotypes of children and both parents are extremely rare. Finally, it should be emphasized that even when my co-authors and I could estimate causal genetic effects, these effects could be fully mediated by the environment and should not be seen as purely biological or

[2] However, more and more biobanks are emerging, which would allow well-powered studies for other ancestry groups, e.g. BioBank Japan (N=200,000 individuals) and China Kadoorie Biobank (N=500,000 individuals) as discussed in Martin et al. (2019).

immutable. In fact, in most cases the biological and environmental pathways through which SNPs in a given PGS operate are not clear, and it could be that a genetic association arises from purely environmental channels. A famous example in this regard is that in a society where red-haired individuals are prohibited from going to college, a causal genetic effect of a red-hair gene would be fully mediated by the environment.

Gene-Environment Interactions

A simplified empirical framework for studying G×E interactions can be specified as follows:

$$Y_i = \alpha_1 + \alpha_2 G_i + \alpha_3 E_i + \alpha_4 (G_i \times E_i) + \boldsymbol{X_i'\beta} + \xi_i, \quad (4)$$

where Y_i is the outcome of interest, E_i is the environment of interest, G_i is the measure of genes, and $\boldsymbol{X_{ij}}$ is a set of individual level controls and the vector of the first 40 principal components (PCs) of the genetic relatedness matrix. When genes (G_i) are measured by polygenic scores, for α_2 to reflect the causal genetic effect on the outcome Y_i, one needs to use a within-family design by controlling for family fixed effects or include parents' genes in the specification. For the causal effect of the environment on the outcome to be identified (α_3), there needs to be exogenous variation independent of genes in the measure of environment. For the interaction effect (α_4) to measure causal gene-environment interplay one needs independent variation in G_i and E_i and the absence of so-called gene-environment correlation, or *rGE* (Biroli, Galama, Von Hinke, et al., 2022). rGE occurs when genes are correlated with environmental conditions. For example, Biroli & Zünd (2020) find that individuals with higher genetic predisposition for drinking tend to self-select into areas with easier access to alcohol selling establishments. Presence of rGE biases the gene-environment interaction estimates (Biroli, Galama, von Hinke, et al., 2022) and may lead to collider bias when chosen environments are heritable (Akimova, Breen, Brazel, & Mills, 2021). It might also lead to spurious and misinterpreted GxE interactions. This issue can be addressed by using exogenous environments where their variation is independent of the genetic variation (Biroli, Galama, von Hinke, et al., 2022).

In practice, it is very hard to come across the data which has a sufficient number of siblings or children with both parents genotyped and at the same time exogenous independent variation in the environmental exposure of interest. This is one of the areas where the contributions of this thesis lie.

1.5 Contributions of this thesis

The studies presented in this thesis contribute to two strands of genoeconomics literature. The first strand concerns gene-environment interplay using within-family and regional variation in environmental exposures, which provides new empirical insights into existing economic theories all the while these theories have been informative about expectations of the sign of the gene-environment interaction. The second strand of research this thesis contributes to is the methods in social-science genetics, more specifically, studies addressing uncertainty in polygenic scores (Clifton, Collister, Liu, Littlejohn, & Hunter, 2022; Lambert et al., 2021; Pain et al., 2021; Schultz et al., 2021; Turley et al., 2021; Wand et al., 2021; Ware et al., 2017).

In Chapter 2, my-co-authors and I bring new evidence on complementarities between early life investments and endowments using genetic data. Earlier literature shows that within families, on average, firstborns complete more education than their laterborn siblings. Chapter 2 studies whether this birth order effect is amplified by individuals' genetic endowments. Our family-fixed effects approach exploits exogenous variation in birth order and genetic endowments among 14,850 siblings in the UK Biobank. We find that those with higher genetic endowments benefit disproportionally more from being firstborn compared to those with lower genetic endowments, providing a clean example of how nature and nurture interact in producing human capital. Since parental investments are a dominant channel driving birth order effects, our results are consistent with complementarity between endowments and family investments in human capital formation (Equation 2).

I further investigate whether this complementarity extends to the prevailing policy environment in young adulthood in Chapter 3. In particular, using access to contraception as an environment fostering human capital investments among women, I study whether the introduction of the contraceptive pill interacts with women's genetic endowments to enhance their educational attainment. Since the late 1960s, access to contraception has been providing women with broader opportunities to invest in their education and career. This paper is the first to investigate whether access to oral contraception has differential effects on educational attainment according to one's genetic endowment for education. Following earlier work on the topic, my empirical strategy leverages region-by-cohort variation in pill diffusion. I use the UK Biobank, restricting my attention to 145,502 women, and show that exposure to the pill is associated with more years of education. The positive association of the pill diffusion with years of education is concentrated among women with lower genetic endowments for education. This finding suggests the existence of a compensating mechanism: an environment in which contraception is more widely available was most productive for women with a lower genetic predisposition towards education, reducing inequalities in educational attainment. This study highlights the various ways in which women across the PGS distribution gained from the diffusion of the pill in terms of years of education.

While working on Chapters 2, my co-authors and I came across multiple methodological puzzles. Perhaps, the most notable observation is that ranking of individuals was not stable across polygenic scores constructed for the same trait using different GWAS discovery samples and construction methods. So, we asked what are the implications of this rank discordance for our research? Hence, Chapter 4 investigates the differences in rank concordance of polygenic scores. Recently, polygenic scores are increasingly used to identify individuals at risk of developing disease and are advocated as screening tools for personalised medicine and education. This study empirically assesses rank concordance between PGSs created with different construction methods and discovery samples, focusing on cardiovascular disease (CVD) and educational attainment (EA). We find Spearman rank correlations between 0.17-0.93 for CVD, and 0.40-0.83 for EA, indicating highly unstable rankings across different PGSs for the same trait. Potential consequences for personalised medicine and gene-environment (G×E) interplay are illustrated using data from the UK Biobank. The simulations presented in the study show how rank discordance mainly derives from a limited discovery sample size and reveal a tight link between the explained variance of a PGS and its ranking precision. We conclude that PGS-based ranking is highly dependent on the choice of PGS, such that current PGSs do not have the desired precision to be used routinely for personalised intervention. In terms of implications for gene-environment interplay research, using PGSs in a continuous form provides a more robust measure of genetic endowments as compared to the stratified forms (e.g., quartiles, quintiles, of the PGS etc.). Researchers are also recommended to report the construction of PGSs transparently and use a PGS with the highest predictive power for a trait of interest.

1.6 Individual contributions and status of the chapters

The introductory Chapter 1 was written independently, with the drafts benefitting greatly from the comments of my supervisors. Chapters 2 and 4 are based on collaborative work with colleagues at Erasmus University Rotterdam and University of Bristol, who all contributed in various ways. Chapter 3 is my solo-authored chapter. In this section, I will highlight my own contribution as well as the contribution of my colleagues to the work presented in this thesis.

Chapter 2 was the first project on gene-environment interplay I worked on. The original research idea came from C.A. Rietveld and H. van Kippersluis. Initially, we were interested in how the family size and birth order moderate the effect of polygenic scores on human capital. Diving into the family size literature, we learned that birth order offsets the family size effects on education within families. It was possible to identify siblings and other relatives from the kinship matrix in the UK Biobank, the primary dataset we were using for the project. Therefore, we narrowed the topic down to the interaction between birth order and polygenic scores in shaping

human capital within and between families. F. Meddens and I cleaned the genetic data, conducted the GWASs and constructed the polygenic scores. I was responsible for the main empirical analysis and wrote the first draft. H. van Kippersluis developed the theoretical approach, set up ORIV and permutation analysis. All authors were involved in further developing and revising the manuscript. Chapter 2 benefited greatly from numerous conference and seminar presentations.

Chapter 3 is my individual work, inspired by discussions during mentoring meetings with F. Meddens and studies conducted by O. Marie and E. Zwiers, and E. De Cao and colleagues. Early presentations and drafts of this chapter were improved with the comments of F. Meddens and C.A. Rietveld. The most recent version of the draft for this chapter benefited greatly from the valuable feedback of my supervisors, H. van Kippersluis and S. von Hinke. This chapter has also benefited greatly from numerous seminars and conferences.

Chapter 4 was sparked by my observations of the differences in the ranking of individuals when using different polygenic scores for our projects on gene-environment interactions. I presented these differences at the meeting of NORFACE project in Bristol in 2019. Our team grew curious about what the implications of these differences are for our research and for research on personalized approaches. Moreover, given our methodological interest, we were also curious about the driving causes of these differences across PGSs. Is it a difference in discovery sample? Does it matter which statistical method you use to construct a PGS? Or are the PGSs simply all noisy, and therefore lack of concordance arises naturally because of measurement error in the PGSs? Subsequently, we conceptualized the project. F. Meddens and I conducted the GWASs in UKB and the meta-analyses with other GWAS summary statistics, constructed the PGSs, and prepared the illustrative applications. S. von Hinke helped develop the illustrations and set up the code for empirical analysis. R. Dias-Pereira performed the G×E analyses. H. van Kippersluis and C.A. Rietveld conducted the simulations to analyse the causes of rank discordance. All authors contributed to preparing and critically reviewing the manuscript and the supplementary information file.

During my PhD I also contributed to the following research project not included in this thesis:

Van Kippersluis, H., Biroli, P., Galama, T. J., von Hinke, S., Meddens, S. F. W., Muslimova, D., ... & Rietveld, C. A. (2022). Overcoming Attenuation Bias in Regressions using Polygenic Indices: A Comparison of Approaches. *bioRxiv*. (currently resubmitted to *Nature Communications*).

Table 1.1 - Publication status of the chapters

Chapter	Title	Reference	Presentations	Publication status
2	Complementarities in human capital production: Evidence from genetic endowments and birth order	Muslimova, van Kippersluis, Rietveld, von Hinke, & Meddens	Frontiers of Using Genetic Data in Economics (Chicago, 2023), ESSGN (Bologna, 2023), EEA (2022), Dial Final Conference (2021), NORFACE Meeting (Online, 2020), IGSS (2020), NORFACE Meeting (2019), DIAL Mid-term Conference (2019), RSF Summer Institute in Social Science Genomics (2019)	Under review
3	Diffusion of the pill and women's education: The role of gene-environment interactions	Muslimova	ESE Female Network with Martha Bailey (2022), ESSGN (2022), Health Economics Internal Seminar (2022), IGSS (2021), Tinbergen Institute PhD Jamboree (2021) ESSGN (2023)	Manuscript in preparation
4	Rank concordance of polygenic indices	Muslimova, Dias Pereira, von Hinke, van Kippersluis, Rietveld, & Meddens		Nature Human Behaviour (2023)

1.7 Glossary

A large number of genetic terms are used throughout this thesis. The glossary below is intended to clarify these for a less familiar reader based on Biroli et al. (2022), Mills, Barban, & Tropf (2020), and Wootton & Sallis (2020).

Allele – Alternative forms of a gene found in the same place on a chromosome.

Candidate gene –A predefined loci of interest hypothesized to be associated with a particular phenotype based on the loci's biological function.

Chromosome – A DNA molecule that is a part of the genome. Every cell in the human body contains 23 pairs of chromosomes (22 autosomal and 1 sex chromosome).

DNA – Deoxyribonucleic acid, a molecule that encodes the genetic instructions used in the development and functioning of all known living organisms.

Effect allele – The allele to which the effect estimate refers, regardless of whether this estimate is increasing or decreasing and regardless of whether this allele is coding or non-coding.

Evocative (reactive) gene-environment correlation – Occurs when an individual's (partially) genetically driven characteristics evoke different environmental reactions.

Gene - Basic unit of heredity. Sequence of DNA bases that provides instructions for building a particular protein or proteins.

Gene-environment correlation (rGE) – Occurs when an individual's genotype influences or is associated with the exposure to the environment.

Gene-environment interaction (G×E) – An interplay between genetic predispositions and environmental factors in which the effect of the genes on an outcome is modified by the environment and vice versa.

Genetic nurture – Parental genes influencing off-spring outcomes through environmental pathways.

Genetic variation – Differences in DNA among individuals.

Genotype – An individual's combination of alleles at a particular locus.

Genome-wide association study (GWAS) – A study in which millions of polymorphisms from the whole genome are individually tested for association with a phenotype.

GWAS meta-analysis – Meta-analysis of genome-wide association study (GWAS) results from different samples.

Genome-wide significance – The significance level at which an association is considered statistically significant in a genome-wide association study (GWAS) (5×10^{-8}).

Haplotype – Set of single-nucleotide polymorphisms (SNPs) on a single chromatid, one half of a duplicated chromosome, of a chromosome pair that are associated statistically.

Heritability – The proportion of observed differences in a trait among individuals of a population that is due to genetic differences among these individuals.

Linkage disequilibrium (LD) – The correlation between adjacent nucleotides in the DNA resulting from the co-inheritance of alleles.

Locus – A locus is a specific position of a DNA sequence on a (pair of) chromosome(s).

Major allele – Nucleotide of a SNP that is more common in the population.

Minor allele – Nucleotide of a SNP that is less common in the population.

Nucleotide – The basic component molecules of DNA. Human DNA is composed of a sequence of about 3 billion pairs of nucleotide molecules.

Phenotype – An observable characteristics or traits of an individual (e.g., morphology, behaviour).

Polygenic indices/scores – The best linear genetic predictor of a phenotype, constructed as a linear combination of single-nucleotide polymorphisms (SNPs) weighted by their association with the phenotype estimated in a genome-wide association study (GWAS).

Polygenic trait - A trait influenced by many genetic variants, with each having a small effect. *Polymorphism* - Locations in the DNA where the nucleotides differ between individuals. *Population stratification* - The presence of a systematic difference in allele frequencies between subpopulations within a population.

SNP – Single nucleotide polymorphism. A DNA sequence variation occurring commonly within a population (e.g. 1%) in which a single nucleotide – A, T, C, G – in the genome differs between members of a biological species.

Within-family analysis – An analysis that examines the effect of genetic predispositions among family members. This analysis is used to establish causal relationships between genotypes and outcomes of individuals free of confounding by population stratification or genetic nurture.

Chapter 2
Complementarities in human capital production: Evidence from genetic endowments and birth order[3]

Joint work with Hans van Kippersluis, Cornelius A. Rietveld, Stephanie von Hinke, S. Fleur W. Meddens

Abstract

On average, firstborns complete more education than laterborns. We study whether this effect is amplified by individuals' endowments measured using genetic information. Our family-fixed effects approach allows exploiting exogenous variation in birth order and genetic endowments among 14,850 siblings in the UK Biobank. We find that those with higher genetic endowments benefit disproportionally more from being firstborn compared to those with lower genetic endowments, providing a clean example of how nature and nurture interact in producing human capital. Since parental investments are a dominant channel driving birth order effects, our results are consistent with complementarity between endowments and investments in human capital formation.

[3] This chapter is based on Muslimova, D., van Kippersluis, H., Rietveld, C. A., von Hinke, S., & Meddens, S. F. W.. (2021) Complementarities in human capital production: Evidence from genetic endowments and birth order. *arXiv preprint arXiv:2012.05021.*

2.1 Introduction

It is increasingly accepted that important life outcomes such as educational attainment are influenced by a complex interplay between nature (i.e., genetic variation) and nurture (i.e., environments in which one grows up; Rutter, Moffitt, and Caspi, 2006; Heckman, 2007). Empirically estimating the separate contributions of genetic endowments and environments, and their possible interaction, is however complicated by the endogenous nature of both. Indeed, environmental characteristics are partially heritable and typically cluster together: e.g., higher educated parents tend to have higher incomes. Similarly, although several studies have found specific genetic variants to be associated with human capital outcomes such as educational attainment (Lee et al., 2018; Okbay et al., 2016, 2022; Rietveld et al., 2013b), genetic variation is only random conditional on parental genotypes (e.g., Lawlor, Harbord, Sterne, Timpson, & Davey Smith, 2008; Davey Smith & Ebrahim, 2003). Not controlling for the latter implies that the 'genetic-effect' may in fact reflect 'genetic nurture' – that is, the parental genotype can shape the environment in which children grow up, thereby producing a spurious association between the child's genetic variants and their outcomes (e.g., Belsky et al., 2018; Kong et al., 2018).

This study exploits exogenous variation in both genetic endowments for education and the family environment to analyse the importance of gene-environment interactions for educational attainment. Genetic endowments are measured using a so-called "polygenic score" (PGS), also referred to as a "polygenic index" (Becker, Burik, et al., 2021). A PGS is a highly predictive index that is constructed as the sum of all measured genetic variants, weighted by the strength of their relationship with educational attainment as estimated in an independent sample (Dudbridge, 2013; Lee et al., 2018). The PGS can (currently) explain up to ~15% of the variance in educational attainment (Okbay et al., 2022). Our measure of the environment is an individual's birth order, which is consistently negatively correlated with educational attainment in developed countries (see e.g. Bagger et al., 2021; Behrman et al., 1986; Black, Devereux, & Salvanes, 2005; Booth & Kee, 2009; De Haan, 2010; De Haan, Plug, & Rosero, 2014; Kantarevic et al., 2006).

We overcome endogeneity issues by exploiting within-family variation in both birth order and genetic endowments. Indeed, siblings' birth order is randomly determined within families (e.g., Damian and Roberts, 2015), and genetic variants are randomly assigned across siblings ("Mendel's Law of Independent Assortment"). Hence, within-family, genetic variants are unrelated to birth order by construction.[4] With exogenous variation in both genes and environments, this provides a compelling context in which we can fundamentally improve our understanding of the nature-nurture interplay in shaping educational attainment. That is, does the environment only complement or also moderate genetic advantages? Answering this question

[4] A systematic relationship between birth order and genetic endowments could arise when parents base their fertility decisions on the observed genetic endowments of their offspring, i.e., a stopping rule depending on the "quality" of children (e.g., Eirnæs & Pörtner, 2004). We find no such evidence in our sample (see Section 2.4).

constitutes our main contribution.

We choose birth order as a measure of the environment not merely for being conveniently uncorrelated with genetic endowments, but particularly because birth order has been shown to proxy for parental investments. Indeed, the theoretical literature on the 'quantity-quality trade-off' (Becker, 1960; Becker & Lewis, 1973; Becker & Tomes, 1976; Galor & Weil, 2000) posits that with each additional child, it is more expensive to maintain the same 'quality' children (i.e., with the same level of education or health), implying that parents invest less in laterborn children. Moreover, with parental preferences for fairness in investments over equality in outcomes (see e.g., Berry, Dizon-Ross, & Jagnani, 2020), parents distribute their resources equally over their children, leading to natural dilutions in time investments for laterborn children compared to their firstborn siblings (Blake, 1989; Downey, 2001).[5]

Indeed, firstborns have undivided attention until the arrival of the second child (Breining, Doyle, Figlio, Karbownik, & Roth, 2020), and the empirical literature suggests that a dominant channel through which birth order affects educational attainment is parental time investments (see e.g., Birdsall, 1991; Black, Grönqvist, & Öckert, 2018; Breining et al., 2020; De Haan, 2010; Del Bono et al., 2016; Monfardini & See, 2012; Pavan, 2016; Price, 2008). Using the American Time Use Survey (ATUS), Price (2008) shows that firstborns receive 20-30 minutes more daily quality time compared to their younger siblings (see also Black et al., 2018; Monfardini & See, 2012), with the gap being largest at early ages. For laterborn children, mothers postpone prenatal care, breastfeed less, are more likely to smoke when not breastfeeding (Lehmann, Nuevo-Chiquero, & Vidal-Fernandez, 2018), and fathers take shorter periods of parental leave (Sundström & Duvander, 2002). Hotz and Pantano (2015) additionally show that parents have less stringent parenting strategies for laterborn children. Hence, whilst we do not dismiss other potential channels through which birth order may affect educational attainment[6], parental investments are a prominent channel through which these effects arise, consistent with the evidence that parental investments are an important input into the child's human capital production (e.g., Del Boca, Flinn, & Wiswall, 2014).

As such, economic theories of human capital production suggest a specific direction for the nature-nurture interaction in our study. Becker & Tomes (1986, eq. 4) specify

[5] We do not use actual parental investments because we do not observe these directly. In addition, realized parental investments are endogenous to the child's endowments (Almond & Mazumder, 2013; G. S. Becker & Tomes, 1986; Jere R. Behrman, Pollak, & Taubman, 1982; Breinholt & Conley, 2019; Rosenzweig & Wolpin, 1995; Sanz-De-Galdeano & Terskaya, 2019).

[6] Eirnæs & Pörtner (2004) distinguish three additional environmental channels through which birth order effects may arise: (i) younger children may benefit from the interaction with their older siblings; (ii) firstborns benefit from a lower maternal age and better maternal immune system (e.g., Behrman, 1988; Black et al., 2016); and (iii) in some societies the oldest son (or older children more generally) are favored as they are the first to become economically independent. Other channels could be that later born children are exposed to more family disruptions (Björklund, Ginther, & Sundström, 2021), or experiencing illness at younger ages due to older siblings bringing home viruses (Daysal, Ding, Rossin-Slater, & Schwandt, 2021).

a human capital production function of the form:

$$H_t = f(x_{t-1}, s_{t-1}, E_t), \quad (1)$$

where H denotes human capital, x represents parental investments, s denotes public investments, and E refers to genetic and cultural endowments (with index t for time). A critical assumption made to explain intergenerational persistence of earnings and assets is that there are complementarities between endowments and investments (Becker & Tomes, 1986; eq. 5):

$$\frac{\partial^2 H_t}{\partial x_{t-1} \partial E_t} > 0. \quad (2)$$

In other words, children with higher (genetic) endowments E benefit more from parental investments x_{t-1}.

The complementarity assumption is not specific to Becker & Tomes (1986). In fact, it is also embedded in so-called "Ben-Porath neutrality" where the stock of human capital raises the productivity of investments in human capital (Ben-Porath, 1967; Heckman, 1976; Rosen, 1976). Moreover, it is the key building block for the concept of dynamic complementarity in skill production (Cunha & Heckman, 2007), where skills produced at a given age raise the productivity of investments at later ages. Hence, complementarity between the child's endowments and parental investments is a central assumption underlying seminal economic theories of human capital production and it provides a clear prediction for the sign of the Gene-by-Environment (G×E) interaction term. Our analysis context is therefore a rare example in which economic theory helps formulate hypotheses about fundamental interactions between genetic variation and the environment. In a sample of siblings, being firstborn should show a positive interaction with genetic endowments.

Our empirical analysis exploits data from 14,850 full siblings from the UK Biobank, a population-based sample from the United Kingdom (Fry et al., 2017). To measure participants' genetic endowments, we construct polygenic scores for educational attainment based on the results from our own tailor-made genome-wide association study (GWAS) that uses the UK Biobank sample but excludes all siblings and their relatives.[7] We adopt a family fixed effects approach to exploit within-family variation in genetic endowments and birth order to study the G, E and G×E effects on educational attainment, and we apply Obviously-Related Instrumental Variables estimation (ORIV; Gillen, Snowberg, & Yariv, 2019) to reduce random measurement error in the polygenic score. We mainly focus on firstborns versus laterborns, since the literature suggests that birth order effects are particularly salient at this margin (e.g., Breining et al., 2020).

We confirm earlier findings that firstborns have a higher level of education than

[7] See Appendix A for definitions and explanations of the genetic terms used here.

laterborns, and that one's genetic endowments are a strong predictor of educational attainment. We also confirm that genetic endowments do not differ systematically across birth order within a family. This finding corroborates that birth order effects are due to environmental influences (see also Isungset, Freese, Andreassen, & Lyngstad, 2022). Our main finding is that birth order and genetic endowments interact: being firstborn and having higher genetic endowments for education exhibit a positive interaction, meaning that those with a higher polygenic score benefit disproportionally more from being firstborn compared to those with a lower polygenic score. This finding is a clean example of how genetic endowments and the environment interact in producing important life outcomes such as educational attainment. Moreover, our empirical results are consistent with the existence of complementarity between endowments and investments in human capital production.

These results are quantitatively meaningful. Handy & Shester (2022) estimate that the rise in the fraction of laterborn children was responsible for 20-35% of the stagnation in college completion among US baby-boom cohorts born between 1946 and 1974. We find that one's genetic endowments amplify the effect of birth order: a firstborn sibling with a polygenic score two standard deviations above the mean on average enjoys a full year of schooling extra compared to a laterborn sibling with an average polygenic score. Hence, the random inheritance of genetic endowments across siblings alongside the arbitrary order of birth have important monetary and non-monetary consequences across siblings (e.g., Heckman, Humphries, & Veramendi, 2018) as well as their offspring (e.g., Havari & Savegnago, 2022; Mogstad & Torsvik, 2022).

Our paper speaks to three main literatures. First, we contribute to an emerging literature on gene-environment interactions (G×E), which addresses how the environment moderates the effect of genetic variants, and vice versa. Previous studies have typically examined interactions between polygenic scores and endogenous environments such as socio-economic status (see e.g., Barth, Papageorge, & Thom, 2020; Bierut, Biroli, Galama, & Thom, 2018; Ronda et al., 2020), childhood trauma (e.g., Mullins et al., 2016; Peyrot et al., 2014), or a partner's death (e.g., Domingue, Liu, Okbay, & Belsky, 2017). Interpretation of these findings is not straightforward, because individuals with certain genetic predispositions may self-select into different environments (known as gene-environment correlation, or *rGE*; see Jencks, 1980; Schmitz & Conley, 2017a). In these analyses, therefore, the 'environmental effect' could be reflecting the effect of one's genotype through *rGE*, and the 'genetic effect' could be reflecting the rearing environment shaped by parental genotype (i.e., genetic nurture). A handful of studies use exogenous variation in environments to study G×E in educational attainment. For example, Conley & Rauscher (2013) analyse how random differences in the prenatal environment alter the genetic effects on education; Schmitz & Conley (2017b) use the Vietnam War conscription as a natural experiment; Barcellos, Carvalho, & Turley (2021) use a UK compulsory schooling reform; von Hinke & Sorensen (2022) explore an unanticipated peak in pollution;

Muslimova (2022) uses legal access to contraceptive pill in the UK, and Van den Berg, Von Hinke, & Wang (2022) use the unexpected temporary post-war derationing of sugar confectionery to study G×E effects on individuals' educational attainment.[8] We push this literature one step further by not only considering exogenous variation in the environment, but also in genetic endowments by exploiting within-family variation in polygenic scores.

A second strand of literature that we speak to is the literature on birth order effects. This literature consistently finds that in developed countries, laterborn children have lower educational attainment. Birth order effects have also been found for other outcomes, though sometimes with mixed results, such as early life cognitive skills and intelligence (Black, Devereux, & Salvanesz, 2011; Fenson et al., 1994; Keller, Groesch and Trob, 2015), health (Black, Devereux, & Salvanes, 2016; Pruckner et al., 2019), personality and leadership skills (Black et al., 2018), and delinquency (Breining et al., 2020). We contribute to this literature by studying heterogeneity in the birth order effect on educational attainment with respect to genetic endowments. The potential interaction between birth order and genetic endowments is not merely an important source of heterogeneity in the treatment effect, but one that – if present – carries over to the next generation, potentially exacerbating intergenerational inequalities (Barclay, Lyngstad, & Conley, 2021; Havari & Savegnago, 2022).

Finally, our study offers new empirical insights on the human capital production process (e.g., Becker and Tomes, 1986; Cunha & Heckman, 2007; Cunha et al., 2010; Todd & Wolpin, 2003). Empirically testing complementarity in human capital formation requires independent variation in initial endowments and subsequent investments (Almond & Mazumder, 2013; Johnson & Jackson, 2019), and is therefore extremely challenging (Almond, Currie, & Duque, 2018). Indeed, the previous literature has almost exclusively relied on early-life outcomes such as birthweight as a measure of endowments (e.g., Datar, Kilburn, & Loughran, 2010; Figlio, Guryan, Karbownik, & Roth, 2014). However, such early life outcomes are affected by prenatal investments (Aizer & Cunha, 2012; Rosenzweig & Schultz, 2015), meaning they partially capture parental choices and are therefore endogenous. Furthermore, parents respond to children's endowments (e.g., Adhvaryu & Nyshadham, 2016; Aizer & Cunha, 2012; Almond & Mazumder, 2013; Becker & Tomes, 1986; Bharadwaj, Eberhard, & Neilson, 2018; Datar et al., 2010; Frijters, Johnston, Shah, & Shields, 2013; Giannola, 2020; Hsin & Felfe, 2014), with recent studies suggesting that parental investments also respond to the genetic endowments of children (Breinholt & Conley, 2019; Fletcher, Wu, Zhao, & Lu, 2020; Houmark, Ronda, & Rosholm, 2020; Sanz-De-Galdeano & Terskaya, 2023). Hence, measures of children's endowments are rarely clean of parental investments, and parental investments are rarely independent of endowments, posing a formidable empirical challenge to accurately identify

[8] Studies with other outcomes that exploit exogenous environments include, e.g., Barban, Cao, & Francesconi (2021), Barcellos, Carvalho, & Turley (2018), Dias Pereira, Rietveld, & van Kippersluis (2022), Schmitz & Conley, (2016). For a recent review of the G×E literature in economics and social science, see Dias Pereira et al. (2022).

complementarities in human capital production.[9]

We contribute to this literature by exploiting random within-family variation in genetic endowments that is fixed at conception and is therefore clean from parental investments. Furthermore, we proxy for parental investments using individuals' birth order, which is strongly associated with parental investments, but uncorrelated with genetic endowments. Essentially, employing birth order as a proxy for parental investments exploits the natural reduction in the time and money available with the arrival of a laterborn child, and is thus independent of the child's endowments. There could be other channels through which birth order may impact educational attainment, such as through interaction with younger siblings (Eirnæs & Pörtner, 2004). However, unless these other channels have completely opposite interaction effects with genetic endowments, we provide a promising setting to empirically test a necessary condition for complementarity in human capital production.

More generally, this paper shows how economic theory can inform empirical G×E analyses, and how genetic data can be leveraged to assess economic theories (Biroli, Galama, von Hinke, et al., 2022). We show that the analysis of G×E within-family, exploiting exogenous G as well as E, provides a way to test for complementarities more generally, which is not restricted to birth order effects, but extends to other (exogenous) parental investments and policy changes (e.g. Biroli, Galama, von Hinke, et al., 2022), minimum school leaving ages (Barcellos et al., 2021), public health investments (Muslimova, 2022; Van den Berg, Von Hinke, & Vitt, 2022; etc.). Our findings further provide one of the first pieces of causal evidence of how genetic variation (here measured by the polygenic score for years of education) and the environment (here measured by birth order) jointly shape and interact in producing important life outcomes such as educational attainment. While this finding was long anticipated (e.g., Heckman, 2007; Rutter et al., 2006), finding credible and independent sources of variation in genes and environments is rare given how tightly genetic and environmental influences are entangled (e.g., Koellinger & Harden, 2018). Showing evidence of an interaction between genetic variation and environments is therefore a step forward in our fundamental understanding of how nature and nurture jointly shape human capital, while also providing an antidote against arguments of genetic or environmental determinism.

2.2 Empirical Strategy

We analyse gene-environment (G×E) interactions between genetic endowments for educational attainment and birth order. The empirical specification is rooted in the

[9] A recent set of studies has examined rare cases where exogenous variation exists in both initial endowments as well as later-life investments, with mixed evidence. Some studies find evidence consistent with complementarity (Adhvaryu et al., 2019; Duque et al., 2018; Gunnsteinsson et al., 2014; Johnson & Jackson, 2019), whereas others find weaker evidence or even substitutability between endowments and investments (Lubotsky & Kaestner, 2016; Malamud et al., 2016; Rossin-Slater & Wüst, 2020). See Appendix B for a detailed overview.

human capital production function (1), where we take adult educational attainment as a proxy for human capital by the end of childhood, as in Cunha & Heckman (2008) and Cunha et al. (2010).

Following Todd & Wolpin (2003) and Cunha & Heckman (2008), we specify a linear production function, where years of completed education for individual *i* of family *j* (denoted by $\boldsymbol{Y_{ij}}$), is a function of initial endowments, the environment shaped by parents, and unobserved parental characteristics. Empirically, we measure initial endowments by the polygenic score for educational attainment (denoted by $\boldsymbol{G_{ij}}$)[10], the environment is measured by an indicator for being firstborn ($\boldsymbol{E_{ij}}$), and parental characteristics are subsumed into the family fixed effect ($\boldsymbol{\delta_j}$). We then allow for an interaction term between $\boldsymbol{G_{ij}}$ and $\boldsymbol{E_{ij}}$, leading to the following specification:

$$Y_{ij} = \alpha_1 + \alpha_2 G_{ij} + \alpha_3 E_{ij} + \alpha_4 \left(G_{ij} \times E_{ij}\right) + \alpha_5 X_{ij} + \delta_j + \xi_{ij}, \qquad (3)$$

where $\boldsymbol{X_{ij}}$ is a set of individual level controls, including gender, month, and year of birth dummies (Black et al., 2005; Handy & Shester, 2022). It also includes the vector of the first 40 principal components (PCs) of the genetic relatedness matrix.[11] Finally, $\boldsymbol{\xi_{ij}}$ is the error term; we use heteroskedasticity-robust standard errors, clustered at the family level. The parameter α_2 captures the association between the polygenic score for education and years of schooling, whilst α_3 estimates the average advantage in years of schooling for firstborn children compared with their laterborn siblings. α_4 shows the extent to which the polygenic score and being first born complement each other's effect on education and is therefore informative about the existence of putative G×E effects.[12] Following Black et al. (2005, 2011), Heiland (2009), Lehmann, Nuevo-Chiquero, & Vidal-Fernandez (2018), and Amin et al. (2021), we compare specifications with and without family fixed effects. The inclusion of family fixed effects theoretically ensures that variation in the polygenic score and birth order is random, ensuring polygenic score and birth order are orthogonal to each other, and we verify this empirically in Section 2.4.

Our polygenic score incorporates two sources of measurement error. First, since the GWAS underlying the construction of a polygenic score is based on a finite sample (see Section 2.3), our estimated polygenic score that is based on the GWAS estimates is a noisy proxy for the true (latent) polygenic score (Van Kippersluis et al., 2022). Second, GWASs typically do not control for parental genotypes (Howe et al., 2022). Trejo & Domingue (2019) show that not controlling for parental genotypes causes the polygenic scores to be measured with error. Both sources of measurement error lead

[10] For ease of interpretation, we standardize $\boldsymbol{G_{ij}}$ to have mean zero and standard deviation one. See Section 2.3 and Appendix A for more information about its construction.

[11] Genetic principal components (PCs) can be used to control for subtle forms of population stratification (i.e., correlations between allele frequencies and environmental factors across subpopulations in the sample) in a between-family (population-level) analysis (Price et al., 2006; Rietveld et al., 2014). Although not strictly necessary to include in the within-family analysis due to the inclusion of family fixed effects, we keep the PCs in all specifications to facilitate a clean comparison between the between-family and within-family results.

[12] In section 2.4, we study the interaction effect using more flexible approaches than the linear interaction presented here.

to an attenuation bias in the coefficient of the polygenic score in within-family analyses.[13] While we cannot solve the attenuation bias arising from the omission of parental genotype in the GWAS, we follow DiPrete, Burik, & Koellinger (2018) and Van Kippersluis et al. (2022) in applying Instrumental Variables (IV) to tackle the classic measurement error problem. More specifically, we split our discovery GWAS sample into two equal halves and construct two polygenic scores based on the two discovery samples. To ensure that these subsamples are unrelated, we randomly select only one individual from each cousin cluster. Even though the two polygenic scores individually have lower predictive power, the measurement error in the two is plausibly orthogonal and so they can be used as instrumental variables for each other. Using these two polygenic scores, we apply Obviously-Related Instrumental Variables (ORIV; Gillen et al., 2019); see Appendix C for more details.

2.3 Data

We use data from the UK Biobank, a population-based cohort with approximately 500,000 individuals aged between 40-69 at the time of interview in 2006-2010 and living within a radius of 35 km from one of the 22 assessment centres in England, Wales, and Scotland (Fry et al., 2017). It contains survey data, biomarker and DNA samples, physical measurements, and linkage to inpatient registers and death records (Sudlow et al., 2015). Because participation in the UK Biobank is voluntary, it is not a representative sample of the UK population (see Fry et al. (2017) for a detailed analysis).

We apply the following sample selection criteria. The original data include 502,498 consented individuals. We follow the literature and remove those of non-European descent (92,892 observations), as well as twins and multiple births (9,310 observations), and individuals with missing or conflicting information regarding the number of siblings and/or family size (3,801 observations). In doing so, we arrive at a sample of 396,494 individuals. We further restrict this sample to individuals with at least one sibling[14] in the UK Biobank and without missing values on any of the variables included in our analysis (i.e., years of education, birth order, family size, year and month of birth, principal components of the genetic relatedness matrix,

[13] The reason why classical measurement error as a result of finite-sample bias leads to attenuation bias is well known. It is more subtle why the exclusion of parental genotype in the discovery GWAS leads to an attenuation bias in within-family studies. The reason is that a polygenic score constructed on basis of a GWAS that did not control for parental genotype will reflect both direct genetic effects arising from the individual's genotype as well as indirect genetic effects arising from the omitted parental genotypes. The latter effects are known as 'genetic nurture' (e.g., Kong et al., 2018). When applied within families, the differences in the polygenic score arising from parental genotype are spurious since parental genotype is the same across siblings. Hence, part of the differences across siblings in the polygenic score reflect genetic nurture (i.e., indirect genetic effects) and can be considered measurement error attenuating the resulting within-family estimates (Trejo & Domingue, 2019).

[14] Siblings are identified based on the genetic data because there is no direct information about sibling status in the UK Biobank. We also use the genetic data to identify other relatives up to the third degree. See Appendix B for the full procedure followed to identify siblings and relatives.

gender, and the polygenic score for education). Since the UK Biobank did not specifically target families, this leads to a final sample size of N = 14,850 individuals within 7,281 full-sibling clusters.

We follow the literature (see e.g., Lee et al., 2018; Okbay et al., 2016, 2022; Rietveld et al., 2013) and convert individuals' qualifications to equivalent years of education using the International Standard Classification of Education (ISCED).[15] The average years of education for the sibling sample is 15 years (see Table 2.1). We construct individuals' birth order based on their response to a question of how many older siblings they have. If a respondent reports zero older siblings, the birth order is set to one. For individuals with missing information on the number of older siblings, we determine birth order based on family size and birth year of the individual and his/her siblings if all of them are present in the UK Biobank. This adds information on birth order for 1,752 siblings in our analysis sample. Table 2.1 shows that 39.4% of our sample are firstborns, with an average birth order of 1.91 (where we have censored birth order at 5 for the 245 respondents with birth order beyond 5). Around 37% of our sample is lastborn, and the average family size is 3 (i.e., the average number of siblings is 2).

Table 2.1. Descriptive statistics analysis sample (N = 14,850).

Variable	Mean	S.D.	Min.	Max.
Years of education	15.058	4.951	7.000	20.000
Firstborn (1/0)	39.4%			
PGS for years of education	0.000	1.000	-4.049	3.912
Birth order	1.913	0.998	1.000	5.000
Second born	41.50%			
Third born	11.18%			
Fourth born	4.33%			
Fifth- or laterborn	3.60%			
Family size	2.987	1.530	2.000	14.000
Last child (1/0)	36.9%			
Male (1/0)	42.5%			

Notes: PGS = Polygenic score; S.D. = Standard deviation; Min. = Minimum; Max. = Maximum.

Our measure of genetic endowment for education is the polygenic score for education. A polygenic score is a weighted sum of genetic variants called Single Nucleotide Polymorphisms (SNPs, see Appendix A for details). The SNP weights are determined by the association between a SNP and years of education (Dudbridge,

[15] Years of education ranges from 7 to 20, where College or University degree is equivalent to 20 years, National Vocational Qualification (NVQ), Higher National Diploma (HND), or Higher National Certificate (HNC) to 19 years, other professional qualifications to 15 years, having an A or AS levels or similar to 13 years, O levels, (General) Certificate of Secondary Education ((G)CSE) to 10 years, and if none of the above to the lowest level of 7 years.

2013) in an independent (discovery) sample:

$$PGS_i = \sum_{k=1}^{K} \beta_k x_{ik}, \tag{4}$$

where PGS_i is the value for the polygenic score for individual i, β_j is the regression coefficient of SNP k ($k = 1, \ldots, K$) from our own tailor-made GWAS (see below), and x_{ik} is the genotype of individual i for SNP k (coded as 0, 1 or 2, indicating the number of "effect" alleles). We use the LDpred software (Vilhjálmsson et al., 2015) to correct for the correlation structure across SNPs. The polygenic scores are standardized within the sibling sample to have mean 0 and standard deviation 1.

The polygenic score measures the genetic predisposition towards educational attainment within the environmental context and demographic characteristics of the discovery GWAS sample (Domingue, Trejo, Armstrong-Carter, & Tucker-Drob, 2020). It is therefore preferable to select discovery and analysis samples from the same environmental context, especially when analysing gene-environment interactions. At the same time, the discovery sample should be independent of the analysis sample to avoid overfitting (Dudbridge, 2013). Navigating this trade-off, we therefore construct the polygenic score by using the weights from our own tailor-made GWAS that uses the UK Biobank sample, but without the siblings (our analysis sample) and their relatives. The GWAS discovery sample comprises 389,419 individuals from the UK Biobank; we use the summary statistics from these analyses to create the polygenic scores on the sample of 14,850 siblings. This tailor-made polygenic score alleviates the differences between the discovery and the analysis samples in terms of demographics and environmental context, as well as measurement, i.e., the variables of interest are measured in the same way (Elam, Clifford, Shaw, Wilson, & Lemery-Chalfant, 2019; Tropf et al., 2017). Moreover, running our own GWAS enables the construction of two independent polygenic scores, obtained by splitting the discovery GWAS sample into two equal halves, which can be used in ORIV. This approach has been shown to outperform a single polygenic score that is based on meta-analysing multiple cohorts (Van Kippersluis et al., 2022).[16]

[16] To check if our results are sensitive to the choice of polygenic scores, we constructed a polygenic score based on the meta-analysed GWAS results of 23andMe summary statistics and our own UK Biobank discovery sample GWAS. As expected, this polygenic score is more predictive for educational attainment than the polygenic score constructed on the basis of the UK Biobank alone. However, this polygenic score is based on several discovery cohorts from very different environmental contexts and does not allow us to use ORIV since we do not have access to all underlying samples to allow us to create multiple polygenic scores. The results are very similar, with the interaction term being slightly smaller but not significantly different from our main results (see Appendix F).

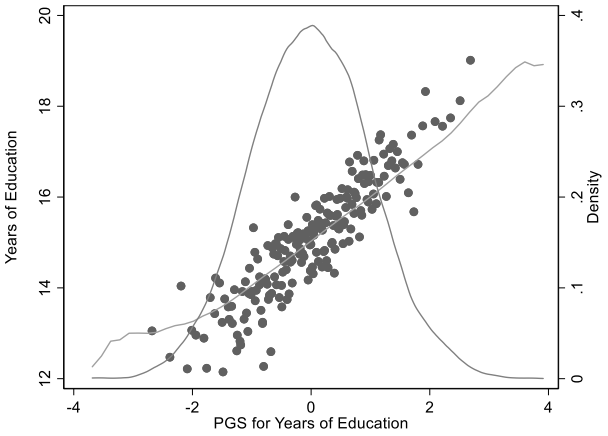
2.4 Results

Predictive power of the polygenic score for educational attainment

Figure 2.1 shows that our polygenic score for years of education is approximately normally distributed. For the scatterplot, we divided the polygenic score in 200 bins; the dots represent the average years of education for each bin. The line through the dots is obtained from a local polynomial regression of years of education on our polygenic score. In line with the literature the polygenic score is positively correlated with years of education ($r = 0.23$, $p < 0.001$). The average difference between those two standard deviations below the mean of the polygenic score, and those two standard deviations above the mean is almost 4 years of completed education, highlighting the substantial predictive power of the polygenic score. Furthermore, Figure 2.1 suggests the relationship is approximately linear.

Table 2.2 shows that the incremental R^2 of the polygenic score (i.e., the additional variance explained by the polygenic score after controlling for gender, month, and year of birth, and the first 40 principal components of the genetic relatedness matrix) is 5.4% in the analysis without family fixed effects (i.e., 0.113-0.059=0.054; Columns 1 and 2). In the specifications with family fixed effects (Columns 3 and 4), the incremental (within) R^2 for the polygenic score is reduced to 1%. This reduction in predictive power is well-established in the literature (see e.g., Koellinger & Harden, 2018; Kong et al., 2018; Lee et al., 2018; Rietveld et al., 2013; Selzam et al., 2019), and reflects the fact that family fixed effects account for the shared family environment and parental genotype, which was not accounted for in the first specification. In terms of the effect sizes, we observe that a one standard deviation increase in the polygenic score is associated with an increase of 1.155 years of education. With family fixed effects, the effect size is reduced to 0.646.

Figure 2.1. The relationship between the standardized polygenic score and years of education in the analysis sample.



Notes: Plotted using 200 bins of the polygenic score.

Table 2.2. Results of the regressions of years of education on the polygenic score (PGS).

	Without family fixed effects		With family fixed effects	
	(1)	(2)	(3)	(4)
PGS for years of education		1.155***		0.646***
		(0.038)		(0.071)
Constant	15.889***	15.249***	13.757***	13.628***
	(1.690)	(1.666)	(1.685)	(1.821)
R^2	0.059	0.113	0.046	0.056
N	14,850	14,850	14,850	14,850

Notes: Robust standard errors in parentheses, clustered by family; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Coefficients for the control variables (year and month of birth, gender and the first 40 principal components of the genetic relatedness matrix) are not displayed, but available upon request from the authors.

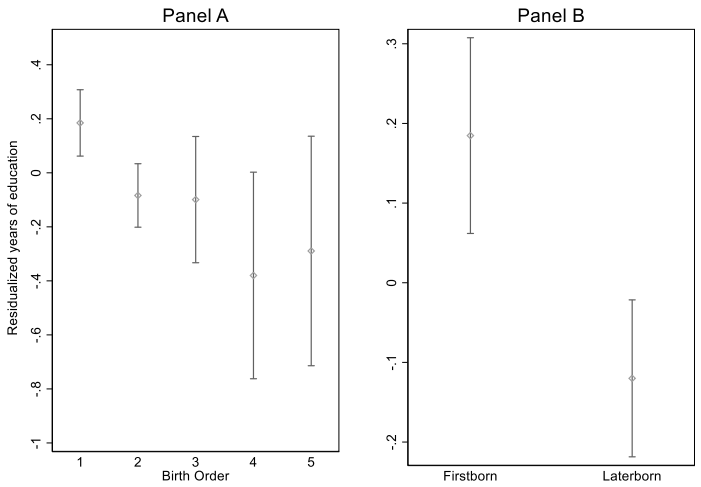
The relationship between birth order and educational attainment

Figure 2.2 shows differences in years of education (residualised with respect to year and month of birth, family size, gender, and 40 first principal components of the genetic relatedness matrix) by birth order (panel A). Not taking into account family fixed effects, we find that years of education for laterborn children is lower (albeit

with larger confidence intervals) than that for firstborns. When pooling all laterborns (panel B), the difference between firstborns and laterborns becomes more pronounced.

Table 2.3 confirms the birth order effects in the specifications with and without family fixed effects. We observe a consistent gap of 0.3-0.4 years of schooling between first- and laterborn children. The direction and magnitude of the effect is robust to using the binary indicator or the categorical variable for birth order. The within-family effect size for 5th born children does not reach statistical significance due to the relatively small number of observations (see Table 2.1).

Figure 2.2. The relationship between birth order and years of education in the analysis sample.



Notes: 95% confidence intervals. Years of education is residualised with respect to year and month of birth, family size, gender, and the first 40 principal components of the genetic related matrix.

Table 2.3. Regressions of years of education on different specifications of birth order.

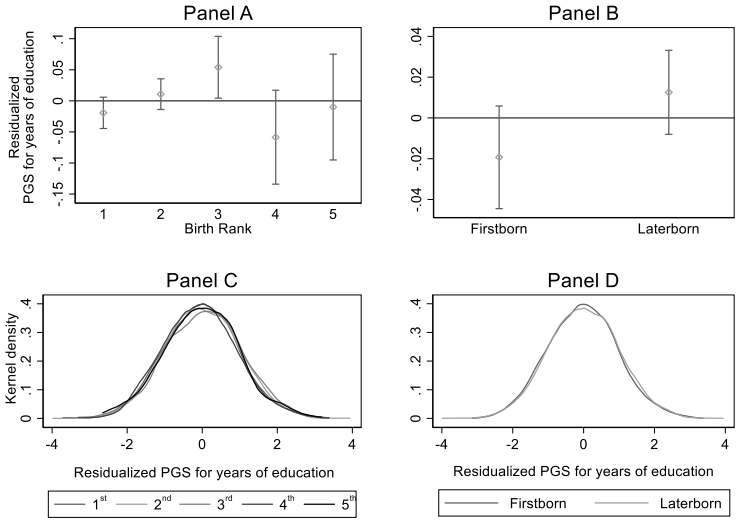
	Without family fixed effects		With family fixed effects	
	(1)	(2)	(3)	(4)
Firstborn	0.357***		0.418***	
	(0.087)		(0.109)	
2nd born		-0.310***		-0.450***
		(0.090)		(0.124)
3rd born		-0.431***		-0.743***
		(0.146)		(0.245)
4th born		-0.822***		-0.808**
		(0.228)		(0.366)
5th born		-0.892***		-0.463
		(0.279)		(0.498)
Constant	16.258***	16.473***	12.711***	13.002***
	(1.746)	(1.737)	(1.681)	(1.678)
R^2	0.069	0.069	0.048	0.048
N	14,850	14,850	14,850	14,850

Notes: Robust standard errors in parentheses, clustered by; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Coefficients for the control variables (year and month of birth, gender and the first 40 principal components of the genetic relatedness matrix) are not displayed, but available upon request from the authors.

The relationship between birth order and the polygenic score for years of education

To measure causal gene-environment interactions, birth order and the polygenic score need to be orthogonal. Figure 2.3 provides the first impression of the relationship between the two measures, where the polygenic score for years of education is again residualised using our standard set of controls. Panel A illustrates that the educational attainment polygenic score does not reveal any systematic pattern across birth order. Although Panel A suggests that 3rd born children have somewhat higher genetic endowments, this result is not confirmed in the within-family analysis. The same holds when focusing the comparison on firstborns versus laterborns in Panel B: the observed differences are small and not statistically significant. Furthermore, the distributions of the polygenic score by birth order are overlapping almost perfectly (Panels C and D).

Figure 2.3. The relationship between birth order and the polygenic score for years of education in the analysis sample.



Notes: 95% confidence intervals. The polygenic score (PGS) for years of education is residualised with respect to year and month of birth, family size, gender, and 40 first principal components of the genetic related matrix.

Consistent with the graphical evidence, Table 2.4 shows a slight difference of 0.04 standard deviations in the polygenic score between firstborn and laterborn children without family fixed effects. However, the difference becomes negligible and statistically insignificant within families. These results indicate that firstborns on average do not have different genetic endowments for educational attainment compared to their laterborn siblings (as expected, based on Mendel's Law of Independent Assortment).

The evidence is however not sufficient to claim that birth order is unrelated to any genetic endowments. To check if there are systematic differences by birth order in other, possibly related, polygenic scores, we analyse the polygenic scores from the Polygenic Index Repository (Becker, Burik, et al., 2021) for all anthropometric, health, health behaviour, and personality traits. We find no systematic difference by birth order in any of the 29 available polygenic scores (see Appendix G for details). These results corroborate that there is no systematic association between birth order and genetic endowments, and hence no evidence for gene-environment correlation (*rGE*).

Table 2.4. Results of the regressions of polygenic score for educational attainment on birth order.

	Without family fixed effects		With family fixed effects	
	(1)	(2)	(3)	(4)
Firstborn	-0.037**		-0.002	
	(0.018)		(0.018)	
2nd born		0.034*		0.008
		(0.019)		(0.020)
3rd born		0.078**		0.009
		(0.031)		(0.039)
4th born		-0.036		0.018
		(0.045)		(0.057)
5th born		-0.013		0.061
		(0.057)		(0.080)
Constant	0.648*	0.605	0.207	0.226
	(0.356)	(0.355)	(0.377)	(0.378)
R^2	0.017	0.018	0.013	0.013
N	14,850	14,850	14,850	14,850

Notes: Robust standard errors in parentheses, clustered by; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Coefficients for the control variables (year and month of birth, gender and the first 40 principal components of the genetic relatedness matrix) are not displayed, but available upon request from the authors.

The gene-environment interaction

Table 2.5 presents the gene-environment interaction results. Comparison of Columns 1 and 4 in Table 2.5 to the estimates presented in Table 2.2 and Table 2.3 shows that including both the educational attainment polygenic score and the indicator for being firstborn does not change their main effects, again indicating independence between the polygenic score and birth order. Firstborns enjoy on average 0.40 – 0.42 extra years of schooling compared to laterborns. Furthermore, a one standard deviation increase in the polygenic score is estimated to raise years of education by 1.14 years (without family fixed effects) and 0.65 years (with family fixed effects).

The design without family fixed effects shows a positive interaction between the polygenic score and being firstborn (Column 2), which is borderline statistically significant at conventional thresholds. When we use Obviously-Related Instrumental Variables (ORIV) regression (Column 3), the interaction term becomes larger in magnitude and statistically significant at the 5% threshold, suggesting that measurement error in the polygenic score attenuated the main effect of the polygenic score and the interaction term. In the family fixed effects specification, we find that the interaction effect is significant in both the OLS and the ORIV specifications. The

measurement error correction again strengthens the interaction term in terms of magnitude.

The positive and statistically significant interaction term provides strong evidence for gene-environment interactions in education and is consistent with the existence of complementarity between endowments and investments in human capital production: the effect of being firstborn (associated with more parental investments) is complementary to a higher value for the polygenic score for years of education. In other words: those with a higher polygenic score benefit more from being firstborn and its associated environmental benefits (e.g., higher parental investments). The magnitude of the coefficients suggests that for those with a polygenic score two standard deviations below the mean, there is no advantage of being firstborn. In contrast, for those with a high polygenic score, being firstborn increases one's years of education. For example, firstborns with a polygenic score two standard deviations above the mean completed on average 0.82 (OLS) to 1 (ORIV) more years of education compared to their laterborn siblings with a similar polygenic score.[17]

Table 2.5. Results of the regressions of years of education on the gene-environment interaction.

	Without family fixed effects			With family fixed effects		
	(1) OLS	(2) OLS	(3) ORIV	(4) OLS	(5) OLS	(6) ORIV
Firstborn	0.400***	0.400***	0.428***	0.419***	0.415***	0.428***
	(0.078)	(0.078)	(0.078)	(0.109)	(0.109)	(0.109)
PGS for years of education	1.144***	1.099***	1.376***	0.646***	0.567***	0.757***
	(0.040)	(0.049)	(0.063)	(0.071)	(0.078)	(0.102)
Firstborn × PGS for years of education		0.119*	0.222**		0.204**	0.285***
		(0.072)	(0.097)		(0.080)	(0.102)
Constant	15.516***	15.517***	14.742***	12.578***	12.544***	12.184***
	(1.710)	(1.716)	(1.762)	(1.819)	(1.838)	(4.419)
R^2	0.121	0.121		0.058	0.059	
Cragg-Donald *F*-statistic			3568.570			2971.412
N	14,850	14,850	14,850	14,850	14,850	14,850

Notes: Robust standard errors in parentheses, clustered by family; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Coefficients for the control variables (year and month of birth, gender and the first 40 principal components of the genetic relatedness matrix) are not displayed, but available upon request from the authors. We do not report the R^2 for the ORIV specifications in Column (3) and (6) given the differences in its interpretation and computation for the instrumental variable type of regressions.

[17] We also investigated potential gender differences in the interaction of interest. This analysis relies on families with a mixed gender composition to ensure variation in gender within families, reducing the analysis sample considerably. While we do find that gender is an important predictor of educational attainment for the cohorts examined here, with men having 1.46 more years of education than women, the interaction term does not differ significantly by gender.

Robustness checks

In this section, we check the robustness of our results against potential non-linearities in the functional form of the polygenic score as well as birth order, against the addition of further control variables, and against potentially endogenous fertility choices. While the linear form adopted in the previous section seems justified based on the visual inspection of Figure 2.1, we explore robustness of our results by allowing for possible non-linearities. Table 2.6 compares the within-family specification in continuous form (replicated in Column 1 for comparison), with those where we specify the polygenic score as binary (above and below the mean, Column 2), in quartiles (Column 3) and in squared form (Column 4). We observe positive interaction terms across all specifications. In line with our main findings, the effect of being firstborn is insignificant for those with a lower polygenic score, and the effects are concentrated among those in the top quartile of the polygenic score distribution. We also find that the main result in Column 1 is robust to the inclusion of a quadratic version of the polygenic score.

Table 2.7 reports the sensitivity of the results to an alternative specification of birth order. Column 1 replicates the main result from Table 2.5 for comparison. In Column 2, we include dummies for each birth rank with firstborns as the reference category. All point estimates of the main effects and the interaction terms are as expected: on average, secondborns have lower educational attainment compared to firstborns and benefit less from having a high polygenic score. The estimates for higher ranks are rather imprecisely estimated due to the relatively small sample sizes for those with a birth rank of three or higher, but the point estimates consistently suggest that laterborns benefit less from having a high polygenic score.

In Table 2.8, we explore robustness of our results to potentially endogenous fertility decisions. A possible correlation between our measure of endowments and birth order could arise when fertility decisions are based on the genetic endowments of the children, known in the literature as the "child stopping rule" (Black et al., 2005; Pavan, 2016). Whereas earlier we show that such a correlation does not exist, here we explicitly control for a possible child-stopping rule by including a dummy variable for being lastborn, which is set to one if an individual's birth order is equal to the total number of children in his/her family. To facilitate an easy comparison, Column 1 in Table 2.8 replicates the results from Column 5 in Table 2.5. As can be seen from Column 2, the lastborn dummy is not statistically significant and does not meaningfully affect our results, suggesting that potential endogenous fertility decisions do not change any of our conclusions. In Columns 3 and 4, we report the within-family results based on data from the families with at most three siblings (N = 11,364) and two siblings only (N = 7,918) to check if our results are sensitive to the exclusion of relatively large families with possibly different characteristics. Even though this sample restriction is endogenous, it is reassuring that the point estimates are virtually identical in these restricted samples.

Table 2.6. Results of the regressions of years of education on the gene-environment interaction; Robustness to non-linearities in the polygenic score.

	With family fixed effects			
	(1)	(2)	(3)	(4)
Firstborn	0.415***	0.239*	0.186	0.437***
	(0.109)	(0.145)	(0.198)	(0.125)
PGS for years of education	0.567***			0.566***
	(0.078)			(0.078)
Firstborn × PGS for years of education	0.204**			0.205**
	(0.080)			(0.081)
PGS for years of education (>mean)		0.588***		
		(0.136)		
Firstborn × PGS for years of education (>mean)		0.356**		
		(0.169)		
PGS for years of education (2nd quartile)			0.439**	
			(0.190)	
PGS for years of education (3rd quartile)			0.810***	
			(0.197)	
PGS for years of education (4th quartile)			1.230***	
			(0.208)	
PGS for years of education (2nd quartile) × Firstborn			0.155	
			(0.276)	
PGS for years of education (3rd quartile) × Firstborn			0.190	
			(0.255)	
PGS for years of education (4th quartile) × Firstborn			0.595**	
			(0.232)	
PGS for years of education (squared)				0.018
				(0.047)
PGS for years of education (squared) × Firstborn				-0.022
				(0.061)
Constant	12.544***	12.087***	11.821***	12.518***
	(1.838)	(1.796)	(1.868)	(1.838)
R^2	0.059	0.053	0.057	0.059
N	14,850	14,850	14,850	14,850

Notes: Robust standard errors in parentheses, clustered by family; * $p<0.10$, ** $p<0.05$, *** $p<0.01$; Coefficients for the control variables (year and month of birth, gender and the first 40 principal components of the genetic relatedness matrix) are not displayed, but available upon request from the authors.

Table 2.7. Regressions of years of education; Robustness to non-linearities in birth order.

	With family fixed effects	
	(1)	(2)
PGS for years of education	0.567***	0.767***
	(0.078)	(0.086)
Firstborn	0.415***	
	(0.109)	
Firstborn × PGS for years of education	0.204**	
	(0.080)	
2nd born		-0.450***
		(0.124)
3rd born		-0.741***
		(0.244)
4th born		-0.832**
		(0.362)
5th born		-0.470
		(0.499)
2nd born × PGS for years of education		-0.212**
		(0.083)
3rd born × PGS for years of education		-0.152
		(0.148)
4th born × PGS for years of education		-0.316
		(0.254)
5th born × PGS for years of education		-0.089
		(0.300)
Constant	12.544***	12.882***
	(1.838)	(1.842)
R^2	0.059	0.060
N	14,850	14,850

Notes: Robust standard errors in parentheses, clustered by family; * $p<0.10$, ** $p<0.05$, *** $p<0.01$; Coefficients for the control variables (year and month of birth, gender and the first 40 principal components of the genetic relatedness matrix) are not displayed, but available upon request.

Table 2.8. Results of the regressions of years of education on the gene-environment interaction; Robustness to fertility choices.

	With family fixed effects			
	(1)	(2)	(3)	(4)
Firstborn	0.415***	0.392***	0.439***	0.466***
	(0.109)	(0.128)	(0.132)	(0.178)
PGS for years of education	0.567***	0.568***	0.589***	0.570***
	(0.078)	(0.078)	(0.090)	(0.109)
Firstborn × PGS for years of education	0.204**	0.203**	0.197**	0.183*
	(0.080)	(0.080)	(0.087)	(0.102)
Lastborn		-0.049		
		(0.136)		
Male	1.265***	1.265***	1.429***	1.571***
	(0.099)	(0.099)	(0.109)	(0.130)
Male × PGS for years of education				
Firstborn × PGS for years of education × Male				
Constant	12.544***	12.509***	13.065***	12.578***
	(1.838)	(1.840)	(2.106)	(2.148)
R^2	0.059	0.059	0.066	0.073
N	14,850	14,850	11,364	7,918

Notes: Column 1 replicates our main results from Table 2.5; Column 2 additionally controls for a dummy indicating whether the individual is the lastborn; Column 3 restricts the sample to families with less than four siblings; Columns 4 restricts the sample to families with two siblings only. Robust standard errors in parentheses, clustered by family; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Coefficients for the control variables (year and month of birth, family size, and the first 40 principal components of the genetic relatedness matrix) are not displayed, but available upon request from the authors.

2.5 Discussion

A large literature shows consistently higher educational attainments for firstborn children. Using within-family data we move beyond the existing literature by showing that children benefit disproportionally from being firstborn when they have a relatively high genetic endowments, as proxied by their polygenic score for educational attainment. More specifically, firstborns with an average polygenic score enjoy 0.428 years (≈ 5 months) of additional schooling compared to their laterborn siblings, on average. However, firstborns with a polygenic score that is one standard

deviation above the mean enjoy an additional 0.285 years (≈ 3.5 months) of education, compared to their laterborn siblings with the same genetic endowment. In contrast, for individuals with below-average polygenic scores, being firstborn does not provide an advantage in terms of educational attainment. Since previous literature suggests that birth order effects on children's education are at least partly driven by parental investments, a plausible interpretation of the positive and significant interaction term is the existence of the complementarity between endowments and investments in human capital production.

An alternative interpretation of our finding that birth order effects are concentrated among those with higher polygenic scores could be that the additional investments associated with being firstborn are higher for those with higher polygenic scores. That is, the positive interaction effect could also be explained by parents investing more in the firstborn, or altering fertility decisions, when the firstborn child has a higher polygenic score. While we cannot fully rule out this explanation, we believe this explanation is less plausible for three reasons. First, Breinholt & Conley (2019) and Houmark et al. (2020) show that parenting during infancy is not driven by genetic make-up because these endowments are not clearly expressed yet, and parental investment responses to polygenic scores do not arise before age six. This is long after the typical arrival of subsequent children, and so the most precious time of undivided attention for the firstborn is unlikely to be influenced by – at that time unobserved – differences in polygenic scores.[18] Second, whereas it is established that parents do respond to the polygenic scores of children at later ages (e.g. Sanz-De-Galdeano & Terskaya, 2023), we control for this with the main effect of the polygenic score in our specification. Only when – for some reason – parents respond more to the polygenic score when the child is firstborn, our interpretation of the interaction term would be challenged. We cannot rule this out, but deem it less plausible, as the additional investments associated with being firstborn decrease with age (Price, 2008), and become increasingly modest at the ages where genetic endowments are more clearly expressed (Breinholt & Conley, 2019). Third, recent evidence exploiting within-family differences in polygenic scores for educational attainment suggests that parents compensate for, rather than reinforce, genetic differences in education between siblings. More specifically, Fletcher et al. (2020) show that the association between the education polygenic score and educational attainment is stronger for siblings with the lower polygenic score. This is consistent with parental preferences for equality among siblings, and compensating parental investments, and does not support the alternative interpretation that parents invest more in the firstborn only when the firstborn has a higher polygenic score. Our findings therefore suggest that the additional investments associated with being firstborn are driven by less restrictive time and budget constraints and are independent of the child's genetic endowment.

[18] For the few early-life parental investments we observe in our data, we do not find evidence of any response to the polygenic score. Appendix D shows that maternal smoking around pregnancy and whether the child was breastfed are all unrelated to the firstborn's polygenic score. If anything, the age gap between first- and secondborns is slightly lower if the firstborn has a higher polygenic score.

Another potential explanation for our findings is that other mechanisms through which birth order effects arise (e.g., parental age at birth, interactions with younger siblings) could interact with genetic endowments. It is reassuring that additionally accounting for parental age at birth and other potential confounders does not affect our results (see Appendix E). A recent study however suggests that older siblings may influence the education of younger siblings, although it is acknowledged that the estimated relationship may also reflect unaccounted for differences between families (Howe et al., 2022). Unfortunately, we cannot test alternative explanations regarding sibling interactions directly because the UK Biobank is very limited in measures of actual parental and sibling interactions. Still, Table 2.7 suggests that our findings are mainly driven by a distinction between first- and all laterborns, with significant differences appearing prominently between the first and secondborns. If interactions with younger siblings would be driving our results, one would expect a more gradual decrease in the magnitude of the interaction terms with rising birth rank, as some second- and thirdborns similarly benefit from interactions with younger siblings.

It is important to emphasize that even if birth order only partially captures parental investment – a premise that should not be overly controversial given the evidence in the literature – then unless these other channels exhibit completely opposite interaction effects, a necessary (but not sufficient) condition for complementarity would be a positive interaction between birth order and genetic endowments. This is indeed what we find. More generally, and independent of the exact environmental mechanism, since we provide evidence that genetic endowments are orthogonal to birth order, our findings provide one of the first pieces of empirical evidence for the long-held theoretical belief that educational attainment is shaped by a complex interplay between genes and environments.

A number of limitations should be acknowledged. First, our specification may not be the perfect empirical translation of the human capital production function. In particular, we do not measure skills or human capital directly (see e.g., Araujo et al. 2016 for a study that measures skills directly). Instead, we follow Cunha & Heckman (2008) and Cunha et al. (2010) who specify adult human capital as a combination of skills accumulated by the end of childhood, and employ a commonly used and convenient proxy: years of education. Moreover, we do not measure parental investments directly, and use an environmental variable closely related to parental investments: birth order. The upside of using birth order rather than a direct measure of parental investments is that birth order is randomly assigned within families, whereas parental investments are known to be endogenous to offspring endowments. Moreover, whereas birth order cannot distinguish between early-life and later-life investments, it does capture a persistent difference across siblings rather than a one-time shock in investments that many other papers rely on (see Almond et al., 2018, and Appendix B). Future studies should employ richer measures to break down the complementarity between genes and advantageous environments into e.g., parental investments and sibling interactions.

A second limitation is that our measure of genetic endowments is imperfect. In particular, a polygenic score captures only common genetic variations in the human genome, and even within the realm of common variations the measure is subject to measurement error. While the use of ORIV reduces concerns about classical measurement error, our family fixed effects estimates of the polygenic score are still subject to attenuation bias due to genetic nurture. Still, since the sign of the bias arising from genetic nurture is known to be negative, our effect size represents in fact a conservative estimate.

The polygenic score should also not be interpreted narrowly as a measure of immutable biological endowments: while within-family analyses allow us to interpret the effect of the polygenic score as a causal effect of genetic variation, it is well-established that the environment may mediate this effect (e.g., Breinholt & Conley, 2019; Houmark, Ronda, & Rosholm, 2020), including the family environment (Fletcher et al., 2020). Thus, a polygenic score measures education-enhancing endowments, and will reflect how *on average* in the discovery sample environments respond to differences in genetic endowments. Importantly though, since the measure is fixed at conception and orthogonal to birth order in the within-family analysis, the measure does not reflect parental investments *of the child's own parents*, nor does it reflect parental genetic (nurture) effects. As a result, the inclusion of environmental responses to genetic variation into the construction of the polygenic score is not a source of concern for our identification strategy but does imply that our findings may be specific to the context studied.

A related limitation regards the external validity of the empirical findings. As mentioned in the data section, there is sample selection into the UK Biobank, with a bias towards healthier and higher-educated individuals (Fry et al., 2017). On top of this, we focus on European-ancestry individuals and the coincidental sampling of siblings even though these were not specifically targeted, further reducing the representativeness of the sample. Finally, we construct our polygenic score on basis of a tailor-made GWAS, again on basis of the same UK Biobank excluding the siblings and their relatives. While the latter choice helps to maintain the same environments across discovery and prediction sample, it may further increase the likelihood that our results are specific to the UK Biobank. Therefore, we call for future studies to replicate our findings in other contexts.

Acknowledgments

This research has been conducted using the UK Biobank Resource (Application Number 41382). The authors gratefully acknowledge funding from NORFACE through the Dynamic of Inequality across the Life Course (DIAL) programme (462-16-100). Research reported in this publication was also supported by the National Institute on Aging of the National Institutes of Health (R56AG058726). C.A.R. and S.v.H. gratefully acknowledge funding from the European Research Council (GEPSI 946647; DONNI 851725). S.F.W.M. gratefully acknowledges funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement (GENIO 101019584). We thank Aysu Okbay and employees and participants of the 23andMe, Inc. for sharing GWAS summary statistics for educational attainment. We also thank Pietro Biroli, Gabriella Conti, Ben Domingue, Monique de Haan, Paul Hufe, Owen O'Donnell, Kjell Salvanes, Dinand Webbink, Herman van de Werfhorst, and seminar participants at the NORFACE DIAL workshop on intergenerational mobility and human capital accumulation, the Vienna University of Economics and Business, DIAL Mid-term Conference (University of Turku), RSF Summer Institute in Social Science Genomics, Health Economics Symposium (Erasmus University Rotterdam), the Applied Microeconomics and Gender Workshop (University of Alicante), and the Dutch Economists Day for providing constructive comments. This work made use of the Dutch national e-infrastructure with the support of the SURF Cooperative (EINF-1107).

2.6 Appendix to Chapter 2

A. Genetic data, GWAS, and Polygenic scores

Genetic Data. A complete human genome consists of 23 pairs of chromosomes, from which the 23rd pair determines the biological sex of a person. One of each pair of chromosomes is inherited from the father, and the other is inherited from the mother. A chromosome is composed of two intertwined strands of deoxyribonucleic acid (DNA), each made up of a sequence of four possible nucleotide molecules: adenine, cytosine, thymine, and guanine. Adenine (A) on one strand is always paired with thymine (T) on the other strand, and cytosine (C) is always paired with guanine (G). These pairs are called base pairs. Every human genome consists of approximately 3 billion base pairs and stretches of base pairs coding for proteins are called genes. There are approximately 20,000 genes in the human genome, with varying lengths in terms of base pairs (Ezkurdia et al., 2014).

Two unrelated human beings share approximately 99.6% of their DNA, and most genetic differences across humans can be attributed to single nucleotide polymorphisms (SNP) (Auton et al., 2015). A SNP is a locus in the DNA at which two different nucleotides can be observed in the population. Each of the two possible nucleotides is called an allele for that SNP. An individual's genotype is coded as 0, 1, or 2, depending on the number of "effect" alleles present. In the human genome, there are at least 85 million SNPs with a "minor" allele prevalence of at least 1% (Auton et al., 2015).

Genome-Wide Association Studies (GWASs) aim to identify genetic variants that are associated with a particular trait of interest by relating each variant to the trait in a hypothesis-free approach. Stringent significance thresholds are used to identify variants that are robustly associated with the trait, with other independent samples used for replication. Using the GWAS approach, thousands of genetic discoveries have been made (Visscher et al., 2017).

Individual SNPs typically explain less than 0.02% of the variance in a behavioural outcome (Chabris et al., 2015; Visscher et al., 2017). It is therefore common to combine multiple SNPs into a polygenic score (Dudbridge, 2013), constructed as a weighted sum of SNPs. Through increases in GWAS sample sizes, the predictive power of the polygenic score for education has increased from 2-3% (Rietveld et al., 2013b), to 6-8% (Okbay et al., 2016), to 11-13% (Lee et al., 2018), to 12-16% (Okbay et al., 2022). In terms of biological pathways, there is evidence that many of the identified genes associate with health, cognitive, and central nervous system traits (Rietveld et al., 2013b). Likewise, the majority of the significant SNPs in Okbay et al. (2016) and Lee et al. (2018) relate to genomic regions responsible for gene expression in a child's brain during the prenatal period.

Methods. *Relatedness.* As a first step, we identify siblings and their relatives using the

kinship matrix provided by the UK Biobank. The kinship matrix is based on genetically identified relatedness and contains relatives of third degree and closer identified using the KING software (Manichaikul et al., 2010). The UK Biobank does not have information about self-reported relatedness (Bycroft et al., 2018). The degree of relatedness between the pairs of individuals is based on the combination of the kinship coefficient and genetic similarity in terms of the identity by state (IBS_0) coefficient. IBS_0 measures the fraction of markers for which the related individuals do not share alleles. We follow the KING manual regarding the thresholds for how to determine family relationship (see Table A.1). The identified number of pairs per relationship type differs slightly from that of Bycroft et al. (2017), because some UK Biobank participants withdrew their consent to analyse their data since then.

For our analyses, we go one step further by separating those who are related to the siblings up to the 3rd degree (kinship coefficient ≥ 0.025), i.e., siblings, parents of siblings, cousins of siblings (See Table A.2). In this way, our holdout sample for polygenic score construction and prediction (i.e., the sibling subsample) is unrelated to the GWAS discovery sample which is used to calibrate the SNP weights that are used to construct the polygenic score.

Table A.1. Thresholds used to determine relatedness between individuals in the UK

	Duplicate / Monozygotic twins	1st degree / Parent-child	1st degree siblings	2nd -3rd degree relatives / cousins	Total
Kinship coefficient	>0.3540	0.1770–0.3540	0.1770–0.3540	0.0442–0.1770	
IBS_0		<0.0012	>0.0012		
N (pairs)	179	6,271	22,659	78,038	107,147

Table A.2. Relatedness to the individuals in the siblings' subsample of UK Biobank.

Relationship to siblings	Unrelated to siblings	Full siblings	2nd-3rd relative of siblings	Parent or child of siblings	Total
N (individuals)	91,055	41,498	10,207	4,740	147,500

Notes: Relatedness to siblings is computed based on the relatedness classification as reported in Table A.1.

GWAS. Our tailor-made GWAS is performed using the fastGWA protocol for Genome-wide Complex Trait Analysis (GCTA) developed by Jiang et al. (2019). fastGWA applies mixed linear modelling (MLM) to the genetic data of the UK

Biobank. fastGWA requires the following steps. First, we generate a sparse genetic relatedness matrix (GRM) using the family relatedness file from the UK Biobank based on the KING software output. Next, we perform an MLM-based GWAS using the SNP data, the sparse GRM, the phenotype file and the minor allele frequency (MAF) filter of 0.001. The phenotype file provides the data on individual years of education residualised with respect to birth year, gender, interaction of birth year and gender, batch, and the first 40 principal components (PCs) of the genetic relatedness matrix. For quality control reasons, some individuals were not included in the GWAS.[19] The eventual GWAS discovery sample includes 389,419 individuals.

We further quality control the resulting GWAS summary statistics using EasyQC tool (Winkler et al., 2014) and meta-analyse our tailor-made GWAS weights with the summary statistics from Okbay et al. (2016). We use these for constructing an alternative polygenic score that is used in the robustness analysis (see footnote 16). Meta-analysis is conducted using the software package METAL (Willer, Li, & Abecasis, 2010).

Polygenic scores. The polygenic scores are constructed while accounting for linkage disequilibrium between SNPs using LDpred (Vilhjálmsson et al., 2015), version 1.06, and Python, version 3.6.6. Linkage disequilibrium pertains to the non-random correlations between SNPs at various loci of a single chromosome. LDpred is a software package based on Python that adjusts the GWAS weights for LD using a Bayesian approach. We follow the steps as outlined in Mills, Barban, & Tropf (2020), including the coordination of the base and target files, computing the LD adjusted weights, and then applying them for polygenic score construction using PLINK (Purcell et al., 2007). We re-weight the SNP effects on the basis of LD and the supposed fraction of causal SNPs, which we set to 1, as is standard practice for behavioural traits (Cesarini & Visscher, 2017). Our hold-out sample for constructing polygenic scores consists of 49,866 siblings and their relatives, where the final analysis sample with observations for all variables available is 14,850 individual siblings. The polygenic scores include all SNPs, that is 1,065,078 SNPs after filtering for HapMap3 SNPs at the coordination step. For the split sample GWAS, we first remove all remaining parent-child pairs (N = 5,134) and cousins except one from each cousin cluster (N = 45,099) and split the unrelated discovery sample with all control variables available (N = 340,009) randomly into two samples of 170,005 and 170,004 individuals each and use the same fastGWA procedure as for the full UKB GWAS to obtain SNP weights. The removal of parent-child pairs and cousins ensures that two subsamples do not contain related individuals and are thus independent from each other. We proceed by using LDpred to construct two polygenic scores based on the two sets of summary statistics. Likewise, we include all SNPs (1,065,146 after filtering for HapMap SNPs at the coordination step).

[19] More specifically, we exclude individuals who withdrew consent, have missing gender or whose self-reported gender does not match the genetic sex, are of other than European ancestry, have bad genotyping quality, putative sex chromosome aneuploidy, whose second chromosome karyotypes are different from XX or XY, with outliers in heterozygosity, or have missing information on any of the former criteria.

B. Empirical evidence on complementarity in human capital production

Testing complementarity between children's endowments and parental investments is challenging, since it requires independent variation in initial endowments and later-life investments (Almond & Mazumder, 2013; Johnson & Jackson, 2019). Cunha & Heckman (2007) and Cunha et al. (2010) adopt a structural approach, modelling both skills as well as parental investments as low-dimensional latent variables, and find evidence consistent with (dynamic) complementarity. A number of studies have examined whether the effect of specific interventions or investments varies by initial skills. Aizer & Cunha (2012) correct early life health measures for certain prenatal investments, and find that pre-school enrolment is more productive for children with higher levels of this residualised measure of endowments. Lubotsky & Kaestner (2016) use entrance-age in kindergarten as plausibly exogenous variation in initial cognitive skills, and find some evidence for complementarity, although the effect dies out after the first grade.

A recent set of papers have examined rare cases where there exists exogenous variation in both initial endowments as well as later-life investments. For example, Malamud et al. (2016) study the interaction between exogenous variation in access to better schools and variation in family backgrounds induced by access to abortion in Romania. Their findings do not suggest a meaningful interaction between initial endowments and later-life investments. Rossin-Slater & Wüst (2020) exploit a nurse home visiting program as an exogenous shock to endowments, and staggered access to high quality preschool childcare in Denmark as an exogenous shock to investment, and find that these interventions are substitutes rather than complements. Gunnsteinsson et al. (2014) exploit a unique combination where a tornado struck an area of Bangladesh that was coincidentally involved in a randomized experiment on vitamin A supplementation. Their findings are consistent with complementarity since children treated with Vitamin A supplements were better protected from the consequences of the earthquake. Adhvaryu et al. (2019) exploit local rainfall in the year of birth as exogenous variation in endowments, and randomized cash incentives from Progresa as an exogenous shock to investment. Their main finding is that children from families who received cash transfers were protected better against adverse endowments, consistent with complementarity. Similarly, Duque et al. (2018) also use a combination of adverse weather shocks and conditional cash transfers in Colombia to show that children born under normal weather conditions benefit more from the cash transfers. Finally, Johnson & Jackson (2019) exploit the rollout of Head Start and the implementation of court-ordered school finance reforms (SFRs) that increased spending at public K-12 schools as two exogenous shocks to human capital investment, again finding evidence in favour of complementarity.

C. Obviously-Related Instrumental Variable (ORIV) regression

In this section, we describe Obviously-Related Instrumental Variable (ORIV; Gillen et al., 2019) regression. Suppose we would like to predict an outcome variable of interest, *Y*, using a polygenic score, i.e., estimate the following model:

$$Y = \alpha + \beta PGS^* + \varepsilon, \qquad (C.1)$$

where α is a constant, β is the effect of a true polygenic score PGS^* and ε is the error term. We have two estimates of the true polygenic score: $PGS_1 = PGS^* + \vartheta_1$ and $PGS_2 = PGS^* + \vartheta_2$. The covariance between the two measurement error terms ϑ_1, ϑ_2 is zero, $Cov(\vartheta_1, \vartheta_2) = 0$, and they have the same relative variance of the measurement errors ϑ_1, ϑ_2. That is:

$$\frac{\sigma^2_{\vartheta_1}}{\sigma^2_{PGS_1}} = \frac{\sigma^2_{\vartheta_2}}{\sigma^2_{PGS_2}} = \frac{\sigma^2_{\vartheta}}{\sigma^2_{PGS}}, \qquad (C.2)$$

where $\sigma^2_{\vartheta_1}$ and $\sigma^2_{\vartheta_2}$ are the variances of the measurement errors ϑ_1, ϑ_2 respectively, and $\sigma^2_{PGS_1}$ and $\sigma^2_{PGS_2}$ are the variances of respective polygenic scores. If we use PGS_2 as an instrumental variable for PGS_1, the following applies:

$$plim\ \beta_{IV} = \frac{Cov(Y,PGS_2)/V(PGS_2)}{Cov(PGS_1,PGS_2)/V(PGS_2)} = \frac{Cov(\alpha+\beta PGS^*+\varepsilon, PGS^*+\vartheta_2)}{Cov(PGS^*+\vartheta_1, PGS^*+\vartheta_2)} = \beta \frac{\sigma^2_{PGS^*}}{\sigma^2_{PGS^*}} = \beta. \qquad (C.3)$$

ORIV regression as developed by Gillen et al. (2019) estimates a 'stacked' model:

$$\begin{pmatrix} Y \\ Y \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \beta \begin{pmatrix} PGS_{1+} \\ PGS_{2+} \end{pmatrix} + \varepsilon, \qquad (C.4)$$

where one instruments the stack of estimated polygenic scores $\begin{pmatrix} PGS_{1+} \\ PGS_{2+} \end{pmatrix}$ with $\begin{pmatrix} PGS_{2+} & 0_N \\ 0_N & PGS_{1+} \end{pmatrix}$, where *N* is the sample size and 0_N is a *N*x1 vector with zero's. We include a family-stack fixed effect to conduct the within-family comparisons within a stack of the data. Standard errors are clustered at both the family and individual level following Correia (2017, 2019).

D. Early-life parental investments

The only early-life parental investments observed in the UK Biobank are whether the child was breastfed and whether the mother smoked around birth. We also observe the age gap between subsequent siblings. Table D.1 reports the results of regressions explaining the few early-life parental investment as a function of being firstborn, the polygenic score for education, and their interaction. This shows whether mothers change their behaviour depending on whether the child is first- or laterborn and the polygenic score of their children. The results show that the probability of being breastfed (Column 1) and the likelihood of maternal smoking around pregnancy (Column 2) are similar between first- and laterborns. Furthermore, the coefficient of the polygenic score for (the child's) education is not significantly different from zero, and we find no evidence of any differences in maternal investments around pregnancy by the firstborn's polygenic score.

Next, we explore the relationship between the polygenic score and the age gap between siblings. Column 3 presents the estimates from a regression of the age gap in months between every two consecutive siblings on the polygenic score of the older sibling in the pair. These between-family estimates suggest that birth spacing is one month shorter for every standard deviation increase in the polygenic score of the older sibling. However, note that these results derive from a between-family comparison, and therefore should not be interpreted causally. For families for whom we observe three consecutive siblings or more (N = 2,542), we can check this result in a within-family specification. The within-family estimates (Column 4) are very noisy because of the small sample size, and the wide confidence intervals do not allow drawing any firm conclusions here.

Table D.1. Regressions of early life parental investments on the gene-environment interaction.

	Breastfed	Mother smoked around birth	Age gap	
	With family fixed effects		Without family fixed effects	With family fixed effects
	(1)	(2)	(3)	(4)
Firstborn	0.006	0.006		
	(0.010)	(0.006)		
PGS for years of education	-0.001	0.001		
	(0.007)	(0.004)		
Firstborn × PGS for years of education	-0.011	0.004		
	(0.007)	(0.004)		
PGS for years of education of the older sibling			-0.970***	6.417*
			(0.331)	(3.466)
Constant	0.705***	0.270***	50.444***	-24.447
	(0.067)	(0.040)	(6.934)	(31.458)
R^2	0.042	0.020	0.067	0.376
N	11,818	13,156	6,501	2,542

Notes: Robust standard errors in parentheses, clustered by family; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Coefficients for the control variables (year and month of birth, gender, family size, and the first 40 principal components of the genetic relatedness matrix) are not displayed, but available upon request from the authors. Sample sizes vary depending on the availability of early life parental investments and the number of siblings included in the analyses.

E. Additional confounders

In Table E.1 we present further analyses to investigate robustness of our results against the inclusion of additional confounders. Column 1 reproduces our main results (Table 2.5, Column 5). In Column 2 and 3, we employ the correction for missing confounders in gene-environment interaction analyses suggested by Keller (2014) without (column 2) and with (column 3) corrections for measurement error. Specifically, we interact both the dummy for being firstborn and the polygenic score for education with year of birth, month of birth, gender, and first 40 principal components of the genetic relatedness matrix and include all these interactions as additional control variables in the analysis. The direct effect sizes of being firstborn and the polygenic score for years of education are now relative to the reference categories of the control variables (born in 1937, born in January, being female). The standard error of the interaction term increases due to the much larger number of regressors in both columns. Whereas the magnitude of the interaction term drops somewhat in column 2, the point estimates in both columns 2 and 3 do not significantly differ from our baseline specification in column (1), and show a robust interaction effect, at least after correction for measurement error.

We also explore whether our results are robust against the inclusion of parental age as an additional control, as this may be one mechanism through which birth order effects arise (Eirnæs & Pörtner, 2004). In our sample, we have only 5,601 individuals with information about the age of the mother and 3,085 individuals with information about the age of the father (both measured at time of birth of the child). The inclusion of these variables as control variables in Column 4 and Column 5 respectively of Table E.1. and the subsequent drop in the sample size decreases the statistical power of our tests. Indeed, the interaction term loses statistical significance when controlling for father's age at birth. However, the effect sizes of the interaction term are similar to our main results, and if anything, are larger.

Table E.1. Results of the regressions of years of education on the gene-environment interaction and additional control variables.

	With family fixed effects				
	(1)	(2)	(3)	(4)	(5)
Firstborn	0.415***	-10.488***	1.397	0.199	0.021
	(0.109)	(3.777)	(1.589)	(0.174)	(0.220)
PGS for years of education	0.567***	-4.870***	-3.577**	0.576***	0.528***
	(0.078)	(1.286)	(1.644)	(0.127)	(0.155)
Firstborn × PGS for years of education	0.204**	0.150	0.229*	0.265**	0.235
	(0.080)	(0.092)	(0.132)	(0.125)	(0.160)
Mother's age at birth				0.099	
				(0.088)	
Father's age at birth					-0.028
					(0.148)
Constant	12.544***	25.753***	-53.432	12.376***	6.177***
	(1.838)	(3.019)	(4.503)	(2.976)	(2.278)
R^2	0.059	0.076	-	0.062	0.093
N	14,850	14,850	14,850	5,601	3,085

Notes: Robust standard errors in parentheses, clustered by family; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Coefficients for the control variables (year and month of birth, gender and the first 40 principal components of the genetic relatedness matrix) are not displayed, but available upon request from the authors. Column 1 replicates our main results for comparison. Column 2 includes as additional control variables the interactions between firstborn and the polygenic score with year of birth, month of birth, gender and the first 40 principal components. Column (3) provides the ORIV estimation of the Column (2). Column (4) and (5) include controls for maternal age at birth and paternal age at birth, respectively.

F. Replication of main results with a different polygenic score for educational attainment

Columns 1-3 of Table F.1 replicate the main analyses using the polygenic score based on the meta-analysis of the GWAS summary statistics of our own UK Biobank GWAS described in Section 2.3 and the GWAS summary statistics of 23andMe (Okbay et al., 2016). The procedure for constructing the polygenic score is identical to the one used in the main analysis. From Columns 1-2, we can see that the incremental R^2 of the polygenic scores is 1.1% in the within-family specification. The GxE interaction term (Column 3) is slightly smaller than in the main analysis, but not significantly different from our main result.

Table F.1. Regressions of years of education on the gene-environment interaction with the meta-analyzed polygenic scores (PGS).

	With family fixed effects		
	(1)	(2)	(3)
	OLS	OLS	OLS
PGS for years of education		0.664***	0.616***
		(0.069)	(0.077)
Firstborn			0.436***
			(0.109)
Firstborn × PGS for years of education			0.136*
			(0.079)
Constant	13.757***	13.408***	12.348***
	(1.685)	(1.823)	(1.838)
R^2	0.046	0.057	0.060
N	14,850	14,850	14,850

Notes: Robust standard errors in parentheses, clustered by family; * $p<0.10$, ** $p<0.05$, *** $p<0.01$; Coefficients for the control variables (year and month of birth, gender and the first 40 principal components) are not displayed, but available upon request from the authors; The PGS for years of education in Columns 1-3 is based on a meta-analysis of summary statistics from our own UK Biobank GWAS and 23andMe.

G. *Systematic differences in polygenic scores for other traits by birth order.*

In this section, we test if there are any systematic differences in polygenic scores for anthropometric, health, and personality traits by birth order. Polygenic scores for these analyses are obtained from the Polygenic Index Repository. These polygenic scores are constructed using PLINK 2 (Chang et al., 2015a), where SNPs are corrected for linkage disequilibrium using LDpred (Vilhjálmsson et al., 2015; for further technical details, see Becker, Burik, et al., 2021). Table G.1 reports the results of regressions of the single-trait polygenic scores on a dummy for being firstborn and control variables for year and month of birth, gender, the first 40 principal components of the genetic relatedness matrix, and family fixed effects. All associations are very small in magnitude, and only 2 out of 28 individually reach statistical significance at the 5% level. After a Bonferroni correction for multiple testing (i.e., with a Bonferroni-corrected p-value significance threshold of 0.0018 (=0.05/28)), being firstborn is not significantly associated with any of the selected polygenic scores.

Table G.1. Regressions explaining selected single-trait polygenic scores from the Polygenic Index Repository (Becker et al., 2021) on a dummy for being firstborn within family.

Trait	Coefficient	SE	*p*-value
Anthropometric			
1. Body mass index	0.002	(0.019)	0.903
2. Height	-0.018	(0.018)	0.323
Health and health behaviors			
3. Alcohol misuse	0.050	(0.019)	0.008
4. Asthma	-0.021	(0.019)	0.263
5. Asthma/eczema/rhinitis	-0.016	(0.019)	0.393
6. Attention deficit hyperactivity disorder (ADHD)	0.000	(0.019)	0.985
7. Cannabis use	0.045	(0.019)	0.015
8. Cigarettes per day	0.002	(0.019)	0.904
9. Depressive symptoms	0.004	(0.019)	0.822
10. Drinks per week	0.026	(0.019)	0.165
11. Ever smoker	0.026	(0.019)	0.163
12. Hay fever	0.002	(0.019)	0.936
13. Migraine	0.012	(0.019)	0.520
14. Near-sightedness	0.006	(0.019)	0.730
15. Physical activity	0.039	(0.019)	0.957
16. Self-rated health	0.001	(0.019)	0.044
Personality and well-being			
17. Adventurousness	0.001	(0.019)	0.959
18. Extraversion	-0.009	(0.019)	0.623
19. Left out of social activity	-0.014	(0.019)	0.478
20. Life satisfaction, family	-0.026	(0.019)	0.162
21. Life satisfaction, friends	-0.004	(0.019)	0.840
22. Morning person	-0.000	(0.019)	0.987
23. Narcissism	-0.030	(0.019)	0.121
24. Neuroticism	0.007	(0.019)	0.692
25. Openness	-0.014	(0.019)	0.458
26. Religious attendance	-0.023	(0.019)	0.218
27. Risk tolerance	0.049	(0.019)	0.009
28. Subjective well-being	-0.020	(0.019)	0.273

Notes: N=14,835; Robust standard errors (SEs) in parentheses; The Bonferroni corrected significance threshold is 0.05/28 = 0.0018. Coefficients for the control variables (year and month of birth, gender and the first 40 principal components) are not displayed, but available upon request from the authors.

Chapter 3
Diffusion of the pill and women's education: The role of gene-environment interactions

Abstract

Access to contraception provides women broader opportunities to invest in their education and career. This paper is the first to investigate whether access to oral contraception has differential effects on educational attainment according to one's genetic endowment for education. This is informative of the existence of complementarities in human capital formation. I use the UK Biobank, restricting my attention to 145,502 women, and show that exposure to the pill is associated with more years of education. The positive association of the pill diffusion with years of education is concentrated among women with lower genetic endowment for education. This finding suggests the existence of a compensating mechanism: an environment in which contraception is more widely available was most productive for women with a lower genetic predisposition towards education, reducing inequalities in educational attainment.

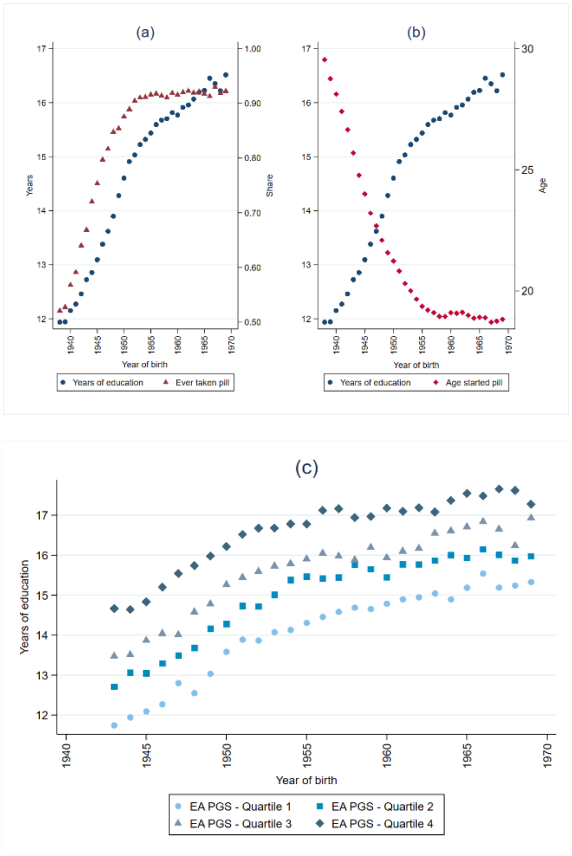
3.1 Introduction

The contraceptive pill has been legalized more than 60 years ago, but there is still a sharp divide in access to contraception across countries. For example, only half of EU member states reimburse the pill[20], raising financial barriers for a lot of women (European Parliament, 2020). In the US, after the court overturned *Roe v. Wade*, which guaranteed women a constitutional right to abortion since 1973 (Guldi, 2008; N. Sun, 2022), some justices called for reconsidering *Griswold, Lawrence, and Obergefell*, which legalized access to contraception (Kolhatkar, 2022). In the fear of this and the increased demand for emergency contraception (Rosman & Cherelus, 2022)[21], the US House passed a bill to protect federal rights for access to contraception, yet, with 96% of Republicans voting against (Sotomayor & Caldwell, 2022). Given the vulnerable state of access to birth control policies, it is pivotal to bring more evidence to understand if access to the birth control pill enhances educational opportunities for young women, as well as which women are most affected by the policy.

In this paper, I first estimate the effect of gaining access to the birth control pill on UK women's education[22]. In the UK, the pill was first introduced in 1961 for married women and only in 1967 for unmarried women, and was made available by the National Health Service (NHS) free of charge. Barban, de Cao, & Francesconi (2021) show that the introduction of the pill in the UK led to changes in fertility decisions, postponing motherhood, a decrease in family size, and an increase in childlessness. Based on the same data as in Barban, de Cao, & Francesconi (2021), Figures 3.1a and 3.1b provide evidence for a strong correlation between the educational attainment of women and the uptake of the pill[23] with a dramatic increase in both for the cohorts born between 1940 and 1960. For women born in early 1940s, the share of those ever taken the pill was around 50 percent, while for the women born in 1960s, this proportion climbed to 90 percent (Figure 3.1a). Likewise, with the expansion of pill access to unmarried women, the average age at first pill came down rapidly and plateaued at 18 years of age for later cohorts (Figure 3.1b). All the while, average years of education surged by more than 4 years comparing the cohorts born during WWII and those born in the 1960s. This observation begs the question: Does this correlation stem from a causal effect of access to the pill on education?

[20] The French government decided to raise the age of women eligible for free contraception and related medical care from 15-18 to everyone under 25 years old due to declining pill take up related to unaffordability (Willsher, 2021).
[21] For example, the demand for a morning after pill, Restart, has increased by 600% after the overturn of Roe vs. Wade, with 75% of people buying more than one dose (Rosman & Cherelus, 2022).
[22] Although the literature on the effects of access to birth control pill on education of women is abundant, the evidence is mainly based on the US data. These studies are discussed briefly later in the paper.
[23] More information on the nationwide historical statistics for education in the UK is available in the report of House of Commons (Bolton, 2012) and for the use of contraception at Statista (Clark, 2022).

Figure 3.1. Cohort trends in pill take up and years of education.



Notes: Cohort of birth on horizontal axes. Panel (a) depicts average years of education attained (left-hand side) and the proportion of women who have ever taken the pill (right-hand side) against the birth cohorts. Panel (b) shows the average years of education attained (left-hand side) and the average age at starting the pill (right hand side) against birth cohort. Panel (c) shows the average years of education attained by above and below mean EA PGS (23andMe-based) against birth cohort. Source: author's calculations based on the UK Biobank data, 242,128 observations.

I address this question using data on over 145,000 women from the UK Biobank (Bycroft et al., 2018). This is a unique dataset that covers the period of the initial diffusion of the pill in the United Kingdom and has information on genetic, socio-economic and health outcomes of women in older age. The data allow me to observe women who turned 18 up to 20 years after the first legalization of the pill in 1961. Diffusion of the pill is measured following the approach of Barban et al. (2021), exploiting the plausibly exogenous region-by-birth-cohort variation in access to the

pill in the UK between 1960-1970. More specifically, the pill diffusion variable is computed as the proportion of other (i.e., excluding the respondent herself) childless women taking the pill at the age of 18 and above in the administrative county of birth when the respondent was 18. Hence, by construction, the measure of pill diffusion excludes own pill take up behaviour of the women and focuses on the exposure effects of other women taking the pill in the geographical area in which they were born.

A second main question addressed in this paper is whether the effect of pill diffusion on educational attainment differs across women with higher and lower genetic endowments for education. The genetic endowments are measured using a polygenic score for years of education (EA PGS) – PGSs are indices constructed from the sum of all measured genetic variants (Single Nucleotide Polymorphisms – SNPs), weighted by the strength of the association between these variants and the outcome of interest in an external sample (Dudbridge, 2013). They can be interpreted as the best linear genetic predictor of a certain outcome. Figure 3.1c shows how the cohort trends in years of education of UKB women unfolded depending on their EA PGS quartile. The initial gap of 3.5 years between the highest and the lowest PGS quartiles observed among those born in 1940s narrowed to 3 years for women born in late 1960s. Whereas one of the explanations for this convergence is the raising of school leaving age (RoSLA) introduced in the UK in 1972 affecting those born in the late 1950s (Barcellos et al., 2018, 2021), Figure 3.1c shows converging patterns for women born as early as the 1950s, well before the cohorts affected by RoSLA 1972 and after the cohorts affected by the RoSLA in 1947, leaving access to contraception as another potential explanation for the reduction in the gap across women with higher and lower polygenic scores.

I believe addressing the second question is meaningful for at least three reasons. First, exploring how the effect of a sudden change in environments varies by genetic endowments provides an advance in our fundamental understanding of how the interplay between genes and environments shapes important life outcomes. Indeed, Barban et al. (2021) find that the effect of the pill diffusion meaningfully interacted with the genetic propensity for fertility, and here I extend their work to explore whether the effect of pill diffusion also interacted with genetic propensity for education in shaping human capital outcomes. Second, genetic endowments are fixed at conception and even though they are firmly established to meaningfully affect life outcomes (e.g., Okbay et al. 2022), they are beyond the control of an individual. Hence, arguably, genetic endowments provide a source of advantage that constitutes inequality of opportunity (Barcellos et al. 2022) and exploring whether pill diffusion increased or decreased genetic inequalities is therefore relevant for understanding how access to birth control enhances or aggravates inequality of opportunity. Finally, enhanced access to birth control is associated with an improvement in women's opportunities to invest in their human capital. When conceptualizing the channels through which the pill affects the income of women, Bailey, Hershbein, & Miller (2012) hypothesized a "formal human-capital investment mechanism" operating through more schooling, enrolment in longer educational programmes, and higher

fraction of college education due to the pill (Ananat & Hungerman, 2012; Goldin & Katz, 2002). Hence, studying the interaction between genetic endowments and these enhanced opportunities is a way to test the theoretical hypothesis that endowments and conducive environments to human capital investments are complementary (e.g., Becker & Tomes, 1986).

The results suggest that pill diffusion affects the educational attainment of women. For every 10-percentage point increase in the pill diffusion, a woman is estimated to gain 0.08 years of education. In other words, when the environment changes from no access to the pill to full access, a woman gains 0.8 additional years of education on average. Further, there are important heterogeneities in the effect of the pill diffusion: the positive association of the pill diffusion with years of education is concentrated among women with lower genetic endowments for years of education. This finding is consistent with the narrowing gap in years of education we observed in Figure 3.1c and suggests a compensating mechanism between genetic endowments and conducive environments in terms of human capital formation.

The paper contributes to the following key strands of the literature. The first strand concerns studies on the power of the pill. Goldin & Katz (2002) show that the access to the contraceptive pill increased the age at first marriage, decreased the likelihood of having first children by the age of 23 (see also Guldi, 2008), lowered the opportunity cost of obtaining higher education (see also Ananat & Hungerman, 2012; Bailey, 2006; Bailey et al., 2012) and increased the share of women in traditionally non-female professions such as doctors and lawyers in the US (see also Steingrimsdottir, 2016). Further, the pill increased labour force participation of women, number of hours worked and led to a convergence in the gender gap (Bailey, 2006; Bailey et al., 2012), More recently, Marie & Zwiers (2021) confirm a decrease in untimely births and early marriages, and an increase in the number of women who completed their education, worked, and were better off financially.[24]

Meanwhile, the heterogeneities in the effect of contraceptive policies on women have been somewhat under the radar. In their additional analyses, Bailey et al. (2012) show that women with middle and higher IQ were affected more by access to the pill in terms of education, labour market outcomes, and presence in more male dominated occupations. Similarly, Steingrimsdottir (2016) finds that the pill affected high ability women by allowing them to switch to higher paying and male dominated jobs. I contribute to this literature by revisiting these heterogeneities in the effect of the pill on women's educational attainment using polygenic scores for education, which are strong predictors of education and intelligence (Okbay et al., 2022), and are fixed at conception.

[24] An additional body of work focusing on the intergenerational effects of the pill finds that over the long-term, the pill diffusion leads to an increase in the share of children with highly educated and non-single mothers (Ananat & Hungerman, 2012). A recent study by Myers (2022) questions the results of Bailey (2006), Bailey et al. (2012), and Goldin & Katz (2002), however, further evidence (e.g. Aleman et al., 2022) suggests there remains considerable evidence that birth control policies have strong effects on socio-economic outcomes of women (Myers, 2017, 2022).

The second main contribution is to the literature on gene-environment interplay. This paper is closest to the studies on the interplay between exogenous socio-economic shocks and polygenic scores in shaping educational attainment.[25] The two most closely related studies are those by Barban et al. (2021) and Muslimova et al. (2021). Barban et al. (2021) show how the effect of access to the contraceptive pill on long-term fertility outcomes of women, such as sexual debut, family size, childlessness, and age at motherhood, are moderated by the polygenic scores for the respective outcomes. Like Barban et al. (2021), I exploit the regional diffusion of the pill in the UK Biobank. While Barban et al. (2021) zoom in on fertility outcomes and genetic drivers, I focus on human capital effects of pill diffusion and how these effects differ across the PGS distribution. Muslimova et al. (2021) find evidence for complementarities between genetic endowments for education and parental investments proxied by birth order using UKB data. While Muslimova et al. (2021) focus on early life investments, in this paper I investigate whether there are complementarities between genetic endowments for education and the environmental *opportunity for* human capital investments in late adolescence triggered by pill diffusion in one's local area.

The rest of the paper is organized as follows. I begin with an overview of theories that inform my hypotheses. Then I turn my attention to the empirical methods and data. Next, I expand on the results. Lastly, I discuss the implications of the results and possible future directions of this research. Additional analyses are presented in the Appendices.

3.2 Theoretical background

In the framework provided by Goldin & Katz (2002), the effect of the contraceptive pill on education and career outcomes of women works via direct and indirect routes. The direct route implies reduction in the cost of marriage delay and career investment, making women with greater career prospects more attractive potential partners. In the indirect effect, Goldin & Katz (2002) account for social multiplier effect happening due to a thickening marriage market and the resulting improvement in the matches for career oriented women. So, Goldin & Katz (2002)'s theory predicts that the introduction of the pill increases the likelihood of women to engage in professional careers, postpones marriage and age at first birth. The pill was also expected to increase assortative mating on earning's potential of partners and compatibility. In addition to the fertility and marriage market channels, Bailey (2006) argues that even if the pill did not have an effect on completed fertility, it did provide a costless instrument in timing for birth and thus the labour supply of women. An unplanned pregnancy would disrupt human capital investments and make women's

[25] For a history and review of existing gene-environment interplay studies in social sciences, see Schmitz & Conley (2017), Mills et al. (2020), and Pereira et al. (2022).

labour force participation and career investments highly uncertain (Bailey, 2006; Goldin & Katz, 2002).[26] Hence, theory strongly suggests that access to contraception would unambiguously increase educational attainment.

The social genomics literature suggests some theoretical explanations for why there could be *an interaction* between the exposure to the pill and the genetic endowments for education. As summarized in Mills et al. (2020), the diathesis-stress theory states that genetic differences associated with risky behaviours are expressed more strongly in high-risk environments, while being muted or attenuated in low-risk environments. In some sense, before the introduction of the pill, the environment was 'risky' since risky sexual behaviours more often led to unplanned pregnancies. The diathesis-stress theory therefore predicts a reduction in genetic differences in fertility after the introduction of the pill, and thereby potentially a reduction of genetic inequalities in educational attainment via the fertility channel. Alternatively, so-called social compensation theory states that genetic differences associated with positive outcomes are most pronounced in positive stable environments (Bronfenbrenner & Ceci, 1994; Bronfenbrenner & Morris, 2007). The introduction of the contraceptive pill in 1961 in the UK decreased the likelihood of unplanned motherhood and provided women broader opportunities to invest in their education, hence, creating an environment that is more conducive to realizing their educational potential. This implies that according to the social compensation theory, genetic effects on education could be more pronounced for women with more exposure to the pill.

Economic theory can also be used to hypothesize on the expected sign of the gene-environment interaction estimates. Gene-environment interaction studies between genetic propensity for education and policies/environments promoting investment in education provide a way to test for complementarities in human capital formation. For example, the assumption of "Ben-Porath neutrality" essentially imposes that the stock of human capital raises the productivity of investments in human capital (Ben-Porath, 1967; Heckman, 1976; Rosen, 1976). Hence, women with higher (genetic) endowments for education would benefit more from delaying fertility and investing in their education. Bailey et al. (2012) provide early evidence for this potential complementarity between pill diffusion and ability by finding that women of average and middle intelligence scores benefited more from the introduction of the pill in terms of increased education and improved labour market outcomes.

Complementarity in skill formation is closely related to the concept of dynamic complementarity, where skills produced at an earlier stage raise the productivity of future investments (Cunha & Heckman, 2007). While this paper does not look at investments at different points during a woman's life cycle, Galama & Van Kippersluis

[26] Ananat & Hungerman (2012) focus on intergenerational effects of the pill diffusion and hypothesize that the effect on the following generation depends on whether women just delay having children or not have them at all. They further refer to the framework of Akerlof, Yellen, & Katz (1996), who argue that the introduction of the pill might have led to a decline in shotgun marriages, more single parents, better matched partners and hence fewer divorces.

(2022) distinguish between two relevant aspects of dynamic complementarity: 1) *investments are more productive when the stock of skills is higher*, 2) *individuals optimally choose to invest more when their stock of skills is higher*. The first aspect is interpreting complementarity as a property of the production function and would suggest that women with a higher polygenic score for education would benefit more from the additional investments enabled by the introduction of the pill. The second feature is informative of the endogenous response to access to the contraceptive pill given the educational endowments of women, i.e., whether endowments affect the extent to which women actively seek to take up the pill and be exposed to areas with higher pill diffusion. So, the second feature would imply that women's endowments affect their decisions to invest in their human capital by taking the pill and delaying fertility. In the social science genetics literature, such an endogenous response is known as "active gene-environment correlation (active rGE)", which biases the gene-environment interaction estimates (Biroli, Galama, von Hinke, et al., 2022) and lead to collider bias (Akimova et al., 2021). Hence, gene-environment correlations are not only informative for features of economic theories of complementarity, but they are also very important for interpretation of the gene-environment interaction estimates.

In sum, social science theories unambiguously predict a positive effect of the pill introduction on educational attainment, yet, are conflicting on whether the effect would be stronger or weaker among individuals with higher polygenic indices for educational attainment. Moreover, economic theory also presents a cautionary tale that individuals with higher genetic predisposition may actively seek out environments conducive to their educational development, including potentially the contraceptive environment. In the empirical analyses, I therefore analyse both a possible G × E interaction on educational attainment, as well as the possibility of rGE where genetic endowments would be correlated to pill diffusion.

3.3 Data

I use the UK Biobank (UKB), a population-based genotyped sample of 502,488 individuals from the United Kingdom (Sudlow et al., 2015). Of 9.2 million invited participants aged 40-69 between 2006-2010, 5.5% eventually went through the assessment (Fry et al., 2017). Fry et al. (2017) show that participants are not representative of the UK population, as they tend to be older, more often female, and live in areas of higher socio-economic status. The analysis sample and the main variables are discussed below.

There are 273,375 women in the UK Biobank, with 242,128 who pass the quality control protocol[27] and consent to their data being used for research. After accounting for availability of the data on pill take up, 145,502 women remain. I restrict my attention to this sample.

Pill diffusion. The environmental component, the *pill diffusion* variable, is measured by early exposure to the contraceptive pill as in Barban et al. (2021) exploiting the rapid and largely unexpected change in access to the birth control pill introduced in the UK between 1960-1970 as well as the regional variation in the take up of the pill.[28] The construction of the variable is based on the following data fields in the UK Biobank: n_2794 "Age started taking oral contraceptive pill", n_2784 "Ever taken oral contraceptive pill".[29] Combining these data fields with the age at first birth I compute the proportion of other childless women taking the pill at childbearing age (18-45) in the local area of birth when the respondent was 18 excluding the respondent herself. Using individuals' north and east coordinates of birth, I merge in geographic identifiers for individuals' administrative county of birth using the 1951 shapefiles from Vision of Britain (Baker, 2022; Great Britain Historical GIS Project, 2017). There are 271 administrative counties in England and Wales, allowing me to construct local measures of pill diffusion for each woman in the UK Biobank, based on her administrative county and cohort of birth, whilst simultaneously ensuring there is a sufficient number of other childless women of childbearing age in her local area.[30] The average size of an administrative county in the UKB is 691 square kilometres (standard deviation = 1,125).

Polygenic scores. My measure of genetic differences between women is a polygenic score. A polygenic score is a weighted sum of genetic variants, Single Nucleotide Polymorphisms (SNPs, see Appendix A for details). The SNP weights are obtained from the association between SNPs and the traits of interest from Genome-Wide Association Studies (GWAS) in an independent (discovery) sample and are used as in Equation 1 to construct the polygenic score:

[27] The quality control protocol includes verifying the self-reported gender against the genetically identified one, removing individuals of non-European background based on self-reported ethnic background and principal component 1 of genetic relatedness matrix, removing individuals with bad genotype data, and removing individuals with putative sex chromosome aneuploidy (i.e., individuals carrying sex chromosome configurations that are not either XX or XY).

[28] The variation in the take up could result from the marriage status, religiosity and related social norms (Bailey et al., 2012; Marie & Zwiers, 2021). Unfortunately, there is no data on marriage status of women at the time of the pill introduction or religiosity in the UK Biobank to identify the exact sources of variation in the pill take up.

[29] Out of 273,382 women in the UKB, 60,839 (22%) women, mostly from older cohorts with slightly lower educational attainment, either did not know the answer or preferred not to answer the questions related to the pill take up. I classify these answers as missing.

[30] I also observe individuals' *district* of birth, however, with over 1400 districts in England and Wales at the time, several do not have a sufficiently large number of women for each birth cohort to approximate regional pill diffusion. Furthermore, since I only observe birth coordinates (and not coordinates of where the women lived at age 18), using a slightly larger area of birth is more robust to local residential moves as well as measurement error in reporting of birth location, since the latter is recorded at one kilometre resolution.

$$PGS_i = \sum_{j=1}^{J} \beta_j x_{ij}, \quad (1)$$

where PGS_i is the value for the polygenic score for individual *i,* β_j is the regression coefficient of SNP *j (j=1, ..., J)* from a respective GWAS (see below for details), and x_{ij} is the genotype of individual *i* for SNP *j,* which is coded as 0, 1, or 2, indicating the number of effect alleles[31]. Since SNPs that are close to each other on a certain chromosome tend to be inherited together (i.e., they are in 'linkage disequilibrium'), the univariate regression coefficient in the GWAS will be biased. Therefore, once the univariate weights are available, the LDpred software (Vilhjálmsson et al., 2015) is used to correct for the correlation structure across the SNPs. To ease the interpretation, the scores are standardized with mean zero and standard deviation 1.

Importantly, the polygenic scores do not reflect an immutable biological relationship, but they represent the best linear genetic predictors of an outcome within the environmental and demographic context of the GWAS sample (Mills et al., 2020). Moreover, a polygenic score may not even solely reflect genetic effects. After all, a GWAS typically does not control for parental genotype, and so a certain genetic variant of the child may simply proxy for a conducive environment shaped by parental genotype, known as *genetic nurture* (Biroli, Galama, Von Hinke, et al., 2022; Kong et al., 2018). In empirical analyses, estimated coefficients for polygenic scores without family fixed effects or without controlling for parental genotypes therefore reflect both direct effects, driven by alleles transmitted from parents, and indirect genetic effects, stemming from the genetic variants in parents that affect their children's outcomes (Kong et al., 2018).

I use the EA polygenic score based on the 23andMe summary statistics (Lee et al., 2018).While this generates a predictive score avoiding any overlap with the UKB sample, the limitation of the score is that the demographic and social background of the discovery sample is different than that of the UKB. Specifically, the EA3 23andMe summary statistics are mostly based on US population of European ancestry, born between 1901 and 1985 (Mean=1961). By comparison, the UKB is based on the UK population, born between 1934 and 1970 (Mean=1951). The share of women in both cohorts is the same (see Supplementary Table 16 in Lee et al. (2018)).[32]

To explore if there is gene-environment correlation with genetic propensities for other cognition related traits, I also use the polygenic scores for the sibling subsample

[31] The effect allele, nucleotide of a SNP, is the allele to which the effect estimate refers (Wootton & Sallis, 2020).

[32] In complementary within-family (Appendix D) and Obviously Related Instrumental Variable (ORIV) (Appendix **Error! Reference source not found.**) analyses, I use the polygenic scores for years of education based on the meta-analysis of the 23andMe summary statistics (Lee et al., 2018) and the UKB discovery cohort, where siblings and their relatives (identified based on the genetic data) were preserved as a holdout sample (Muslimova et al., 2021). The resulting limitation of this score is that it leads to a substantially smaller analysis sample. However, it has important advantages. First, the discovery and holdout sample are socially and demographically similar. Secondly, with the direct access to the individual genetic data, I can randomly split the discovery sample into two independent subsample, conduct the GWAS, and create two independent scores to use ORIV to account for measurement error in the polygenic score (Muslimova et al., 2021; Van Kippersluis et al., 2022).

of the UKB from the Polygenic Score Repository (Becker, Burik, et al., 2021). All GWASs are based both on the sample of men and women, exclude the 23rd pair of chromosomes determining sex from the analysis, and routinely include sex as a control variable.[33]

Education. We follow the literature (see e.g., Lee et al., 2018; Okbay et al., 2016; Rietveld et al., 2013) and convert individuals' qualifications to equivalent years of education using the International Standard Classification of Education (ISCED). *Years of education* ranges from 7 to 20, where College or University degree is equivalent to 20 years, National Vocational Qualification (NVQ), Higher National Diploma (HND), or Higher National Certificate (HNC) to 19 years, other professional qualifications to 15 years, having an A or AS levels or similar to 13 years, O levels, (General) Certificate of Secondary Education ((G)CSE) to 10 years, and if none of the above to the lowest level of 7 years.

Covariates. In all baseline specifications, I control for month of birth dummies, year of birth and year of birth squared, and family size computed based on the number of siblings the respondent has plus herself. To control for population stratification such as correlations between allele frequencies and environmental factors across subpopulations in the sample, I use 40 first principal components of the genetic relatedness matrix[34] (Price et al., 2006; Rietveld, Conley, Eriksson, Esko, Medland, Vinkhuyzen, Yang, Boardman, et al., 2014). I use administrative county IDs defined earlier as geographic fixed effects.

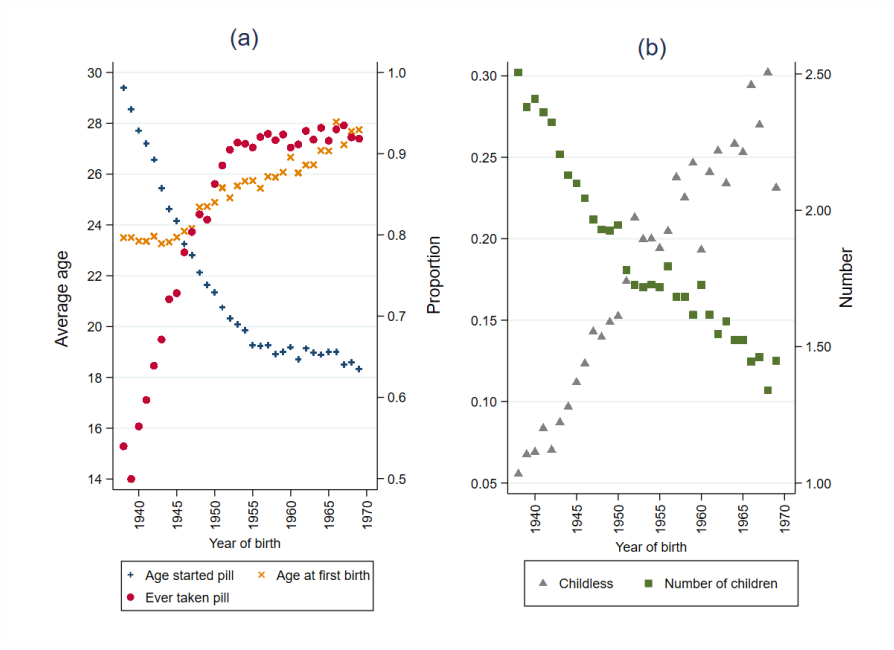
Cohort trends. While available for married women since 1961, unmarried women got access to the pill only in 1967 in England and Wales, and 1968 in Scotland. Figure 3.2 shows how the pill revolution drove the remarkable change in the environment young women faced in the UK over birth years between 1940-1970 using the data from the UKB. Figure 3.2 compares the trends in the pill take-up with the average age at which women had their first child in the analysis sample. While the share of women who ever took the pill at any age has been increasing to reach 90% for later cohorts, women were changing their behaviour by taking the pill earlier in their lives and delaying motherhood (Panel A). Meanwhile, the average number of live births declined over the cohorts and childlessness increased (Panel B). This is comparable to the nationwide trend in the crude birth rates in the UK for the same period (Clark, 2022). Moreover, while educational attainment was generally increasing for cohorts covered by the UKB, Figure 3.3 provides evidence for a clear gap in years of education between men and women for the first UKB cohorts. For example, women born in 1938 had one year less of schooling compared to men on average. This gap started

[33] While one might question if GWAS based on a training sample of women only would provide better weights, Okbay et al. (2022) show that the genetic correlation between sex-stratified GWASs is almost one. Besides, limiting the discovery sample to women only would reduce the sample size by half and, hence, the power of the GWAS.

[34] Genetic principal components (PCs) are conventionally used to control for population stratification (i.e., correlations between allele frequencies and environmental factors across subpopulations in the sample) in a population-level (e.g. between family) analysis (Price et al., 2006; Rietveld, Conley, Eriksson, Esko, Medland, Vinkhuyzen, Yang, Boardman, et al., 2014).

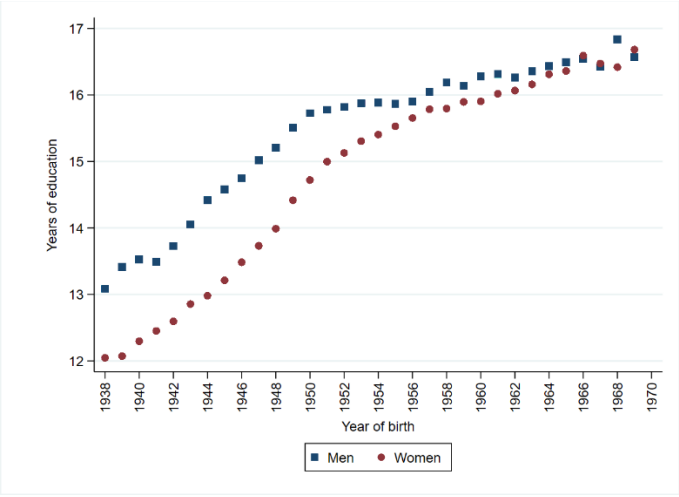
narrowing for women born from 1950s onwards, with the latest cohort of UKB women having the same level of education on average as men. Here I investigate whether this narrowing of the gap is driven by the liberalization of contraception.

Figure 3.2. Cohort trends in pill take up and fertility in the UK Biobank.



Notes: Panel (a) depicts average ages at first pill and first birth (left-hand side axis) against the proportion of women who have ever taken the pill (right-hand side axis) over the cohorts of birth. Panel (b) depicts the proportion of childless women (left-hand side) and the average number of children (right-hand side) over the same birth cohorts. Source: author's calculations based on the UK Biobank data, 145,502 observations.

Figure 3.3. Cohort trends in completed years of education for women and men in the UKB.



Notes: Average years of education per year of birth computed for each gender. Year of birth on horizontal axis. Source: author's calculations based on the UK Biobank data, based on 446,254 participants, 204,126 men and 242,128 women who passed the quality control protocol.

3.4 Empirical approach

Effect of the pill on education

To estimate the effect of the pill diffusion on educational outcomes of women, I exploit the administrative county-by-birth cohort variation in the pill diffusion as in the approach of Goldin & Katz (2002) and Barban et al. (2021). Thus, I consider the following general specification in Equation 2:

$$Y_{ic} = \alpha_1 + \alpha_2 E_{ic} + \boldsymbol{X}'_{\boldsymbol{ij}}\boldsymbol{\theta} + \zeta_c + \nu_{ic}, \quad (2)$$

where for individual *i* from an administrative county *c*, Y_{ic} is the educational outcome of interest (i.e., *years of education*), E_{ic} is the measure of environment (*pill diffusion*), $\boldsymbol{X}'_{\boldsymbol{ic}}$ includes a vector of controls for month of birth dummies, year of birth and year of birth squared, family size, and the first 40 principal components of genetic relatedness matrix to control for population stratification. ζ_c is the administrative county fixed effect, and ν_{ic} is the error term clustered at the administrative county level. Parameter α_2 captures the effect of the exposure to the pill diffusion on years

of education conditional on abovementioned controls. Here, pill diffusion ranges from 0 to 1.

Note that the effect of the pill diffusion on educational attainment is a reduced-form effect, with at least two possible channels through which pill diffusion might affect women's education. Firstly, pill diffusion might increase women's own pill take up (i.e., a quasi-first stage) and change fertility behaviour, thereby enabling them to stay longer in school. Secondly, with more women around taking the pill and staying longer in school, the indirect effect of the pill diffusion might be reflected in peer effects and societal shift towards better educated women. The main analysis will not distinguish between these channels, but I will explore them in later sections.

Identification

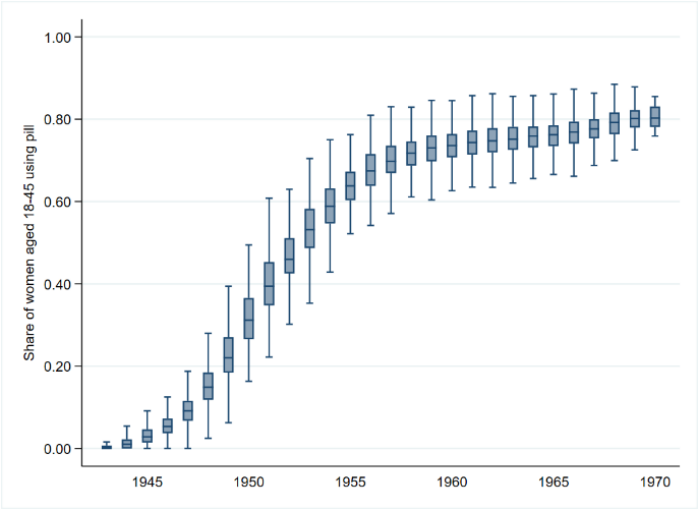
For parameter α_2 to capture the effect of the pill diffusion, there should be sufficient variation in pill diffusion within the cells defined by the control variables (i.e., conditional on the control variables). Furthermore, conditional on the control variables, pill diffusion should be as good as exogenous. I will discuss each of these in turn.

Figure 3.4 presents a box plot showing the geographic variation in pill diffusion across administrative counties in England and Wales over time. This illustrates two main issues. First, it shows the extent of variation in pill diffusion across administrative counties in England and Wales. The variation is particularly salient in the middle of the birth cohort distribution, for women who would turn 18 around the legalization of access to the pill in 1960s, i.e., women born in the late 1940s and throughout the 1950s. Second, it documents a substantial increase in pill diffusion among women of childbearing age over the cohorts present in the UK Biobank, with an increase in the share of women taking the pill from 0 to 80%.

However, the figure suggests that the increase in the share is highly correlated to year of birth. Table 3.1 shows to what extent variation in pill diffusion is absorbed by each control variable. The variation is measured by the standard deviation of pill diffusion gradually residualised for the control variables shown in the second half of Table 3.1. Initial variation in the pill diffusion is almost unchanged when residualised for family size (column 2), month of birth (column 3), and the 40 first principal components of the genetic relatedness matrix (column 4). However, when residualizing for individual year of birth dummies, the standard deviation of pill diffusion is substantially reduced, from 0.285 to 0.036 (Column 5). This suggests that including individual year of birth dummies alongside a full set of administrative county of birth dummies (i.e., a two-way fixed effects specification) will leave very little variation in the treatment of interest. To strike a balance between preserving sufficient variation in the pill diffusion, yet sufficiently controlling for the cohort effects, I control for

quadratic in individuals' year of birth.[35] Although this approach also leads to a reduction in the variation of pill diffusion, it still leaves double the variation in pill diffusion within the administrative county of birth as compared to using the year of birth dummies. I explore the robustness of my results to different specifications for year of birth in sensitivity analyses below.

Figure 3.4. Administrative county-level variation in pill diffusion by birth cohort.



Notes: Share of women of childbearing age (18-45) taking the pill by year of birth. Administrative counties (271 in total) are identified based on the eastings and northings of birth location (Baker, 2022). Source: author's calculations based on the UK Biobank data, 145,502 observations.

[35] Since including year of birth as dummies absorbs a substantial amount of variation in pill diffusion, I do not use them as would be desirable by the conventional two-way fixed effects framework.

Table 3.1. Analysis of variation in pill diffusion remaining after controls.

	Pill Diffusion					
	(1)	(2)	(3)	(4)	(5)	(6)
Within administrative county standard deviation	0.287	0.286	0.286	0.285	0.036	0.061
N	145,502	145,502	145,502	145,502	145,502	145,502
Controls						
Family size		✓	✓	✓	✓	✓
Month of birth			✓	✓	✓	✓
40 PCs				✓	✓	✓
Birth year fixed effects					✓	
YoB & YoB²						✓

Notes: Standard deviation is the within administrative county variation in pill diffusion.

Table 3.2. Results of the regressions of own pill take up at age 18 and fertility outcomes on pill diffusion

	(1)	(2)	(3)	(4)	(5)
	Took pill by 18	Age at first birth	First birth by 21	Number of live births	Childlessness
Pill diffusion	0.343***	0.149	-0.063***	0.155***	0.011
	(0.014)	(0.261)	(0.012)	(0.049)	(0.017)
R^2	0.057	0.064	0.054	0.057	0.024
N	145,501	97,801	145,501	145,501	145,501
Outcome mean	0.139	26.116	0.062	1.743	0.182
Outcome SD	0.346	4.282	0.241	1.088	0.386
Controls					
Constant	✓	✓	✓	✓	✓
Administrative county of birth FEs	✓	✓	✓	✓	✓
Family size	✓	✓	✓	✓	✓
Month of birth	✓	✓	✓	✓	✓
40 PCs	✓	✓	✓	✓	✓
YoB & YoB²	✓	✓	✓	✓	✓

Notes: Robust standard errors clustered at the administrative county level in parentheses; * p<0.10, ** p<0.05, *** p<0.01. Coefficients for the control variables are not displayed, but available upon request from the authors.

Table 3.2 reports the effects of the pill exposure on a woman's own pill take up at age 18 and her fertility outcomes. The core set of controls used throughout the analyses are included. I observe that women who experienced more exposure to the contraceptive pill were more likely to start taking the pill themselves at age 18 ($p<0.01$), were less likely to have their child before the age of 21 ($p<0.01$). Surprisingly, these women tend to have more children. First birth before the age of 21 occurred among 6.21% of women in the analysis sample. The results are inconclusive with respect to the childlessness and age at first birth, although the direction of association is in line with the previous literature on the effect of the pill on decreasing birth rates and delay in motherhood. Overall, these analyses suggest that including year of birth in a quadratic way leaves sufficient variation in our treatment of interest and produces plausible results in a quasi-first stage of pill diffusion on fertility.

The second identifying assumption is that pill diffusion is as good as random conditional on the controls, in particular administrative county of birth and a quadratic polynomial in year of birth. This assumption would be violated if the pill diffusion is correlated to other concurring area-specific trends in educational attainment, not sufficiently picked up by the quadratic polynomial in year of birth and administrative county of birth fixed effects. One such highly non-linear effect is presented by the Raising of School Leaving Age (RoSLA) reform that affected cohorts born after September 1957. In the sensitivity analysis, I show that the results are not affected by including a dummy for being affected by the RoSLA reform. There could be other area-specific trends in education correlated with pill diffusion. This is why in the sensitivity analyses, I use broader 5-year cohort fixed effects (as in Barban et al., 2021) as well as 3-year cohort fixed effects to show that the results are robust to these different specifications. Finally, a threat to identification could be if individuals with better education prospects self-select into regions with higher pill diffusion. Exploiting the genetic data in the UK Biobank enables a unique opportunity to directly assess this threat to identification; in Section 3.5 and Appendix B, I show that pill diffusion is not related to any genetic markers available in the UKB.

Gene-Environment Interplay (G×E)

To explore gene-environment interplay, I apply the gene-environment interaction framework (Barban et al., 2021; Biroli, Galama, von Hinke, et al., 2022; Muslimova et al., 2021) with administrative-county-by-birth cohort variation in the exposure to the pill following Goldin & Katz (2002). Basically, I expand Equation 2 into Equation 3 below.

$$Y_{ic} = \gamma_1 + \gamma_2 G_{ic} + \gamma_3 E_{ic} + \gamma_4 (G_{ic} \times E_{ic}) + \boldsymbol{X}'_{\boldsymbol{ic}}\boldsymbol{\theta} + \zeta_c + \nu_{ic}, \quad (3)$$

where the only difference is the main effect of the polygenic score (G_{ic}) as well as its interaction with the pill diffusion (E_{ic}). The coefficient on the latter (γ_4) shows the

extent to which the polygenic score and exposure to pill diffusion modify each other's effect on education.

Identifying assumptions. For the parameter γ_4 to be identified we need exogenous variation in G_{ij} and E_{ij}, and for E_{ij} to be independent of G_{ij}. The identification of the effect of the pill diffusion on educational attainment (parameter γ_3) is discussed extensively in the earlier section on *Pill diffusion and education of women.* Next, for genes to be interpreted causally, one needs to either employ family fixed effects or control for parental genes, as genes are randomly distributed only conditional on the parent's genome. When parental genes are not accounted for, there could be bias arising from genetic nurture effects (Kong et al., 2018) since parental genes might correlate with the environment in which the children grow up. While the UK Biobank has a siblings' subsample available, using family fixed effects on top of controlling for administrative county of birth would substantially limit the sample size, as well as the variation in pill diffusion, deriving it only from within-sister differences in pill diffusion. Therefore, I report the baseline specification without family fixed effects, interpreting the polygenic scores as an advantage in years of education predicted by the genes of women. However, the caveat remains that the polygenic scores in this baseline specification cannot be interpreted as causal with respect to years of education and can still reflect genetic nurture. I discuss the comparison between baseline results to family fixed effects specification in Section 3.6.

Gene-Environment Correlations (rGE)

The presence of *rGE* would not only reduce the variation in genetic endowments at any point in individuals' environments (Arold et al., 2022), but also make the interpretation of the interaction unclear. To investigate gene-environment correlations and their implication for the *G*×*E* results, I employ the following generic specification in Equation 4:

$$E_{ic} = \beta_1 + \beta_2 G_{ic} + \boldsymbol{X'_{ic}\mu} + \delta_c + u_{ic}, \quad (4)$$

where for an individual i from an administrative county c, E_{ic} is the measure of environment (*pill diffusion*), G_{ic} is the measure of genetic endowments (*polygenic score for a trait*), and $\boldsymbol{X'_{ic}}$ includes a vector of controls for month, year of birth and year of birth squared, family size, and the first 40 principal components of genetic relatedness matrix. δ_c is the administrative county of birth fixed effect, and u_{ic} is the error term. I test for rGE using a variety of PGSs, including those for anthropometric, health and health behaviour, education and cognition, and personality related traits. I use heteroskedasticity-robust standard errors, clustered at the administrative county level.

3.5 Results

Descriptive Statistics

Table 3.3 compares women born just ten years apart but facing completely different reproductive policy environments. For a woman born in 1945, when she turned 18 in 1963, only 1 percent of other childless women in her district of birth were taking the pill. 10 years forward and a woman born in 1955 reaches her adolescence in a much more liberal environment, with 36 percent of women in her birth district taking the pill. Table 3.3 also shows noticeable differences in fertility, and even more remarkable differences in educational attainment between these two cohorts. Women gained two years on average in education, more women were obtaining university degrees and different qualifications, and the proportion of women without any qualifications dropped substantially, from 25 to 9 percent.

Table 3.3. Comparison of women before and after diffusion of the contraceptive pill.

	Born in 1945	Born in 1955	Difference	p-value	Full sample
Education					
Years of education	13.51	15.61	2.10	0.0000	15.05
University degree	0.25	0.37	0.12	0.0000	0.33
A-level	0.20	0.34	0.14	0.0000	0.30
GCSE	0.45	0.56	0.11	0.0000	0.54
CSE	0.04	0.19	0.16	0.0000	0.17
Vocational	0.10	0.18	0.08	0.0000	0.17
Professional	0.29	0.35	0.06	0.0000	0.32
No qualifications	0.25	0.09	-0.16	0.0000	0.12
Pill take up & exposure					
Age started pill	24.07	19.41	-4.66	0.0000	20.86
Ever taken pill	0.75	0.91	0.16	0.0000	0.82
Pill diffusion	0.01	0.36	0.35	0.0000	0.30
Fertility					
Childless	0.11	0.19	0.08	0.0000	0.18
Age at first child	24.96	26.49	1.53	0.0000	26.12
Number of children	1.94	1.71	-0.23	0.0000	1.74
Observations	6,177	5,669			145,502

Notes: Two-sample t-test with unequal variances. Based on the UKB data.

Pill diffusion and education of women

Table 3.4 depicts the estimates of the regression of years of education on the pill diffusion with administrative county of birth fixed effects. When not controlling for birth year cohorts, specifications (1) to (3) show that for every 10 percentage points increase in the pill diffusion, a woman gains 0.35 years of education on average ($p<0.01$). This estimate is however smaller, 0.09 percentage points, after having accounted for the year of birth trends. Still, the last column suggests that going from no diffusion to full diffusion increases women's education by almost one additional year.

Table 3.4. Results of the regressions of pill diffusion on years of education.

	Years of Education			
	(1)	(2)	(3)	(4)
Pill diffusion	3.507***	3.508***	3.519***	0.860***
	(0.072)	(0.072)	(0.074)	(0.229)
R^2	0.054	0.054	0.057	0.059
N	145,502	145,502	145,502	145,502
Outcome mean	15.050	15.050	15.050	15.050
Outcome SD	4.880	4.880	4.880	4.880
Controls				
Constant	✓	✓	✓	✓
Administrative county of birth FEs	✓	✓	✓	✓
Family size	✓	✓	✓	✓
Month of birth		✓	✓	✓
40 PCs			✓	✓
YoB & YoB^2				✓

Notes: Robust standard errors in parentheses, clustered at the administrative county of birth level; * p<0.10, ** p<0.05, *** p<0.01. Coefficients for the control variables are not displayed, but available upon request from the author. Pill diffusion is measured as a share of women taking the pill between age 18 and 45 in the administrative county of birth of the observed woman when she was 18. Pill diffusion varies between 0 and 1, so for the ease of interpretation, I divide it by 10 for every 10-percentage point increase in the exposure.

Genetic differences in pill consumption (rGE)

In this section, I explore gene-environment correlations between pill diffusion and the polygenic score for years of education. Given the premise from the literature that pill uptake behaviour might have differed by IQ of women (Bailey et al., 2012), I expand the selection of polygenic scores to add those available for cognitive performance, reading, and match ability from the Polygenic Index Repository (Becker, Burik, et al., 2021). Testing the potential systematic differences in pill take up behavior at the individual level and the pill diffusion at the aggregated level of administrative county of birth would allow addressing two questions: (1) whether women born in the areas with higher pill take up are different in their genetic predispositions for cognitive and educational traits, (2) whether the individual age at first pill and ever taking pill are different across these genetic predispositions.

I regress each measure related to contraceptive pill uptake on a polygenic score, controlling for the same set of controls as in the earlier analysis: year of birth and year of birth squared, month of birth, family size, and the 40 first principal components of the genetic relatedness matrix, and present the coefficients in Figure 3.5. I benchmark these results against the same specifications with family fixed effects for a smaller sibling subsample (N=6,478).

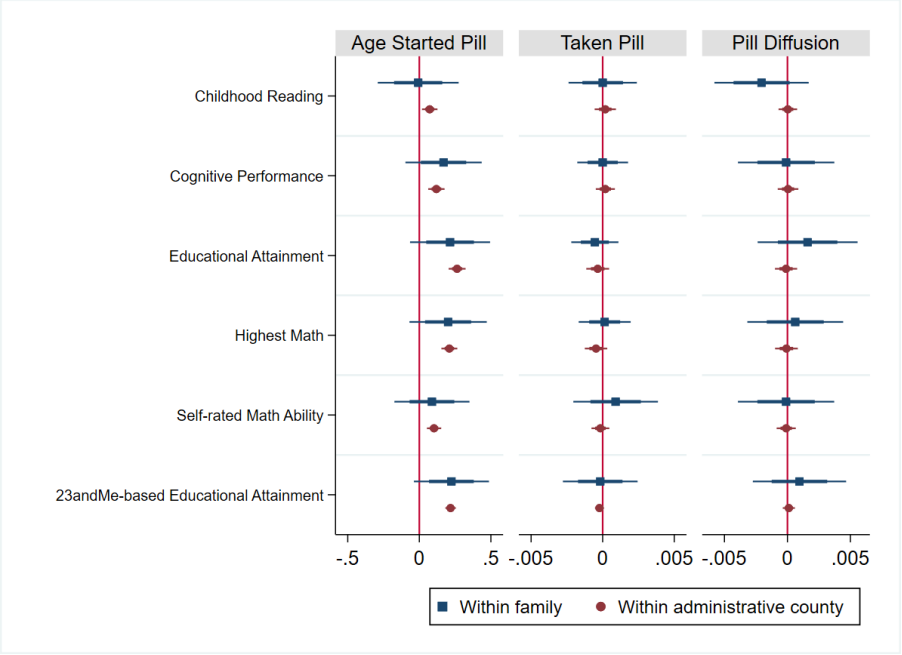
At the conventional thresholds of significance level of 0.05, there seem to be some differences where polygenic scores drive the age at starting the pill. Specifically, almost all cognition and education PGSs are significantly associated with starting the pill at a slightly later age: one standard deviation increase in the EA PGS brings about 2.5 months (0.2 years) delay in starting the first pill (Figure 3.5, Column 1). Adding family fixed effects only widens the confidence intervals, while the effect sizes are still economically significant. In contrast, the effect sizes for pill diffusion are very small (<0.002) in both the within-administrative county of birth and within-family analysis, and not significantly different from zero with Bonferroni correction for multiple testing[36] ($p<0.05/37 = 0.0014$, thin confidence intervals in Figure 3.5) and without it (thicker confidence intervals).

Overall, these findings suggest that those with higher genetic predisposition towards education behave differently, resulting in a later age at which they started to take the pill. In other words, the age at which one started to take the pill (i.e., at the individual level) is endogenous, and related to one's genetic predisposition, and thereby a poor starting point for a study on *G×E* interplay. In contrast, *rGE* is unlikely to bias the *G×E* results where *pill diffusion* is the measure of the environment since the coefficient sizes are neither economically meaningful nor statistically significant[37]. Hence, using pill diffusion as my environmental basis seems a safer choice to study *G×E* interplay.

[36] In total 37 polygenic scores are tested including those used in rGE analysis beyond education and cognition related scores (see Appendix B).

[37] This result hold when testing a set of polygenic scores related to anthropometric, health and health behaviour, fertility, and personality outcomes. See Appendix B for details.

Figure 3.5. Systematic differences in contraceptive pill take up by Education and Cognition PGS.

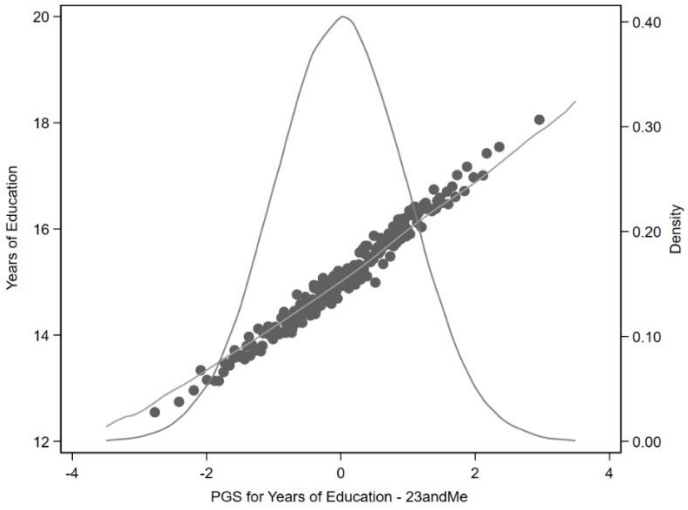


Notes: This figure depicts coefficients of regressions of the contraceptive pill related outcomes on trait-specific polygenic scores (vertical axis). Coefficients are plotted with a 95% confidence interval (thick blue and red lines) and the Bonferroni corrected 99.9% confidence intervals (thin blue and red lines). Regression specifications control for family size, month of birth dummies, year of birth and year of birth squared and the first 40 principal components of the genetic relatedness matrix. Within-family specifications include 6,478 observations. Within-administrative county of birth specifications include 50,691 observations. Regressions with 23andMe-based Educational Attainment PGS are constructed by the author and are based in 23andMe summary statistics (Okbay et al., 2016). Other PGSs are obtained from Polygenic Index Repository, Batch 1(Becker, Burik, et al., 2021). I show the results for PGS repository based Educational Attainment PGSs and 23andMe-based PGS used in the GxE analysis to illustrate their comparability. All PGSs are standardized within the analysis sample. Source: author's calculations based on the UK Biobank data.

Gene-Environment Interactions (G×E)

Gene-environment interaction analyses explore if genetic endowments moderate the effect of the environment on an outcome, specifically, the effect of the pill diffusion on human capital formation of women. The distribution of the polygenic score used in this analysis is depicted in Figure 3.6. For the analysis sample, the variance in the years of education explained by the EA PGS in addition to control variables, the incremental R^2, is at 3.5% (Table 3.5). Figure 3.6 shows that the relationship between EA PGS and Years of Education is approximately linear, which supports the current empirical specification. The figure also clearly shows an average difference of four years in the years of education between women with an EA PGS two standard deviations below and two standard deviations above the mean.

Figure 3.6. The relationship between standardized polygenic score and years of education in the analysis sample.



Notes: PGS for years of education for UKB women (N=145,502) based on 23andMe, Inc. summary statistics ($N_{discovery\_sample}$=365,536). PGSs are standardized within the analysis sample. Plotted using 200 bins of polygenic scores. Source: author's calculations based on the UK Biobank data.

Table 3.5 confirms that the polygenic score for years of education (EA PGS) is positively associated with years of education. Specifically, Column 2 shows that a one standard deviation increase in EA PGS brings an advantage of about 0.909 years of education ($p<0.01$), conditional on the controls. More importantly, the results show

that there is a meaningful interaction between the pill diffusion and the EA PGS, i.e., women with a lower polygenic score benefited more from the diffusion of the pill in their surroundings by 0.327 years (~4 months) of education on average as compared to those who were less exposed to the pill. This implies a cushioning effect of the exposure to the pill against genetic disadvantages pertaining to education. In other words, the results suggest that women with a lower genetic predisposition towards education benefited from the diffusion of the pill in terms of their educational attainment.

Table 3.5. Results of the regressions of years of education on the gene-environment interaction.

	Years of education		
	All women		
	(1)	(2)	(3)
Pill diffusion (E)	0.860***	0.872***	0.869***
	(0.229)	(0.229)	(0.229)
PGS for years of education (G)		0.909***	1.052***
		(0.017)	(0.025)
G×E			-0.327***
			(0.039)
Controls			
Constant	✓	✓	✓
Administrative county of birth FEs	✓	✓	✓
Family size	✓	✓	✓
Month of birth	✓	✓	✓
40 PCs	✓	✓	✓
YoB & YoB²	✓	✓	✓
R^2	0.058	0.093	0.093
N	145,502	145,502	145,502
Outcome mean	15.050	15.050	15.050
Outcome SD	4.880	4.880	4.880

Notes: Robust standard errors in parentheses, clustered at the administrative county of birth level; Coefficients for the control variables (year of birth, year of birth squared and month of birth, and the first 40 principal components of the genetic relatedness matrix) are not displayed, but available upon request from the authors. * p < 0.10, ** p < 0.05, *** p < 0.01; The polygenic score for years of education is based on the 23andMe summary statistics. Pill diffusion is measured as a share of women taking the pill between age 18 and 45 in the administrative county of birth of the observed woman when she was 18. Pill diffusion varies between 0 and 1.

3.6 Sensitivity analyses

In this section, I investigate the robustness of the main results presented in Table 3.5 to a number of alternative specifications, including different formulations of the cohort effects, accounting for the 1974 Raising of the School Leaving Age (RoSLA), and allowing for the effect of the PGS to differ by birth cohorts. Column 1 of Table 3.6 replicates the results presented in Column 1 from Table 3.5, henceforth, I refer to them as *the baseline results*.

Specification of cohort effects

Firstly, I explore the sensitivity of the results to the use of alternative specifications of year of birth, where I include 3-year or 5-year birth cohort fixed effects as in Barban et al. (2021), instead of the quadratic in year of birth. Columns 2 and 3 of Table 3.6 show that broader birth cohort controls of 3-year and 5-year span result in a larger coefficient of the pill diffusion effect on years of education compared to the baseline result presented in Column 1. Importantly, however, the interaction coefficient is virtually identical in these specifications.

Raising of the School Leaving Age. Secondly, there is evidence that the 1957 RoSLA affected the educational outcomes of UKB respondents and that the effect of this policy was moderated by the polygenic score for educational attainment (Barcellos et al., 2018, 2021). This would imply that the interaction term between the pill diffusion and the EA PGS might be overestimated and actually reflect the RoSLA effect. Column 4 of Table 3.6 shows that accounting for RoSLA by way of a dummy that equals one for those born in or after September 1957 and its interaction with the EA PGS doesn't change the estimate of the pill diffusion. It does reduce the interaction coefficient between EA PGS and pill diffusion by 30%, although it is still economically and statistically ($p<0.01$) meaningful.

Predictive power of the polygenic score

Thirdly, there is evidence that the predictive power of the polygenic score for educational attainment has been decreasing for later (i.e., younger) cohorts (Conley, Laidley, Boardman, & Domingue, 2016). I therefore apply a sensitivity check in the fashion of Keller (2014) by interacting the year of birth controls with the EA PGS. Table 3.6 presents the results of the sensitivity of the gene-environment interaction estimates to the potential issues discussed above.

When controlling for the reduction in the polygenic score penetrance over cohorts in Column 5, the effect size of the interaction term between pill diffusion and EA PGS slightly reduces, and very saliently the standard error is three times that of the baseline GxE estimate. Apparently, it is very hard to disentangle the interaction between the EA PGS and pill diffusion from its interaction with year of birth controls.

Therefore, in Table 3.7 to avoid issues with multicollinearity, I stratify the results by High EA PGS (EA PGS>0) and Low EA PGS (EA PGS<0) with a cut off at the mean, and also by EA PGS quartiles, whilst controlling for birth cohort trends. Columns 1 and 2 confirm that the results at the mean in the baseline model are mainly driven by the positive effect of the pill diffusion on the education of the Low EA PGS group. A further stratification by quartiles of EA PGS in Columns 3 to 6 confirm that the effect of the pill diffusion is more meaningful for women in the 1st and 2nd quartiles of polygenic score, with the latter group being particularly strongly affected. These flexible specifications build confidence that the effects of pill diffusion indeed are driven predominantly by individuals with below-mean polygenic scores.

Within-family estimates

As discussed earlier in the paper, in empirical analyses, estimated coefficients for polygenic scores without family fixed effects or without controlling for parental genotypes reflect both direct and indirect genetic effects (Kong et al., 2018). While due to the lack of family trio in the UKB, it is not feasible to control for parental genotypes, it is possible to conduct a within-family analysis on a subsample of siblings, 6,482 sisters in this case. Such family fixed effects specification yields the estimates for pill diffusion of a comparable magnitude to the baseline specification reported in Table 3.5, but less precise due to the significantly smaller sample. The effect size of G × E interaction estimate is larger (β = -0.503) but lacks precision. The estimates of EA PGS are naturally reduced as compared to the baseline since the part of PGS relationship with years of education driven by genetic nurture is now accounted for. Details of the sampling for this analysis and results are reported in Appendix D.

Measurement error in the polygenic score

Since GWASs are based on finite samples, the estimated PGSs are typically a noisy proxy for a true PGS (Dudbridge, 2013). I employ Obviously Related Instrumental Variable approach to address such measurement error resulting in attenuation bias and show that the G × E interaction estimates between the pill diffusion and EA PGS if anything go up (β = -0.681, p<0.01). However, this approach is applied only to a smaller sister's subsample of the UKB (N=6,482). Complete details of this analysis are reported in Appendix C.

Table 3.6. Sensitivity of G×E on years of education to alternative specifications.

	Years of Education				
	(1)	(2)	(3)	(4)	(5)
Pill diffusion (E)	0.869***	2.811***	3.030***	0.828***	0.865***
	(0.229)	(0.371)	(0.215)	(0.229)	(0.228)
PGS for years of education (G)	1.052***	1.053***	1.053***	1.035***	0.988***
	(0.025)	(0.025)	(0.025)	(0.027)	(0.114)
G×E	-0.327***	-0.329***	-0.329***	-0.216***	-0.178
	(0.039)	(0.039)	(0.039)	(0.057)	(0.124)
R^2	0.093	0.094	0.093	0.094	0.093
N	145,502	145,502	145,502	145,502	145,502
Outcome mean	15.050	15.050	15.050	15.050	15.050
Outcome SD	4.880	4.880	4.880	4.880	4.880
Controls					
Constant	✓	✓	✓	✓	✓
Administrative county of birth FEs	✓	✓	✓	✓	✓
Family size	✓	✓	✓	✓	✓
Month of birth	✓	✓	✓	✓	✓
40 PCs	✓	✓	✓	✓	✓
YoB & YoB²	✓			✓	
YoB & YoB² (demeaned)					✓
YoB×G & YoB²×G (demeaned)					✓
RoSLA & RoSLA×G				✓	
3-year birth cohort dummies		✓			
5-year birth cohort dummies			✓		

Notes: Robust standard errors in parentheses, clustered at the administrative county of birth level; Coefficients for the control variables (year and month of birth, and the first 40 principal components of the genetic relatedness matrix) are not displayed, but available upon request from the authors. * p < 0.10, ** p < 0.05, *** p < 0.01; The polygenic score for years of education is based on the 23andMe summary statistics. Pill diffusion is measured as a share of women taking the pill between age 18 and 45 in the administrative county of birth of the observed woman when she was 18. Pill diffusion varies between 0 and 1. In specification 3, RoSLA is equal 1 if a woman is subject to raising of school leaving age policy in the UK, that is born after September 1957.

Table 3.7. Results of the regressions of years of education on pill diffusion by quartiles of polygenic score for education.

	(1)	(2)	(3)	(4)	(5)	(6)
	Low EA PGS	High EA PGS	EA PGS 1st quartile	EA PGS 2nd quartile	EA PGS 3rd quartile	EA PGS 4th quartile
Pill diffusion	1.367***	0.380	0.992**	1.740***	0.606	0.185
	(0.306)	(0.332)	(0.445)	(0.455)	(0.431)	(0.482)
R^2	0.063	0.056	0.066	0.063	0.058	0.055
N	72,725	72,776	36,376	36,375	36,375	36,375
Outcome mean	15.050	15.050	15.050	15.050	15.050	15.050
Outcome SD	4.880	4.880	4.880	4.880	4.880	4.880
Controls						
Constant	✓	✓	✓	✓	✓	✓
Administrative county of birth FEs	✓	✓	✓	✓	✓	✓
Family size	✓	✓	✓	✓	✓	✓
Month of birth	✓	✓	✓	✓	✓	✓
40 PCs	✓	✓	✓	✓	✓	✓
YoB & YoB^2	✓	✓	✓	✓	✓	✓

Notes: Robust standard errors clustered at the administrative county of birth level in parentheses; * p<0.10, ** p<0.05, *** p<0.01. Coefficients for the control variables are not displayed, but available upon request from the authors. The polygenic score for years of education is based on the 23andMe summary statistics. Pill diffusion is measured as a share of women taking the pill between age 18 and 45 in the administrative county of birth of the observed woman when she was 18. Pill diffusion varies between 0 and 1.

3.7 Discussion & Conclusions

Overturning *Roe v. Wade* on June 24th, 2022 opened a debate on the access of women to birth control, including the contraceptive pill (Kolhatkar, 2022; Sotomayor & Caldwell, 2022). Despite the battery of studies providing evidence for positive effects of the access to the contraceptive pill on women's long-term educational and labour market outcomes, and their children's outcomes since its legalization more than 60 years ago (Ananat & Hungerman, 2012; Bailey, 2006; Bailey et al., 2012; Goldin & Katz, 2002; Marie & Zwiers, 2021), the contraceptive pill remains unaffordable and inaccessible to a lot of women of childbearing age (European Parliament, 2020; Willsher, 2021).

In terms of findings, I show that the diffusion of the pill in the UK increased the completed years of education of women: for every 10-percentage points increase in the diffusion of the pill, a woman would gain 0.09 years of education, implying that a full access to the pill would result in almost one additional year of education. Although it is somewhat difficult to compare this finding to the existing literature due to the differences in the nature of the data, the liberalization of the pill policy, and the educational systems between countries, the direction of the effect is in the same ballpark as for example the findings of Bailey et al. (2012) that women with Early Legal Access to contraception in the US gained 0.18 years more schooling by their late forties and the findings of Ananat & Hungerman (2012) on the resulting increase in the fraction of college-graduated women by 0.0226.

This paper is the first to shed light on the heterogeneities in the effect of the pill diffusion on education by the genetic endowments for education. Earlier studies have shown that women of different ability levels experienced returns from access to the pill differently (Bailey et al., 2012; Steingrimsdottir, 2016). More recently, Barban et al. (2021) document heterogeneities by fertility related genetic risk scores in the effect of the pill diffusion on completed fertility of women. I contribute to this literature by showing that gains in the years of education due to the diffusion in the contraceptive pill in the UK were higher among women with lower genetic endowments for education as compared to women with higher genetic endowments. Although this finding contradicts the evidence from Bailey et al. (2012), Steingrimsdottir (2016) and the theoretical expectations and evidence from the literature on complementarities between endowments and investments (Ben-Porath, 1967; Heckman, 1976; Muslimova et al., 2021; Rosen, 1976), is in line with a number of other papers providing evidence for policy environments compensating individuals with lower genetic endowments (e.g. Arold, Hufe, & Stoeckli, 2022; Cheesman et al., 2022) and thereby, reducing inequality of opportunity due to genetic endowments (Barcellos et al. 2022). From a policy perspective, the negative interaction between the genetic endowments and the pill diffusion has important implications in terms of reducing the inequalities in educational outcomes, which are shown to carry over to the next generations by improving the family environment and outcomes of children of the women who have better access to the contraceptive pill (Ananat & Hungerman, 2012).

This paper is not without its limitations. For the GxE to be identified, one needs generational genetic data (i.e., parents and children) and a measure of the environment that is exogenous. Controlling for family fixed effects would allow obtaining conservative estimates for gene-environment interplay (Biroli, Galama, von Hinke, et al., 2022; Muslimova et al., 2021). However, limiting the sample to sisters only would significantly reduce the power, as is evident from the sensitivity analyses reported in Appendix D.

To address endogeneity of the contraceptive pill take up, I use exposure to cohort-by-birth county level variation in the pill diffusion as in (Barban et al., 2021). However, unlike the US reproductive policy environment with substantial differences between

states (Bailey, 2006; Bailey et al., 2012; Goldin & Katz, 2002; Myers, 2022), less regional variation in the UK after legalization of the pill may hinder identification. Further, as discussed in Barban et al. (2021), the pill diffusion might be not the only driver of changes in fertility given contemporaneous legalization of abortion, advance of feminism via the UK Equal Pay Act in 1970 and the Sex Discrimination Act in 1975. Given the nature of these societal changes, it would not be surprising if they also affected the education of women, although these changes were implemented later than the introduction of the pill, which had an unusually rapid diffusion among women.

A third limitation is that I could not establish the precise mechanism through which pill diffusion affected educational attainment. In Appendix E, I present some tentative evidence on mechanisms, and even though Murphy (1993, cited in Barban et al., 2021) document that the pill was one of the main determinants of fertility in the UK during mid-1960s and mid-1970s, my findings suggest that fertility was not the only, or even main, channel through which pill diffusion affected years of schooling. This suggests that other, more indirect, mechanisms also played a role, for example peer effects or social norms. The evidence is far from conclusive however, as the fertility outcomes are self-reported, and the resulting measurement error could have attenuated its coefficients. Moreover, data that would allow isolating contemporaneous societal changes to provide a cleaner and more robust interpretation is not available in the UK Biobank.

Further, the UK Biobank sample differs from the UK population by having higher levels of education, income and possibly other characteristics (Fry et al., 2017). While such nature of the data might threaten the external validity of the results, it does not necessarily affect their internal validity. One could address issues about external validity this by applying UK Population Census based weights, however, this would lead to inconsistencies between the GWASs and polygenic scores obtained externally (computed without weighting) and my analysis sample. This could threaten the internal validity of the results unless weighting is applied at every stage of the analysis.

Ample possibilities exist for important follow-up studies, such as the effect of the pill on partners and men in general, intergenerational effects of the pill and its interplay with genetic endowments, gene-environment interactions, and correlations with the pill consumption behaviour in other outcomes than years of education such as health and income to name a few. Moreover, it is important to disentangle the effect of abortion on women's outcomes and heterogeneities in this effect with respect to genetics of education and mental health risks. In times when the reproductive policies are continuously challenged, it is important to continue bringing robust evidence for how they could meaningfully and positively impact our society.

Acknowledgements

This research has been conducted using the UK Biobank Resource (Application Number 41382). The author gratefully acknowledges the funding from the NORFACE through the Dynamic Inequality across the Life Course (DIAL) programme (462-16-100) and the European Research Council (GEPSI 946647). The author thanks Aysu Okbay and employees and research participants of 23andMe, Inc. for sharing GWAS summary statistics for educational attainment. The author also thanks Fleur Meddens, Hans van Kippersluis, Stephanie von Hinke, Pilar Garcia-Gomez, Thomas Buser, Esmee Zwiers, Melinda Mills, Marco Francesconi and Elisabetta De Cao, participants of IGSS 2021 Conference (University of Colorado), ESSGN 2022 Conference (University of Bologna), and Health Economics Internal Seminars (Erasmus University Rotterdam). The author would also like to acknowledge Samuel Baker for constructing the geographic identifiers and making these data available for the UK Biobank. This work made use of the Dutch national e-infrastructure with the support of the SURF Cooperative (EINF-1107).

3.8 Appendix to Chapter 3

A. Genetic data, GWAS, and Polygenic scores

This section provides a brief background on genetics and genetic data.

Human genome. Human genome consists of 23 pairs of chromosomes, one from each parent with the 23rd pair determining the biological sex of a person. A chromosome is composed of two intertwined strands of deoxyribonucleic acid (DNA), each made up of a sequence of four possible nucleotide molecules: adenine (A) and thymine (T), cytosine (C) and guanine (G). Human genome consists of approximately 3 billion such base pairs, which in turn form approximately 20,000 genes coding for proteins (Ezkurdia et al., 2014). Most genetic differences across humans (<1%) stem from single nucleotide polymorphisms (SNP) (Auton et al., 2015). A SNP is a locus in the DNA at which two different nucleotides, alleles for that SNP, can be observed in the population. An individual's genotype is coded as 0, 1, or 2, depending on the count of "effect" alleles (Auton et al., 2015).

GWAS & PGS. Genome-Wide Association Studies (GWASs) identify SNPs that are associated with a particular trait of interest by relating each SNP to the trait in a hypothesis-free approach using stringent significance thresholds, with independent holdout samples used for replication. By now, the GWAS approach has resulted in numerous genetic discoveries (Visscher et al., 2017).

Since individual SNPs usually explain a very small portion (<0.02%) of the variance in behavioural outcomes (Chabris et al., 2015; Visscher et al., 2017), it is a common practice to combine multiple SNPs into a polygenic risk score - PGS (Dudbridge, 2013), constructed as a sum of SNPs weighted by their GWAS effect size. The PGS for years of education, the outcome of interest of this paper, resulting from the most recent GWAS (Okbay et al., 2022) explains between 12-16% of the variation in years of education, with roughly half of the variation explained attributable to direct genetic effects. There is evidence that many of the genes identified via education GWAS are associated with health, cognitive, and central nervous system traits (Rietveld et al., 2013b)and are related to genomic regions responsible for gene expression in a child's brain during the prenatal period (Lee et al., 2018; Okbay et al., 2016).

The main analysis of this study uses 23andMe-based polygenic score (Okbay et al., 2016), which is constructed using LDpred (Vilhjálmsson et al., 2015) version 1.06, and Python, version 3.6.6. LDpred allows accounting for linkage disequilibrium between SNPs, the non-random correlations between SNPs at various loci of a single chromosome, using a Bayesian approach. The polygenic scores include all SNPs, that is 1,065,078 SNPs after filtering for HapMap3 SNPs at the coordination step.

B. Extended rGE

In this section, I report if polygenic scores for a broader set of traits are related to pill take up behaviour. The polygenic scores for this broader rGE analyses are obtained from the Polygenic Index Repository. These polygenic scores are constructed using PLINK 2 (Chang et al., 2015), where SNPs are corrected for linkage disequilibrium using LDpred (Vilhjálmsson et al., 2015; for further technical details, see Becker, Burik, et al., 2021).

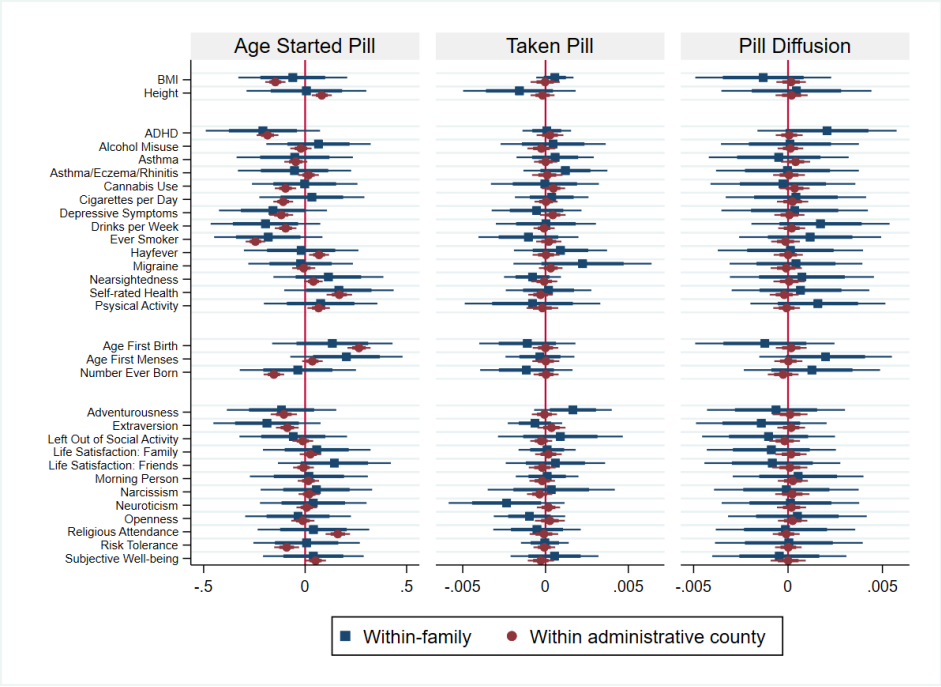
Figure B clusters the polygenic scores by the following traits: Anthropometric, Health and Health Behaviours, Fertility, and Personality. Figure B documents some salient genetic differences in age at starting the pill in within administrative county of birth specifications. For example, women with higher BMI PGS tend to start taking the pill earlier, women with higher Height PGS tend to start taking the pill later, women with higher ADHD PGS also tend to start taking the pill earlier. Likewise, women with higher PGSs for drinks per week, cannabis use, smoking, extraversion and adventurousness start taking the pill at an earlier age. The coefficient sizes of most of these associations are not sensitive to the inclusion of family fixed effects, yet the confidence intervals widen given the smaller sample size for sisters (N=6,478). This finding is in line with the research on the genetics of reproductive behaviour. Specifically, Mills et al (2021) find that the genetic variants for age at first sex and birth are linked to externalizing behaviour such as ADHD and substance use disorders. While substance use disorders have already been linked to earlier age at first sex and birth, we observe that women with externalizing behaviour were also trying to protect themselves with the availability of the pill. Interestingly, one another personality trait stands out here: genetic propensity for religious attendance. Women with higher genetic propensity for religious attendance would start taking the pill later (Figure B, Column 1). This is in line with the literature on the role of religion in access and adoption of contraception (Bailey, 2006; Marie & Zwiers, 2021). However, the effect size of this association is notably smaller after inclusion of the family fixed effects with wide confidence intervals. Overall, this analysis confirms that the age at starting the pill is not independent of the genetic endowments.

Ever taking pill were somewhat associated with PGSs for migraine, adventurousness, and neuroticism. For example, women with higher PGS for neuroticism seems to have slightly higher likelihood of taking the pill. A similar positive association holds for adventurousness. On the contrary, higher PGS for neuroticism reduces the likelihood of taking the pill. However, the effect sizes for these associations are close to zero due to lack of variation in this measure and not statistically significant when corrected for multiple testing (Figure B, Column 2).

Finally, there is not enough evidence for the association between the administrative county level pill diffusion with any of the polygenic scores, the estimates are very small (<0.002), and not statistically significant neither before nor after the Bonferroni correction for multiple testing (Figure B, Column 3). These results confirm that the

employed measure of environmental exposure in the main analysis (pill diffusion) is orthogonal to genetic differences. Hence, rGE unlikely to bias the G×E estimates.

Figure B. Systematic differences in contraceptive pill consumption behaviour by other PGSs.



Notes: This figure depicts coefficients of regressions of the contraceptive pill related outcomes on trait specific polygenic scores (vertical axis). Coefficients are plotted with a 95% confidence interval (thick blue and red lines) and the Bonferroni corrected 99.9% confidence intervals (thin blue and red lines). Regression specifications control for family size, month of birth, year of birth and year of birth squared and the first 40 principal components of the genetic relatedness matrix. Within-family specifications include 6,478 observations. Within-administrative county of birth specifications include 50,691 observations. PGSs are obtained from Polygenic Index Repository, Batch 1 (Becker, Burik, et al., 2021). All PGSs are standardized within the analysis sample. Source: author's calculations based on the UK Biobank data.

C. Obviously-Related Instrumental Variable (ORIV) regression

Polygenic scores are prone to measurement error because the underlying GWAS is based on a finite sample, hence, the estimated PGS is a noisy proxy for a true latent PGS (Dudbridge, 2013). There are two known methods to address such measurement error. First, the measurement error correction developed by Becker, Burik, et al. (2021), where the authors correct for measurement error using external to the PGS SNP-based heritability estimates for a trait. The second method is the Obviously-Related Instrumental Variable (ORIV; Gillen et al., 2019) regression, which uses two independent PGSs where the measurement error is independent (plausibly orthogonal) by construction (Van Kippersluis et al., 2022).

This paper uses the ORIV method given its implementation simplicity in standard software (e.g. STATA), independence of the estimates of the SNP-based heritability, and advantages in terms of being less biased and more precise in the presence of assortative mating (Van Kippersluis et al., 2022), the correlation in genotypes involved in aetiology of specific traits between spouses, including those in the UK Biobank (N=7,780 pairs) for systolic blood pressure, waist-to-hip ratio and educational attainment (Robinson et al., 2017).

To construct the two independent PGSs needed for the ORIV method, I use the split sample GWASs. The discovery sample of these GWASs are based on the UKB, where I first remove all siblings, their relatives, remaining parent-child pairs (N = 5,134) and cousins except one randomly selected from each cousin cluster (N = 45,099). Then I split the unrelated discovery sample (N = 340,009) randomly into two samples of 170,005 and 170,004 individuals each and use the fastGWA procedure (Jiang et al., 2019) to obtain SNP weights. I proceed by using LDpred software to construct two polygenic scores based on the two sets of summary statistics. Likewise, we include all SNPs (1,065,146 after filtering for HapMap SNPs).

Table C presents the ORIV corrected G×E estimates using the split-sample EA PGS (Van Kippersluis et al., 2022). Column 2 shows that the estimates for the EA PGS and the G×E corrected for the attenuation bias are indeed larger than the ones estimated using the OLS specification. However, the effect size of the pill diffusion is not precisely estimated due to a significantly smaller sample size as compared to the baseline analysis.

D. Within-family G×E

In this section, I replicate the key results for a smaller sample of sisters in the UKB. First, I identify siblings and their relatives using the UKB kinship matrix, which is based on genetically identified relatedness using the KING software (Manichaikul et al., 2010). The degree of relatedness is identified from the kinship coefficient and genetic similarity in terms of the identity by state (IBS0) coefficient, which measures the fraction of markers for which the related individuals do not share alleles. Complete details and thresholds can be found in Bycroft et al. (2017). Among 145,502 UKB women, 6,482 women have at least one sibling who also participated in the UKB survey, forming 3,240 sibling clusters. The average years of education of this subsample (μ=14.993, st.dev.= 4.857) is not significantly different than that of the women in the main analysis sample (μ=15.050, st.dev.= 4.880).

Column 1 of Table D reports the baseline results for comparison. Column 2 reports the results of the same specification as in Column 1 but for the sisters' subsample, showing a significant increase in the standard errors of the estimates of interest due to the noticeably smaller sample size. The change in the effect size of the pill diffusion for this reduced sample signals that these sample might be different in terms of the exposure to the pill. Column 3 reports the results for within-family analysis without controlling for administrative county of birth fixed effects. The within-family results obviously lack power, but the effect size for the estimate of the pill diffusion is quite comparable to the baseline specification. The effect size for the G×E if anything is slightly larger than that of the baseline and direction of the interaction is in line with the main results. The estimates for the EA PGS are smaller which is to be expected once controlling for family fixed effects (genetic nurture).

Table D. Comparison of the baseline results to the results of the regressions of years of education on the gene-environment interaction within-family.

	Years of education		
	All women	Sisters	Sisters
	(1)	(2)	(3)
Pill diffusion (E)	0.869***	1.350	0.979
	(0.229)	(1.204)	(1.257)
PGS for years of education (G)	1.052***	1.007***	0.844***
	(0.025)	(0.113)	(0.178)
G×E	-0.327***	-0.177	-0.503
	(0.039)	(0.219)	(0.315)
Controls			
Constant	✓	✓	✓
Administrative county of birth fixed effects	✓	✓	
Family fixed effects			✓
Family size	✓	✓	✓
Month of birth	✓	✓	✓
40 PCs	✓	✓	✓
YoB & YoB²	✓	✓	✓
R^2	0.093	0.079	0.040
N	145,502	6,482	6,482
Outcome mean	15.050	14.993	14.993
Outcome SD	4.880	4.857	4.857

Notes: Robust standard errors in parentheses, clustered at the administrative county of birth level in Column 1 and 2, and at the family level in Column 3; Coefficients for the control variables are not displayed, but available upon request from the authors. * p < 0.10, ** p < 0.05, *** p < 0.01; The polygenic score for years of education is based on the 23andMe summary statistics. Pill diffusion is measured as a share of women taking the pill between age 18 and 45 in the administrative county of birth of the observed woman when she was 18. Pill diffusion varies between 0 and 1.

E. Potential mechanisms

In Section 3.4 on Empirical Approach, I discuss that pill diffusion might affect women's education through two channels, namely, own pill take-up and fertility, and through peer effects and societal shift towards better educated women. In this section, I analyse these possible mechanisms using the available measures of fertility. Moreover, I construct a variable *education diffusion,* which similarly to pill diffusion measures the proportion of women with college education in the administrative county of birth excluding the respondent herself when she was 18.

In Table 3.2, I show that pill diffusion is meaningfully associated with fertility behaviour of women. Below, I test whether fertility behaviour and its interaction with the EA PGS is one of the potential mechanisms for the main results we find. Table E reports the regression of measures of fertility and education diffusion on pill diffusion and the G×E interaction with the polygenic score for education. First, the results show that women with higher EA PGS tend to delay motherhood, have fewer live births, less likely to have children and have children before age 21. Column 1 documents a statistically significant interaction between the pill diffusion and EA PGS, however, there is no conclusive evidence for the effect of the pill diffusion on the age at first birth. Column 2 shows that pill diffusion is positively associated with the number of live births. One potential explanation might be that pill diffusion clustered in the areas with higher fertility rate at least among the UKB respondents. There is no difference in this association by EA PGS. Column 4 shows that pill diffusion leads to a decrease in likelihood in having a child by the age of 21, and that for women with lower PGS this effect is higher. Column 5 documents a positive association between pill diffusion and education diffusion, signalling that the societal shift towards better educated women could also be a potential channel. This association is slightly stronger in terms of the effect size for women with higher EA PGS.

Table E. Results of the regressions of fertility outcomes on pill diffusion by polygenic score for education.

	Age at first birth	Number of live births	Childlessness	First birth by 21	Education diffusion
	(1)	(2)	(3)	(4)	(5)
Pill diffusion (E)	0.159	0.155***	0.011	-0.063***	0.025**
	(0.258)	(0.049)	(0.017)	(0.012)	(0.010)
PGS for years of education (G)	0.533***	-0.025***	0.011***	-0.025***	-0.004***
	(0.022)	(0.005)	(0.002)	(0.001)	(0.000)
G×E	0.355***	-0.002	0.005	0.017***	0.009***
	(0.046)	(0.009)	(0.003)	(0.002)	(0.001)
R^2	0.089	0.093	0.025	0.025	0.919
N	97,801	145,502	145,502	145,502	145,502
Outcome mean	26.116	1.743	0.182	0.062	0.139
Outcome SD	4.282	1.088	0.386	0.241	0.104
Controls					
Constant	✓	✓	✓	✓	✓
Administrative county of birth FEs	✓	✓	✓	✓	✓
Family size	✓	✓	✓	✓	✓
Month of birth	✓	✓	✓	✓	✓
40 PCs	✓	✓	✓	✓	✓
YoB & YoB^2	✓	✓	✓	✓	✓

Notes: Robust standard errors in parentheses, clustered at the administrative county of birth level; Coefficients for the control variables (year and month of birth, and the first 40 principal components of the genetic relatedness matrix) are not displayed, but available upon request from the authors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; The polygenic score for years of education is based on the 23andMe summary statistics. Pill diffusion is measured as a share of women taking the pill between age 18 and 45 in the administrative county of birth of the observed woman when she was 18. Pill diffusion varies between 0 and 1. Education diffusion is measured as a share of college educated women in the administrative district of birth of the observed woman when she was 18.

Chapter 5
Conclusions

In summary, this thesis investigates the complex ways in which genetic endowments, measured by polygenic scores (PGSs), and environments interact with each other and the methodological nuances we need to consider when using the polygenic scores. Chapter 2 shows that genetic endowments interact positively with birth order within families, which could be signalling complementarity between endowments and parental investments strongly correlated with birth order. Chapter 3 finds evidence for possible compensating interaction between genetic endowments for education and diffusion of the birth control pill, which provided a lot of women with opportunities to invest in education. This chapter suggest that certain policy environments can compensate individuals with lower genetic endowments and possibly reducing inequalities among individuals with different endowments. Chapter 4 investigates rank concordance of PGSs created with different construction methods and discovery samples and finds that the rankings of individuals for the same trait are highly unstable, which has important implications for personalised interventions that are based on PGSs, whilst we find limited impact for gene-environment interplay estimates. We show with simulations that the main driver of rank discordance across current-day PGSs is measurement error stemming from a finite GWAS discovery sample and provide some practical recommendations for researchers.

Looking at the big picture, Benjamin et al. (2012) outlined four promises of genoeconomics a decade ago: (1) gaining better insights into biological mechanisms underlying the behaviour of interest and, thus, learning about the nature of preferences, (2) having access to a direct measure of a construct rather than a realized approximation, (3) early detection of diseases and targeted interventions, (4) using genes as control variables or instruments. In Chapter 4 of this thesis, my co-authors and I investigate caveats in using polygenic scores to further explore the third promise of targeted interventions. We show that one such caveat is rank discordance of polygenic scores, implying that targeted interventions based on polygenic scores are currently overly optimistic. Furthermore, I believe that Chapters 2 and 3 illustrate that gene-environment interaction studies could be realizing a *fifth* promise of genoeconomics, that is learning more about the heterogeneities in the effects of policy interventions and environmental exposures and being able to test and gain more insight into existing economic theories.

Understanding heterogeneities could be informative for more comprehensive policy designs, where different groups of the population differentially benefit from their

implementation. Specifically, in Chapter 2, I present evidence in support of complementarities in human capital formation between genetic endowments and parental investments by showing that individuals with higher polygenic score for years of education benefit more from the additional parental investments that are associated with being firstborn. To the best of my knowledge, this study is the first to use genetic data and birth order to study complementarities between human capital endowments and parental investments. The existing literature using alternative measures of endowments and investments produced mixed results. In Chapter 3, I show that investment opportunities resulting from the legal access to contraceptive pill tend to interact negatively with the genetic endowments for education, that is women with lower polygenic scores for education benefit more from the access to the pill compared to women with higher polygenic scores. In both cases, on average, all individuals attain more education if they are firstborn or if they have access to the pill, however, there are heterogeneities by genetic endowments in how much they benefit.

This distinction in the sources of variation in private investments speaks to the production function in Equation 1, where private investments respond to changes in environmental conditions. The difference in the interaction that I find (positive in case of birth order, negative in case of exposure to the pill) might be due to the *policy* change in the case of pill diffusion driving women's choices to invest in education *vs* a change in the *family environment* in the case of birth order driving parental investments. The emerging literature on the interplay between public investments and genetic propensity for education also finds a negative interaction with schools (see Cheesman et al., 2022), quality of teachers (see Arold, Hufe, & Stoeckli, 2022), and month of birth (potentially proxying for (public) teacher investments; Biroli, Galama, von Hinke, et al., 2022). It is not implausible that public investments compensate for genetic differences reducing genetic inequalities, whereas family investments may be complementary to genetic endowments. Another difference between the two studies is that in Chapter 2, I investigate the differences in parental investments driven by birth order differences, which are particularly salient in childhood. Meanwhile, in Chapter 3, I investigate the investment opportunities driven by the legal access to the contraceptive pill and the resulting pill diffusion when women were 18, at the age when they were deciding about their further investments into their education. Hence, the timing of investments and their target beneficiaries are different in two studies, and this could also explain the opposite interaction effects. Finally, the two chapters are also different in that chapter 2 detects a complementarity of the production itself: how does the exogenous reduction in parental time investment for later-born children interact with genetic endowments in producing human capital? Here, the environment E exogenously induces a variation in private parental investments x, and hence we can directly study the possible complementarity between parental investments x and genetic endowments G as in equation (2). Chapter 3, on the other hand, explores whether women exposed to more opportunities for family planning (i.e., E being a conducive environment for

investment in human capital) respond to this opportunity depending on their genetic endowments? The latter is partially a choice-based response to policy, which is not designed to directly affect the production of human capital. Technically, here I am identifying a complementarity between an environment *E* and genetic endowments *G*, which may run through the individual's choice of higher private investments *x*.[40] Possibly there exists complementarity between endowments and investments in the production function, but when the environment is more conducive to investments, it is those with lower polygenic scores that tend to grab these opportunities more. Future research should further test these interesting hypotheses.

The results of this thesis highlight the important role of gene-environment interactions in human capital formation and serve as examples of the ever-growing research opportunities brought by increasing availability of genetic data. Gene-environment interplay has important implications for inequalities in education: a positive interaction reflects how genetic and environmental differences amplify educational inequalities, while a negative interaction reflects how the two factors could compensate each other and ameliorate the inequalities, even though in both cases, individually, positive genetic and environmental influences might be complementing each other. Given genetic inheritance and intergenerational effects of birth order differences (Barclay, Lyngstad, & Conley, 2021; Havari & Savegnago, 2022) and maternal education (Currie & Moretti, 2003) and birth control pill effects (Ananat & Hungerman, 2012), these inequalities can easily carry over to the next generations. Importantly, this strand of research also brings evidence against environmental and genetic determinism and makes a case in favour of nature and nurture being able to complement or compensate each other.

Future research in the field would benefit from availability of richer genetic, socioeconomic and geographic data to tackle the existing methodological challenges. As has been mentioned, efforts to expand the genoeconomics literature to other ancestry groups by means of availability of data for well-powered analysis should be encouraged. This would improve inclusivity of this field of research and improve its generalizability.

[40] Although endogenous responses can reflect the presence of (active) rGE, I leverage the diffusion of the pill resulting from the policy of the introduction of the pill, which is exogenous and show that pill diffusion does not exhibit gene-environment correlation. Moreover, the endogenous response here would be through *x*, private investments, which is a mechanism through which the interaction between *G* and *E* may arise, but would not introduce any bias in the estimation of the interaction between *G* and *E*.

Bibliography

Abdellaoui, A., Hottenga, J.-J., De Knijff, P., Nivard, M. G., Xiao, X., Scheet, P., … Boomsma, D. I. (2013). Population structure, Migration, and Diversifying Selection in the Netherlands. *European Journal of Human Genetics*, *21*, 1277–1285. https://doi.org/10.1038/ejhg.2013.48

Acemoglu, D., & Autor, D. (2012). What Does Human Capital Do? A Review of Goldin and Katz's The Race between Education and Technology. *Journal of Economic Literature*, *50*(2), 426–463. https://doi.org/10.1257/JEL.50.2.426

Adhvaryu, A., Fenske, J., & Nyshadham, A. (2019). Early Life Circumstance and Adult Mental Health. *Journal of Political Economy*, *127*(4), 1516–1549. https://doi.org/10.1086/701606

Adhvaryu, A., & Nyshadham, A. (2016). Endowments at Birth and Parents' Investments in Children. *The Economic Journal*, *126*(593), 781–820. https://doi.org/10.1111/ecoj.12186

Adhyaru, B. B., & Jacobson, T. A. (2018). Safety and efficacy of statin therapy. *Nature Reviews Cardiology*, *15*(12), 757–769. https://doi.org/10.1038/s41569-018-0098-5

Aizer, A., & Cunha, F. (2012). The Production of Human Capital: Endowments, Investments and Fertility. *National Bureau of Economic Research*. https://doi.org/10.3386/w18429

Akerlof, G. A., Yellen, J. L., & Katz, M. L. (1996). An Analysis of Out-of-Wedlock Childbearing in the United States. *The Quarterly Journal of Economics*, *111*(2), 277–317.

Akimova, E. T., Breen, R., Brazel, D. M., & Mills, M. C. (2021). Gene-environment Dependencies Lead to Collider Bias in Models with Polygenic Scores. *Scientific Reports*, *11*(1), 1–9.

Alemán, C., Busch, C., Ludwig, A., & Santaeulàlia-Llopis, R. (2022). *A Stage-Based Identification of Policy Effects*. *BSE Working Paper No. 1369*.

Almond, D., Currie, J., & Duque, V. (2018). Childhood Circumstances and Adult Outcomes: Act II. *Journal of Economic Literature*, *56 (4)*, 1360–1446. https://doi.org/10.1257/jel.20171164

Almond, D., & Mazumder, B. (2013). Fetal Origins and Parental Responses. *Annual*

Review of Economics, *5*(1), 37–56. https://doi.org/10.1146/annurev-economics-082912-110145

Amin, V., Behrman, J. R., Fletcher, J. M., Carlos, |, Flores, A., Flores-Lagunes, A., & Kohler, | Hans-Peter. (2021). Genetic risks, adolescent health, and schooling attainment. *Health Economics*, *30*(11), 2905–2920. https://doi.org/10.1002/hec.4421

Ananat, E. O., & Hungerman, D. M. (2012). The Power of the Pill for the Next Generation: Oral Contraception's Effects on Fertility, Abortion, and Maternal and Child Characteristics. *Review of Economics and Statistics*, *94*(1), 37–51.

Aragam, K. G., Dobbyn, A., Judy, R., Chaffin, M., Chaudhary, K., Hindy, G., … Natarajan, P. (2020). Limitations of contemporary guidelines for managing patients at high genetic risk of coronary artery disease. *Journal of the American College of Cardiology*, *75*(22), 2769–2780. https://doi.org/10.1016/j.jacc.2020.04.027

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., Schady, N., Araujo, M. C., Carneiro, P., … Schady, N. (2016). Teacher Quality and Learning Outcomes in Kindergarten. *The Quarterly Journal of Economics*, *131*(3), 1415–1453. https://doi.org/10.1093/QJE/QJW016

Arold, B., Hufe, P., & Stoeckli, M. (2022). Genetic Endowments, Educational Outcomes and the Mediating Influence of School Investments. *CESifo Working Paper*, *9841*.

Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., … Schloss, J. A. (2015). A Global Feference for Human Genetic Variation. *Nature*, *526*(7571), 68–74.

Bagger, J., Birchenall, J., Mansour, H., & Urzua, S. (2013). Education, Birth Order, and Family Size. *National Bureau of Economic Research*. https://doi.org/10.3386/w19111

Bailey, M. J. (2006). More Power to the Pill : The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply. *The Quarterly Journal of Economics*, *121*(1), 289–320.

Bailey, M. J., Hershbein, B., & Miller, A. R. (2012). The Opt-In Revolution? Contraception and the Gender Gap in Wages. *American Economic Journal: Applied Economics*, *4*(3), 225–254.

Baker, S. (2022). The Biobank Historical Geographic Information System (BIO-HGIS). Version 1.5.

Barban, N., de Cao, E., & Francesconi, M. (2021). Gene-Environment Effects on Female Fertility. *CESifo Working Paper*, *9337*(October).

Barcellos, S. H., Carvalho, L. S., & Turley, P. (2018). Education Can Reduce Health Differences Related to Genetic Risk of Obesity. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(42), E9765–E9772.

Barcellos, S. H., Carvalho, L., & Turley, P. (2021). The Effect of Education on the Relationship between Genetics, Early-Life Disadvantages, and Later-Life SES. *SSRN*, (28750). https://doi.org/10.3386/W28750

Barclay, K. J., Lyngstad, T. H., & Conley, D. C. (2021). The Production of Inequalities within Families and across Generations: The Intergenerational Effects of Birth Order on Educational Attainment. *European Sociological Review*, *37*(4), 607–625. https://doi.org/10.1093/ESR/JCAB005

Barro, R. J. (2001). Human Capital and Growth. *American Economic Review*, *91*(2), 12–17. https://doi.org/10.1257/AER.91.2.12

Barth, D., Papageorge, N. W., & Thom, K. (2020). Genetic Endowments and Wealth Inequality. *Journal of Political Economy*, *128*(4), 1474–1522. https://doi.org/10.1086/705415

Becker, G. (1960). Underinvestment in College Education? *The American Economic Review*, *50*(2), 346–354.

Becker, G. S., & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, *4*(3), 1–47.

Becker, G., & Tomes, N. (1976). Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy*, *84*(4), 143–162.

Becker, Gary S. (1964). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education.* (3rd ed.). The University of Chicago Press.

Becker, Gary S, & Lewis, H. G. (1973). On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy*, *81*(2), S279–S288.

Becker, J., Burik, C. A. P., Goldman, G., Wang, N., Jayashankar, H., Bennett, M., … Okbay, A. (2021). Resource Profile and User Guide of the Polygenic Index Repository. *Nature Human Behaviour*, *5*, 1744–1758.

Becker, J., P Burik, C. A., Goldman, G., Wang, N., Jayashankar, H., Bennett, M., … Okbay, A. (2021). Resource profile and user guide of the polygenic index repository. *Nature Human Behaviour*, *5*, 1744–1758. https://doi.org/10.1038/s41562-021-01119-3

Behrman, J. R. (1988). Nutrition, Health, Birth Order and Seasonality. Intrahousehold Allocation among Children in Rural India. *Journal of Development Economics*, *28*(1), 43–62. https://doi.org/10.1016/0304-3878(88)90013-2

Behrman, Jere R., Pollak, R. A., & Taubman, P. (1982). Parental Preferences and Provision for Progeny. *Journal of Political Economy*, *90*(1), 52–73. https://doi.org/10.1086/261039

Behrman, Jere Richard, Taubman, P., Behrman, J., & Taubman, P. (1986). Birth Order, Schooling, and Earnings. *Journal of Labor Economics*, *4*(3), S121-45. Retrieved from https://econpapers.repec.org/RePEc:ucp:jlabec:v:4:y:1986:i:3:p:s121-45

Belsky, D. W., Domingue, B. W., Wedow, R., Arseneault, L., Boardman, J. D., Caspi, A., … Harris, K. M. (2018). Genetic Analysis of Social-class Mobility in Five Longitudinal Studies. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(31), E7275–E7284. https://doi.org/10.1073/pnas.1801238115

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy*, *75*(4, Part 1), 352–365.

Benjamin, D., Cesarini, D., Laibson, D. I., & Turley, P. (2020). Social science genetics: a primer and progress report. *Journal of Economic Literature*.

Benjamin, D. J., Cesarini, D., Chabris, C. F., Glaeser, E. L., Laibson, D. I., Gunason, V., … Lichtenstein, P. (2012). The Promises and Pitfalls of Genoeconomics. *Annual Review of Economics*, *4*, 627–662. https://doi.org/10.1146/ANNUREV-ECONOMICS-080511-110939

Berry, J., Dizon-Ross, R., & Jagnani, M. (2020). Not Playing Favorites: An Experiment on Parental Fairness Preferences. *NBER Working Paper Series*, *26732*. https://doi.org/10.3386/w26732

Bharadwaj, P., Eberhard, J. P., & Neilson, C. A. (2018). Health at Birth, Parental Investments, and Academic Outcomes. *Journal of Labor Economics*, *36*(2), 349–394. https://doi.org/10.1086/695616

Bierut, L., Biroli, P., Galama, T., & Thom, K. (2018). Childhood Socioeconomic Status Moderates Genetic Predisposition for Peak Smoking. *BioRxiv*, 336834. https://doi.org/10.1101/336834

Birdsall, N. (1991). Birth Order Effects and Time Allocation. *Research in Population Economics*, *7*, 191–213. Retrieved from https://europepmc.org/article/med/12317028

Biroli, P., Galama, T. J., von Hinke, S., van Kippersluis, H., Rietveld, C. A., & Thom, K. (2022). The Economics and Econometrics of Gene-environment Interplay. *ArXiv*. https://doi.org/10.48550/arXiv.2203.00729

Biroli, P., & Zünd, C. L. (2020). *Genes, Pubs, and Drinks: Gene-environment interplay and alcohol licensing policy in the UK*.

Björklund, A., Ginther, D., & Sundström, M. (2021). Family structure and child outcomes in the USA and Sweden. *NBER Working Paper 29524*. https://doi.org/10.1007/s00148-006-0094-7

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *Quarterly Journal of Economics*, *120*(2), 669–700. https://doi.org/10.1162/0033553053970179

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2011). Older and Wiser? Birth Order and IQ of Young Men. *CESifo Economic Studies*, *57*(1), 103–120. https://doi.org/10.1093/cesifo/ifq022

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2016). Healthy(?), Wealthy, and Wise: Birth Order and Adult Health. *Economics and Human Biology*, *23*, 27–45. https://doi.org/10.1016/j.ehb.2016.06.005

Black, S. E., Grönqvist, E., & Öckert, B. (2018). Born to Lead? The Effect of Birth Order on Noncognitive Abilities. *Review of Economics and Statistics*, *100*(2), 274–286. https://doi.org/10.1162/rest_a_00690

Blake, J. (1989). Number of Siblings and Educational Attainment. *Science*, *245*(4913), 32–36. https://doi.org/10.1126/SCIENCE.2740913

Bolton, P. (2012). *Education: Historical statistics*. Retrieved from https://researchbriefings.files.parliament.uk/documents/SN04252/SN04252.pdf

Booth, A. L., & Kee, H. J. (2009). Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment. *Journal of Population Economics*, *22*(2), 367–397. https://doi.org/10.1007/s00148-007-0181-4

Branigan, A. R., Mccallum, K. J., & Freese, J. (2013). Variation in the Heritability of Educational Attainment: An International Meta-Analysis. *Social Forces*, *92*(1), 109–140. https://doi.org/10.1093/SF/SOT076

Breinholt, A., & Conley, D. (2019). Child-driven Parenting: Differential Early Childhood Investmnt by Offsrping Genotype. *Working Paper*.

Breining, S., Doyle, J., Figlio, D. N., Karbownik, K., & Roth, J. (2020). Birth Order and Delinquency: Evidence from Denmark and Florida. *Journal of Labor Economics*, *38*(1), 95–142. https://doi.org/10.1086/704497

Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-Nurture Reconceptualized in Developmental Perspective: A Bioecological Model. *Psychological Review*, *101*(4), 568–586.

Bronfenbrenner, U., & Morris, P. A. (2007). The Bioecological Model of Human

Development. In *Handbook of Child Psychology* (pp. 793–828).

Bulik-Sullivan, B., Loh, P. R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., … O'Donovan, M. C. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics 2015 47:3*, *47*(3), 291–295. Retrieved from https://www.nature.com/articles/ng.3211

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., … Marchini, J. (2018). The UK Biobank Resource with Deep Phenotyping and Genomic Data. *Nature*, *562*(7726), 203–209.

Cesarini, D., & Visscher, P. M. (2017). Genetics and Educational Attainment. *Npj Science of Learning*, *2*(1), 1–8.

Chabris, C. F., Lee, J. J., Cesarini, D., Benjamin, D. J., & Laibson, D. I. (2015). The Fourth Law of Behavior Genetics. *Current Directions in Psychological Science*, *24*(4), 304–312.

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the Challenge of Larger and Richer Datasets. *GigaScience*, *4*(1), 1–16.

Cheesman, R., Borgen, N. T., Lyngstad, T. H., Eilertsen, E. M., Ayorech, Z., Torvik, F. A., … Ystrom, E. (2022). How Genes, Schools, and Residential Areas Shape Achievement: a Population-wide Gene-environment Interaction Study. *PsyArXiv Preprints*. https://doi.org/10.31234/osf.io/r9vzp

Choi, S. W., Mak, T. S.-H., & O'Reilly, P. F. (2020). Tutorial: a guide to performing polygenic risk score analyses. *Nature Protocols*, *15*(9), 2759–2772. https://doi.org/10.1038/s41596-020-0353-1

Clark, D. (2022, January 3). Crude birth rate in the United Kingdom from 1938 to 2020. Retrieved July 25, 2022, from Statista website: https://www.statista.com/statistics/281416/birth-rate-in-the-united-kingdom-uk/

Clifton, L., Collister, J. A., Liu, X., Littlejohn, T. J., & Hunter, D. J. (2022). Assessing agreement between different polygenic risk scores in the UK Biobank. *MedRxiv*. https://doi.org/10.1101/2022.02.09.22270719

Conley, D., Laidley, T. M., Boardman, J. D., & Domingue, B. W. (2016). Changing Polygenic Penetrance on Phenotypes in the 20th Century Among Adults in the US Population. *Scientific Reports*, *6*(July), 6–10.

Conley, D., & Rauscher, E. (2013). Genetic Interactions with Prenatal Social Environment: Effects on Academic and Behavioral Outcomes. *Journal of Health and Social Behavior*, *54*(1), 109–127. https://doi.org/10.1177/0022146512473758

Correia, S. (2017). Linear Models with High-dimensional Fixed Effects: An Efficient and Feasible Estimator. *Working Paper*.

Correia, S. (2019). REGHDFE: Stata Module to Perform Linear or Instrumental-variable Regression Absorbing Any Number of High-dimensional Fixed Effects. *Statistical Software Components*.

Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, *97*(2), 31–47.

Cunha, F., & Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, *43*(4), 738–782. https://doi.org/10.3368/jhr.43.4.738

Cunha, F., Heckman, J. J., Schennach, S. M., Attanasio, O., Becker, G., Cattan, S., … Schennach, S. M. (2010). Estimationg The Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, *78*(3), 883–931. https://doi.org/10.3982/ECTA6551

Currie, J., & Moretti, E. (2003). Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings. *The Quarterly Journal of Economics*, *118*(4), 1495–1532. https://doi.org/10.1162/003355303322552856

Damian, R. I., & Roberts, B. W. (2015). The Associations of Birth Order with Personality and Intelligence in a Representative Sample of U.S. High School Students. *Journal of Research in Personality*, *58*, 96–105. https://doi.org/10.1016/j.jrp.2015.05.005

Datar, A., Kilburn, M. R., & Loughran, D. S. (2010). Endowments and Parental Investments in Infancy and Early Childhood. *Demography*, *47*(1), 145–162. https://doi.org/10.1353/dem.0.0092

Davey Smith, G, & Ebrahim, S. (2003). "Mendelian Randomization": Can Genetic Epidemiology Contribute to Understanding Environmental Determinants of Disease? *International Journal of Epidemiology*, *32*(1), 1–22. https://doi.org/10.1093/ije/dyg070

Daysal, N. M., Ding, H., Rossin-Slater, M., & Schwandt, H. (2021). *Germs in the Family: The Long-Term Consequences of Intra-Household Endemic Respiratory Disease Spread*. https://doi.org/10.3386/W29524

De Haan, M. (2010). Birth Order, Family Size and Educational Attainment. *Economics of Education Review*, *29*(4), 576–588. https://doi.org/10.1016/j.econedurev.2009.10.012

De Haan, M., Plug, E., & Rosero, J. (2014). Birth Order and Human Capital Development: Evidence from Ecuador. *Journal of Human Resources*, *49*(2),

359–392. https://doi.org/10.3368/jhr.49.2.359

de Vlaming, R. (2022). *Genetic-Nuture and Assortative-Mating-Effects Simulator*. Retrieved from https://github.com/devlaming/gnames

de Vlaming, R., Okbay, A., Rietveld, C. A., Johannesson, M., Magnusson, P. K. E., Uitterlinden, A. G., … Koellinger, P. D. (2017). Meta-GWAS Accuracy and Power (MetaGAP) calculator shows that hiding heritability is partially due to imperfect genetic correlations across studies. *PLOS Genetics*, *13*(1), e1006495.

Del Boca, D., Flinn, C., & Wiswall, M. (2014). Household Choices and Child Development. *Review of Economic Studies*, *81*(1), 137–185. https://doi.org/10.1093/restud/rdt026

Del Bono, E., Francesconi, M., Kelly, Y., Sacker, A., Machin, S., Salvanes, K., … McNally, S. (2016). Early Maternal Time Investment and Early Child Outcomes. *The Economic Journal*, *126*(596), F96–F135. https://doi.org/10.1111/ECOJ.12342

Dias Pereira, R, Biroli, P., Galama, T., von Hinke, S., van Kippersluis, H., Rietveld, C. A., & Thom, K. (2022). Gene-by-Environment Interplay in the Social Sciences. *Mimeo, Erasmus University Rotterdam*.

Dias Pereira, Rita, Rietveld, C. A., & van Kippersluis, H. (2022). The Interplay between Maternal Smoking and Genes in Offspring Birth Weight. *Journal of Human Resources*, *57*(5).

Ding, Y., Hou, K., Burch, K. S., Lapinska, S., Privé, F., Vilhjálmsson, B., … Pasaniuc, B. (2022). Large uncertainty in individual polygenic risk score estimation impacts PRS-based risk stratification. *Nature Genetics*, *54*, 30–39. Retrieved from https://www.biorxiv.org/content/10.1101/2020.11.30.403188v3

DiPrete, T. A., Burik, C. A. P., & Koellinger, P. D. (2018). Genetic instrumental variable regression: explaining socioeconomic and health outcomes in nonexperimental data. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(22), 4970–4979. https://doi.org/10.1073/pnas.1707388115

Domingue, B., Trejo, S., Armstrong-Carter, E., & Tucker-Drob, E. (2020). *Interactions Between Polygenic Scores and Environments: Methodological and Conceptual Challenges*. https://doi.org/10.31235/osf.io/u7sh4

Domingue, B. W., Liu, H., Okbay, A., & Belsky, D. W. (2017). Genetic Heterogeneity in Depressive Symptoms Following the Death of a Spouse: Polygenic Score Analysis of the U.S. Health and Retirement Study. *American Journal of Psychiatry*, *174*(10), 963–970. https://doi.org/10.1176/appi.ajp.2017.16111209

Downey, D. B. (2001). Number of siblings and intellectual development: The

resource dilution explanation. *American Psychologist*, *56*(6), 497–504. https://doi.org/10.1037/0003-066X.56.6-7.497

Dudbridge, F. (2013). Power and Predictive Accuracy of Polygenic Risk Scores. *PLoS Genetics*, *9*(3), 1003348.

Duncan, L. E., & Keller, M. C. (2011). A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *American Journal of Psychiatry*, *168*(10), 1041–1049. https://doi.org/10.1176/APPI.AJP.2011.11020191

Duncan, L., Pollastri, A., & Smoller, J. (2014). Mind the gap: Why many geneticists and psychological scientists have discrepant views about gene–environment interaction (G× E) research. *American Psychologist*, *69*(3), 249. Retrieved from https://psycnet.apa.org/record/2014-13696-001

Duque, V., Rosales-Rueda, M., Sanchez, F., Currie, J., Almond, D., Bailey, M., … Shah, M. (2018). How Do Early-Life Shocks Interact with Subsequent Human-Capital Investments? Evidence from Administrative Data. *IZA World of Labor Conference*.

Eirnæs, M., & Pörtner, C. C. (2004). Birth order and the intrahousehold allocation of time and education. *Review of Economics and Statistics*, *86*(4), 1008–1019. https://doi.org/10.1162/0034653043125176

Elam, K. K., Clifford, S., Shaw, D. S., Wilson, M. N., & Lemery-Chalfant, K. (2019). Gene Set Enrichment Analysis to Create Polygenic Scores: a Developmental Examination of Aggression. *Translational Psychiatry*, *9*(1), 212. https://doi.org/10.1038/s41398-019-0513-7

European Parliament. (2020). 2022 European Contraception Policy Atlas. Retrieved from https://www.epfweb.org/node/669

Ezkurdia, I., Juan, D., Rodriguez, J. M., Frankish, A., Diekhans, M., Harrow, J., … Tress, M. L. (2014). Multiple Evidence Strands Suggest that There May be as Few as 19 000 Human Protein-coding Genes. *Human Molecular Genetics*, *23*(22), 5866–5878.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., … Stiles, J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development*, *59*(5), i. https://doi.org/10.2307/1166093

Figlio, D., Guryan, J., Karbownik, K., & Roth, J. (2014). The Effects of Poor Neonatal Health on Children's Cognitive Development? *American Economic Review*, *104*(12), 4205–4230. https://doi.org/10.1257/aer.104.12.3921

Fletcher, J., Wu, Y., Zhao, Z., & Lu, Q. (2020). The Production of Within-family

Inequality: Insights and Implications of Integrating Genetic Data. *BioRxiv*, 2020.06.06.137778. https://doi.org/10.1101/2020.06.06.137778

Frijters, P., Johnston, D. W., Shah, M., & Shields, M. A. (2013). Intrahousehold Resource Allocation: Do Parents Reduce or Reinforce Child Ability Gaps? *Demography*, *50*(6), 2187–2208. https://doi.org/10.1007/s13524-013-0224-2

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., … Allen, N. E. (2017). Comparison of Sociodemographic and Health-related Characteristics of UK Biobank Participants with Those of the General Population. *American Journal of Epidemiology*, *186*(9), 1026–1034.

Galor, O., & Weil, D. N. (2000). Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. *American Economic Review*, *90*(4), 806–828. https://doi.org/10.1257/aer.90.4.806

Giannola, M. (2020). Parental Investment, and Intra-household Inequality in Child Development: Theory, Measurement and Evidence from a Lab-in-the-field Experiment. *Mimeo, University College London.*

Gillen, B., Snowberg, E., & Yariv, L. (2019). Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study. *Journal of Political Economy*, *127*(4), 1826–1863.

Glasziou, P. P., Irwig, L., Heritier, S., Simes, R. J., & Tonkin, A. (2008). Monitoring cholesterol levels: measurement error or true change? *Annals of Internal Medicine*, *148*(9), 656–661.

Goff, D. C., Lloyd-Jones, D. M., Bennett, G., Coady, S., D'Agostino, R. B., Gibbons, R., … Wilson, P. W. F. (2014). 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the american college of cardiology/american heart association task force on practice guidelines. *Journal of the American College of Cardiology*, *63*(25 PART B), 2935–2959. https://doi.org/10.1016/j.jacc.2013.11.005

Goldin, C., & Katz, L. (2018). The Race Between Education and Technology. In *Inequality in the 21st Century* (1st ed.). Taylor & Francis.

Goldin, C., & Katz, L. F. (2002). The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions. *Journal of Political Economy*, *110*(4), 730–770.

Great Britain Historical GIS Project. (2017). Great Britain Historical GIS. Retrieved from University of Portsmouth website: https://www.visionofbritain.org.uk/data/#

Grundy, S. M., Stone, N. J., Bailey, A. L., Beam, C., Birtcher, K. K., Blumenthal, R. S., … Yeboah, J. (2019). 2018

AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: executive summary. *Journal of the American College of Cardiology*, *73*(24), 3168–3209. https://doi.org/10.1016/j.jacc.2018.11.002

Guldi, M. (2008). Fertility effects of abortion and birth control pill access for minors. *Demography 2008 45:4*, *45*(4), 817–827.

Gunnsteinsson, S., Adhvaryu, A., Christian, P., Labrique, A., Sugimoto, J., Ahmed, A., … Venkataramani, A. (2014). *Resilience to Early Life Shocks Evidence from the Interaction of a Randomized Controlled Trial and a Natural Experiment*. Retrieved from http://sites.bu.edu/neudc/files/2014/10/paper_47.pdf

Handy, C., & Shester, K. (2022). Birth Order and the Decline in College Completion Among the Baby Boom Generation. *Economic Inquiry*, *60*(4), 1626–1643.

Havari, E., & Savegnago, M. (2022). The Intergenerational Effects of Birth Order on Education. *Journal of Population Economics*, *35*(1), 349–377. https://doi.org/10.1007/S00148-020-00810-5/FIGURES/3

Heckman, J. J. (1976). A Life-Cycle Model of Earnings, Learning, and Consumption. *Journal of Political Economy*, *84*(4, Part 2), S9–S44.

Heckman, J. J. (2007). The Economics, Technology, and Neuroscience of Human Capability Formation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(33), 13250–13255. https://doi.org/10.1073/pnas.0701362104

Heckman, J. J., Humphries, J. E., & Veramendi, G. (2018). Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking. *Journal of Political Economy*, *126*(S1), S197–S246. https://doi.org/10.1086/698760

Heiland, F. (2009). Does the Birth Order Affect the Cognitive Development of a Child? *Applied Economics*, *41*(14), 1799–1818. https://doi.org/10.1080/00036840601083220

Hotz, V. J., & Pantano, J. (2015). Strategic Parenting, Birth Order, and School Performance. *Journal of Population Economics*, *28*(4), 911–936. https://doi.org/10.1007/s00148-015-0542-3

Houmark, M. A., Ronda, V., & Rosholm, M. (2020). The Nurture of Nature and the Nature of Nurture: How Genes and Investments Interact in the Formation of Skills. *IZA Discussion Paper*, (13780).

Howe, L. J., Nivard, M. G., Morris, T. T., Hansen, A. F., Rasheed, H., Cho, Y., … Davies, N. M. (2022). Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nature Genetics*, *54*(5), 581–592. https://doi.org/10.1038/S41588-022-01062-7

Hsin, A., & Felfe, C. (2014). When Does Time Matter? Maternal Employment, Children's Time With Parents, and Child Development. *Demography*, *51*(5), 1867–1894. https://doi.org/10.1007/s13524-014-0334-5

Inouye, M., Abraham, G., Nelson, C. P., Wood, A. M., Sweeting, M. J., Dudbridge, F., … Samani, N. J. (2018). Genomic risk prediction of coronary artery disease in 480,000 adults: implications for primary prevention. *Journal of the American College of Cardiology*, *72*(16), 1883–1893. https://doi.org/10.1016/j.jacc.2018.07.079

Isungset, A. M., Freese, J., Andreassen, O. A., & Lyngstad, T. H. (2022). Birth Order Differences in Education Originate in Postnatal Environments. *PNAS Nexus*, *1*, 1–11. https://doi.org/10.1093/pnasnexus/pgac051

Jencks, C. (1980). Heredity, Environment, and Public Policy Reconsidered. *American Sociological Review*, *45*(5), 723–736. https://doi.org/10.2307/2094892

Jiang, L., Zheng, Z., Qi, T., Kemper, K. E., Wray, N. R., Visscher, P. M., & Yang, J. (2019). A Resource-efficient Tool for Mixed Model Association Analysis of Large-scale Data. *Nature Genetics*, *51*(12), 1749–1755.

Johnson, R. C., & Jackson, C. K. (2019). Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending. *American Economic Journal: Economic Policy*, *11*(4), 310–349. https://doi.org/10.1257/pol.20180510

Johnston, J., & Matthews, L. J. (2022). Polygenic embryo testing: understated ethics, unclear utility. *Nature Medicine*, *28*(3), 446–448. Retrieved from https://www.nature.com/articles/s41591-022-01743-0

Kantarevic, J., Mechoulan, S., Kantarevic, J., & Mechoulan, S. (2006). Birth Order, Educational Attainment, and Earnings: An Investigation Using the PSID. *Journal of Human Resources*, *41*(4). Retrieved from https://econpapers.repec.org/RePEc:uwp:jhriss:v:41:y:2006:i:4:p755-777

Keller, K., Troesch, L. M., & Grob, A. (2015). First-born siblings show better second language skills than later born siblings. *Frontiers in Psychology*, *6*(JUN). https://doi.org/10.3389/FPSYG.2015.00705

Keller, M. C. (2014). Gene × environment Interaction Studies Have Not Properly Controlled for Potential Confounders: The Problem and the (Simple) Solution. *Biological Psychiatry*, *75*(1), 18–24.

Khera, A. V., Chaffin, M., Aragam, K. G., Haas, M. E., Roselli, C., Choi, S. H., … Kathiresan, S. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics*, *50*(9), 1219–1224. https://doi.org/10.1038/s41588-018-0183-z

Koellinger, P. D., & Harden, K. P. (2018). Using Nature to Understand Nurture: Genetic Associations Show How Parenting Matters for Children's Education. *Science*, *359*(6374), 386–387. https://doi.org/10.1126/science.aar6429

Kolhatkar, S. (2022). Restrictions on Contraception Could Set Women Back Generations. *The New Yorker*.

Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., … Stefansson, K. (2018). The Nature of Nurture: Effects of Parental Genotypes. *Science*, *359*(6374), 424–428.

Kumar, A., Im, K., Banjevic, M., Ng, P. C., Tunstall, T., Garcia, G., … Rabinowitz, M. (2022). Whole-genome risk prediction of common diseases in human preimplantation embryos. *Nature Medicine*, *28*(3), 513–516. https://doi.org/10.1038/s41591-022-01735-0

Lambert, S. A., Gil, L., Jupp, S., Ritchie, S. C., Xu, Y., Buniello, A., … Inouye, M. (2021). The polygenic score catalog as an open database for reproducibility and systematic evaluation. *Nature Genetics*, *53*(4), 420–425. https://doi.org/10.1038/s41588-021-00783-5

Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N., & Davey Smith, G. (2008). Mendelian Randomization: Using Genes as Instruments for Making Causal Inferences in Epidemiology. *Statistics in Medicine*, *27*(8), 1133–1163. https://doi.org/10.1002/sim.3034

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., … Turley, P. (2018). Gene Discovery and Polygenic Prediction from a Genome-wide Association Study of Educational Attainment in 1.1 Million Individuals. *Nature Genetics*, *50*(8), 1112–1121.

Lehmann, J. Y. K., Nuevo-Chiquero, A., & Vidal-Fernandez, M. (2018). The Early Origins of Birth Order Differences in Children's Outcomes and Parental Behavior. *Journal of Human Resources*, *53*(1), 123–156. https://doi.org/10.3368/jhr.53.1.0816-8177

Lemieux, T. (2006). The "Mincer Equation" Thirty Years After Schooling, Experience, and Earnings. In S. Grossbard (Ed.), *Jacob Mincer A Pioneer of Modern Labor Economics* (pp. 127–145). https://doi.org/10.1007/0-387-29175-X_11

Lloyd-Jones, D. M., Leip, E. P., Larson, M. G., D'Agostino, R. B., Beiser, A., Wilson, P. W. F., … Levy, D. (2006). Prediction of lifetime risk for cardiovascular disease by risk factor burden at 50 years of age. *Circulation*, *113*(6), 791–798. https://doi.org/10.1161/CIRCULATIONAHA.105.548206

Lubotsky, D., & Kaestner, R. (2016). Do `Skills Beget Skills'? Evidence on the Effect of Kindergarten Entrance Age on the Evolution of Cognitive and Non-

cognitive Skill Gaps in Childhood. *Economics of Education Review*, *53*, 194–206. https://doi.org/10.1016/j.econedurev.2016.04.001

Malamud, O., Pop-Eleches, C., Urquiola, M., Malamud, O., Pop-Eleches, C., & Urquiola, M. (2016). Interactions Between Family and School Environments: Evidence on Dynamic Complementarities? *NBER Working Paper Series*, *22112*. Retrieved from https://econpapers.repec.org/RePEc:nbr:nberwo:22112

Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W.-M. M. (2010). Robust Relationship Inference in Genome-wide Association Studies. *Bioinformatics*, *26*(22), 2867–2873.

Marie, O., & Zwiers, E. (2021). Moral Barriers to Birth Control Access : How the Pill Changed Dutch Women ' s Lives – When Religion Did Not Get in the Way. *Tinbergen Institute Discussion Paper Series*.

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., & Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics 2019 51:4*, *51*(4), 584–591. https://doi.org/10.1038/s41588-019-0379-x

Mega, J. L., Stitziel, N. O., Smith, J. G., Chasman, D. I., Caulfield, M. J., Devlin, J. J., … Sabatine, M. S. (2015). Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. *The Lancet*, *385*(9984), 2264–2271. https://doi.org/10.1016/S0140-6736(14)61730-X

Mills, M. C., Barban, N., & Tropf, F. C. (2020). An Introduction to Statistical Genetic Data Analysis. *Cambridge: MIT Press.*

Mincer, J. (1984). Human Capital and Economic Growth. *Economics of Education Review*, *3*(3), 195–205. https://doi.org/10.1016/0272-7757(84)90032-3

Mogstad, M., & Torsvik, G. (2022). Family Background, Neighborhoods and Intergenerational Mobility. *Handbook of Family Economics*. https://doi.org/10.2139/SSRN.3859488

Monfardini, C., & See, S. G. (2012). Birth Order and Child Outcomes: Does Maternal Quality Time Matter? *CHILD Working Papers Series*. Retrieved from https://ideas.repec.org/p/cca/wchild/3.html

Mullins, N., Power, R. A., Fisher, H. L., Hanscombe, K. B., Euesden, J., Iniesta, R., … Lewis, C. M. (2016). Polygenic Interactions with Environmental Adversity in the Aetiology of Major Depressive Disorder. *Psychological Medicine*, *46*(4), 759–770. https://doi.org/10.1017/S0033291715002172

Murphy, M. (1993). The Contraceptive Pill and Women's Employment as Factors in Fertility change in Britain 1963–1980: A Challenge to the Conventional View.

Population Studies: A Journal of Demography , *47*(2), 221–243. https://doi.org/10.1080/00324720310001469 86

Muslimova, D. (2022). Diffusion of the Pill and Women's Education: The Role of Gene-environment Interactions. *Working Paper*.

Muslimova, D., Pereira, R. D., von Hinke Kessler Scholder, S., Van Kippersluis, H., Rietveld, C. A. N., & Meddens, S. F. W. (2023). Rank concordance of polygenic indices. *Nature Human Behaviour*, 1-10.

Muslimova, D., van Kippersluis, H., Rietveld, C. A., von Hinke, S., & Meddens, F. (2021). Complementarities in Skill Production: Evidence From Genetic Endowments and Birth Order. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3748468

Myers, C. K. (2017). The Power of Abortion Policy: Reexamining the Effects of Young Women's Access to Reproductive Control. *Journal of Political Economy*, *125*(6), 2178–2224.

Myers, C. K. (2022). Confidential and legal access to abortion and contraception in the United States, 1960-2020. *Journal of Population Economics*, 1–57.

Ni, G., Zeng, J., Revez, J. A., Wang, Y., Zheng, Z., Ge, T., … Wray, N. R. (2021). A comparison of ten polygenic score methods for psychiatric disorders applied across multiple cohorts. *MedRxiv*. https://doi.org/10.1101/2020.09.10.20192310

Nikpay, M., Goel, A., Won, H. H., Hall, L. M., Willenborg, C., Kanoni, S., … Farrall, M. (2015). A comprehensive 1,000 genomes-based genome-wide association meta-analysis of coronary artery disease. *Nature Genetics*, *47*(10), 1121–1130. https://doi.org/10.1038/NG.3396

Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. J., Pers, T. H., Rietveld, C. A., … Benjamin, D. J. (2016). Genome-wide Association Study Identifies 74 Loci Associated with Educational Attainment. *Nature*, *533*(7604), 539–542.

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., … Young, A. I. (2022). Polygenic Prediction of Educational Attainment Within and Between Families from Genome-wide Association Analyses in 3 Million Individuals. *Nature Genetics*.

Pain, O., Glanville, K. P., Hagenaars, S. P., Selzam, S., Furtjes, A. E., Gaspar, H. A., … Lewis, C. M. (2021). Evaluation of polygenic prediction methodology within a reference-standardized framework. *PLoS Genetics*, *17*(5 May), 1–22. https://doi.org/10.1371/journal.pgen.1009021

Pavan, R. (2016). On the Production of Skills and the Birth-order Effect. *Journal of Human Resources*, *51*(3), 699–726. https://doi.org/10.3368/jhr.51.3.0913-5920R

Pereira, R. D., Biroli, P., Galama, T., von Hinke, S., Kippersluis, H. Van, Rietveld, C. A., & Thom, K. (2022). Gene-Environment Interplay in the Social Sciences. *ArXiv*, 1–59.

Pereira, R. D., van Kippersluis, H., & Rietveld, C. A. (2022). The interplay between maternal smoking and genes in offspring birth weight. *Journal of Human Resources*, 1020-11266R2

Peyrot, W. J., Milaneschi, Y., Abdellaoui, A., Sullivan, P. F., Hottenga, J. J., Boomsma, D. I., & Penninx, B. W. J. H. (2014). Effect of Polygenic Risk Scores on Depression in Childhood Trauma. *British Journal of Psychiatry*, *205*(2), 113–119. https://doi.org/10.1192/bjp.bp.113.143081

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment Interaction and Correlation in the Analysis of Human Behavior. *Psychological Bulletin*, *84*(2), 309–322. https://doi.org/10.1037/0033-2909.84.2.309

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal Components Analysis Corrects for Stratification in Genome-wide Association Studies. *Nature Genetics*, *38*(8), 904–909.

Price, J. (2008). Parent-child Quality Time: Does Birth Order Matter? *Journal of Human Resources*, *43*(1), 240–265. https://doi.org/10.3368/jhr.43.1.240

Privé, F., Arbel, J., & Vilhjálmsson, B. J. (2021). LDpred2: better, faster, stronger. *Bioinformatics*, *36*(22–23), 5424–5431. https://doi.org/10.1093/BIOINFORMATICS/BTAA1029

Privé, F., Vilhjálmsson, B. J., Aschard, H., & Blum, M. G. B. (2019). Making the most of clumping and thresholding for polygenic scores. *The American Journal of Human Genetics*, *105*(6), 1213–1221. https://doi.org/10.1016/J.AJHG.2019.11.001

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: a Tool Set for Whole-genome Association and Population-based Linkage Analyses. *American Journal of Human Genetics*, *81*(3), 559–575.

Rietveld, C. A., Conley, D., Eriksson, N., Esko, T., Medland, S. E., Vinkhuyzen, A. A. E., … Koellinger, P. D. (2014). Replicability and Robustness of Genome-Wide-Association Studies for Behavioral Traits. *Psychological Science*, *25*(11), 1975–1986. https://doi.org/10.1177/0956797614545132

Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., … Koellinger, P. D. (2013b). GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment. *Science*, *340*(6139), 1467–1471.

Robinson, M. R., Kleinman, A., Graff, M., Vinkhuyzen, A. A. E., Couper, D., Miller,

M. B., … Visscher, P. M. (2017). Genetic Evidence of Assortative Mating in Humans. *Nature Human Behaviour*, *1*, 16.

Romer, P. (1990). Endogenous Technological Change. *Journal of Political Economy*, *98*(5).

Ronda, V., Agerbo, E., Bleses, D., Mortensen, P. B., Borglum, A., Hougaard, D. M., … Rosholm, M. (2020). Family Disadvantage, Gender and the Returns to Genetic Human Capital. *IZA Discussion Paper*, *13441*.

Rosen, S. (1976). A Theory of Life Earnings. *Journal of Political Economy*, *84*(4, Part 2), S45–S67.

Rosenzweig, M. R., & Schultz, T. P. (2015). Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight. *Https://Doi.Org/10.1086/261179*, *91*(5), 723–746. https://doi.org/10.1086/261179

Rosenzweig, M. R., & Wolpin, K. I. (1995). Sisters, Siblings, and Mothers: The Effect of Teen-Age Childbearing on Birth Outcomes in a Dynamic Family Context. *Econometrica*, *63*(2), 303. https://doi.org/10.2307/2951628

Rosman, K., & Cherelus, G. (2022). Women on Why They're Stocking Up on the Morning-After Pill - The New York Times. *The New York Times*. Retrieved from https://www.nytimes.com/2022/06/27/style/plan-b-stockpile-roe-abortion.html

Rossin-Slater, M., & Wüst, M. (2020). What is the Added Value of Preschool for Poor Children? Long-term and Intergenerational Impacts and Interactions with an Infant Health Intervention. *American Economic Journal: Applied Economics*, *12*(3), 255–286. https://doi.org/10.1257/app.20180698

Rutter, M., Moffitt, T. E., & Caspi, A. (2006). Gene-environment Interplay and Psychopathology: Multiple Varieties but Real Effects. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*(3–4), 226–261. https://doi.org/10.1111/j.1469-7610.2005.01557.x

Sanz-De-Galdeano, A., & Terskaya, A. (2023). Sibling Differences in Genetic Propensity for Education: How Do Parents React? *Forthcoming in Review of Economics & Statistics*.

Schmitz, L., & Conley, D. (2016). The Long-Term Consequences of Vietnam-Era Conscription and Genotype on Smoking Behavior and Health. *Behavior Genetics*, *46*(1), 43–58. https://doi.org/10.1007/s10519-015-9739-1

Schmitz, L., & Conley, D. (2017a). Modeling Gene-Environment Interactions With Quasi-Natural Experiments. *Journal of Personality*, *85*(1), 10–21.

Schmitz, L., & Conley, D. (2017b). The effect of Vietnam-era Conscription and Genetic Potential for Educational Attainment on Schooling Outcomes. *Economics of Education Review*, *61*(October), 85–97.

Schultz, L. M., Merikangas, A. K., Ruparel, K., Jacquemont, S., Glahn, D. C., Gur, R. E., … Almasy, L. (2021). Stability of polygenic scores across discovery genome-wide association studies. *BioRxiv*. https://doi.org/https://doi.org/10.1101/2021.06.18.449060

Selzam, S., Ritchie, S. J., Pingault, J. B., Reynolds, C. A., O'Reilly, P. F., & Plomin, R. (2019). Comparing Within- and Between-Family Polygenic Score Prediction. *American Journal of Human Genetics*, *105*(2), 351–363. https://doi.org/10.1016/j.ajhg.2019.06.006

Shero, J., van Dijk, W., Edwards, A., Schatschneider, C., Solari, E. J., & Hart, S. A. (2021). The practical utility of genetic screening in school settings. *Npj Science of Learning*, *6*(1), 1–10. https://doi.org/10.1038/s41539-021-00090-y

Slob, E. A. W., & Rietveld, C. A. (2021). Genetic predispositions moderate the effectiveness of tobacco excise taxes. *PLoS ONE*, *16*(11), e0259210. https://doi.org/10.1371/JOURNAL.PONE.0259210

Smith-Woolley, E., Pingault, J.-B., Selzam, S., Rimfeld, K., Krapohl, E., von Stumm, S., … Plomin, R. (2018). Differences in exam performance between pupils attending selective and non-selective schools mirror the genetic differences between them. *Npj Science of Learning*, *3*(1), 1–7. https://doi.org/10.1038/s41539-018-0019-8

Sotomayor, M., & Caldwell, L. A. (2022). House Passes Protection for Birth-control Access; Senate Support is Unclear. *The Washington Post*. Retrieved from https://www.washingtonpost.com/politics/2022/07/21/house-birth-control-access/

Steingrimsdottir, H. (2016). Reproductive Rights and the Career Plans of U.S. College Freshmen. *Labour Economics*, *43*, 29–41.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., … Collins, R. (2015). UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Medicine*, *12*(3), e1001779.

Sun, J., Wang, Y., Folkersen, L., Borné, Y., Amlien, I., Buil, A., … Lage, K. (2021). Translating polygenic risk scores for clinical use by estimating the confidence bounds of risk prediction. *Nature Communications*, *12*(1), 1–9. https://doi.org/10.1038/s41467-021-25014-7

Sun, N. (2022). Overturning Roe v Wade: reproducing injustice. *British Medical Journal*, *377*, 1588.

Sundström, M., & Duvander, A. Z. E. (2002). Gender Division of Childcare and the Sharing of Parental Leave among New Parents in Sweden. *European Sociological Review*, *18*(4), 433–447. https://doi.org/10.1093/ESR/18.4.433

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and Limitations of Genome-wide Association Studies. *Nature Reviews Genetics*, *20*, 467–484. https://doi.org/10.1038/s41576-019-0127-1

The International Schizophrenia Consortium. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature Letters*, *460*, 748–752. https://doi.org/10.1038/nature08185

Todd, P. E., & Wolpin, K. I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal*, *113*(485), F3–F33.

Torkamani, A., Wineinger, N. E., & Topol, E. J. (2018). The Personal and Clinical Utility of Polygenic Risk Scores. *Nature Reviews Genetics*, *19*(9), 581–590.

Trejo, S., & Domingue, B. (2019). Genetic Nature or Genetic Nurture? Quantifying Bias in Analyses Using Polygenic Scores. *BioRxiv*, 524850. https://doi.org/10.1101/524850

Tropf, F. C., Lee, S. H., Verweij, R. M., Stulp, G., Van Der Most, P. J., De Vlaming, R., … Mills, M. C. (2017). Hidden heritability due to heterogeneity across seven populations. *Nature Human Behaviour*, *1*(10), 757–765. https://doi.org/10.1038/s41562-017-0195-1

Turley, P., Meyer, M. N., Wang, N., Cesarini, D., Hammonds, E., Martin, A. R., … Visscher, P. M. (2021). Problems with using polygenic scores to select embryos. *The New England Journal of Medicine*, *385*(1), 79–85.

Van den Berg, G.J., Von Hinke, S., & Vitt, N. (2022). Early life exposure to measles and later-life outcomes: Evidence from the introduction of a vaccine. *University of Bristol Mimeo.*

Van den Berg, Gerard J., Von Hinke, S., & Wang, R. A. H. (2022). Prenatal Sugar Consumption and Late-Life Human Capital and Health: Analyses Based on Postwar Rationing and Polygenic Scores. *IZA Discussion Paper*, *No. 15544*.

Van Kippersluis, H., Biroli, P., Pereira, R., Galama, T., von Hinke, Stephanie Meddens, S. F. W., Muslimova, D., … Rietveld, C. A. (2022). Overcoming Attenuation Bias in Regressions using Polygenic Indices: A Comparison of Approaches. *BioRxiv*. https://doi.org/10.1101/2021.04.09.439157

Vilhjálmsson, B. J., Yang, J., Finucane, H. K., Gusev, A., Lindström, S., Ripke, S., … Price, A. L. (2015). Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *American Journal of Human Genetics*, *97*(4), 576–592.

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *American Journal of Human Genetics*, *101*(1), 5–22.

von Hinke, S., & Sorensen, E. (2022). The Long-term Effects of Early Life Pollution Exposure: Evidence from the London Smog. *University of Bristol Mimeo*.

Von Stumm, S., & Plomin, R. (2021). Using DNA to predict intelligence. *Intelligence*, *86*, 101530. https://doi.org/10.1016/J.INTELL.2021.101530

Wand, H., Lambert, S. A., Tamburro, C., Iacocca, M. A., O'Sullivan, J. W., Sillari, C., … Wojcik, G. L. (2021). Improving reporting standards for polygenic scores in risk prediction studies. *Nature*, *591*(7849), 211–219. https://doi.org/10.1038/s41586-021-03243-6

Ware, E. B., Schmitz, L. L., Faul MPH, J., Gard, A. M., Mitchell, C., Smith MPH, J. A., … Kardia, S. L. (2017). Heterogeneity in polygenic scores for common human traits. *BioRxiv*. https://doi.org/10.1101/106062

Willer, C. J., Li, Y., & Abecasis, G. R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, *26*(17), 2190–2191. https://doi.org/10.1093/bioinformatics/btq340

Willsher, K. (2021). France to offer free contraception to women under 25. *The Guardian*. Retrieved from https://www.theguardian.com/world/2021/sep/09/france-free-contraception-women

Winkler, T. W., Day, F. R., Croteau-Chonka, D. C., Wood, A. R., Locke, A. E., Mägi, R., … Loos, R. J. F. (2014). Quality control and conduct of genome-wide association meta-analyses. *Nature Protocols*, *9*(5), 1192–1212. https://doi.org/10.1038/nprot.2014.071

Witte, J. S., Visscher, P. M., & Wray, N. R. (2014). The contribution of genetic variants to disease depends on the ruler. *Nature Reviews Genetics*, *5*(11), 765–776. Retrieved from www.nature.com/reviews/genetics

Wootton, R. E., & Sallis, H. M. (2020). Let's call it the effect allele: a suggestion for GWAS naming conventions. *International Journal of Epidemiology*, *45*(5), 1717–1743.

Yengo, L., Yang, J., & Visscher, P. M. (2018). Expectation of the intercept from bivariate LD score regression in the presence of population stratification. *BioRxiv*. https://doi.org/10.1101/310565

Summary

Human capital, a deciding factor in economic productivity at the micro and macro level, is shaped by a complex interplay between nature (one's genes) and nurture (one's environment). This thesis presents three applications of how gene–environment interplay studies can contribute to testing of economic theories as well as understanding heterogeneities in treatment effects in the formation of human capital. Chapter 2 brings new evidence on complementarities between early life investments and genetic endowments. It studies whether the within-family effect of birth order is amplified by individuals' genetic endowments, measured by a polygenic score (PGS) for years of education, among 14,850 siblings in the UK Biobank. It finds that those with higher genetic endowments benefit disproportionally more from being firstborn compared to those with lower genetic endowments, providing a clean example of how nature and nurture interact in producing human capital. I further investigate whether such a complementarity extends to the prevailing policy environment in young adulthood in Chapter 3 by using access to contraception as an environment fostering human capital investments among women. I restrict my attention to 145,502 women in the UK Biobank and show that exposure to the pill is associated with more years of education. The positive association is concentrated among women with lower genetic endowments for education suggesting the existence of a compensating mechanism: an environment in which contraception is more widely available is most productive for women with a lower genetic predisposition towards education, reducing inequalities in educational attainment. Chapter 4 takes a more fundamental approach by analysing the reliability of the PGS as a measure of genetic endowments in gene-environment interaction studies in which the PGS is used to genetically rank individuals. Simulations show how rank discordance mainly derives from a limited discovery sample size and reveal a tight link between the explained variance of a PGS and its ranking precision. PGS-based ranking is highly dependent on the choice of PGS, such that current PGSs do not have the desired precision to be used routinely for personalised intervention. In regard to gene-environment interplay research, using PGSs in a continuous form provides more robust inference than stratification into subsamples based on the PGS distribution. Researchers are also recommended to report the construction of PGSs transparently and use a PGS with the highest predictive power for the trait of interest. To conclude, this thesis provides new evidence on how environments conducive to investments in human capital exacerbate or alleviate genetic inequalities in education and provides insights into the reliability of PGS-based rankings exploited in gene-environment interplay research. Importantly, this thesis adds to the evidence against environmental and genetic determinism and makes a case in favour of nature and nurture being able to complement and compensate each other.

Samenvatting

Menselijk kapitaal, een bepalende factor voor economische productiviteit op micro- en macroniveau, wordt gevormd door een complexe interactie tussen natuur (iemands genen) en opvoeding (iemands omgeving). Dit proefschrift presenteert drie studies over hoe de interactie tussen genen en omgeving kan bijdragen aan het testen van economische theorieën en het begrijpen van heterogeniteit in effecten bij de vorming van menselijk kapitaal. Hoofdstuk 2 presenteert nieuw bewijs over de complementariteit tussen investeringen die vroeg in het leven gedaan worden en genetische aanleg. Hier wordt onderzocht of het effect van de geboortevolgorde binnen een gezin wordt versterkt door de genetische eigenschappen van individuen, gemeten door een polygenetische score (PGS) voor opleidingsniveau, op basis van data van 14.850 broers en zussen in de UK Biobank. De resultaten laten zien dat degenen met een hogere genetische aanleg voor opleidingsniveau meer baat hebben bij het eerstgeboren zijn dan degenen met een lagere genetische aanleg, wat een goed voorbeeld is van hoe natuur en opvoeding elkaar kunnen versterken bij het produceren van menselijk kapitaal. In hoofdstuk 3 onderzoek ik verder of een dergelijke complementariteit er ook is met betrekking tot beleidsomstandigheden tijdens adolescentie. Specifiek kijk ik naar toegang tot anticonceptie en mijn analyse op basis van data van 145.502 vrouwen in de UK Biobank laat zien dat toegang tot de pil gepaard gaat met meer jaren onderwijs. De positieve associatie is het sterkst voor vrouwen met een lagere genetische aanleg voor opleidingsniveau, wat suggereert dat er een compensatiemechanisme bestaat: een omgeving waarin anticonceptie op grotere schaal beschikbaar is, is het meest nuttig voor vrouwen met een lagere genetische aanleg voor opleidingsniveau, waardoor ongelijkheden in opleidingsniveau worden verminderd. Hoofdstuk 4 kijkt op een fundamentelere manier naar de betrouwbaarheid van de polygenetische score als maatstaf voor genetische aanleg in gen-omgeving interactie onderzoek waarin de PGS wordt gebruikt om individuen te rangschikken. Simulaties laten zien hoe rangschikkingsdiscordantie voornamelijk veroorzaakt wordt door een onvoldoende steekproefgrootte en dat er een nauw verband is tussen de verklarende kracht van een PGS en de nauwkeurigheid van de rangorde. Op PGS gebaseerde rangschikking is sterk afhankelijk van welke PGS wordt gebruikt, zodat de huidige polygenetische scores niet de gewenste precisie hebben om routinematig te gebruiken voor gepersonaliseerde interventies. Met betrekking tot onderzoek naar interacties tussen genen en omgeving, leidt het gebruik van polygenetische scores in een continue vorm tot robuustere conclusies dan wanneer individuen worden verdeeld in verschillende groepen op basis van de PGS verdeling. Onderzoekers wordt ook aanbevolen om de gebruikte constructie methode van een PGS transparant te rapporteren en een PGS te gebruiken met de grootste voorspellende kracht voor het kenmerk dat onderzocht wordt. Samengevat levert dit proefschrift allereerst nieuw bewijs over hoe omgevingen die bevorderlijk zijn voor investeringen in menselijk kapitaal genetische

ongelijkheden in het onderwijs kunnen vergroten of verkleinen. Daarnaast verschaft het inzicht in de betrouwbaarheid van rangordes die gebaseerd zijn op een PGS verdeling in onderzoek naar interacties tussen genen en omgeving. Belangrijk is dat dit proefschrift bewijs aandraagt tegen zogenaamd omgevings- en genetisch determinisme doordat het laat zien dat natuur en opvoeding elkaar kunnen aanvullen en compenseren.

About the Author

Dilnoza Muslimova (1991) was born in Shymkent city, Kazakhstan, on October 5, 1991. She grew up in Tashkent, Uzbekistan. Dilnoza holds a BSc degree in Economics with Finance (2014) from Westminster International University Tashkent and a Research Master's degree in Economics and Business (2018) from the University of Groningen. In 2018, Dilnoza started as a Ph.D. candidate at the Tinbergen Institute and the Erasmus School of Economics under the supervision of Prof.dr. Hans van Kippersluis and Prof. Stephanie von Hinke and mentorship of Dr. Fleur Meddens.

Dilnoza's research and teaching focus on using genetic data to answer questions in the field of applied microeconomics. Particularly, she is interested in questions and methods related to gene-environment interactions in human capital formation. Dilnoza is continuing her career as a Postdoctoral Scholar at the Erasmus School of Economics.

Portfolio

Working papers

- Muslimova, D., Van Kippersluis, H., Rietveld, C. A., Von Hinke, S., & Meddens, S. F. W. (2020). Dynamic complementarity in skill production: Evidence from genetic endowments and birth order. *arXiv preprint arXiv:2012.05021*.

- Muslimova, D., Pereira, R., von Hinke, S., Van Kippersluis, H., Rietveld, C. A., & Meddens, S. F. W. (2023). Rank concordance of polygenic indices. *Nature Human Behaviour*.

- van Kippersluis, H., Biroli, P., Galama, T. J., von Hinke, S., Meddens, S. F. W., Muslimova, D., ... & Rietveld, C. A. (2021). Using obviously-related instrumental variables to increase the predictive power of polygenic scores. *bioRxiv*.

- Slob, E.A.W., Muslimova, D., Rietveld, C.A. (2022). Polygenic indices for censored outcomes: the problem for G×E analysis and a solution. *Working paper.*

Work in progress

- Muslimova, D. (2022). Diffusion of the pill and women's education: The role of gene-environment interactions. *Working paper.*

Academic positions

- **Postdoctoral Researcher in Genes, Policy, & Social Inequality**
  2022 - 2026
  Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands

- **Doctoral Researcher in Gene-environment Interplay** *2018 - 2021*
  Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands
  Tinbergen Institute, The Netherlands

- **Research Assistant** *2017 - 2018*
  University of Groningen, Groningen, The Netherlands

- **Research Assistant** *2014 - 2016*
  Westminster International University in Tashkent (WIUT), Tashkent, Uzbekistan

- **Visiting researcher** *2015*
  International Food Policy Research Institute, Washington, DC, US

Education

- **MPhil in Economics and Business** *2016 - 2018*
  University of Groningen, Groningen, The Netherlands

- **Postgraduate Certificate in Teaching & Leaning in Higher Education** *2014 – 2015*
  Westminster International University in Tashkent, Tashkent, Uzbekistan

- **BSc in Economics with Finance** *2010 - 2014*
  Westminster International University in Tashkent, Tashkent, Uzbekistan

Conference presentations

2023: Frontiers of Using Genetic Data in Economics (Chicago, US, by invitation), European Social Science Genomics Network (University of Bologna, Italy).

2022: European Economics Association (Bocconi University, Italy), European Social Science Genomics Network (University of Bologna, Italy), Health Economics Internal Seminars (Erasmus University Rotterdam, The Netherlands).

2021: Tinbergen Jamboree (Tinbergen Institute, The Netherlands), DIAL Final Conference (Brussels, Belgium), Integrating Genetics and Social Science Conference (University of Colorado, US)

2020: NORFACE Meeting (Online), DIAL Workshop (Online), Integrating Genetics and Social Science Conference (University of Colorado, US)

2019: NORFACE Meeting (University of Bristol, UK), RSF Summer Institute in Social Genomics (Santa Barbara, US), DIAL Mid-term Conference (University of Turku, Finland)

On the job training

2022-2023: University Teaching Qualification (BKO), Erasmus University Rotterdam

2019: Summer Institute in Social Science Genomics (University of Southern California, Santa Barbara, US), Good Practices for Research with Python (SURF Utrecht, The Netherlands)

2018: Linux for Scientists (Erasmus University Medical Center, The Netherlands), Genome-Wide Data Analysis (Tinbergen Institute, The Netherlands), Economics of Health Inequality (Tinbergen Institute, The Netherlands)

Teaching

- Economics and Genetics, TA, Coordinator/Lecturer, Erasmus School of Economics
- Research Project, TA, Erasmus School of Economics
- Introduction to Econometrics, TA, Erasmus School of Economics
- Bachelor and Master Thesis Supervisor, Erasmus School of Economics
- Economics, TA, Westminster International University in Tashkent
- Introduction to Statistics, TA,Westminster International University in Tashkent

Recognitions

2016: Faculty of Economics & Business Merit-based Scholarship (University of Groningen), Holland Scholarship for Excellent Students (NUFFIC), WIUT Staff Performance Appraisal Award for 2015-2016 a.y.: Teacher of the Year 2016 (KPI 4.6/5)

2015: WIUT Staff Performance Appraisal Award for 2014-2015 a.y, Swiss Scientific Foundation Conference Grant for the 4th European Research Conference on Microfinance, Geneva

2014: WIUT Best Thesis in Economics (Grade: 96/100), Bachelor thesis ranked in Top 15 (of 100) papers at New Economic Talent competition, CERGE-EI

2010: WIUT Merit-based Scholarship for Undergraduate Studies, ranked 1st among 1000+ applicants

Academic activities

- Organizer of the ESE-ESHPM Health Economics Symposium

Refereeing

- Journal of Health Economics, Health Economics Journal

TI Publication List

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and Vrije Universiteit Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. For a full list of PhD theses that appeared in the series we refer to List of PhD Theses – Tinbergen.nl. The following books recently appeared in the Tinbergen Institute Research Series:

764 D.M. KUNST *Essays on Technological Change, Skill Premia and Development*

765 A.J. HUMMEL, *Tax Policy in Imperfect Labor Markets*

766 T. KLEIN, *Essays in Competition Economics*

767 M. VIGH, *Climbing the Socioeconomic Ladder: Essays on Sanitation and Schooling*

768 YAN XU, *Eliciting Preferences and Private Information: Tell Me What You Like and What You Think*

769 S. RELLSTAB, *Balancing Paid Work and Unpaid Care over the Life-Cycle*

770 Z. DENG, *Empirical Studies in Health and Development Economics*

771 L. KONG, *Identification Robust Testing in Linear Factor Models*

772 I. NEAMȚU, *Unintended Consequences of Post-Crisis Banking Reforms*

773 B. KLEIN TEESELINK, *From Mice to Men: Field Studies in Behavioral Economics*

774 B. TEREICK, *Making Crowds Wiser: The Role of Incentives, Individual Biases, and Improved Aggregation*

775 A. CASTELEIN, *Models for Individual Responses*

776 D. KOLESNYK, *Consumer Disclosures on Social Media Platforms: A Global Investigation*

777 M.A. ROLA-JANICKA, *Essays on Financial Instability and Political Economy of Regulation*

778 J.J. KLINGEN, *Natural Experiments in Environmental and Transport Economics*

779 E.M. ARTMANN, *Educational Choices and Family Outcomes*

780 F.J. OSTERMEIJER, *Economic Analyses of Cars in the City*

781 T. ÖZDEN, *Adaptive Learning and Monetary Policy in DSGE Models*

782 D. WANG, *Empirical Studies in Financial Stability and Natural Capital*

783 L.S. STEPHAN, *Estimating Diffusion and Adoption Parameters in Networks New Estimation Approaches for the Latent-Diffusion-Observed-Adoption Model*

784 S.R. MAYER, *Essays in Financial Economics*

785 A.R.S. WOERNER, *Behavioral and Financial Change – Essays in Market Design*

786 M. WIEGAND, *Essays in Development Economics*

787 L.M. TREUREN, *Essays in Industrial Economics - Labor market imperfections, cartel stability, and public interest cartels*

788 D.K. BRANDS, *Economic Policies and Mobility Behaviour*

789 H.T.T. NGUYỄN, *Words Matter? Gender Disparities in Speeches, Evaluation and Competitive Performance*

790 C.A.P BURIK *The Genetic Lottery. Essays on Genetics, Income, and Inequality*

791 S.W.J. OLIJSLAGERS, *The Economics of Climate Change: on the Role of Risk and Preferences*

792 C.W.A. VAN DER KRAATS, *On Inequalities in Well-Being and Human Capital Formation*

793 Y. YUE, *Essays on Risk and Econometrics*

794 E.F. JANSSENS, *Estimation and Identification in Macroeconomic Models with Incomplete Markets*

795 P.B. KASTELEIN, *Essays in Household Finance: Pension Funding, Housing and Consumption over the Life Cycle*

796 J.O. OORSCHOT, *Extremes in Statistics and Econometrics*

797 S.D.T. HOEY, *Economics on Ice: Research on Peer Effects, Rehiring Decisions and Worker Absenteeism*

798 J. VIDIELLA-MARTIN, *Levelling the Playing Field: Inequalities in early life conditions and policy responses*

799 Y. XIAO, *Fertility, parental investments and intergenerational mobility*

800 X. YU, *Decision Making under Different Circumstances: Uncertainty, Urgency, and Health Threat*

801 G. GIANLUCA, *Productivity and Strategies of Multiproduct Firms*

802 H. KWEON, *Biological Perspective of Socioeconomic Inequality*

803 D.K. DIMITROV, *Three Essays on the Optimal Allocation of Risk with Illiquidity, Intergenerational Sharing and Systemic Institutions*

804 J.B. BLOOMFIELD, *Essays on Early Childhood Interventions*

805 S. YU, *Trading and Clearing in Fast-Paced Markets*

806 M.G. GREGORI, *Advanced Measurement and Sampling for Marketing Research*

807 O.C. SOONS, *The Past, Present, and Future of the Euro Area*

808 D. GARCES URZAINQUI *The Distribution of Development. Essays on Economic Mobility, Inequality and Social Change*

809 A.C. PEKER, *Guess What I Think: Essays on the Wisdom in Meta-predictions*

810 A. AKDENIZ, *On the Origins of Human Sociality*

811 K. BRÜTT, *Strategic Interaction and Social Information: Essays in Behavioural Economics*

812 P.N. KUSUMAWARDHANI, *Learning Trends and Supply-side Education Policies in Indonesia*

813 F. CAPOZZA, *Essays on the Behavioral Economics of Social Inequalities.*

Human capital, a deciding factor in economic productivity at the micro and macro level, is shaped by a complex interplay between nature (one's genes) and nurture (one's environment).

This thesis provides new evidence on how environments conducive to investments in human capital, such as being born first in the family or having access to contraceptive pill for women, exacerbate or alleviate genetic inequalities in education. Moreover, this thesis provides insights into the reliability of ranking individuals based on polygenic scores, which have been gaining attention as measures of genetic variation in gene-environment interplay research and personalized interventions. Importantly, this thesis adds to the evidence against environmental and genetic determinism and makes a case in favour of nature and nurture being able to complement and compensate each other.

Dilnoza Muslimova (1991) was born in Shymkent city, Kazakhstan. She grew up in Tashkent, Uzbekistan. Dilnoza holds a BSc degree in Economics with Finance from Westminster International University Tashkent and a Research Master degree in Economics and Business from the University of Groningen. In 2018, Dilnoza started as a Ph.D. candidate at the Tinbergen Institute and the Erasmus School of Economics. Dilnoza's research and teaching focus on using genetic data to answer questions in the field of applied microeconomics. Particularly, she is interested in questions and methods related to gene-environment interactions in human capital formation. Dilnoza is continuing her career as a Postdoctoral Scholar at the Erasmus School of Economics.

tinbergen institute

Tinbergen Institute

Erasmus University Rotterdam
Erasmus School of Economics

W: www.eur.nl/ese
www.tinbergen.nl