Genetic Endowments, Educational Outcomes and the Mediating Influence of School Quality*

Benjamin W. Arold, Paul Hufe & Marc Stoeckli

March 27, 2023

Abstract

Genetic endowments are fixed at conception and matter for the educational attainment of individuals. Do better schools mitigate or magnify the outcomes of this genetic lottery? We analyze the interplay of genetic endowments and school quality for educational attainment in the United States. Our results suggest that higher-quality schools are substitutes for genetic endowments: a 1 SD increase in school quality reduces the positive association of educational attainment with a 1 SD increase in the relevant polygenic index from 0.36 to 0.30 years—a decrease of 17%. High-quality schools increase the probabilities that genetically disadvantaged students complete high school and college, but not post-graduate degrees. These increases are underpinned by relative gains in language ability, patience, risk aversion, and health.

JEL-Codes: I29; I21; J24
Keywords: Polygenic indexes, School resources, Skill formation

---

*__Arold__: ETH Zurich (aroldb@ethz.ch); __Hufe__ (corresponding author): University of Bristol, CESifo, HCEO, IFS, and IZA, Priory Road Complex, Bristol, BS81TU, UK (paul.hufe@bristol.ac.uk); __Stoeckli__: LMU Munich and ifo Institute (marc.stoeckli@econ.lmu.de). This paper was previously circulated under the title "Genetic Endowments, Educational Outcomes and the Mediating Influence of School Investments." We gratefully acknowledge funding from Deutsche Forschungsgemeinschaft (DFG) through NORFACE project "IMCHILD: The impact of childhood circumstances on individual outcomes over the life-course" (PE 1675/5-1). This paper has benefited from discussions with Pietro Biroli, Dalton Conley, Benjamin Domingue, Jeremy Freese, Chao Fu, Colin Green, Stephanie von Hinke, Ingo Isphording, Chris Karbownik, Hans van Kipperluis, Nicolas Papageorge, Hans Sievertsen, Pietro Spini, Ludger Woessmann, and Weilong Zhang. We are also grateful to conference participants at IIPF 2019, IWAEE 2019, SMYE 2021, the ifo Conference on Genes, Social Mobility, and Inequalities across the Life-Course, the CESifo Venice Summer Institute on Economics of Education and Equality of Opportunity, the European Social Science Genomics Network (ESSGN) Conference, the IZA Education Conference 2022, and the 2023 Conference on Frontiers in Economic Analysis with Genetic Data at the University of Chicago. Furthermore, we have received valuable feedback from seminar audiences in Cologne, Copenhagen, Harvard, Munich, Princeton, and at USC. All remaining errors are our own.

1 INTRODUCTION

Education is a key determinant of life outcomes, both for individuals and for societies as a whole (Acemoglu and Autor, 2011; Hanushek and Woessmann, 2008; Krueger and Lindahl, 2001). At the same time, a broad literature has shown that genetic endowments are crucial for educational success. In heritability studies, genetic endowments account for 40% of the variation in years of education (Branigan et al., 2013; Lee et al., 2018). High heritability, however, does not imply that the impact of genes on education is immutable. On the contrary, growing evidence shows that the impact of genetic endowments on life outcomes varies with social environments like families, neighborhoods, and schools (Cesarini and Visscher, 2017; Koellinger and Harden, 2018). This observation raises important questions for researchers and policymakers alike: can schools moderate the link between genetic endowments and educational outcomes; and if yes, do they magnify or reduce genetic inequality in educational attainment?

Answers to these questions are important to address concerns about educational inequality. Since genetic endowments are fixed at conception and remain constant over the life course of individuals, genetic inequality contradicts the widespread goal of providing equal educational opportunities to all members of society (Alesina et al., 2018; Lergetporer et al., 2020). In line with this policy goal, there is a long-standing literature in economics that inquires "the question of how well schools reduce the inequity of birth" (Coleman et al., 1966, p.36). This literature has focused on inequalities by socioeconomic status (SES), race, and gender. However, detailed evidence of how well schools address inequality due to genetic endowments is scant due to the long-term unavailability of genetic endowments at the individual level.

In this paper, we address this gap. We use novel, individual-level data on genetic endowments to study the interplay of genes and school quality in the production of educational attainment. That is, we investigate whether high-quality schools magnify or reduce genetic inequality in educational success. Notably, the effect sign is a priori unclear, since different branches of related literature provide support for effects in both directions. On the one hand, behavioral geneticists have found that advantaged family environments and genetic endowments complement each other (Turkheimer et al., 2003; Woodley of Menie et al., 2021)—a result often referred to as the Scarr-Rowe interaction. Hence, if school environments worked similarly to family environments, high-quality schools would magnify genetic inequality. On the other hand, recent evidence on heterogeneity in school effects suggests that school quality and student socioeconomic background substitute each other (Jackson et al., forthcoming). Hence, if disadvantages based on genetic endowments were similar to disadvantages based on students' socioeconomic backgrounds, high-quality schools would reduce genetic inequality.

To study the interplay of genetic endowments and school quality in the production of educational attainment, we use data from the National Longitudinal Study of Adolescent to Adult Health (Add Health). Add Health is a 5-wave panel study that follows a representative sample

of US high school students from 1994/95 to the present. To measure genetic endowments, we leverage recent advances in molecular biology and use a polygenic index for educational attainment (PGI^{EA}, Becker et al., 2021; Dudbridge, 2013; Lee et al., 2018). PGI^{EA} is an individual measure of the genetic propensity to attain education.[1] It offers important advantages over traditional proxies for "innate ability", such as student test scores and IQ tests (Brinch and Galloway, 2012; Hanushek and Woessmann, 2008, 2012; Heckman et al., 2010). More specifically, it is a fine-grained DNA-based measure that is fixed at conception and cannot be modified thereafter. To measure the quality of school environments, we use information from headmaster surveys and construct school quality indicators based on the following observable characteristics: teacher experience, teacher turnover, teacher education, and class size. To the best of our knowledge, Add Health is the only (publicly available) data set that offers detailed information on schooling environments from both survey and administrative sources for a genotyped sample of reasonable size.

For identification, we combine a between-family comparison with the control function approach suggested by Altonji and Mansfield (2018). We discuss the underlying identification assumptions in detail and provide tests for their satisfaction. A first identification challenge arises from the fact that the genetic endowments of children are a function of the genetic endowments of their parents. Therefore, our parameters of interest may be confounded by *genetic nurture effects*, i.e., genetic endowments of children may be correlated with other family characteristics that co-determine educational attainment. In response, we show that the relevant point estimates from the between-family design can be replicated in a smaller sibling sample that allows us to control for genetic nurture by including family-fixed effects. This result suggests that any remaining confounding from genetic nurture effects is small and unlikely to overturn our main results. A second identification challenge arises from the fact that children sort into schools based on family background. Therefore, our parameters of interest may be confounded by *selection effects*, i.e., school quality may be correlated with other family characteristics that co-determine educational attainment. In response, we follow the control function approach suggested in Altonji and Mansfield (2018), i.e., we use group-level averages of observable characteristics to remove all cross-group variation that results from sorting into school catchment areas. Furthermore, we calculate different summary statistics to quantify the potential magnitude of confounding by unobservables (Cinelli and Hazlett, 2020; Oster, 2019). The results suggest that any residual confounding from selection effects is small and unlikely to overturn our main results. Lastly, gene-environment interactions can only be identified if genetic endowments and the environmental variable of interest are distributed independently of each other. This prerequisite may be violated if children sort into schools based on their genetic endowments. In response, we show that we cannot reject the equality of PGI^{EA} distributions at different levels of school quality. This result rules out selection into schools based on ge-

[1] In addition, PGI^{EA} is highly predictive of several life outcomes that are closely related to educational attainment. These outcomes include earnings, wealth and (non-)cognitive skills (Barth et al., 2020; Buser et al., 2021a; Demange et al., 2021; Houmark et al., 2020; Lee et al., 2018; Muslimova et al., 2020; Papageorge and Thom, 2020).

netic endowments. In summary, all our tests point to the satisfaction of crucial identification assumptions and provide support for the between-family design adopted in this paper.

Our results can be summarized as follows. First, genetic endowments and school quality are highly predictive of years of education: a one-standard-deviation increase in PGI^{EA} (school quality) increases educational attainment by ≈ 0.36 (0.14) years.[2] Second, genetic endowments and school quality are substitutes in the production of educational attainment: a one-standard-deviation increase in school quality reduces the positive association between educational attainment and PGI^{EA} from 0.36 to 0.30 years—a decrease of $\approx 17\%$. This result implies that improvements in school quality may reduce genetic inequality in educational attainment.

We perform a series of robustness checks to evaluate whether our results are conflated by competing mechanisms. We first show that our measures for school quality do not pick up the effects of other school characteristics that may correlate with student outcomes. These characteristics comprise school-level policies such as the prevalence of ability grouping and retention as well as the demographic composition of teachers. Next, we demonstrate that our results are not driven by gene-environment interactions that reflect family instead of school environments. To that end, we estimate very flexible models including higher-order polynomials of PGI^{EA} and our measures for school quality and account for all possible interactions with a broad set of family background characteristics (Biroli et al., 2022; Keller, 2014). In all these robustness analyses, our main results remain unaffected.

We also analyze mechanisms that underpin the substitutability of genetic endowments and school quality. First, to uncover which types of skills drive our results, we analyze the associations of PGI^{EA} and school quality with a range of intermediate outcomes that are highly predictive of educational attainment. These intermediate outcomes include cognitive skills, economic preferences, personality measures, and health. We find substitutability of genetic endowments and school quality for the formation of verbal intelligence, risk-aversion, patience, and subjective health. Hence, we consider these intermediate outcomes plausible transmission channels for the overall result regarding educational attainment. Second, educational attainment summarizes information from various educational stages, where each stage requires a different mix of skills (Cunha et al., 2006, 2010). Therefore, we repeat our analysis by replacing total educational attainment with binary variables indicating whether respondents have obtained a given educational degree. We find substitutability of genetic endowments and school quality for high school and college graduation, but there is no substitutability for post-graduate degrees. This pattern stands in notable contrast to the gene-environment interaction for family SES. Consistent with evidence from recent studies (Buser et al., 2021a; Papageorge and Thom, 2020), we show that the interaction of genetic endowments and family socioeconomic status switches its sign as individuals progress through the educational system and that there is a substantial complementarity for the attainment of post-graduate degrees. The contrast between these pat-

[2] These increases correspond to 16% (6%) of a standard deviation.

terns indicates that gene-environment interplay varies across different sources of investments in children: relative to their high PGI^{EA} peers, genetically disadvantaged children consistently gain more from having attended a high-quality school than from having a high-SES family background.

Our study contributes to two strands of literature. First, we contribute to the literature on gene-environment interactions. There is a large literature showing that the association between genetic endowments and life outcomes varies with family SES (Figlio et al., 2017; Houmark et al., 2020; Papageorge and Thom, 2020; Ronda et al., 2022). However, evidence on gene-environment interactions in the school context is scant. Houmark et al. (2022) show that genetic gaps in achievement among Danish students increase across grades 2–8, particularly among children from low SES families. In contrast to their study, we do not focus on the dynamics of genetic gradients across the educational biography of students but investigate whether school quality can moderate these gradients. Barcellos et al. (2021) use a regression discontinuity design to show relative gains in educational attainment for low PGI^{EA} students after a compulsory schooling reform in the UK.[3] In contrast to their paper, we do not focus on variation in the length of schooling but on variation in the quality of schools. Trejo et al. (2018) investigate whether the socioeconomic composition of schools moderates the strength of genetic gradients in educational and occupational attainment. Using data from Add Health and the Wisconsin Longitudinal Study, they find inconclusive evidence for the presence of gene-environment interplay. In contrast to their study, we do not focus on school contexts, e.g., the socioeconomic composition of students, but on quality indicators that reflect school practices and policies: teacher experience, teacher turnover, teacher education, and class size.[4] This shift of focus also has methodological implications. The socioeconomic composition of schools directly reflects endogenous sorting that leads to gene-environment correlations. Since the independent distribution of genetic endowments and environmental variables is a prerequisite for the identification of their interaction, the interpretation of gene-environment interactions with school composition variables is not straightforward. To the contrary, we show that our measure of school quality is independently distributed of PGI^{EA} and we strengthen identification by using the control function approach of Altonji and Mansfield (2018) to account for cross-group variation that results from endogenous sorting into schools.

Second, we contribute to the literature on school quality. A broad literature demonstrates that school quality raises educational attainment, wages, and health, and reduces crime (Beuermann and Jackson, 2022; Beuermann et al., 2022; Deming, 2011; Deming et al., 2014). Specifically, teacher quality, which is a core component of our school quality measure, has been shown to improve student outcomes in the short- and long-term (Chetty et al., 2014a,b; Jackson, 2019; Rivkin et al., 2005; Rockoff, 2004). However, it is less known which student groups benefit the

[3]However, these relative gains in educational attainment did not translate in corresponding relative income gains as high PGI^{EA} students experienced higher returns to education.

[4]See also Jennings et al. (2015) and Raudenbush and Willms (1995) for the conceptual distinction of school context and school practice when estimating school effects.

most from school quality. On the one hand, disadvantaged students may benefit more, as they have more room for improvement; on the other hand, they may benefit less, as they do not have the means to take advantage of better schools (Cunha et al., 2010). In line with the ambiguity of theoretical predictions, the empirical evidence is mixed. For example, Walters (2018) show that disadvantaged students benefit more from admission to charter schools in Boston, while Dustan et al. (2017) show that disadvantaged students benefit less from admission to elite public high schools in Mexico City. Many studies in this literature, analyze samples of applicants to charter or elite schools, which are not representative of disadvantaged students in general. A notable exception is a study by Jackson et al. (forthcoming) who use a sample of all schools and students in Chicago's public school district to demonstrate that disadvantaged students gain relatively more from higher-quality schools. In contrast to these previous studies, which have evaluated heterogeneity by socioeconomic status, we focus on genetic endowments as the source of disadvantage. The focus on genetic endowments is highly relevant since they are fixed at conception and cannot be influenced by individual choice. Therefore, they constitute a source of inequality that is perceived as unfair by many people in the United States and other Western societies (Almås et al., 2020). Furthermore, they have substantial explanatory power for educational attainment that is not captured by standard measures of disadvantage like family SES.[5] Using a representative sample of schools from all over the United States, we show that genetically disadvantaged students gain relatively more from high-quality schools than their advantaged peers.

Furthermore, our paper relates to the literature on educational inequality and intergenerational educational mobility (Blanden et al., 2023; Chetty et al., 2014c,d; Hanushek et al., 2021; Rossin-Slater and Wüst, 2020). This literature shows that educational inequality is strongly persistent across generations, part of which can be explained by the intra-family transmission of genetic endowments. Our results show that investments in school quality can address this channel of intergenerational persistence in the United States.

The remainder of this paper is structured as follows. In section 2, we provide an introduction to the measurement of genetic endowments. In section 3, we detail our empirical strategy. After introducing our data sources in section 4, we present results in section 5. Section 6 concludes the paper.

2 MEASURING GENETIC ENDOWMENT

The "First Law of Behavior Genetics" states that all human traits are heritable; that is, genetic endowments explain the expression of each trait, at least to some extent (Turkheimer, 2000).

[5]See Supplementary Figure S.1 where we show that PGI^{EA} is correlated with family SES but that PGI^{EA} distributions have significant overlap at different levels of family SES. This pattern suggests that genetic endowments are a distinct dimension of student advantage that cannot be captured by family SES alone.

The empirical challenge is to identify specific sequences in the genome that are related to the traits of interest.[6] Recent advances in molecular genetics have enabled a novel method of genetic discovery: genome-wide association studies (GWAS). GWAS exploit the most common type of genetic variation between humans, so-called single-nucleotide polymorphisms (SNPs). SNPs occur when a single nucleotide—the basic building block of DNA molecules—differs at a specific position in the genome. Humans have around ten million of these SNPs. GWAS estimate separate linear regressions that relate the SNP of individual i at genome location j to an outcome of interest y:

$$y_i = \psi_j^y SNP_{ij} + \delta C_i + \varepsilon_i. \tag{1}$$

$SNP_{ij} \in \{0, 1, 2\}$ is a count variable and indicates the number of minor alleles that individual i possesses at location j. Minor alleles are the less frequent genetic variation within a population. As humans inherit one of each chromosome from each parent, they possess either zero, one, or two minor alleles at each location j. C_i is a vector of control variables to filter out spurious correlations due to non-biological differences across population groups. A particular SNP coefficient ψ_j^y is considered genome-wide significant if the null hypothesis of non-association is rejected at a level of $p < 5 \times 10^{-8}$ (Chanock et al., 2007). The p-value is set low to account for multiple hypothesis testing.

The association of single SNPs with y is small, but jointly they can explain a substantial share of observed outcome differences between individuals (Lee et al., 2018). In particular, the estimated SNP coefficients can be used to construct polygenic indexes (PGI). PGIs are scalar measures of an individual's genetic predisposition for an outcome y relative to the population. Formally, PGI_i^y are constructed as a linear aggregation of all SNP_{ij} using the GWAS coefficients $\widehat{\psi}_j^y$ as weights:

$$PGI_i^y = \sum_j \widehat{\psi}_j^y SNP_{ij}. \tag{2}$$

To avoid overfitting, equation (1) is estimated in a discovery sample, whereas the PGI_i^y is constructed in a hold-out sample (Wray et al., 2014).

PGIs are now available for a variety of outcomes. These include, for example, the body mass index and height (Yengo et al., 2018), attention deficit hyperactivity disorder (Demontis et al., 2019), major depressive disorder (Howard et al., 2019), intelligence (Savage et al., 2018), smoking (Liu et al., 2019), and sleep duration (Jansen et al., 2019). For our analysis, we rely on the PGI for educational attainment from Lee et al. (2018) which we denote by PGI^{EA}. It is based on information from 1.1 million individuals and explains around 12.7% of the variance in educational attainment in the United States.

The interpretation of PGIs is not trivial. First, PGIs are not pure measures of biological influ-

[6] Human genetic information is stored in 23 chromosome pairs that consist of deoxyribonucleic acid (DNA) molecules. These chromosomes, in turn, contain 20,000 to 25,000 genes—specific DNA sequences that provide instructions for building proteins. More than 99% of the sequences are identical in all humans.

ence. For example, GWAS coefficients may capture environmental factors such as population stratification across geographic regions (Abdellaoui et al., 2019). To address this concern, we follow standard practice in the literature and always control for the first 20 principal components of the genetic data in our empirical analysis. Second, PGIs are noisy measures of genetic endowments.[7] For example, GWAS coefficients are estimated in finite samples leading to measurement errors in the PGI weights. Furthermore, the explanatory power of PGIs depends on the context of their application. If a PGI is applied in one context, while the underlying GWAS was estimated in another context, the predictive power of the PGI will be attenuated.[8] To address the concern of attenuation bias due to measurement error, we use the procedure of Becker et al. (2021) and provide robustness checks based on measurement-error-corrected estimates for the coefficients of PGI^{EA} and the gene-environment interaction. Reassuringly, these analyses suggest that our conclusions about the relative strength of gene-environment interactions are not affected by measurement error in PGI^{EA}.

3 EMPIRICAL STRATEGY

3.1 Empirical Model

Consider a model in which the skills θ of child i at age a are determined by prior skill levels θ_{ia-1}, family inputs I^F_{ia}, school inputs I^S_{ia}, and genetic endowments G_i. There are three phases of skill accumulation:

$$\theta_{ia} = \begin{cases} f_a(G_i) & \text{, for child age } a = -1, \\ f_a(I^F_{ia}, \theta_{ia-1}, G_i) & \text{, for child age } a = 0, ..., 5, \\ f_a(I^S_{ia}, I^F_{ia}, \theta_{ia-1}, G_i) & \text{, for child age } a = 6, ..., A. \end{cases} \tag{3}$$

Skills at conception are determined by genetic endowments only. For children aged $a = 0, ..., 5$, i.e., in the period after conception and before attending school, families are the only source of investments in children. Family inputs may include health behaviors during pregnancy, monetary investments such as buying books, and time investments such as reading to the child. For $a = 6, ..., A$, schools are an additional source of investments in children. School-based inputs

[7] The heritability of educational attainment is estimated to be 40% whereas PGI^{EA} explains 12.7% of the variance in educational attainment (Branigan et al., 2013).

[8] For example, educational attainment in a country without compulsory schooling likely correlates with a different set of genetic endowments than in a country with high-quality compulsory schools. In our study, this concern is limited: we apply PGI^{EA} to a sample from the United States, while the underlying GWAS predominantly draws on samples from the United States and other industrialized countries with comparable education systems (Lee et al., 2018).

may include the quality of instruction by teachers and school policies to regulate behavior.

Completed education Y is a function of individual skills accumulated at age $a = A$:

$$Y_i = m(\theta_{iA}). \tag{4}$$

By recursively substituting equations (3) and (4) across ages $a = 1, ..., A$, we obtain a model in which educational attainment is determined by initial genetic endowments, the full history of family inputs, and the full history of school inputs:

$$Y_i = h(I^S_{iA}, ..., I^S_{i6}, I^F_{iA}, ..., I^F_{i1}, G_i). \tag{5}$$

We are interested in the complementarity of school inputs and genetic endowments at a particular child age a:

$$\kappa = \frac{\partial^2 h(I^S_{ia}, \overline{I^S_{ia-1}}, ..., \overline{I^S_{i6}}, \overline{I^F_{ia}}, \overline{I^F_{ia-1}}, ..., \overline{I^F_{i1}}, G_i)}{\partial I^S_{ia} \partial G_i}. \tag{6}$$

If $\kappa < 0$, genetic endowments and school inputs at age a are *substitutes* in the production of educational attainment, i.e., high-quality schools are more productive for individuals with relatively disadvantageous genetic endowments. Reversely, if $\kappa > 0$, genetic endowments and school inputs at age a are *complements* in the production of educational attainment, i.e., high-quality schools are more productive for individuals with relatively advantageous genetic endowments.

In this study, we measure genetic endowments G using the polygenic score PGI^{EA}. Furthermore, we focus on school inputs I^S_{ia} during adolescence ($14 \leq a \leq 18$) by constructing a summary index Q^{S} that proxies for the quality of high schools attended by individuals. Then, we estimate κ using a linear regression model with an interaction term:

$$Y_i = \alpha PGI^{EA}_i + \beta Q^S_i + \kappa(PGI^{EA}_i \times Q^S_i) + \mathbf{X}_i(a)\gamma + \epsilon_i, \tag{7}$$

where $\mathbf{X}_i(a)$ denotes a vector of control variables to condition on the history of family and school inputs up to age $a = 14$.

3.2 Conditions for Identification

The parameter of interest κ is identified if the following conditions are met: (i) exogenous variation in PGI^{EA}, (ii) exogenous variation in Q^{S}, and (iii) independent variation in PGI^{EA} and Q^{S} (Almond and Mazumder, 2013; Biroli et al., 2022; Johnson and Jackson, 2019; Nicoletti and

Rabe, 2014). In the following, we discuss each of these conditions, potential threats to their satisfaction, and how we address these threats in this paper.

(i) Exogenous variation in PGI^{EA}. Genetic endowments are not exogenous to family characteristics as the genetic endowments of children are drawn from the genetic pool of their biological parents. As a consequence, PGI^{EA} is a function of maternal and paternal genetic endowments that may correlate with family inputs $I^F_{i1}, ..., I^F_{ia}$. Hence, when estimating equation (7), α and κ may be confounded by *genetic nurture effects* (Kong et al., 2018).[9] In particular, there may be a positive correlation between advantageous genetic endowments and family environments in which children receive more investments. Genetic nurture can be controlled by estimating a sibling fixed effects model that relies on within-family variation in PGI^{EA} only (Houmark et al., 2020; Kweon et al., 2020; Selzam et al., 2019); in a non-transmitted genes design, in which both maternal and paternal genetic endowments are included in control vector $\mathbf{X}_i(a)$; or in an adoption design, in which offspring are biologically unrelated to their parents.[10] All approaches, however, are very data-demanding. For example, the sibling design requires a large sample of biological siblings with sequenced DNA data to construct PGI^{EA}. Therefore, it can only be applied to a limited set of existing data sets.

In this study, we estimate a between-family model using an extensive set of pre-determined family background characteristics to control for genetic nurture effects. This approach is standard in the literature and aims to approximate condition (i) while maximizing statistical power to estimate the gene-environment interaction κ (Domingue et al., 2020). We formally assess the residual potential for confounding through genetic nurture effects by comparing the estimates of α from the between-family model to a sibling fixed effects model that we estimate on a subset of our data ($N = 677$, Table 2). Reassuringly, both point estimates are precisely estimated; yet we cannot reject the null hypothesis that they are equal ($p = 0.863$). This result suggests that after conditioning on $\mathbf{X}_i(a)$, residual genetic nurture is close to zero and unlikely to overturn our main findings.

(ii) Exogenous variation in Q^S. School quality is not exogenous to family characteristics as parents choose schools for their children (Altonji et al., 2005; Beuermann et al., 2022; Rothstein, 2006). As a consequence, Q^S is a function of family and child characteristics that may correlate with family inputs $I^F_{i1}, ..., I^F_{ia}$. Hence, in estimation model (7), β and κ may be confounded by *selection effects* (Altonji et al., 2005; Altonji and Mansfield, 2018). In particular, there may be a positive correlation between school quality and family environments in which children receive more investments. Sorting into schools can be controlled in quasi-experimental settings,

[9] A related concept is passive gene-environment correlation (Plomin et al., 1977). In contrast to genetic nurture, passive gene-environment correlation is a narrower concept that focuses only on the intentional reaction of the environment to the genetic endowments of individuals. We will only refer to genetic nurture in the following.

[10] See Demange et al. (2020) for a detailed comparison of all three approaches.

e.g., by using variation from admission lotteries (Angrist et al., 2016; Cullen et al., 2006), or the geographic design of catchment areas (Laliberté, 2021). Existing data sets that avail such variation, however, do not contain sequenced DNA data which is needed to measure genetic endowments at the individual level.

In this study, we use an extensive set of pre-determined family background characteristics to account for selection into schools. Furthermore, we apply the control function approach proposed by Altonji and Mansfield (2018). In their paper, Altonji and Mansfield (2018) consider settings where individuals sort into treatments, e.g., when families move to neighborhoods based on school quality. They show that group-level averages of observable characteristics are correlated with unobservable characteristics. Based on this insight they suggest that controlling for a limited number of group-level averages may remove all cross-group variation in both observable and unobservable characteristics. Agrawal et al. (2019) extend this result to settings where the variable of interest is an interaction between group-level factors, such as Q^S, and observed individual characteristics, such as PGI^{EA}. We provide supporting evidence for their conjecture in our setting by showing that school-level averages for only five student characteristics remove all correlation between educational attainment and an extensive set of measures for family background, including parental education and wages (Table 2). In addition, we formally assess the sensitivity of our results to residual confounding by calculating summary statistics for selection on unobservables (Cinelli and Hazlett, 2020; Oster, 2019). Reassuringly, these summary statistics consistently point to a low potential for selection on unobservables (Supplementary Figure S.2). These results suggest that after conditioning on $\mathbf{X}_i(a)$, residual confounding due to selection into schools is low and unlikely to overturn our main findings. We further support this assertion in Appendix C. In this appendix, we show analytically and based on simulations that any residual confounding due to positive selection into schools is likely to attenuate our estimate for the gene-environment interaction κ towards zero.

(iii) Independent variation in PGI^{EA} and Q^S. In addition to conditions (i) and (ii), PGI^{EA} and Q^S have to be distributed independently of each other. A strong correlation between PGI^{EA} and Q^S would imply little variation in PGI^{EA} at different levels of Q^S and vice versa. As a consequence, there would be insufficient variation to identify $\alpha, \beta,$ and κ separately from each other.

To verify that condition (iii) is satisfied, we present empirical evidence that PGI^{EA} and Q^S are indeed distributed independently of each other. This conclusion holds both unconditionally and conditional on $\mathbf{X}_i(a)$ (Figure 2 and Supplementary Figure S.3.). We note that this finding does not imply that parents choose schools randomly. To wit, in Supplementary Table S.1 we show that there is a positive gene-environment correlation concerning peer characteristics like socioeconomic background and educational achievement. However, consistent with existing literature, this pattern suggests that parents select schools based on factors like proximity and peer quality but not necessarily based on the school practices that underpin our measures of school quality (Abdulkadiroğlu et al., 2020; Beuermann et al., 2022).

In summary: an ideal design to estimate the complementarity parameter κ would combine a sibling fixed effects model with experimental variation in school characteristics among children of the same biological parents. We are not aware of any data set that simultaneously includes genetic data at the individual level, a large set of siblings, and quasi-experimental within-family variation in school assignments. Therefore, we approximate the ideal-type conditions with the best data available to us. Within this setting, causal identification of α, β, and κ relies on rather strong assumptions. We provide extensive empirical evidence that supports the satisfaction of these assumptions and the validity of our research design. Yet, in the absence of clear-cut quasi-experimental variation, we choose to err on the side of caution and speak of *associations* instead *causal effects* in the remainder of the paper.

4 DATA

We use data from the National Longitudinal Study of Adolescent to Adult Health (Add Health), a 5-wave panel study that focuses on the determinants of health-related behaviors and health outcomes. Add Health is a nationally representative sample of adolescents enrolled in grades 7–12 in 1994/95. Initial information (wave 1, $N = 20,745$) was collected from a stratified sample of 80 high schools across the United States as well as from their associated feeder schools. In addition to in-depth interviews with adolescents, questionnaires were administered to school representatives, parents, and roughly 90,000 students of the sampled schools. Follow-up in-home questionnaires were collected in 1996 (wave 2, $N = 14,738$), 2001/02 (wave 3, $N = 15,179$), and 2008/09 (wave 4, $N = 15,701$). In the most recent wave (2016/18, $N = 12,300$), Add Health respondents are between 33 and 43 years old.

In the following, we describe our main variables of interest. Detailed descriptions of all variables used in our analysis are disclosed in Supplementary Material D.

Outcomes. We measure educational attainment Y_i by the total number of years of education after age 27. In each wave, respondents were asked about their highest level of education at the time of the interview. For each individual, we use the most recent information and transform education levels into years of education, following the mapping suggested by Domingue et al. (2015).[11]

To analyze the mechanisms behind our headline results, we additionally use a series of measures for (non-)cognitive skills, health, and academic degrees. First, measures for (non-)cognitive

[11] Numeric values in parentheses: eighth grade or less (8), some high school (10), high school graduate (12), GED (12), some vocational/technical training (13), some community college (14), some college (14), completed vocational/technical training (14), associate or junior college degree (14), completed college (16), some graduate school (17), completed a master's degree (18), some post-baccalaureate professional education (18), some graduate training beyond a master's degree (19), completed post-baccalaureate professional education (19), completed a doctoral degree (20).

skills and health serve as proxy variables for θ_{iA} and allow us to analyze the dimensions of skill development that drive the main findings on educational attainment. We proxy cognitive skills using the Picture Vocabulary Test (PVT), a test of receptive hearing vocabulary that is a widely-used measure of verbal ability and scholastic aptitude. We proxy non-cognitive skills by self-reported measures of general risk aversion and patience (Falk et al., 2018) and self-reported information on the Big Five personality traits (Almlund et al., 2011). Regarding health, we use quality-adjusted life years (QALY), which we derive from self-assessed health measures as well as a summary index of diagnosed health conditions. Second, academic degrees allow us to investigate at which educational stage our results emerge. We focus on whether respondents finished high school, obtained a college degree, or completed a post-graduate degree.

Genetic endowments. Add Health obtained saliva samples from consenting participants in wave 4. After quality control procedures, genotyped data is available for 9,974 individuals and 609,130 SNPs. Add Health uses this data to calculate different PGIs using summary statistics from existing GWAS. We use a PGI for educational attainment, referred to as PGI^{EA}, that is based on the GWAS by Lee et al. (2018).[12]

Lee et al. (2018) perform a meta-analysis of 71 quality-controlled cohort-level GWAS. Their meta-analysis produced association statistics for around 10 million SNPs, of which 1,271 reached genome-wide significance. Genes near these genome-wide significant SNPs are relevant for the central nervous system, and many of them encode proteins that carry out neurophysiological functions such as neurotransmitter secretion or synaptic plasticity. They are relevant for brain-development processes before and after birth.

PGI^{EA} is highly predictive of educational attainment and has been widely used in existing studies. Lee et al. (2018) suggest that PGI^{EA} is a better predictor for years of education than household income. Including the score in a regression of years of education on a set of controls yields an incremental R^2 of 0.127 in the Add Health sample. Among other uses, PGI^{EA} has been used to study the formation of early childhood skills (Belsky et al., 2016; Houmark et al., 2020), educational attainment (Domingue et al., 2015), earnings (Papageorge and Thom, 2020), wealth accumulation (Barth et al., 2020), and social mobility (Belsky et al., 2018).

School quality. In waves 1 and 2, Add Health administered detailed questionnaires to head-masters of Add Health schools. The schools are also linked to administrative data from the Common Core of Data (CCD) and the Private School Survey (PSS). We use these sources to construct an indicator for Q^S using a principal component analysis that includes the following

[12] Lee et al. (2018) construct PGI^{EA} for two prediction cohorts, Add Health and the Health and Retirement Study (HRS). PGI^{EA} is based on results from the meta-analysis in which these two cohorts were excluded from the discovery sample. PGI^{EA} was generated from HapMap3 SNPs using the software LDpred—a Bayesian method that weights each SNP by the posterior mean of its conditional effect given other SNPs.

school-level inputs (component weights in parenthesis): (i) the share of teachers with a master's degree (+0.63), (ii) the share of teachers with school-specific tenure of more than five years (+0.62), (iii) the share of new teachers in the current school year (−0.36), and (iv) the average class size (−0.31). As an alternative to the principal component analysis, we also construct an indicator for Q^S using the aggregation method proposed by Anderson (2008) and Kling et al. (2007).

We focus on school inputs for which there is strong evidence in the existing literature on how they enter the education production function. Clotfelter et al. (2010) and Jacob et al. (2018) demonstrate that academic credentials, which we proxy by the share of teachers with a master's degree, are positively associated with teacher effectiveness. Papay and Kraft (2015) and Rockoff (2004) show that teaching experience, which we proxy by the share of teachers with more than five years of tenure, correlates with teacher performance and student achievement. Hanushek et al. (2016), Hill and Jones (2018), Hwang et al. (2021), and Ronfeldt et al. (2013) demonstrate that a high teacher turnover, which we proxy by the share of new teachers, impairs teaching quality and student achievement. Finally, the large literature on class size reductions (Angrist and Lavy, 1999; Angrist et al., 2019; Chetty et al., 2011; Fredriksson et al., 2013; Krueger, 1999; Leuven and Løkken, 2020) finds either positive or zero associations with student achievement.

Control variables. Add Health provides extensive information on the environments to which respondents were exposed during childhood and in school. We approximate the identification prerequisites discussed in section 3 by choosing a vector $\mathbf{X}_i(a)$ that includes an extensive set of predetermined variables for child and family characteristics as well as school-level averages of observable characteristics to implement the control function approach by Altonji and Mansfield (2018).

We control for child characteristics by including variables for child age at the time of the survey (in months), biological sex, and their interaction as well as an indicator for firstborns. Furthermore, we follow standard practice in the literature and account for population stratification in genetic endowments by including the first 20 principal components of the full matrix of genetic data.

We control for family background characteristics by including maternal and paternal education (in years), the family's religious affiliation (Christian/non-Christian), parental birthplace (US/non-US), and maternal age at birth (in years). Furthermore, we include the mean and standard deviation of potential wages for both mothers and fathers across child age 0–14.[13] Furthermore, all estimations include a vector of state-fixed effects. Note that we focus on predeter-

[13]Note that Add Health contains information on actual income. However, actual income may be a bad control as it reflects parental responses to both PGI^{EA} and Q^S. Therefore, we follow the procedure of Shenhav (2021) and combine data from the 1970 Census and the March Current Population Survey (1975–2000) to construct potential wages for gender/education/census region/race/ethnicity cells and match these potential wages to parents of children aged $a = 1, ..., 14$.

mined variables—variables that are fixed before the period of observation—to avoid smearing through "bad controls" (Angrist and Pischke, 2009). However, in robustness analyses, we expand the vector of controls to include potentially endogenous parental investments and family income. Our results remain unaffected.

To further account for selection into schools, we include a control function based on group characteristics that are observable at the school level. The in-school questionnaires in wave 1 of Add Health collect information about approximately $90,000$ students in Add Health schools. We use these data to construct five variables for the control function: the share of white peers, the share of peers with single mothers, the average years of education of peers' mothers, and the average GPAs in English and Math (standardized by grade and federal state). To avoid mechanical relationships, all school-level averages are calculated excluding the respondent's school cohort.

Analysis sample. We apply the following sample selection criteria. First, we restrict our sample to genotyped respondents of European descent. This is common practice in the literature because GWAS are predominantly conducted on this ancestry group. As a consequence, there is a lack of statistical power to account for population stratification between ancestry groups and estimates of genetic influence would be biased without this restriction (Martin et al., 2017; Ware et al., 2017).

Second, our main subsample comprises individuals who visited an Add Health high school or an associated feeder school in wave 1. However, for a subset of individuals from the feeder schools, we have no information about whether they transferred from their feeder schools to the designated Add Health school. In a robustness check presented in Supplementary Table S.2, we drop these individuals from the sample. In another robustness check, we additionally exclude respondents for whom we do not have information on whether they graduated from an Add Health high school. We note that neither excluding potential movers before high school nor potential movers during high school overturns our main conclusions.

Third, we drop all observations with missing information in Y_i, G_i, Q^{S} , and $\mathbf{X}_i(a)$ by list-wise deletion.

Applying these restrictions, we obtain a sample of $4,036$ individuals from 72 high schools across the United States, for which we provide summary statistics in Table 1. 54% are female, and the average age measured at wave 1 equals $\approx$ 16 years (192 months). The average educational attainment in our sample is 14.7 years which exceeds the average educational attainment in the parental generation by $\approx$ 1.1 years. 96% graduate from high school, which is not surprising given that our sample is restricted to individuals of European descent who attended an Add Health high school or an associated feeder school in wave 1. The 2-year college completion rate equals $\approx$ 50%.

TABLE 1 – Summary Statistics

	N=4,036; Siblings=677; High Schools=72			
	Mean	SD	Min	Max
Educational Attainment				
Years of Education	14.68	2.27	8.00	20.00
High School Degree	0.96	0.20	0.00	1.00
2-year College Degree	0.50	0.50	0.00	1.00
4-year College Degree	0.39	0.49	0.00	1.00
Post-Graduate Degree	0.14	0.35	0.00	1.00
Child and Family Characteristics				
PGI^{EA}	0.00	1.00	-4.18	3.40
Female	0.54	0.50	0.00	1.00
Firstborn	0.48	0.50	0.00	1.00
Age in Months (Wave 1)	192.41	19.62	144.00	256.00
Maternal Age at Birth	25.33	4.84	16.00	46.08
Christian	0.82	0.38	0.00	1.00
Education Mother (in Years)	13.54	2.48	0.00	19.00
Education Father (in Years)	13.56	2.68	0.00	19.00
Foreign-born Mother	0.03	0.17	0.00	1.00
Foreign-born Father	0.03	0.16	0.00	1.00
Potential Wage/Hour Mother	12.57	1.39	9.40	14.27
Potential Wage/Hour Father	15.40	1.32	11.14	17.11
School Quality Indicators				
Q^{S}	0.00	1.00	-2.79	1.82
Teacher w/ MA (%)	51.20	24.10	0.00	95.00
Experienced Teacher (%)	66.66	23.43	0.00	98.00
New Teacher (%)	7.87	7.28	0.00	47.00
Class Size	24.40	4.50	12.00	38.00

Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This table shows summary statistics for the core analysis sample. The sample is restricted to genotyped individuals who (i) are of European descent, and (ii) attended an Add Health high school or an associated feeder school in wave 1. Observations with missing information in any of the displayed variables are dropped by list-wise deletion.

To assess sample representativeness, we compare our analysis sample to the 1974–1983 birth cohorts of non-Hispanic Whites in the American Community Survey (ACS) and the Current Population Survey (CPS) (Supplementary Table S.3). This comparison shows a slight over-representation of females and children of young mothers in our sample. Otherwise, our sample is comparable to the corresponding groups in the ACS and CPS. In robustness analyses, we re-weight our analysis sample to match the ACS and CPS concerning gender composition, educational attainment of parents, and the age of mothers at birth. Our results remain unaffected (Supplementary Table S.2).

5 RESULTS

We present our results in four steps. In section 5.1, we discuss the association of educational attainment, genetic endowments, and school quality in light of the identifying assumptions discussed in section 3. In section 5.2, we present our estimates for the gene-environment interaction κ. After a robustness analysis in section 5.3, we conclude with an analysis of mechanisms in section 5.4. In all analyses, we standardize PGI^{EA}, Q^{S}, and the variables in $\mathbf{X}_i(a)$ so that they have a mean of zero ($\mu = 0$) and a standard deviation of one ($\sigma = 1$).

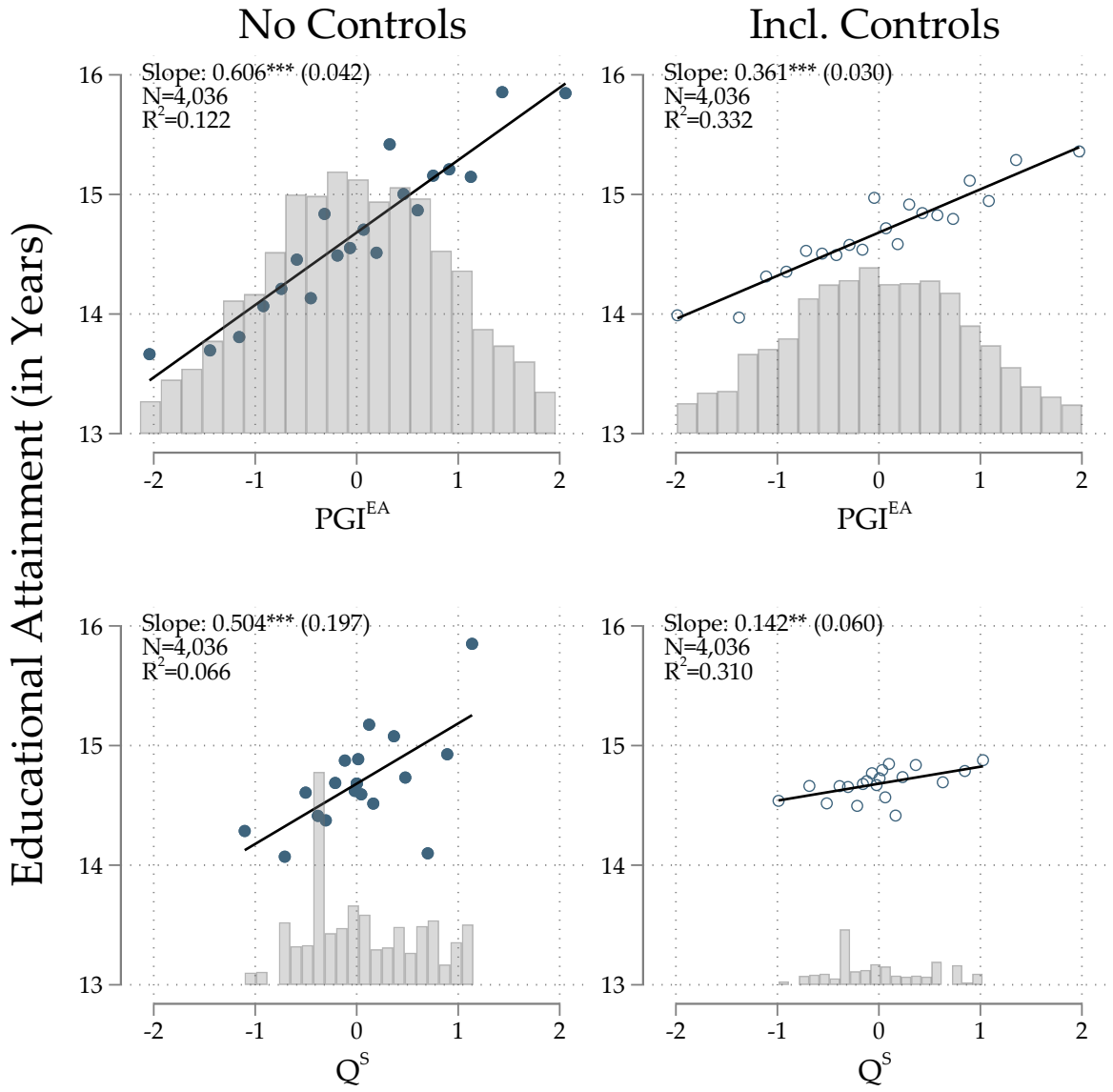
5.1 *The Association of Educational Attainment with Genetic Endowments and School Quality*

Figure 1 displays the association of educational attainment with our measures for genetic endowments PGI^{EA} and school quality Q^{S}. In the left column, we show raw correlations that do not account for the control variables $\mathbf{X}_i(a)$. In the right column, we show associations conditional on $\mathbf{X}_i(a)$.

First, PGI^{EA} is highly predictive of educational attainment. Without controls, an increase in PGI^{EA} by one standard deviation (1 SD) is associated with an increase in educational attainment by 0.606 years. This association does not have a causal interpretation as it may be confounded by genetic nurture effects. When we control for pre-determined child and family characteristics as well as the control function, a 1 SD increase in PGI^{EA} is associated with an increase in educational attainment by 0.361 years. This decrease is consistent with sibling studies showing that genetic nurture effects usually account for 40–50% of the raw association between PGI^{EA} and educational attainment (Kweon et al., 2020; Muslimova et al., 2020; Ronda et al., 2022; Selzam et al., 2019).

Is $\mathbf{X}_i(a)$ sufficient to control for genetic nurture effects? We test whether there remains confounding due to genetic nurture by comparing estimates of the between-family model with a sibling fixed effects model that we estimate on a subsample of our data ($N = 677$). The within-family comparison allows us to perfectly control for genetic nurture effects. Therefore, a strong divergence of between- and within-family estimates would suggest that there is residual genetic nurture that is not picked up by $\mathbf{X}_i(a)$. In Table 2 we show that this is not the case. In the sibling sample, the between-family estimate yields a point estimate of 0.419 after controlling for $\mathbf{X}_i(a)$. The within-family comparison yields a point estimate of 0.445. Both point estimates are precisely estimated and very close to each other. We cannot reject the null hypothesis of their equality at conventional levels of statistical significance ($p = 0.863$). This result suggests that after conditioning on $\mathbf{X}_i(a)$, residual genetic nurture is very low, lending credence to our research design.

FIGURE 1 – Association of Educational Attainment with PGI^{EA} and Q^S



Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows the correlation of completed years of education with PGI^{EA} and Q^S, respectively. We bin scatterplots using 20 quantiles of the variable of interest. Gray bars indicate density distributions of the (residualized) variable of interest. Black lines are fitted from linear regressions of educational attainment on the variable of interest. In the left-column, we only control for state fixed effects. In the right column, we introduce the full set of control variables. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the school level.

Second, Q^S is highly predictive of educational attainment. Without controls, a 1 SD increase in Q^S is associated with an increase in educational attainment by 0.504 years. This association does not have a causal interpretation as it may be confounded by selection effects due to sorting into schools. When we control for pre-determined child and family characteristics as

TABLE 2 – Testing for Genetic Nurture and Selection into Schools

	Years of Education: Between- vs. Within-Family		Predicted Years of Education: w/o vs. w/ Control Function	
	(1)	(2)	(3)	(4)
PGI^{EA}	0.419*** (0.086)	0.445*** (0.142)	–	–
Q^S	–	–	0.263** (0.129)	0.005 (0.055)
Equality of coefficients (p-value)	0.863		0.024	
Child Controls	✓	×	×	×
Family Controls	✓	×	×	×
Control Function	✓	×	×	✓
Sibling Fixed Effect	×	✓	×	×
N	677	677	4,036	4,036
R^2	0.416	0.760	0.084	0.181
Outcome Mean	14.722	14.722	14.681	14.681
Outcome SD	2.277	2.277	1.163	1.163

Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This table shows the associations of PGI^{EA} and Q^S with years of education (left panel) and predicted years of education (right panel). The left panel shows estimates in the sibling sample: Column (1) displays results from a between-family comparison. Column (2) displays results from a within-family comparison. The right panel shows estimates in the core analysis sample: Column (3) displays results without any controls. Column (4) displays results including the control function variables. Predicted education is calculated from a regression of completed years of education on all *Child Controls* and *Family Controls*. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. All right-hand side variables are standardized on the estimation sample ($\mu = 0$, $\sigma = 1$). Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the school level.

well as the control function, a 1 SD increase in Q^S is associated with an increase in educational attainment of 0.142 years. This decrease reflects positive selection into schools—a pattern that has been thoroughly documented in the existing literature for the United States (Deming et al., 2014; Rothstein, 2006). Nevertheless, even when accounting for selection, the association of Q^S and educational attainment remains strong and positive. This result is consistent with prior literature showing positive effects of high school quality on students' educational success (Angrist et al., 2019; Deming, 2014; Hanushek and Rivkin, 2010; Jackson et al., 2020).

Is $\mathbf{X}_i(a)$ sufficient to control for selection into schools? To control for sorting into schools, we follow the control function approach suggested in Altonji and Mansfield (2018) and include school-level averages of observed individual characteristics into $\mathbf{X}_i(a)$. We test the effectiveness of this control function approach as follows: First, we replace educational attainment with

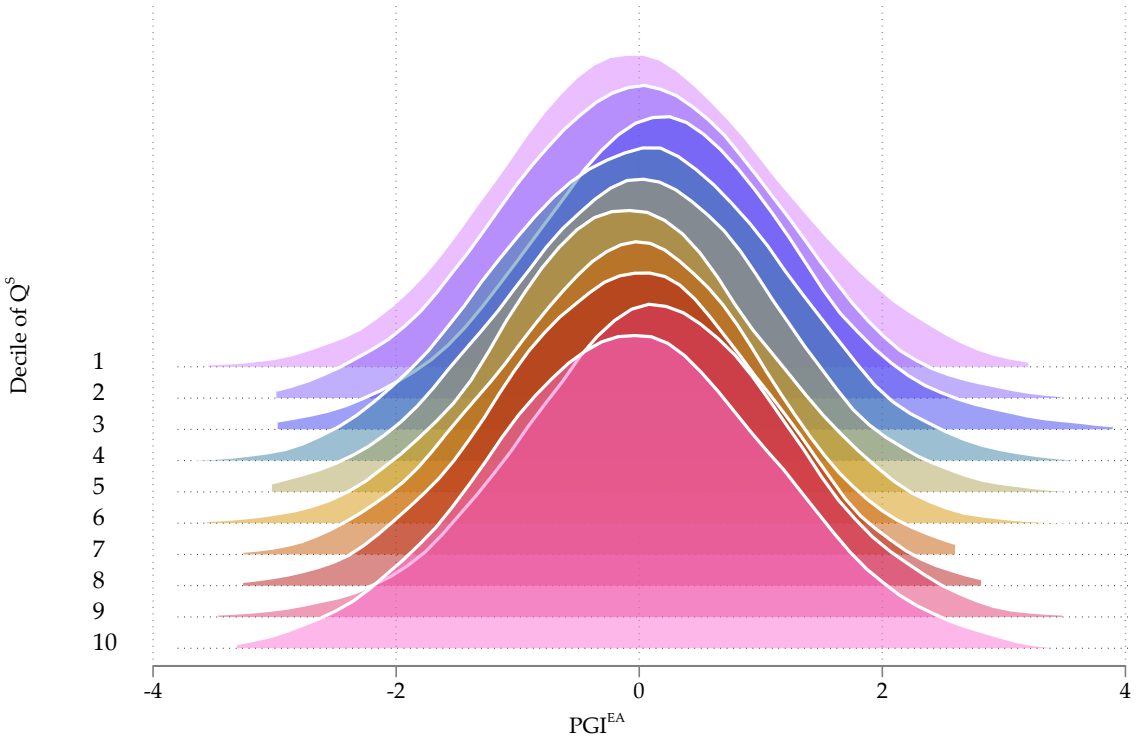
predicted educational attainment. In particular, we predict years of education using the full set of pre-determined child and family characteristics included in $\mathbf{X}_i(a)$. These characteristics are highly predictive of years of education ($R^2 = 0.297$) and other measures of educational attainment and skills (Supplementary Figure S.4). Second, we assess the correlation between predicted educational attainment and Q^S before and after conditioning on the variables of the control function. If the control function approach is effective, any positive correlation between predicted educational attainment and Q^S should be attenuated to zero after conditioning on the variables of the control function. In Table 2, we show that the unconditional association between predicted educational attainment and Q^S is 0.263 indicating significant positive selection into schools. Conditional on the control function, this correlation shrinks to 0.005 and becomes insignificant. This result suggests that the control function approach indeed is effective in controlling for selection into schools. Therefore, conditional on $\mathbf{X}_i(a)$, residual sorting into schools is likely to be low, lending further credence to our research design.

Alternatively, we can assess potential confounding through remaining selection effects by assuming that changes in the coefficient of Q^S due to the introduction of $\mathbf{X}_i(a)$ provide information about the extent of confounding due to unobservables (Altonji et al., 2005; Cinelli and Hazlett, 2020; Oster, 2019). We follow Cinelli and Hazlett (2020) and assess what association unobserved confounders would need to have with both the variable of interest (Q^S) and the outcome of interest (educational attainment) to change our conclusions. In Supplementary Figure S.2, we show that Q^S would remain positive and statistically significant at the 5%-level even if the partial R^2 of unobserved confounders with Q^S and educational attainment were more than ten times higher than the corresponding partial R^2 of paternal education with these variables. Given the decisive role of parental education in school choices, and its strong predictive power for educational outcomes of children, these results lend further confidence that our results are genuine and not a mere reflection of selection into schools based on family background.

In addition to genetic nurture effects and selection effects, a high correlation between PGI^{EA} and Q^S would pose another threat to the identification of the gene-environment interaction. Such a correlation could arise if children were sorted into schools based on their genetic endowments. Figure 2 shows that this concern does not apply in our setting. In this figure, we plot unconditional PGI^{EA} distributions by decile of Q^S. Visual inspection suggests that PGI^{EA} distributions are almost congruent to each other within each decile of the school quality indicator. This conclusion also holds after residualizing PGI^{EA} and Q^S from $\mathbf{X}_i(a)$ (Supplementary Figure S.3). More formally, we compute two-sample Kolmogorov-Smirnov tests for the equality of PGI^{EA} distributions within the deciles of Q^S. Out of 45 pairwise comparisons, only 2 differences are significant at the 10% level. This result is expected by chance and hence we conclude that PGI^{EA} and Q^S are indeed independently distributed.[14]

[14] See also Supplementary Table S.1, where we show that PGI^{EA} is correlated with the socioeconomic composition of schools. To the contrary, we cannot reject the null hypothesis of zero correlation between PGI^{EA} and Q^S.

FIGURE 2 – Distribution of PGI^{EA} by Q^S



Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows unconditional PGI^{EA} distribution by deciles of Q^S. Density distributions are smoothed using the Epanechnikov kernel function with a bandwidth of 0.5.

5.2 The Interplay of Genetic Endowments and School Quality

Table 3 shows our baseline estimates for the interaction of genetic endowments and school quality. In all regressions, we include the vector $\mathbf{X}_i(a)$ to control for genetic nurture and selection into schools. As noted previously, $\mathbf{X}_i(a)$ comprises an extensive set of pre-determined child and family characteristics as well as the control function using school-level averages of observed individual characteristics (see section 4).

The point estimates in column (1) replicate the findings from Figure 1 and show a strong and positive association of PGI^{EA} and Q^S with educational attainment. A 1 SD increase in PGI^{EA} (Q^S) increases educational attainment by $\approx$ 0.36 ($\approx$ 0.13) years. The coefficient of the interaction $PGI^{EA} \times Q^S$ is our estimate of κ. The negative interaction coefficient indicates that genetic endowments and school quality are *substitutes* in the production of educational attainment: a 1 SD increase in school quality reduces the positive association of educational attainment with PGI^{EA} by $\approx$ 17% (= 0.062/0.360).

In column (2), we replicate the analysis for an alternative way of constructing the school quality indicator. Instead of using weights from principal component analysis, we aggregate dimen-

TABLE 3 – Association of Years of Education with PGI^{EA} and School Environments

	Overall		Decomposition of Q^S			
Outcome: Years of Education	PCA (1)	Anderson (2008) (2)	(3)	(4)	(5)	(6)
PGI^{EA}	0.360*** (0.028)	0.361*** (0.029)	0.360*** (0.027)	0.361*** (0.029)	0.361*** (0.030)	0.361*** (0.030)
Q^S	0.128** (0.054)	0.095** (0.047)	–	–	–	–
$PGI^{EA} \times Q^S$	-0.062** (0.026)	-0.059** (0.027)	–	–	–	–
Teacher w/ MA	–	–	0.169*** (0.065)	–	–	–
$PGI^{EA} \times$ Teacher w/ MA	–	–	-0.065** (0.026)	–	–	–
Exp. Teacher	–	–	–	0.083* (0.049)	–	–
$PGI^{EA} \times$ Exp. Teacher	–	–	–	-0.043* (0.026)	–	–
New Teacher	–	–	–	–	-0.001 (0.045)	–
$PGI^{EA} \times$ New Teacher	–	–	–	–	0.036 (0.029)	–
Class Size	–	–	–	–	–	-0.004 (0.038)
$PGI^{EA} \times$ Class Size	–	–	–	–	–	-0.006 (0.030)
Child Controls	✓	✓	✓	✓	✓	✓
Family Controls	✓	✓	✓	✓	✓	✓
Control Function	✓	✓	✓	✓	✓	✓
N	4,036	4,036	4,036	4,036	4,036	4,036
R^2	0.334	0.334	0.334	0.333	0.333	0.332
Outcome Mean	14.681	14.681	14.681	14.681	14.681	14.681
Outcome SD	2.268	2.268	2.268	2.268	2.268	2.268

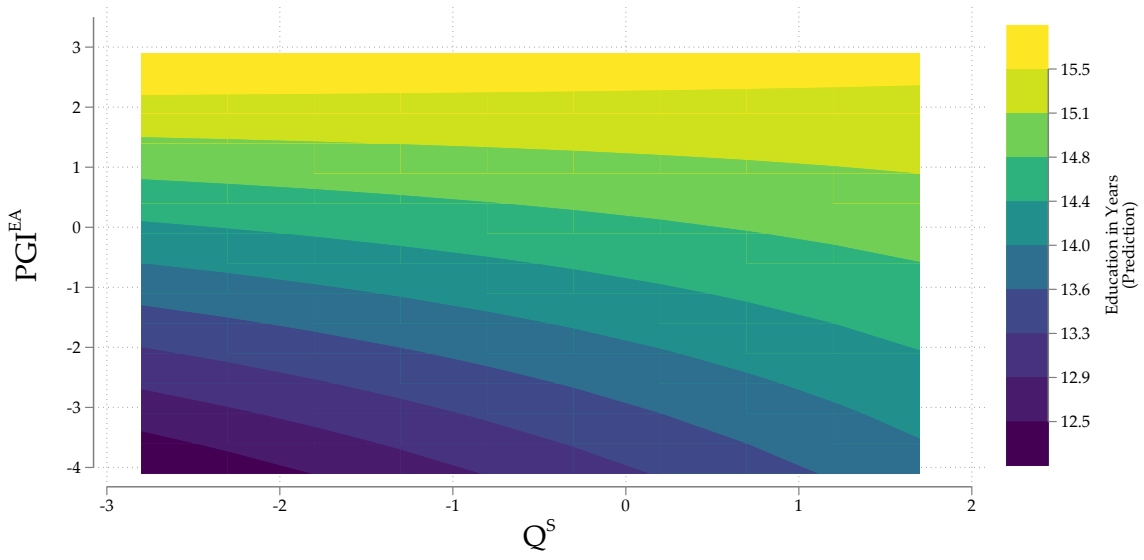
Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This table shows the joint association of PGI^{EA} and Q^S with completed years of education. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. All right-hand side variables are standardized on the estimation sample ($\mu = 0$, $\sigma = 1$). Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the school level.

sions using the method of Anderson (2008) and Kling et al. (2007). Our results are not sensitive to this alternation: a 1 SD increase in school quality reduces the positive association of educational attainment with PGI^{EA} by $\approx 16\%$ (= 0.059/0.361).

In columns (3)-(6), we decompose the overall index Q^S into its underlying components. Columns (2) and (3) show that school-specific teacher tenure and teacher education have positive average effects on educational attainment. These results are consistent with previous literature showing that teaching experience and teacher education tend to have positive effects on student learning (Hill and Jones, 2018; Hwang et al., 2021; Jackson and Bruegmann, 2009; Rockoff, 2004). Furthermore, there is evidence for substitutability, i.e., the positive impact of more experienced and more educated teachers is larger for students in the lower parts of the PGI^{EA} distribution. To the contrary, Columns (4) and (5) show that class size and the percentage of new teachers neither increase average educational attainment nor is there evidence for heterogeneity across the PGI^{EA} distribution. Overall, the decomposition suggests that the results for the overall index Q^S are mostly driven by teachers' school-specific tenure and education levels.

In principle, the negative gene-environment interaction shown in Table 3 could be due to low PGI^{EA} students benefiting from high-quality schools, or high PGI^{EA} students losing from high-quality schools. In Figure 3, we provide evidence for the former, but not for the latter. In this figure, we show predictive margins of years of education for different combi-

FIGURE 3 – Association of Years of Education with PGI^{EA} by Q^S



Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows predictions of completed years of education by PGI^{EA} and Q^S cell. Predictions are calculated using the model estimated in column (1) of Table 3.

nations of PGI^{EA} and Q^S while controlling for $\mathbf{X}_i(a)$. Moving horizontally from left to right at a given PGI^{EA} level, we see that predicted education increases sharply in the lower parts of the PGI^{EA} distribution. On the contrary, in the upper ranges of the PGI^{EA} distribution, predicted education remains unchanged regardless of school quality. This pattern is encouraging

as it suggests that investments in school quality mitigate genetic inequality in educational outcomes without compromising the attainment of genetically advantaged students.

It is interesting to contrast this substitutability result with existing evidence on Scarr-Rowe interactions. In this literature, researchers tend to find (weakly) positive gene-environment interaction between PGI^{EA} and parental socio-economic status, i.e., in the family context genetically advantaged children benefit more from environments associated with higher investment into children.[15] To the contrary, we find that sign of interplay between genetic endowments and high-investment environments reverses in the school context, i.e., genetically disadvantaged children benefit more from environments associated with higher investment into children.

5.3 Robustness Analysis

We check the robustness of our results in two steps. First, we investigate whether Q^S picks up the effects of other school characteristics that may correlate with student outcomes. Second, we test whether our estimates of the gene-environment interaction κ are confounded by the heterogeneity of genetic effects in different family environments.

Other school characteristics. In Table 4, we focus on potential confounders at the school-level. For ease of comparison, we replicate our baseline estimates in column (1).

In columns (2)–(4), we address the concern that our index for school quality may be conflated with other school policies and practices that have an impact on educational attainment. To this end, we sequentially introduce indicators for other school policies as well as their interaction with PGI^{EA}. Indicators for other school policies include the average share of retained students across grades 9-11 (column 2), a binary indicator of whether schools group students by English ability (column 3), and an index for the strictness of school sanctions (column 4).[16] Comparing columns (2)–(4) to our baseline estimate, we see that all coefficients of interest are robust to the inclusion of these indicators.

In columns (5)–(7), we address the concern that our index for school quality may be conflated with the demographic composition of teachers at a school. This possibility arises if teachers sort into particular schools based on student characteristics (Jackson, 2009a). To this end, we

[15] See for example Ronda et al. (2022) and Turkheimer et al. (2003) for positive Scarr-Rowe interactions and Figlio et al. (2017) for a null finding. We are not aware of any studies finding negative effects.

[16] The strictness index is based on headmaster questionnaires. Headmasters were asked about the school's policy in the following domains of behavior: cheating, fighting with or injuring another student, alcohol or drug possession, drinking alcohol or using illegal drugs, smoking, verbally or physically abusing a teacher, and stealing school property. Possible measures are (i) no action, (ii) verbal warning, (iii) minor action, (iv) in-school suspension, (v) out-of-school suspension, and (vi) expulsion. Following Anderson (2008) and Kling et al. (2007) we standardize each response dimension to $\mu = 0$ and $\sigma = 1$ and aggregate them linearly to obtain the strictness index. See Supplementary Material D for details.

sequentially introduce indicators for teacher demographics as well as their interaction with PGI^{EA}. Indicators for teacher demographics include the share of white teachers (column 5), the share of Hispanic teachers (column 6), and the share of female teachers (column 7). Our results remain robust to the inclusion of these variables.

In column (8), we re-estimate the coefficients of interest while accounting for unobserved differences across schools through the introduction of school fixed effects. In this specification, we cannot estimate the effect of Q^S on educational attainment. However, it is reassuring that the estimates for PGI^{EA} and the gene-environment interaction are very close to our benchmark estimates.

TABLE 4 – Robustness to Additional School Characteristics

	Baseline	+ School Policies			+ Teacher Composition			+ School FE
Outcome: Years of Education	(1)	Retention Policy (2)	Ability Groups (3)	Strict. Index (4)	White Teacher (5)	Hispanic Teacher (6)	Female Teacher (7)	(8)
PGI^{EA}	0.360***	0.360***	0.365***	0.362***	0.360***	0.361***	0.359***	0.352***
	(0.028)	(0.028)	(0.028)	(0.029)	(0.028)	(0.028)	(0.027)	(0.029)
Q^S	0.128**	0.130**	0.119**	0.138**	0.125**	0.127**	0.162***	–
	(0.054)	(0.054)	(0.053)	(0.056)	(0.053)	(0.055)	(0.054)	
$PGI^{EA} \times Q^S$	-0.062**	-0.062**	-0.061**	-0.058**	-0.060**	-0.062**	-0.060**	-0.063**
	(0.026)	(0.025)	(0.025)	(0.027)	(0.026)	(0.026)	(0.026)	(0.027)
School Characteristic	–	0.047	-0.058	0.061*	-0.006	0.100***	-0.096***	–
		(0.032)	(0.054)	(0.033)	(0.065)	(0.032)	(0.037)	
$PGI^{EA} \times$ School Characteristic	–	-0.017	0.030	0.019	-0.021	0.007	0.040	–
		(0.029)	(0.030)	(0.024)	(0.032)	(0.022)	(0.033)	
Child Controls	✓	✓	✓	✓	✓	✓	✓	✓
Family Controls	✓	✓	✓	✓	✓	✓	✓	✓
Control Function	✓	✓	✓	✓	✓	✓	✓	✓
N	4,036	4,036	3,971	4,036	4,036	4,036	4,036	4,036
R^2	0.334	0.334	0.335	0.334	0.334	0.335	0.335	0.343

Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This table shows the joint association of PGI^{EA} and Q^S with completed years of education. We control for additional school characteristics and their interaction with PGI^{EA}. The relevant school characteristics are indicated in the column header. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. All right-hand side variables are standardized on the estimation sample ($\mu = 0$, $\sigma = 1$). Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the school level.

Overall, these results suggest that our estimates of the gene-environment interaction κ are not confounded by other school environments that are not reflected in our school quality index.

Family environments and behavioral responses. In Table 5 we focus on potential confounders at the family-level. For example, highly educated parents may try to compensate their children for poor school quality by helping with homework, providing additional educational resources, etc. Since parental education correlates with PGI^{EA}, such parental responses could also conflate our estimates for the interaction between PGI^{EA} and Q^{S}. We address these concerns in the following.

In column (2), we include additional controls for family investments which we omit from our baseline analysis since they may reflect endogenous family responses to genetic endowments of children and school quality. In particular, we include indexes for breastfeeding, parental time investments, and the log of annual family income in the set of controls.[17] Despite a sizable sample reduction, our results remain unaffected. This result supports our previous conclusion that the set of pre-determined family background characteristics and the control function capture most of the relevant information on family environments that are associated with higher investments in children.

In column (3), we include higher-order polynomials for PGI^{EA} and Q^{S} and allow for all possible interactions between both variables. The expansion to higher-order polynomials is motivated by Biroli et al. (2022) who suggest this procedure to control for heterogeneity in family responses to genetic endowments and school quality. Our results are unaffected by this expansion.

In column (4), we test for potential confounding due to gene-environment interactions with family socio-economic status. In particular, we follow Domingue et al. (2020) and Keller (2014) and extend our estimation model by interacting PGI^{EA} and Q^{S} with the full control vector $\mathbf{X}_i(a)$. In doing so, we allow for the possibility that family socioeconomic status interacts with both genetic endowments and school quality. In this model, the interpretation for the base coefficients of PGI^{EA} and Q^{S} changes because the estimated coefficients now reflect effects for the subgroup of individuals that sit at the mean of all control variables in $\mathbf{X}_i(a)$.[18] This change in interpretation explains the lower coefficient for Q^{S} in comparison to the baseline model. Importantly, however, our estimate for the interaction of genetic endowments and school quality

[17] We collect information on a series of activities that the child has engaged in with either their mother or father over the past four weeks. For both parents, these activities include shopping, playing sports, going to church, talking about dates, going to the movies and similar events, talking about personal problems, talking about school work, working together on school work, and talking about other things at school. Following Anderson (2008) and Kling et al. (2007) we standardize each response dimension to $\mu = 0$ and $\sigma = 1$ and aggregate them linearly to obtain an aggregate index of time investment. See Supplementary Material D for details.

[18] Effects are estimated at the mean of the control variables since we standardize all variables in $\mathbf{X}_i(a)$ to have a mean of zero and a standard deviation of one.

remains unaffected.

TABLE 5 – Robustness to Family Environments and Behavioral Responses

Outcome: Years of Education	Baseline (1)	Endogenous Controls (2)	Higher-Order Polynomials (3)	Full Interaction (4)
PGI^{EA}	0.360***	0.325***	0.386***	0.352***
	(0.028)	(0.031)	(0.036)	(0.028)
Q^S	0.128**	0.169***	0.142**	0.082
	(0.054)	(0.057)	(0.063)	(0.063)
$PGI^{EA} \times Q^S$	-0.062**	-0.072**	-0.069***	-0.065**
	(0.026)	(0.031)	(0.025)	(0.029)
Child Controls	✓	✓	✓	✓
Family Controls	✓	✓	✓	✓
Control Function	✓	✓	✓	✓
Endogenous Controls	×	✓	×	×
Higher-Order Polynomials	×	×	✓	×
Full Interaction	×	×	×	✓
N	4,036	3,355	4,036	4,036
R^2	0.334	0.348	0.334	0.347

Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This table shows the joint association of PGI^{EA} and Q^S with completed years of education. In column (2) we introduce potentially endogenous control variables. Endogenous control variables include an indicator for breastfeeding, an index for maternal time investments, and log family income. In column (3) we control for second-order polynomials of PGI^{EA} and Q^S and allow for all possible interactions of both indicators. In column (4) we control for all possible interactions between PGI^{EA}, Q^S, *Child Controls*, *Family Controls*, and the *Control Function*. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. All right-hand side variables are standardized on the estimation sample ($\mu = 0$, $\sigma = 1$). Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the school level.

Overall, these results suggest that our estimates of the gene-environment interaction κ are not confounded by the heterogeneity in genetic effects across family environments.

Further robustness checks. In the Supplementary Material, we show the results of additional robustness analyses. First, we check whether our results are driven by outlier schools. Therefore, we re-run our analysis 72 times, excluding one school from the sample per iteration. Reassuringly, in each iteration, the results are very close to our benchmark estimates (Supplementary Figure S.5). Second, we check whether our findings are driven by ceiling effects in educational attainment. To this end, we re-run our analysis, artificially censoring educational

attainment step-wise from above. If ceiling effects were driving our results, we would expect the gene-environment interaction to increase across parts of the censoring interval. However, this is not the case. Instead, the corresponding coefficient decreases monotonically (Supplementary Figure S.6). The absence of ceiling effects is further supported by Supplementary Figure S.7. In this figure, we replicate Figure 3 while replacing educational attainment with the PVT, i.e., an educational outcome that is not artificially censored from above. The figure replicates the data pattern of Figure 3 very well and suggests that our results are indeed driven by relative gains of low PGI^{EA} students. Third, we check whether we pick up the relevant genetic variation. In particular, we sequentially control for six alternative PGIs that target other phenotypes than educational attainment as well as their interaction with Q^S. While some PGIs have predictive power over and above PGI^{EA}, our estimates for PGI^{EA}, Q^S and the gene-environment interaction κ remain unaffected (Supplementary Table S.4). Fourth, we re-estimate our baseline model while applying the correction method of Becker et al. (2021) to account for measurement error and resulting attenuation bias in PGI^{EA} and the gene-environment interaction. As expected, the measurement-error-corrected estimates for PGI^{EA} and the gene-environment interaction are substantially higher than in our baseline estimates. However, they increase in roughly equal proportions. As a result, our conclusions about the relative strength of the substitutability of PGI^{EA} and Q^S remain unaffected. In the measurement-error-corrected version, a 1 SD increase in school quality reduces the positive association of educational attainment with PGI^{EA} by $\approx$ 14–15% (Supplementary Table S.5). Lastly, we run a placebo test. That is, we re-estimate our baseline model while permuting the values of Q^S 10,000 times and keeping constant all other variables. In Supplementary Figure S.8, we show that the resulting distribution of t-statistics for the interaction of PGI^{EA} and Q^S is well-behaved and that roughly 1%, 5%, and 10% of the placebo regressions yield t-statistics at the corresponding critical values. Only slightly more than 1% of the placebo regressions yield a t-statistic more extreme than the one from our baseline estimates, suggesting that our results are not driven by chance.

In summary, this battery of additional checks further supports the conclusion that our main findings are genuine.

5.4 Mechanisms

In this section, we analyze the mechanisms underlying the substitutability of genetic endowments and school quality.

Skill formation. In section 3, we formulated educational attainment Y_i as a function of children's skills θ_i at the end of childhood. The skills that influence educational attainment are multidimensional and comprise a broad set of (non-)cognitive skills and health (Almlund et al., 2011; Heckman and Mosso, 2014). Furthermore, current literature shows that each of these skill dimensions is shaped, in part, by genetic influence (Buser et al., 2021a; Demange et al.,

2020, 2021).

We evaluate these potential channels by analyzing the associations of PGI^{EA} and Q^S with a set of intermediate outcomes. In terms of cognitive skills, we use the Picture Vocabulary Test (PVT) as a measure of verbal intelligence. Furthermore, we focus on personality and preferences as two distinct conceptualizations of non-cognitive skills (Becker et al., 2012; Humphries and Kosse, 2017). In particular, we use measures for self-reported risk aversion and patience, and the Big Five personality traits. In terms of health outcomes, we focus on subjective health, measured by quality-adjusted life years (QALY), and objective health, measured by an index that comprises information on whether the respondent is obese, has first-stage hypertension or has high cholesterol. All measures were collected in waves 3 and 4 of Add Health, that is, after respondents had left high school but potentially before they had completed their highest level of education (see Supplementary Material D for details).

Existing literature shows cognitive skills, risk aversion, patience, and health are strong predictors of educational attainment (Burks et al., 2015; Castillo et al., 2018a,b; Jackson, 2009b). Furthermore, openness and emotional stability—the opposite of neuroticism—are positively associated with educational attainment (Becker et al., 2012; Buser et al., 2021b).[19] Based on this evidence, one would expect positive associations of both PGI^{EA} and Q^S with each of these intermediate outcomes. The sign of the gene-environment interaction is a priori unclear. However, given the substitutability of PGI^{EA} and Q^S in the production of educational attainment, we expect similar substitutability patterns for a subset of these intermediate outcomes as well.

Table 6 summarizes the results. In column (1) of Panel (a), we focus on the PVT as a measure of cognitive skills. As expected, our results show positive associations of both PGI^{EA} and Q^S with the PVT. A 1 SD increase in PGI^{EA} (Q^S) is associated with a 0.174 SD (0.123 SD) increase in the PVT. Furthermore, both factors are substitutes for each other. A 1 SD increase in school quality reduces the positive association of PVT and PGI^{EA} by $\approx 28\%$ $(= 0.048/0.174)$.

In columns (2)–(3) of Panel (a), we focus on economic preferences. As expected, we find positive associations of PGI^{EA} and Q^S with both risk aversion and patience. 1 SD increases in PGI^{EA} and Q^S are associated with increases in risk aversion by 0.40 SD and 0.68 SD, respectively. The corresponding increases in patience are 0.86 SD and 0.75 SD. Furthermore, PGI^{EA} and Q^S are substitutes for each other. A 1 SD increase in Q^S reduces the positive associations of risk aversion and patience with the PGI^{EA} by $\approx 105\%$ $(= 0.042/0.040)$ and $\approx 33\%$ $(= 0.029/0.086)$, respectively.

In columns (4)–(5) of Panel (a), we focus on health outcomes. As expected, our results show a positive association of PGI^{EA} with both subjective and objective health. A 1 SD increase in PGI^{EA} increases subjective (objective) health by 0.077 SD (0.042 SD). Furthermore, the nega-

[19] We replicate these findings in our data by showing that each intermediate outcome is highly predictive for educational attainment conditional on our set of controls $\mathbf{X}_i(a)$. See Supplementary Figure S.9.

TABLE 6 – Association of Skill Measures with PGI^{EA} by Q^S

Panel (a)	Cognitive	Preferences		Health	
	PVT	Risk	Patience	Subjective	Objective
	(1)	(2)	(3)	(4)	(5)
PGI^{EA}	0.174***	0.040**	0.086***	0.077***	0.042***
	(0.014)	(0.017)	(0.016)	(0.016)	(0.016)
Q^S	0.123***	0.068**	0.075*	0.038	-0.013
	(0.045)	(0.030)	(0.039)	(0.040)	(0.025)
$PGI^{EA} \times Q^S$	-0.048***	-0.042***	-0.029**	-0.036***	-0.000
	(0.011)	(0.016)	(0.012)	(0.013)	(0.017)
Child Controls	✓	✓	✓	✓	✓
Family Controls	✓	✓	✓	✓	✓
Control Function	✓	✓	✓	✓	✓
N	3,382	3,467	3,467	3,473	4,036
R^2	0.226	0.111	0.104	0.079	0.053

Panel (b)	Personality				
	Open-ness	Conscient-iousness	Extra-version	Agree-ableness	Neuro-ticism
	(1)	(2)	(3)	(4)	(5)
PGI^{EA}	0.075***	-0.008	-0.020	0.042**	-0.086***
	(0.014)	(0.015)	(0.019)	(0.017)	(0.018)
Q^S	0.005	0.045	-0.061**	0.044	-0.037*
	(0.025)	(0.031)	(0.024)	(0.029)	(0.022)
$PGI^{EA} \times Q^S$	-0.014	-0.006	-0.019	-0.002	0.020
	(0.011)	(0.014)	(0.023)	(0.014)	(0.019)
Child Controls	✓	✓	✓	✓	✓
Family Controls	✓	✓	✓	✓	✓
Control Function	✓	✓	✓	✓	✓
N	3,998	4,030	4,029	4,028	4,024
R^2	0.100	0.046	0.035	0.139	0.090

Data: National Longitudinal Study of Adolescent to Adult Health.

Note: Own calculations. This table shows the joint association of PGI^{EA} and Q^S with cognitive skills, preferences, health, and personality. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. All right-hand side variables are standardized on the estimation sample ($\mu = 0$, $\sigma = 1$). Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the school level.

tive coefficient on the interaction of PGI^{EA} and Q^S suggests that this increase is particularly pronounced for low PGI^{EA} students: a 1 SD increase in school quality reduces the positive association of subjective health with the PGI^{EA} by $\approx 47\%$ $(= 0.036/0.077)$.

In Panel (b), we focus on personality traits. We find positive associations of PGI^{EA} with openness and agreeableness, and a negative association of PGI^{EA} with neuroticism. Q^S is associated with decreases in extraversion and neuroticism. However, we find no evidence of an interaction between PGI^{EA} and Q^S in the production of personality traits.

To summarize: we find negative gene-environment interactions between genetic endowments and school quality in the production of cognitive skills, risk aversion, patience, and subjective health. Given their predictive power for educational attainment, these intermediate outcomes are plausible transmission channels for the substitutability of genetic endowments and school quality in the production of educational attainment.

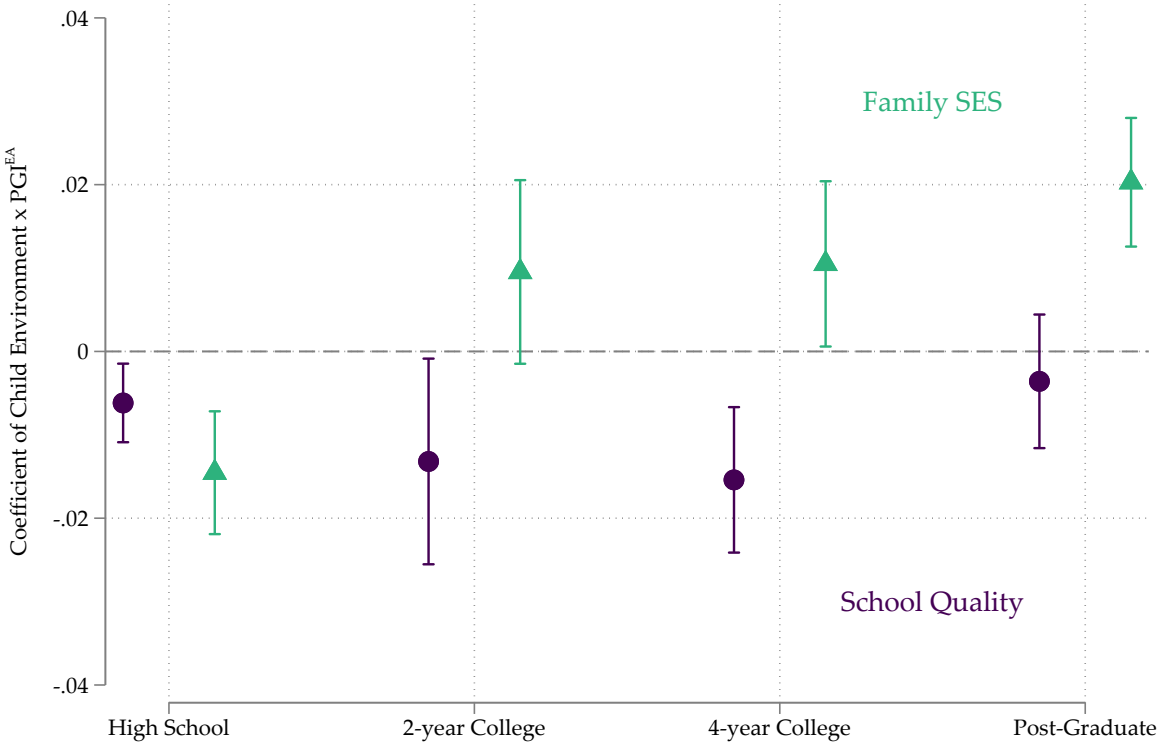
Educational degrees. Total years of education summarizes information from various educational stages, with each stage requiring a different mix of skills θ_i (Cunha et al., 2006, 2010). Therefore, we repeat our analysis by replacing total years of education with binary variables for whether respondents obtained (i) at least a high school degree or GED, (ii) a 2-year college degree, (iii) a 4-year college degree, or (iv) a post-graduate degree.

In Figure 4, we present the resulting point estimates for the gene-environment interaction κ and the associated confidence bands. The circular series suggests that the substitutability of school quality and genetic endowments follows a U-shaped pattern throughout the educational life cycle. For students with low PGI^{EA}, there is a small decrease in the probability of dropping out of high school if they attend high-quality schools, followed by increases in substitutability for 2-year and 4-year college degrees. The substitutability of high-quality schools and genetic endowments levels off at the post-graduate level. This pattern is consistent with the following interpretation: high school graduation is a relatively "inclusive" educational outcome that is accessible to most, including low PGI^{EA} students in low-quality schools. Evidence of this effect is the high school graduation rate of 96% in our sample (Table 1). In contrast, post-graduate education is a relatively "exclusive" educational outcome that is more accessible to students who have advantageous genetic endowments *and* attend high-quality schools. In both cases, there is limited opportunity for high-quality schools to make a difference for low PGI^{EA} students. College education, however, takes a middle ground between these two polar outcomes and therefore offers scope for disadvantageous genetic endowments to be offset by school quality and vice versa. We interpret this pattern as suggestive since the confidence bands are too wide to statistically distinguish among the point estimates for different educational degrees.

The triangular series shows the gene-environment interaction with family socio-economic status (SES) increases over the educational life-cycle of individuals.[20] This pattern replicates recent evidence from the United States and Sweden (Buser et al., 2021a; Papageorge and Thom, 2020) and is consistent with the idea that endowments and investments may be substitutes at the

[20] In particular we use the "social origins score" from Belsky et al. (2018) measured in wave 1. Results for alternative measures of family SES, such as family income or potential wages of either parent, yield similar results.

FIGURE 4 – Association of Education Degrees with PGI^{EA} by Q^S and Family SES



Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows point estimates and 90% confidence bands for the interaction of PGI^{EA} with Q^S , the interaction of PGI^{EA} with an indicator for family SES, and their association with education degrees. For each outcome, coefficients are estimated jointly following the specification of equation (7). *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. Standard errors are clustered at the school level.

early stages of childhood but that their complementarity increases over the life-cycle (Heckman and Mosso, 2014).

More generally, the contrast between the circular and the triangular series shows the interplay of advantageous genetic endowments and conducive environments varies across different sources of investments in children. After high school and relative to their high PGI^{EA} peers, genetically disadvantaged children consistently gain more from having attended a high-quality school than from having a high SES background. There are different potential explanations for this pattern. First, schools may allocate resources differently than parents. For example, Houmark et al. (2020) show that parents magnify skill inequality by investing more in children with higher genetic endowments.[21] Hence, our results are consistent with a model where

[21] In Supplementary Table S.6, we replicate this result in our data using both between-family and within-family designs. This result is in contrast to Sanz-de-Galdeano and Terskaya (forthcoming) who also use AddHealth and find no significant effects of PGI^{EA} on parental investment once the siblings' PGI^{EA} is taken into account. We note

schools allocate resources in a more equalizing way than families. Second, schools may provide different types of investments than families. For example, investments in schools tend to happen in the context of larger groups, whereas investments at home tend to be focused on individuals or smaller sibling groups. Hence, our results are consistent with a model where, relative to their high PGI^{EA} peers, low PGI^{EA} students respond more positively to the type of investments they receive in schools than to the type of investments they receive at home. While we cannot distinguish among these explanations, our results are generally consistent with recent evidence documenting large relative gains of disadvantaged students from attending high-quality schools in the United States (Cohodes et al., 2021; Jackson et al., forthcoming).

6 CONCLUSION

The question of how natural endowments and environmental factors determine life outcomes has a long history of inquiry in philosophy and science (Darwin, 1859; Descartes, 1641; Lamarck, 1838; Locke, 1690). The assumption that life outcomes are the result of genetic and environmental factors initially led to the so-called "nature versus nurture" debate. However, current research has moved beyond this simplistic dichotomy and recognizes that individual life outcomes are the result of a complex interplay between nature and nurture. This insight highlights that the importance of genetic endowments for life outcomes is not immutable. Instead, it opens a path for policy interventions that shape the relevant environment.

In this paper, we contribute to this research agenda by studying the interplay of genetic endowments and school quality. Making use of recent advances in molecular genetics, we link an individual-level index of genetic predispositions for educational success with measures of school quality. In turn, we investigate whether the importance of genetic endowments varies with the quality of high schools.

Our findings suggest that investments in the quality of schools can mitigate the genetic gradient in educational attainment. Furthermore, we show that higher gains in educational attainment for students with lower genetic endowments are mediated by gains in language skills, risk aversion, patience, and subjective health.

The use of genetics in education research has an ugly history. Therefore, many people are wary of the emergence of genetic markers in this context, especially when these markers are used for genetic screening (Martschenko et al., 2019). We emphasize that our results do neither presuppose nor endorse the use of genetic screening for educational interventions. Instead, our results suggest that universal policy reform that increases the quality of schools for *all students* may provide an important step to level the playing field regardless of a student's draw in the

that our analysis differs in several dimensions. Among others, we do not restrict the sample to firstborn children only and we use a more comprehensive index of parental time investments.

genetic lottery.

References

ABDELLAOUI, A., D. HUGH-JONES, L. YENGO, K. E. KEMPER, M. G. NIVARD, L. VEUL, Y. HOLTZ, B. P. ZIETSCH, T. M. FRAYLING, N. R. WRAY, et al. (2019). "Genetic Correlates of Social Stratification in Great Britain". *Nature Human Behaviour* 3 (12), pp. 1332–1342.

ABDULKADIROĞLU, A., P. A. PATHAK, J. SCHELLENBERG, and C. R. WALTERS (2020). "Do Parents Value School Effectiveness?" *American Economic Review* 110 (5), pp. 1502–39.

ACEMOGLU, D. and D. AUTOR (2011). "Skills, Tasks and Technologies: Implications for Employment and Earnings". Ed. by D. CARD and O. ASHENFELTER. Vol. 4. Handbook of Labor Economics. Elsevier, pp. 1043–1171.

AGRAWAL, M., J. G. ALTONJI, and R. K. MANSFIELD (2019). "Quantifying Family, School, and Location Effects in the Presence of Complementarities and Sorting". *Journal of Labor Economics* 37 (S1), S11–S83.

ALESINA, A., S. STANTCHEVA, and E. TESO (2018). "Intergenerational Mobility and Preferences for Redistribution". *American Economic Review* 108 (2), pp. 521–554.

ALMÅS, I., A. W. CAPPELEN, and B. TUNGODDEN (2020). "Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?" *Journal of Political Economy* 128 (5), pp. 1753–1788.

ALMLUND, M., A. L. DUCKWORTH, J. HECKMAN, and T. KAUTZ (2011). "Personality Psychology and Economics". *Handbook of the Economics of Education*. Ed. by E. A. HANUSHEK, S. MACHIN, and L. WOESSMANN. Vol. 4. Elsevier, pp. 1–181.

ALMOND, D. and B. MAZUMDER (2013). "Fetal Origins and Parental Responses". *Annual Review of Economics* 5 (1), pp. 37–56.

ALTONJI, J. G., T. E. ELDER, and C. R. TABER (2005). "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools". *Journal of Political Economy* 113 (1), pp. 151–184.

ALTONJI, J. G. and R. K. MANSFIELD (2018). "Estimating Group Effects Using Averages of Observables to Control for Sorting on Unobservables: School and Neighborhood Effects". *American Economic Review* 108 (10), pp. 2902–46.

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

ANGRIST, J. D., S. R. COHODES, S. M. DYNARSKI, P. A. PATHAK, and C. R. WALTERS (2016). "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice". *Journal of Labor Economics* 34 (2), pp. 275–318.

ANGRIST, J. D. and V. LAVY (1999). "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement". *Quarterly Journal of Economics* 114 (2), pp. 533–575.

ANGRIST, J. D., V. LAVY, J. LEDER-LUIS, and A. SHANY (2019). "Maimonides' Rule Redux". *American Economic Review: Insights* 1 (3), pp. 309–324.

ANGRIST, J. D. and J.-S. PISCHKE (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

BARCELLOS, S. H., L. S. CARVALHO, and P. TURLEY (2021). "The Effect of Education on the Relationship between Genetics, Early-life Disadvantages,and Later-life SES". *National Bureau of Economic Research Working Paper Series* 28750.

BARTH, D., N. W. PAPAGEORGE, and K. THOM (2020). "Genetic Endowments and Wealth Inequality". *Journal of Political Economy* 128 (4), pp. 1474–1522.

BECKER, A., T. DECKERS, T. DOHMEN, A. FALK, and F. KOSSE (2012). "The Relationship Between Economic Preferences and Psychological Personality Measures". *Annual Review of Economics* 4 (1), pp. 453–478.

BECKER, J. et al. (2021). "Resource profile and user guide of the Polygenic Index Repository". *Nature Human Behaviour* 5 (12), pp. 1744–1758.

BELSKY, D. W. et al. (2016). "The Genetics of Success: How Single-nucleotide Polymorphisms Associated with Educational Attainment Relate to Life-course Development". *Psychological Science* 27 (7), pp. 957–972.

BELSKY, D. W. et al. (2018). "Genetic Analysis of Social-class Mobility in Five Longitudinal Studies". *Proceedings of the National Academy of Sciences* 115 (31), E7275–E7284.

BEUERMANN, D. W. and C. K. JACKSON (2022). "The short-and long-run effects of attending the schools that parents prefer". *Journal of Human Resources* 57 (3), pp. 725–746.

BEUERMANN, D. W., C. K. JACKSON, L. NAVARRO-SOLA, and F. PARDO (2022). "What is a Good School, and Can Parents Tell? Evidence on the Multidimensionality of School Output". *Review of Economic Studies* 90 (1), pp. 65–101.

BIROLI, P., T. J. GALAMA, S. von HINKE, H. van KIPPERSLUIS, C. A. RIETVELD, and K. THOM (2022). "The Economics and Econometrics of Gene-Environment Interplay". *arXiv* 2203.00729.

BLANDEN, J., M. DOEPKE, and J. STUHLER (2023). "Educational inequality". Ed. by E. A. HANUSHEK, S. MACHIN, and L. WOESSMANN. Vol. 6. Handbook of the Economics of Education. Amsterdam: Elsevier, pp. 405–497.

BRANIGAN, A. R., K. J. MCCALLUM, and J. FREESE (2013). "Variation in the Heritability of Educational Attainment: An International Meta-analysis". *Social Forces* 92 (1), pp. 109–140.

BRINCH, C. N. and T. A. GALLOWAY (2012). "Schooling in Adolescence Raises IQ Scores". *Proceedings of the National Academy of Sciences* 109 (2), pp. 425–430.

BURKS, S. V., C. LEWIS, P. A. KIVI, A. WIENER, J. E. ANDERSON, L. GÖTTE, C. G. DEYOUNG, and A. RUSTICHINI (2015). "Cognitive skills, personality, and economic preferences in collegiate success". *Journal of Economic Behavior & Organization* 115, pp. 30–44.

BUSER, T., R. AHLSKOG, M. JOHANNESSON, P. KOELLINGER, and S. OSKARSSON (2021a). "Using genes to explore the effects of cognitive and non-cognitive skills on education and labor market outcomes". *Tinbergen Institute Discussion Paper* 2021-088.

BUSER, T., M. NIEDERLE, and H. OOSTERBEEK (2021b). "Can Competitiveness predict Education and Labor Market Outcomes? Evidence from Incentivized Choice and Survey Measures". *National Bureau of Economic Research Working Paper Series* 28916.

CASTILLO, M., J. L. JORDAN, and R. PETRIE (2018a). "Discount Rates of Children and High School Graduation". *Economic Journal* 129 (619), pp. 1153–1181.

CASTILLO, M., J. L. JORDAN, and R. PETRIE (2018b). "Children's Rationality, Risk Attitudes and Field Behavior". *European Economic Review* 102, pp. 62–81.

CESARINI, D. and P. M. VISSCHER (2017). "Genetics and Educational Attainment". *npj Science of Learning* 2 (4).

CHANOCK, S. J., T. MANOLIO, M. BOEHNKE, E. BOERWINKLE, D. J. HUNTER, G. THOMAS, J. N. HIRSCHHORN, G. R. ABECASIS, D. ALTSHULER, J. E. BAILEY-WILSON, et al. (2007). "Replicating Genotype-phenotype Associations". *Nature* (447), pp. 655–660.

CHETTY, R., J. N. FRIEDMAN, N. HILGER, E. SAEZ, D. W. SCHANZENBACH, and D. YAGAN (2011). "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star". *Quarterly Journal of Economics* 126 (4), pp. 1593–1660.

CHETTY, R., J. N. FRIEDMAN, and J. E. ROCKOFF (2014a). "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-added Estimates". *American Economic Review* 104 (9), pp. 2593–2632.

– (2014b). "Measuring the Impacts of Teachers II: Teacher Value-added and Student Outcomes in Adulthood". *American Economic Review* 104 (9), pp. 2633–79.

CHETTY, R., N. HENDREN, P. KLINE, and E. SAEZ (2014c). "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States". *Quarterly Journal of Economics* 129 (4), pp. 1553–1623.

CHETTY, R., N. HENDREN, P. KLINE, E. SAEZ, and N. TURNER (2014d). "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility". *American Economic Review* 104 (5), pp. 141–47.

CINELLI, C. and C. HAZLETT (2020). "Making Sense of Sensitivity: Extending Omitted Variable Bias". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82 (1), pp. 39–67.

CLOTFELTER, C. T., H. F. LADD, and J. L. VIGDOR (2010). "Teacher Credentials and Student Achievement in High School: A Cross-subject Analysis with Student Fixed Effects". *Journal of Human Resources* 45 (3), pp. 655–681.

COHODES, S. R., E. M. SETREN, and C. R. WALTERS (2021). "Can Successful Schools Replicate? Scaling Up Boston's Charter School Sector". *American Economic Journal: Economic Policy* 13 (1), pp. 138–67.

COLEMAN, J. S., E. CAMPBELL, C. HOBSON, J. MCPARTLAND, A. MOOD, F. WEINFELD, and R. YORK (1966). *Equality of Educational Opportunity*. Washington D.C.: National Center for Educational Statistics.

CULLEN, J. B., B. A. JACOB, and S. LEVITT (2006). "The Effect of School Choice on Participants: Evidence from Randomized Lotteries". *Econometrica* 74 (5), pp. 1191–1230.

CUNHA, F., J. J. HECKMAN, L. LOCHNER, and D. V. MASTEROV (2006). "Interpreting the Evidence on Life Cycle Skill Formation". Ed. by E. HANUSHEK and F. WELCH. Vol. 1. Handbook of the Economics of Education. Amsterdam: Elsevier, pp. 697–812.

CUNHA, F., J. J. HECKMAN, and S. M. SCHENNACH (2010). "Estimating the Technology of Cognitive and Noncognitive Skill Formation". *Econometrica* 78 (3), pp. 883–931.

DARWIN, C. R. (1859). *On the Origin of Species*. London: John Murray.

DEMANGE, P. A. et al. (2020). "Parental influences on offspring education: indirect genetic effects of non-cognitive skills". *bioRxiv* 2020.09.15.296236.

DEMANGE, P. A. et al. (2021). "Investigating the Genetic Architecture of Non-Cognitive Skills Using GWAS-by-Subtraction". *Nature Genetics* 53, pp. 35–44.

DEMING, D. J. (2011). "Better Schools, Less Crime?" *Quarterly Journal of Economics* 126 (4), pp. 2063–2115.

– (2014). "Using School Choice Lotteries to Test Measures of School Effectiveness". *American Economic Review* 104 (5), pp. 406–11.

DEMING, D. J., J. S. HASTINGS, T. J. KANE, and D. O. STAIGER (2014). "School Choice, School Quality, and Postsecondary Attainment". *American Economic Review* 104 (3), pp. 991–1013.

DEMONTIS, D., R. K. WALTERS, J. MARTIN, M. MATTHEISEN, T. D. ALS, E. AGERBO, G. BALDURSSON, R. BELLIVEAU, J. BYBJERG-GRAUHOLM, M. BÆKVAD-HANSEN, et al. (2019). "Discovery of the First Genome-wide Significant Risk Loci for Attention Deficit/Hyperactivity Disorder". *Nature Genetics* 51 (1), pp. 63–75.

DESCARTES, R. (1641). "Meditations on First Philosophy". *Descartes: Selected Philosophical Writings*. Ed. by J. COTTINGHAM. Ed. by D. MURDOCH. Ed. by R. STOOTHOFF. Ed. by A. KENNY. Cambridge: Cambridge University Press, pp. 73–122.

DOMINGUE, B. W., D. BELSKY, D. CONLEY, K. M. HARRIS, and J. BOARDMAN (2015). "Polygenic Influence on Educational Attainment: New Evidence From the National Longitudinal Study of Adolescent to Adult Health". *AERA Open* 1 (3), pp. 1–13.

DOMINGUE, B. W., S. TREJO, E. ARMSTRONG-CARTER, and E. M. TUCKER-DROB (2020). "Interactions between Polygenic Scores and Environments: Methodological and Conceptual Challenges". *Sociological Science* 7 (19), pp. 465–486.

DUDBRIDGE, F. (2013). "Power and Predictive Accuracy of Polygenic Risk Scores". *PLOS Genetics* 9 (3), e1003348.

DUSTAN, A., A. DE JANVRY, and E. SADOULET (2017). "Flourish or fail? The risky reward of elite high school admission in Mexico City". *Journal of Human Resources* 52 (3), pp. 756–799.

FALK, A., A. BECKER, T. J. DOHMEN, B. ENKE, D. B. HUFFMAN, and U. SUNDE (2018). "Global Evidence on Economic Preferences". *Quarterly Journal of Economics* 133 (4), pp. 1645–1692.

FIGLIO, D. N., J. FREESE, K. KARBOWNIK, and J. ROTH (2017). "Socioeconomic status and genetic influences on cognitive development". *Proceedings of the National Academy of Sciences* 114 (53), pp. 13441–13446.

FREDRIKSSON, P., B. ÖCKERT, and H. OOSTERBEEK (2013). "Long-term Effects of Class Size". *Quarterly Journal of Economics* 128 (1), pp. 249–285.

HANUSHEK, E. A., B. JACOBS, G. SCHWERDT, R. van der VELDEN, S. VERMEULEN, and S. WIEDERHOLD (2021). "The Intergenerational Transmission of Cognitive Skills: An Investigation of the Causal Impact of Families on Student Outcomes". *National Bureau of Economic Research Working Paper Series* 29450.

HANUSHEK, E. A. and S. G. RIVKIN (2010). "Generalizations about Using Value-added Measures of Teacher Quality". *American Economic Review* 100 (2), pp. 267–271.

HANUSHEK, E. A., S. G. RIVKIN, and J. C. SCHIMAN (2016). "Dynamic Effects of Teacher Turnover on the Quality of Instruction". *Economics of Education Review* 55, pp. 132–148.

HANUSHEK, E. A. and L. WOESSMANN (2008). "The Role of Cognitive Skills in Economic Development". *Journal of Economic Literature* 46 (3), pp. 607–668.

– (2012). "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation". *Journal of Economic Growth* 17 (4), pp. 267–321.

HECKMAN, J., S. H. MOON, R. PINTO, P. SAVELYEV, and A. YAVITZ (2010). "Analyzing Social Experiments as Implemented: A Reexamination of the Evidence from the HighScope Perry Preschool Program". *Quantitative Economics* 1 (1), pp. 1–46.

HECKMAN, J. and S. MOSSO (2014). "The Economics of Human Development and Social Mobility". *Annual Review of Economics* 6 (1), pp. 689–733.

HILL, A. J. and D. B. JONES (2018). "A Teacher Who Knows Me: The Academic Benefits of Repeat Student-teacher Matches". *Economics of Education Review* 64, pp. 1–12.

HOUMARK, M., V. RONDA, E. AGERBO, P. MORTENSEN, and M. ROSHOLM (2022). "Genetic and Socioeconomic Achievement Gaps in Elementary School". *IZA Discussion Paper* 15418.

HOUMARK, M. A., V. RONDA, and M. ROSHOLM (2020). "The Nurture of Nature and the Nature of Nurture: How Genes and Investments Interact in the Formation of Skills". *IZA Discussion Paper* 13780.

HOWARD, D. M., M. J. ADAMS, T.-K. CLARKE, J. D. HAFFERTY, J. GIBSON, M. SHIRALI, J. R. COLEMAN, S. P. HAGENAARS, J. WARD, E. M. WIGMORE, et al. (2019). "Genome-wide Meta-analysis of Depression Identifies 102 Independent Variants and Highlights the Importance of the Prefrontal Brain Regions". *Nature Neuroscience* 22 (3), pp. 343–352.

HUMPHRIES, J. E. and F. KOSSE (2017). "On the Interpretation of Non-cognitive Skills—What is Being Measured and Why it Matters". *Journal of Economic Behavior & Organization* 136, pp. 174–185.

HWANG, N., B. KISIDA, and C. KOEDEL (2021). "A Familiar Face: Student-teacher Rematches and Student Achievement". *Economics of Education Review* 85, p. 102194.

JACKSON, C. K. (2009a). "Student Demographics, Teacher Sorting, and Teacher Quality: Evidence from the End of School Desegregation". *Journal of Labor Economics* 27 (2), pp. 213–256.

– (2019). "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes". *Journal of Political Economy* 126 (5), pp. 2072–2107.

JACKSON, C. K. and E. BRUEGMANN (2009). "Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers". *American Economic Journal: Applied Economics* 1 (4), pp. 85–108.

JACKSON, C. K., S. C. PORTER, J. Q. EASTON, and S. KIGUEL (forthcoming). "Who Benefits From Attending Effective High Schools?" *Journal of Labor Economics*.

JACKSON, C. K., S. C. PORTER, J. Q. EASTON, A. BLANCHARD, and S. KIGUEL (2020). "School Effects on Socioemotional Development, School-based Arrests, and Educational Attainment". *American Economic Review: Insights* 2 (4), pp. 491–508.

JACKSON, M. I. (2009b). "Understanding links between adolescent health and educational attainment". *Demography* 46 (4), pp. 671–694.

JACOB, B. A., J. E. ROCKOFF, E. S. TAYLOR, B. LINDY, and R. ROSEN (2018). "Teacher Applicant Hiring and Teacher Performance: Evidence from DC Public Schools". *Journal of Public Economics* 166, pp. 81–97.

JANSEN, P., K. WATANABE, S. STRINGER, N. SKENE, J. BRYOIS, A. HAMMERSCHLAG, C. DE LEEUW, J. BENJAMINS, A. MUÑOZ-MANCHADO, M. NAGEL, et al. (2019). "Genome-wide Analysis of Insomnia in 1,331,010 Individuals Identifies New Risk Loci and Functional Pathways". *Nature Genetics* 51 (3), pp. 394–403.

JENNINGS, J. L., D. DEMING, C. JENCKS, M. LOPUCH, and B. E. SCHUELER (2015). "Do differences in school quality matter more than we thought? New evidence on educational opportunity in the twenty-first century". *Sociology of Education* 88 (1), pp. 56–82.

JOHNSON, R. C. and C. K. JACKSON (2019). "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending". *American Economic Journal: Economic Policy* 11 (4), pp. 310–49.

KELLER, M. (2014). "Gene × Environment Interaction Studies have not Properly Controlled for Potential Confounders: the Problem and the (Simple) Solution". *Biological Psychiatry* 75 (1), pp. 18–24.

KLING, J. R., J. B. LIEBMAN, and L. F. KATZ (2007). "Experimental Analysis of Neighborhood Effects". *Econometrica* 75 (1), pp. 83–119.

KOELLINGER, P. D. and K. P. HARDEN (2018). "Using Nature to Understand Nurture". *Science* 359 (6374), pp. 386–387.

KONG, A. et al. (2018). "The Nature of Nurture: Effects of Parental Genotypes". *Science* 359 (6374), pp. 424–428.

KRUEGER, A. B. (1999). "Experimental Estimates of Education Production Functions". *Quarterly Journal of Economics* 114 (2), pp. 497–532.

KRUEGER, A. B. and M. LINDAHL (2001). "Education for Growth: Why and for Whom?" *Journal of Economic Literature* 39 (4), pp. 1101–1136.

KWEON, H., C. A. BURIK, R. K. LINNÉR, R. de VLAMING, A. OKBAY, D. MARTSCHENKO, K. P. HARDEN, T. A. DIPRETE, and P. D. KOELLINGER (2020). "Genetic Fortune: Winning or Losing Education, Income, and Health". *Tinbergen Institute Discussion Papers* 20-053.

LALIBERTÉ, J.-W. (2021). "Long-term Contextual Effects in Education: Schools and Neighborhoods". *American Economic Journal: Economic Policy* 13 (2), pp. 336–377.

LAMARCK, J.-B. P. A. d. M. de (1838). *Histoire naturelle des animaux sans vertèbres*. Paris: Verdière.

LEE, J. J. et al. (2018). "Gene Discovery and Polygenic Prediction from a Genome-wide Association Study of Educational Attainment in 1.1 Million Individuals". *Nature Genetics* 50 (8), pp. 1112–1121.

LERGETPORER, P., K. WERNER, and L. WOESSMANN (2020). "Educational inequality and public policy preferences: Evidence from representative survey experiments". *Journal of Public Economics* 188, p. 104226.

LEUVEN, E. and S. A. LØKKEN (2020). "Long-term Impacts of Class Size in Compulsory School". *Journal of Human Resources* 55 (1), pp. 309–348.

LIU, M., Y. JIANG, R. WEDOW, Y. LI, D. BRAZEL, F. CHEN, G. DATTA, J. DAVILA-VELDERRAIN, D. MCGUIRE, C. TIAN, et al. (2019). "Association Studies of up to 1.2 Million Individuals Yield New Insights into the Genetic Etiology of Tobacco and Alcohol Use". *Nature Genetics* 51 (2), pp. 237–244.

LOCKE, J. (1690). *An Essay Concerning Human Understanding*. Hamilton: McMaster University Archive for the History of Economic Thought.

MARTIN, A. R., C. R. GIGNOUX, R. K. WALTERS, G. L. WOJCIK, B. M. NEALE, S. GRAVEL, M. J. DALY, C. D. BUSTAMANTE, and E. E. KENNY (2017). "Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations". *American Journal of Human Genetics* 100 (4), pp. 635–649.

MARTSCHENKO, D., S. TREJO, and B. W. DOMINGUE (2019). "Genetics and Education: Recent Developments in the Context of an Ugly History and an Uncertain Future." *AERA Open* 5 (1), pp. 1–15.

MUSLIMOVA, D., H. van KIPPERSLUIS, C. A. RIETVELD, S. von HINKE, and S. F. W. MEDDENS (2020). "Dynamic Complementarity in Skill Production: Evidence from Genetic Endowments and Birth Order". *Tinbergen Institute Discussion Paper 2020-082/V*.

NICOLETTI, C. and B. RABE (2014). "School Inputs and Skills: Complementarity and Self-productivity". *IZA Discussion Paper* 8693.

OSTER, E. (2019). "Unobservable Selection and Coefficient Stability: Theory and Evidence". *Journal of Business & Economic Statistics* 37 (2), pp. 187–204.

PAPAGEORGE, N. W. and K. THOM (2020). "Genes, Education, and Labor Market Outcomes: Evidence from the Health and Retirement Study". *Journal of the European Economic Association* 18 (3), pp. 1351–1399.

PAPAY, J. P. and M. A. KRAFT (2015). "Productivity Returns to Experience in the Teacher Labor Market: Methodological Challenges and New Evidence on Long-term Career Improvement". *Journal of Public Economics* 130, pp. 105–119.

PLOMIN, R., J. C. DEFRIES, and J. C. LOEHLIN (1977). "Genotype-environment interaction and correlation in the analysis of human behavior". *Psychological Bulletin* 84 (2), pp. 309–322.

RAUDENBUSH, S. W. and J. WILLMS (1995). "The estimation of school effects". *Journal of Educational and Behavioral Statistics* 20 (4), pp. 307–335.

RIVKIN, S. G., E. A. HANUSHEK, and J. F. KAIN (2005). "Teachers, Schools, and Academic Achievement". *Econometrica* 73 (2), pp. 417–458.

ROCKOFF, J. E. (2004). "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data". *American Economic Review* 94 (2), pp. 247–252.

RONDA, V., E. AGERBO, D. BLESES, P. B. MORTENSEN, A. BØRGLUM, D. M. HOUGAARD, O. MORS, M. NORDENTOFT, T. WERGE, and M. ROSHOLM (2022). "Family Disadvantage, Gender and the Returns to Genetic Human Capital". *Scandinavian Journal of Economics* 124 (2), pp. 550–578.

RONFELDT, M., S. LOEB, and J. WYCKOFF (2013). "How Teacher Turnover Harms Student Achievement". *American Educational Research Journal* 50 (1), pp. 4–36.

ROSSIN-SLATER, M. and M. WÜST (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), pp. 255–86.

ROTHSTEIN, J. M. (2006). "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions". *American Economic Review* 96 (4), pp. 1333–1350.

SANZ-DE-GALDEANO, A. and A. TERSKAYA (forthcoming). "Sibling Differences in Educational Polygenic Scores: How Do Parents React?" *Review of Economics and Statistics*.

SAVAGE, J. E. et al. (2018). "Genome-wide Association Meta-analysis in 269,867 Individuals Identifies New Genetic and Functional Links to Intelligence". *Nature Genetics* 50 (7), pp. 912–919.

SELZAM, S., S. J. RITCHIE, J.-B. PINGAULT, C. A. REYNOLDS, P. F. O'REILLY, and R. PLOMIN (2019). "Comparing Within- and Between-Family Polygenic Score Prediction". *American Journal of Human Genetics* 105 (2), pp. 351–363.

SHENHAV, N. (2021). "Lowering Standards to Wed? Spouse Quality, Marriage, and Labor Market Responses to the Gender Wage Gap". *Review of Economics and Statistics* 103 (2), pp. 265–279.

TREJO, S., D. BELSKY, J. BOARDMAN, J. FREESE, K. M. HARRIS, P. HERD, K. SICINSKI, and B. DOMINGUE (2018). "Schools as Moderators of Genetic Associations with Life Course Attainments: Evidence from the WLS and Add Health". *Sociological Science* 5, pp. 513–540.

TURKHEIMER, E. (2000). "Three Laws of Behavior Genetics and What They Mean". *Current Directions in Psychological Science* 9 (5), pp. 160–164.

TURKHEIMER, E., A. HALEY, M. WALDRON, B. D'ONOFRIO, and I. I. GOTTESMAN (2003). "Socioeconomic Status Modifies Heritability of IQ in Young Children". *Psychological Science* 14 (6), pp. 623–628.

WALTERS, C. R. (2018). "The Demand for Effective Charter Schools". *Journal of Political Economy* 126 (6), pp. 2179–2223.

WARE, E. B., L. L. SCHMITZ, J. FAUL, A. GARD, C. MITCHELL, J. A. SMITH, W. ZHAO, D. WEIR, and S. L. KARDIA (2017). "Heterogeneity in Polygenic Scores for Common Human Traits". *bioRxiv 106062*.

WOODLEY OF MENIE, M. A., M. PEÑAHERRERA-AGUIRRE, C. S. DUNKEL, and M. A. SARRAF (2021). "Evidence for the Scarr–Rowe Effect on Genetic Expressivity in the Health and Retirement Study". *Twin Research and Human Genetics* 24 (2), pp. 110–115.

WRAY, N. R., S. H. LEE, D. MEHTA, A. A. VINKHUYZEN, F. DUDBRIDGE, and C. M. MIDDELDORP (2014). "Research Review: Polygenic Methods and their Application to Psychiatric Traits". *Journal of Child Psychology and Psychiatry* 55 (10), pp. 1068–1087.

YENGO, L., J. SIDORENKO, K. E. KEMPER, Z. ZHENG, A. R. WOOD, M. N. WEEDON, T. M. FRAYLING, J. HIRSCHHORN, J. YANG, P. M. VISSCHER, et al. (2018). "Meta-analysis of Genome-wide Association Studies for Height and Body Mass Index in 700000 Individuals of European Ancestry". *Human Molecular Genetics* 27 (20), pp. 3641–3649.

Genetic Endowments, Educational Outcomes and the Mediating Influence of School Quality

Benjamin W. Arold, Paul Hufe & Marc Stöckli

Supplementary Material

A SUPPLEMENTARY TABLES

TABLE S.1 – Gene-Environment Correlations

	PGI^{EA}	Q^S	Educ. Mother (School Av.)	White Student (School Share)	Single Parents (School Share)	GPA English (School Av.)
PGI^{EA}	1.000 (.)					
Q^S	0.023 (1.000)	1.000 (.)				
White Student (School Share)	-0.083 (0.000)	-0.027 (1.000)	1.000 (.)			
Single Parents (School Share)	-0.037 (0.382)	-0.017 (1.000)	-0.567 (0.000)			
Educ. Mother (School Av.)	0.129 (0.000)	0.020 (1.000)	-0.336 (0.000)	1.000 (.)		
GPA English (School Av.)	0.058 (0.004)	0.012 (1.000)	-0.279 (0.000)	0.226 (0.000)	1.000 (.)	
GPA Math (School Av.)	0.077 (0.000)	0.010 (1.000)	-0.272 (0.000)	0.447 (0.000)	0.510 (0.000)	1.000 (.)

Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This table shows correlations between PGI^{EA} and various school characteristics. *p*-values (in parentheses) are Bonferroni corrected to account for multiple hypothesis testing.

TABLE S.2 – Robustness to Sample Composition

	Baseline	Alternative Sample Composition		
Outcome: Years of Education	(1)	Re-Weighted (2)	Excl. (Potential) Movers before High School (3)	Excl. (Potential) Movers during High School (4)
PGI^{EA}	0.360***	0.346***	0.352***	0.342***
	(0.028)	(0.030)	(0.035)	(0.040)
Q^S	0.128**	0.120**	0.162***	0.106
	(0.054)	(0.056)	(0.062)	(0.077)
$PGI^{EA} \times Q^S$	-0.062**	-0.056**	-0.082**	-0.082**
	(0.026)	(0.026)	(0.032)	(0.038)
Child Controls	✓	✓	✓	✓
Family Controls	✓	✓	✓	✓
Control Function	✓	✓	✓	✓
N	4,036	3,970	2,964	2,441
R^2	0.334	0.314	0.351	0.344

Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This table shows the joint association of PGI^{EA} and Q^S with completed years of education. In column (2) we re-weight our analysis sample to match ACS and CPS with respect to gender composition, educational attainment of parents, and the age of mothers at birth—see also Supplementary Table S.3. In column (3) we exclude respondents that visit feeder schools in wave 1 and for whom we do not have information on subsequent high schools. In column (4) we additionally exclude respondents for whom we do not have information on whether they graduated from an Add Health high school. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. All right-hand side variables are standardized on the estimation sample ($\mu = 0$, $\sigma = 1$). Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the school level.

TABLE S.3 – Sample Representativeness

	Population (Cohorts 1974-1983)		Analysis Sample	
	All	Non-Hispanic White	Unweighted	Re-Weighted
Gender				
Male	0.498	0.503	0.462	0.503
Female	0.502	0.497	0.538	0.497
Education Mother				
$\leq$ High School	0.536	0.489	0.507	0.489
$>$ High School; $<$ College Degree	0.281	0.302	0.218	0.301
$\geq$ College Degree	0.183	0.209	0.275	0.210
Education Father				
$\leq$ High School	0.472	0.425	0.514	0.425
$>$ High School; $<$ College Degree	0.255	0.271	0.187	0.271
$\geq$ College Degree	0.273	0.304	0.299	0.303
Age Mother at Birth				
$<$ 25 Years	0.353	0.330	0.503	0.330
$\geq$ 25 Years	0.647	0.670	0.497	0.670
Parental Income				
$<$ \$50,000	0.557	0.491	0.536	0.511
$\geq$ \$50,000; $<$ \$100,000	0.352	0.403	0.386	0.406
$\geq$ \$100,000	0.091	0.106	0.078	0.083
Education Respondent				
$\leq$ High School	0.301	0.225	0.196	0.181
$>$ High School; $<$ College Degree	0.327	0.344	0.408	0.406
$\geq$ College Degree	0.372	0.431	0.396	0.413

Data: National Longitudinal Study of Adolescent to Adult Health, American Community Survey (ACS), Current Population Survey (CPS).
Note: Own calculations. This table shows summary statistics of the core analysis sample in comparison to other population samples. It shows respondents' characteristics for the following samples: (i) the US population from birth cohorts 1974–1983, (ii) the Non-Hispanic White US population from birth cohorts 1974–1983, (iii) the core estimation sample, and (iv) the core estimation sample re-weighted to match (ii) with respect to *Gender*, *Education Mother*, *Education Father*, and *Age Mother at Birth*. Population data on *Gender* and *Education Respondent* from IPUMS ACS 2019 (Ruggles et al., 2020). Population data on *Education Mother*, *Education Father*, *Age Mother at Birth*, and *Parental Income* from IPUMS CPS 1994 (Flood et al., 2020).

TABLE S.4 – Alternative Polygenic Scores

	Baseline	+ Controls for Other Polygenic Scores					
Outcome: Years of Education	(1)	Body Mass Index (2)	ADHD (3)	Depressive Symptoms (4)	Intelligence (5)	Ever Smoker (6)	Sleep Duration (7)
PGI^{EA}	0.360***	0.340***	0.331***	0.358***	0.350***	0.341***	0.359***
	(0.028)	(0.031)	(0.028)	(0.028)	(0.031)	(0.031)	(0.028)
Q^S	0.128**	0.124**	0.122**	0.124**	0.128**	0.125**	0.128**
	(0.054)	(0.054)	(0.053)	(0.055)	(0.054)	(0.053)	(0.054)
$PGI^{EA} \times Q^S$	-0.062**	-0.070**	-0.065**	-0.060**	-0.054**	-0.061**	-0.062**
	(0.026)	(0.029)	(0.025)	(0.026)	(0.028)	(0.027)	(0.026)
Other PGI	–	-0.080***	-0.129***	-0.039	0.020	-0.095***	0.029
		(0.026)	(0.028)	(0.030)	(0.030)	(0.036)	(0.028)
Other PGI $\times Q^S$	–	-0.029	0.003	0.031	-0.015	0.017	-0.004
		(0.028)	(0.028)	(0.029)	(0.027)	(0.033)	(0.029)
Child Controls	✓	✓	✓	✓	✓	✓	✓
Family Controls	✓	✓	✓	✓	✓	✓	✓
Control Function	✓	✓	✓	✓	✓	✓	✓
N	4,036	4,036	4,036	4,036	4,036	4,036	4,036
R^2	0.334	0.335	0.337	0.334	0.334	0.336	0.334

Data: National Longitudinal Study of Adolescent to Adult Health.

Note: Own calculations. This table shows the joint association of PGI^{EA} and Q^S with completed years of education. We control for other PGIs and their interaction with Q^S. The relevant PGIs are indicated in the column header. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. All right-hand side variables are standardized on the estimation sample ($\mu = 0, \sigma = 1$). Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the school level.

TABLE S.5 – Measurement-error-corrected PGI (Becker et al., 2021)

	Coefficient	Standard Error	p-value	Substitutability
Baseline				
PGI^{EA}	0.360	0.031	0.000	
Q^S	0.128	0.062	0.040	
$PGI^{EA} \times Q^S$	-0.062	0.030	0.040	17%
Add Health ($\rho = 1.968$)				
PGI^{EA}	0.589	0.053	0.000	
Q^S	0.110	0.064	0.088	
$PGI^{EA} \times Q^S$	-0.088	0.047	0.060	15%
Health and Retirement Study ($\rho = 1.413$)				
PGI^{EA}	0.566	0.053	0.000	
Q^S	0.112	0.064	0.079	
$PGI^{EA} \times Q^S$	-0.086	0.044	0.049	15%
Wisconsin Longitudinal Study ($\rho = 1.649$)				
PGI^{EA}	0.718	0.069	0.000	
Q^S	0.098	0.067	0.143	
$PGI^{EA} \times Q^S$	-0.099	0.057	0.084	14%
UK Biobank ($\rho = 1.452$)				
PGI^{EA}	0.589	0.055	0.000	
Q^S	0.110	0.065	0.090	
$PGI^{EA} \times Q^S$	-0.088	0.045	0.048	15%

Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This table shows the joint association of PGI^{EA} and Q^S with completed years of education. We apply the correction method of Becker et al. (2021) to account for measurement error in PGI^{EA}. The correction method is based on $\rho = h^2_{SNP}/R^2$, where h^2_{SNP} indicates SNP heritability and R^2 the share of variation in educational attainment explained by PGI^{EA}. For Add Health, we take ρ from Sanz-de-Galdeano and Terskaya (forthcoming), for all other data sets we take ρ from Becker et al. (2021). Standard errors are bootstrapped with 1,000 draws. These standard errors are likely conservative (Becker et al., 2021; Sanz-de-Galdeano and Terskaya, forthcoming).

TABLE S.6 – Association of Parental Investments with PGI^{EA} and Q^S

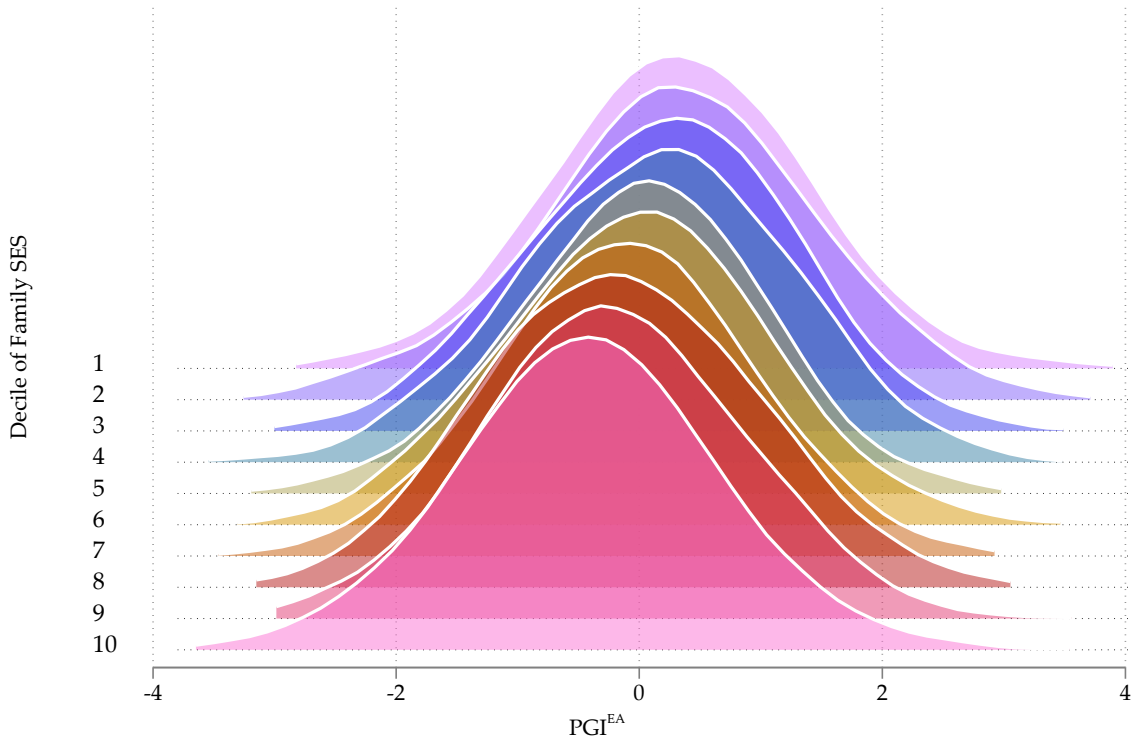
	Between-Family			Within-Family		
Outcome: Parental Investment Index	Both Parents (1)	Mother (2)	Father (3)	Both Parents (4)	Mother (5)	Father (6)
PGI^{EA}	0.063*** (0.015)	0.049*** (0.015)	0.072*** (0.016)	0.092** (0.043)	0.052 (0.041)	0.141*** (0.048)
Q^S	-0.022 (0.015)	-0.031* (0.017)	-0.011 (0.017)	-0.042 (0.036)	-0.052 (0.038)	-0.017 (0.047)
Child Controls	✓	✓	✓	✓	✓	✓
Family Controls	✓	✓	✓	✓	✓	✓
Control Function	✓	✓	✓	✓	✓	✓
Sibling PGI^{EA}	×	×	×	✓	✓	✓
N	4,035	4,035	3,268	670	670	550
R^2	0.086	0.096	0.101	0.119	0.124	0.135
Outcome Mean	0.000	0.000	0.000	0.000	0.000	0.000
Outcome SD	1.000	1.000	1.000	1.000	1.000	1.000

Data: National Longitudinal Study of Adolescent to Adult Health.

Note: Own calculations. This table shows the joint association of PGI^{EA} and Q^S with indexes of parental time investments. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. All right-hand side variables are standardized on the estimation sample ($\mu = 0$, $\sigma = 1$). Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the school level.

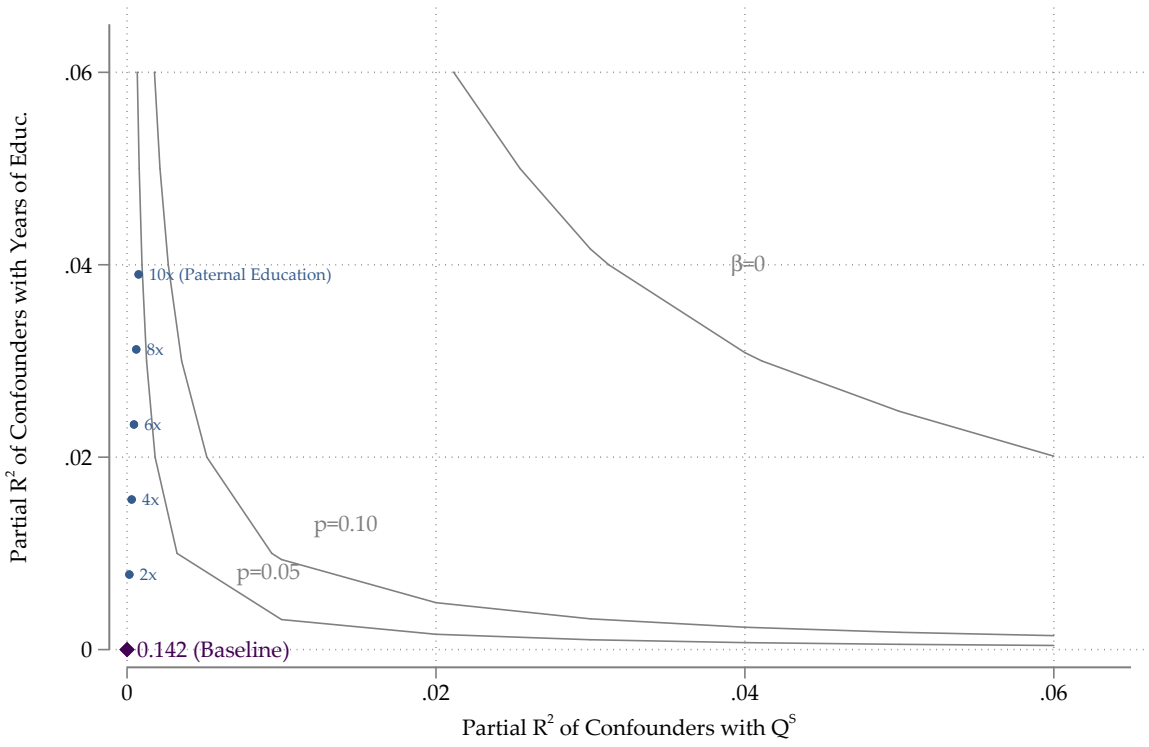
B SUPPLEMENTARY FIGURES

FIGURE S.1 – Distribution of PGI^{EA} by Family SES



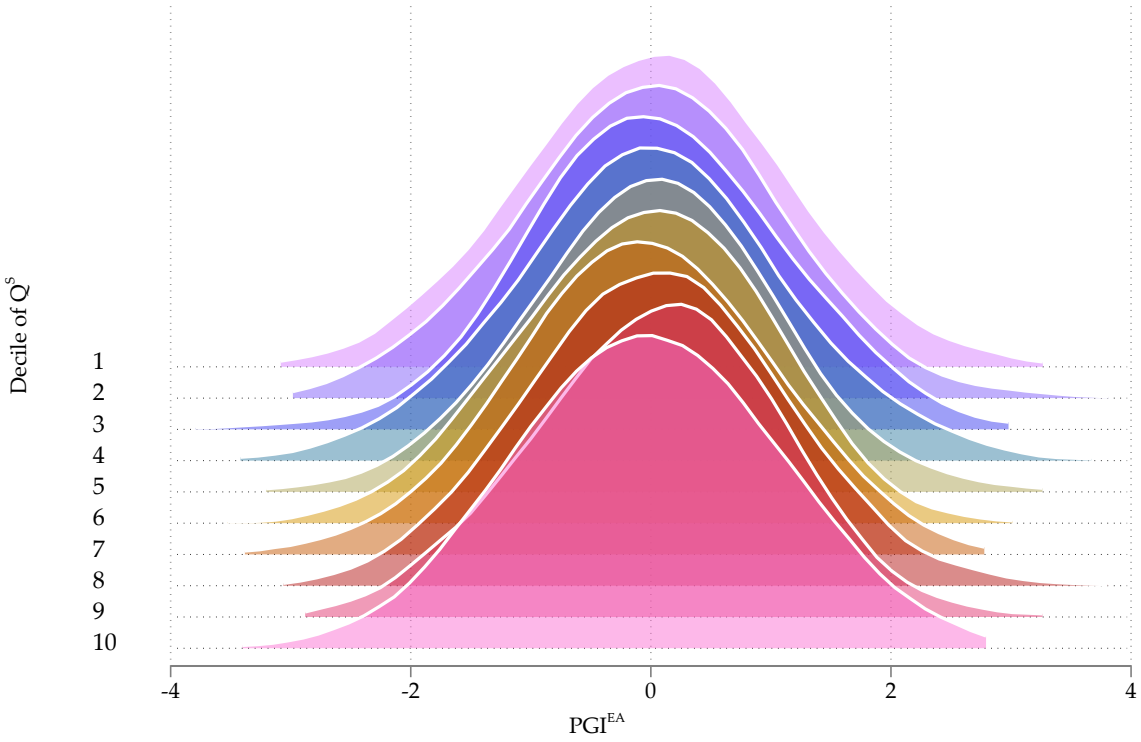
Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows unconditional PGI^{EA} distribution by deciles of family SES. Density distributions are smoothed using the Epanechnikov kernel function with a bandwidth of 0.5.

FIGURE S.2 – Sensitivity to Unobserved Confounders



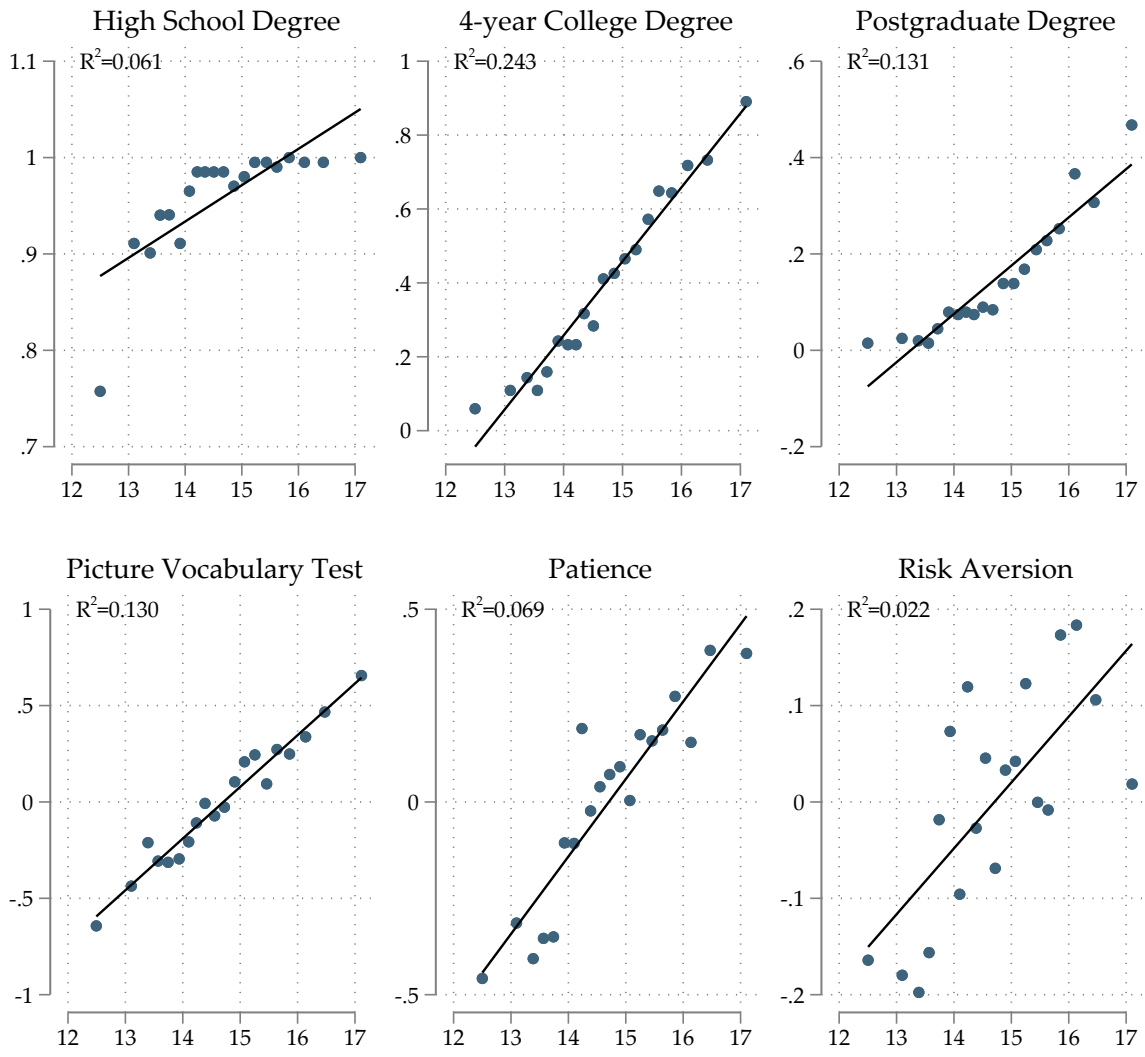
Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows the sensitivity of the point estimate for Q^S to unobserved confounding variables. Following the procedure of Cinelli and Hazlett (2020), we calculate the bias-adjusted treatment effect of Q^S under different assumptions about the partial R^2 of confounding variables with the variables of interest and the partial R^2 of confounding variables with years of education. Each contour line shows p-values for the point estimate of Q^S in a regression of years of education on Q^S and controls under different assumptions about the two partial R^2. Each circle shows analogous values for different multiples of paternal education. The diamond shows baseline estimates from Figure 1. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. Standard errors are clustered at the school level.

FIGURE S.3 – Distribution of PGI^{EA} by Q^S



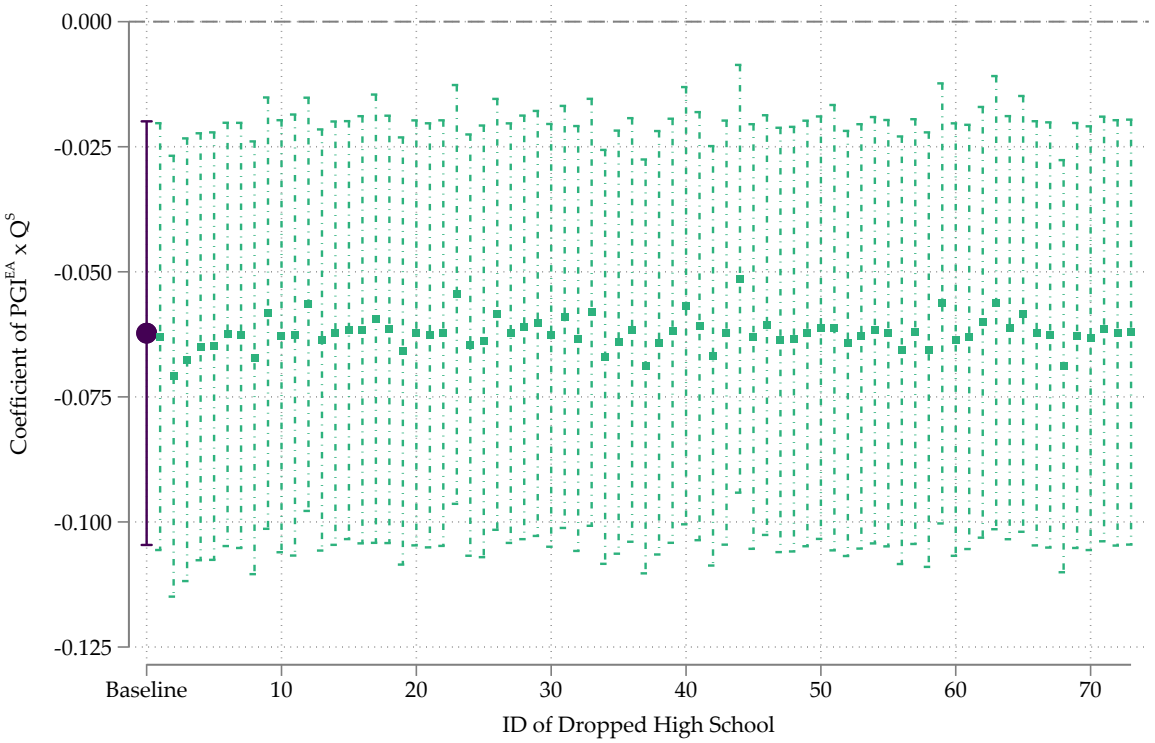
Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows PGI^{EA} distribution by deciles of Q^Safter residualizing PGI^{EA} and Q^S by the full set of control variables. Density distributions are smoothed using the Epanechnikov kernel function with a bandwidth of 0.5.

FIGURE S.4 – Association of Educational Attainment and Skills with Predicted Education



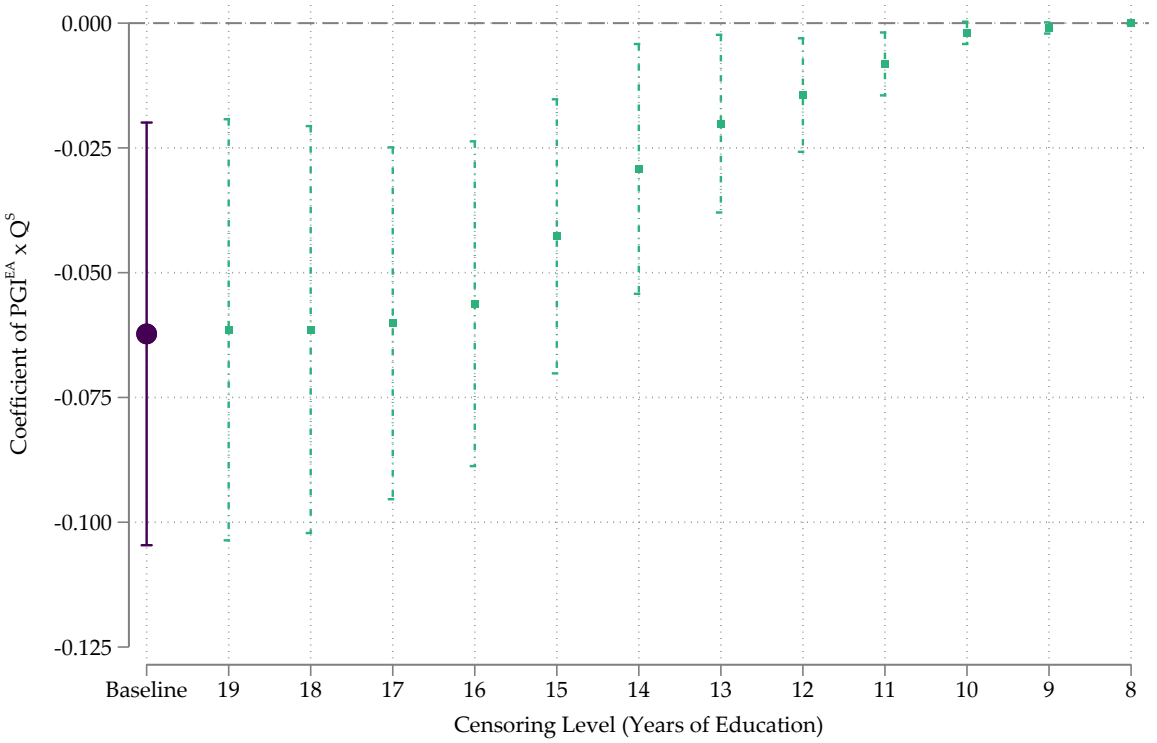
Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows the association of various meausures for educational attainment and (non-)cognitive skills with predicted years of education. We bin scatterplots using 20 quantiles of the variable of interest. Black lines are fitted from linear regressions of the variable of interest on predicted education. Predicted education is calculated from a regression of completed years of education on all *Child Controls* and *Family Controls*. Outcomes are measured in Waves 3, 4 and 5. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects.

FIGURE S.5 – Sensitivity to Outlier Schools



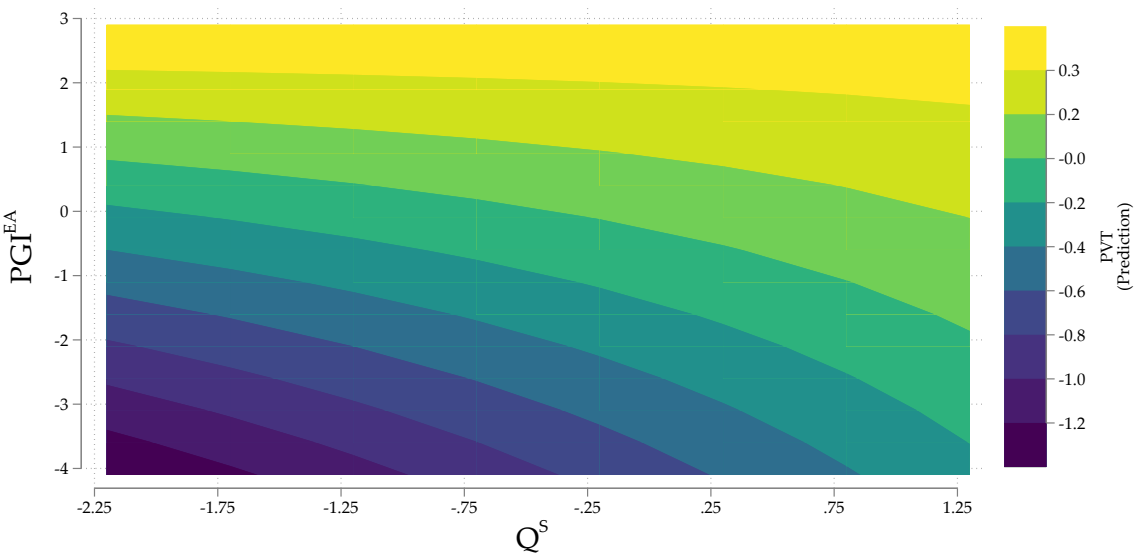
Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows point estimates and 90% confidence bands of the interaction of PGI^{EA} with Q^S, and its association with years of education. Each estimate is derived from a subsample of the data in which we drop one High School, respectively. Estimates follow the specification of equation (7). *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. Standard errors are clustered at the school level.

FIGURE S.6 – Sensitivity to Ceiling Effects



Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows point estimates and 90% confidence bands of the interaction of PGI^{EA} and Q^S, and its association with years of education. Each estimate is derived from the full sample while censoring the outcome variable at different levels from above. Estimates follow the specification of equation (7). *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. Standard errors are clustered at the school level.

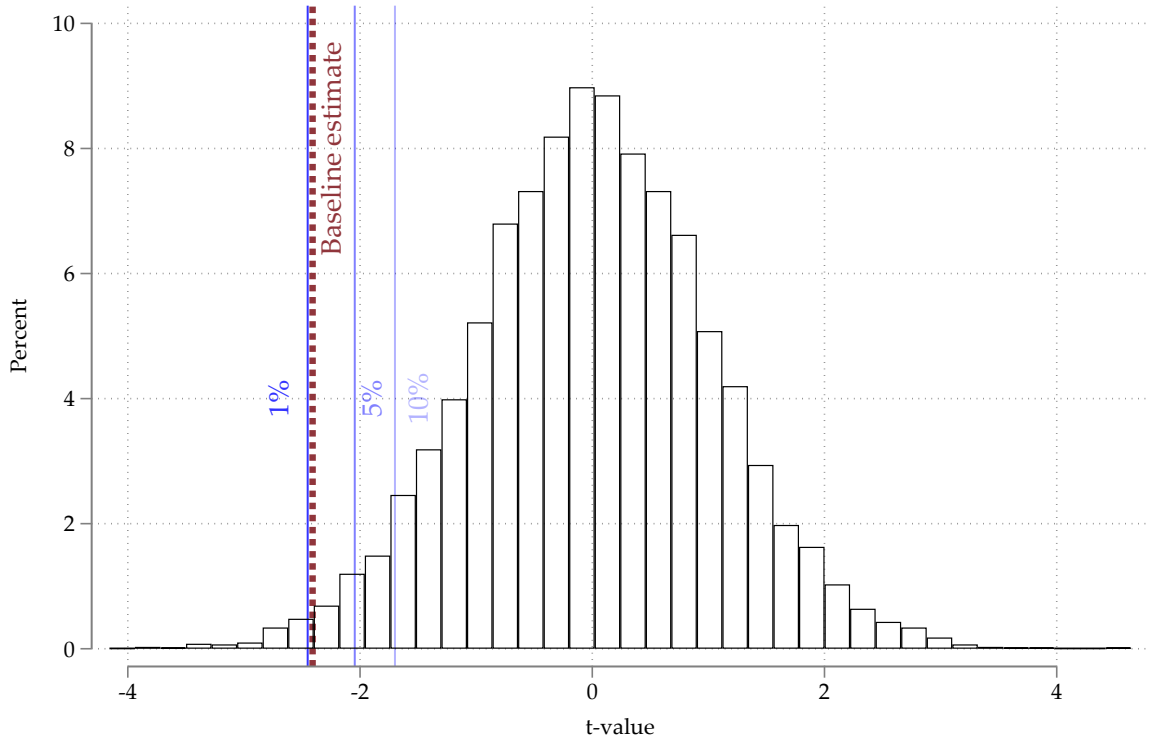
FIGURE S.7 – Association of Picture Vocabulary Test with PGI^{EA} by Q^S



Data: National Longitudinal Study of Adolescent to Adult Health.

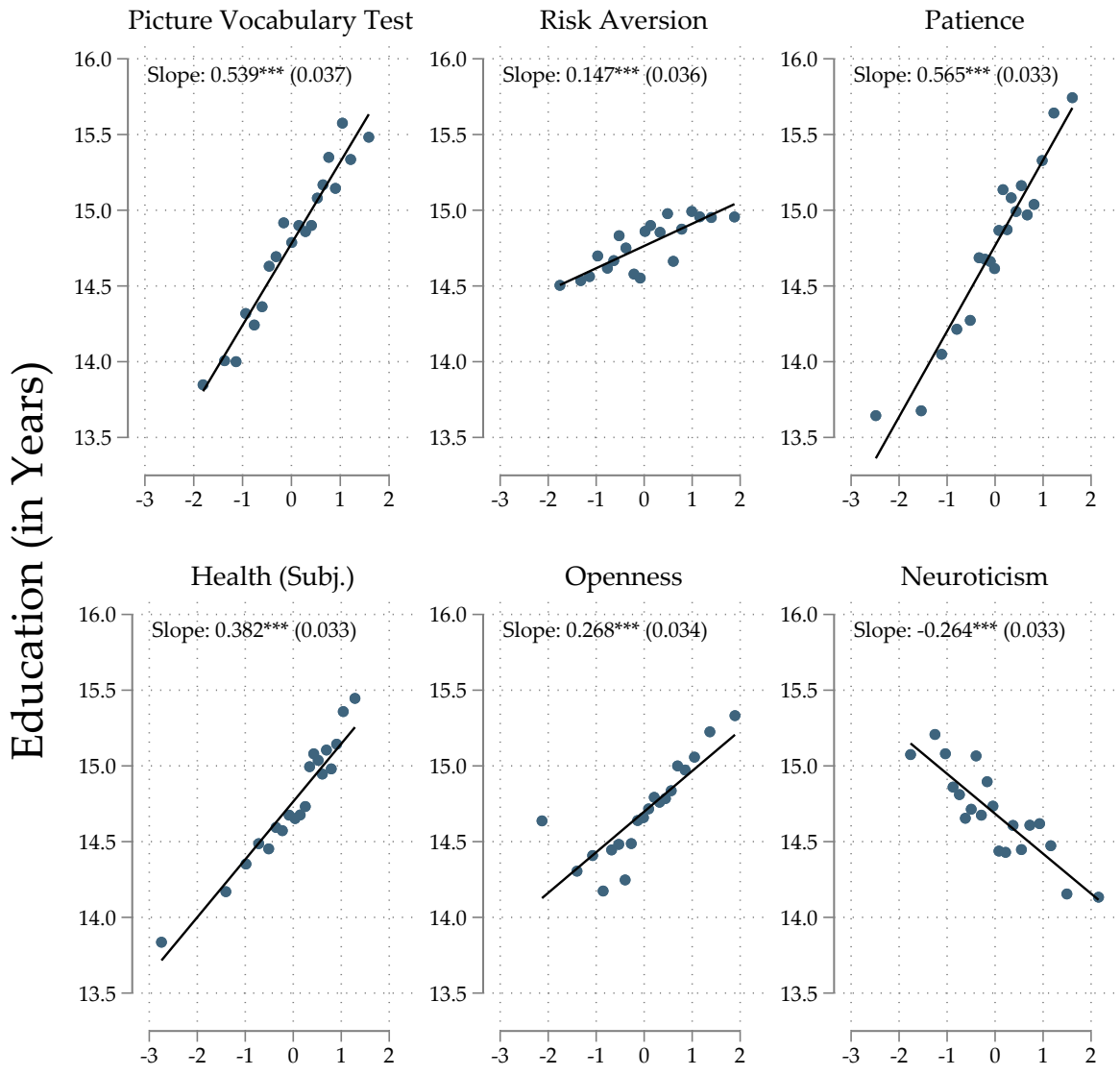
Note: Own calculations. This figure shows predictions of the PVT by PGI^{EA} and Q^S cell. Predictions are calculated using the model estimated in column (1) in Panel (a) of Table 6. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages. Standard errors are clustered at the school level.

FIGURE S.8 – Permutation test for placebo assignments of Q^S



Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows the frequency distribution of *t*-statistics for the interaction of PGI^{EA} and Q^S under 10,000 permutations of Q^S.

FIGURE S.9 – Association of Educational Attainment with (Non-)Cognitive Skills



Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows the association of educational attainment and various measures of (non-)cognitive skills. We bin scatterplots using 20 quantiles of the variable of interest. Black lines are fitted from linear regressions of educational attainment on the variable of interest and the full set of controls. *Child Controls*: Firstborn dummy, linear birth cohort trend (in months) by gender, 20 principal components of the full matrix of genetic data. *Family Controls*: Age of mother at birth, years of education of both mother and father, average potential wages of both mother and father, the standard deviation of potential wages of both mother and father, dummies for non-US born mothers and fathers, a dummy for Christian religion, state fixed effects. *Control Function*: Share white, share single mothers, maternal education (average), peer grades English (average), peer grades Math (average). All control function variables are calculated as leave-cohort-out school averages.

C BIAS IN GENE-ENVIRONMENT INTERACTION

Assume the following population model:

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 (x_{1i} \cdot x_{2i}) + \delta_1 z_i + \delta_2 (x_{1i} \cdot z_i) + \delta_3 (x_{2i} \cdot z_i) + \epsilon_i,$$

with

$$(x_{1i}, x_{2i}, z_i) \sim N\left(0, \begin{bmatrix} 1 & 0 & \sigma^2_{x_1 z} \\ 0 & 1 & \sigma^2_{x_2 z} \\ \sigma^2_{x_1 z} & \sigma^2_{x_2 z} & 1 \end{bmatrix}\right), \; Cov(\epsilon | z, x_1, x_2) = 0.$$

Note we assume $\sigma^2_{x_1 x_2} = 0$ based on the evidence presented in Figure 2. Furthermore, we assume $\sigma^2_{x_1 z} \geq 0, \sigma^2_{x_2 z} \geq 0$ to reflect concerns about i) genetic nurture and ii) selection into schools based on unobservable family background characteristics z.

Since z is unobserved, we estimate the following model:

$$y_i = \tilde{\alpha} + \tilde{\beta}_1 x_{1i} + \tilde{\beta}_2 x_{2i} + \tilde{\beta}_3 (x_{1i} \cdot x_{2i}) + \tilde{\epsilon}_i.$$

What is the bias in the estimated gene-environment interaction $\tilde{\beta}_3$?

By the weak law of large numbers, we know that

$$\hat{\tilde{\beta}} \xrightarrow{p} \beta^* := [\mathbb{E}(X^T X)]^{-1} \mathbb{E}[X^T y].$$

First, under our assumptions about (x_1, x_2), $[\mathbb{E}(X^T X)]$ simplifies to the identity matrix:

$$\mathbb{E}[X^T X] = \begin{bmatrix} 1 & \mathbb{E}[x_1] & \mathbb{E}[x_2] & \mathbb{E}[x_1 x_2] \\ \mathbb{E}[x_1] & \mathbb{E}[x_1^2] & \mathbb{E}[x_1 x_2] & \mathbb{E}[x_1^2 x_2] \\ \mathbb{E}[x_2] & \mathbb{E}[x_1 x_2] & \mathbb{E}[x_2^2] & \mathbb{E}[x_1 x_2^2] \\ \mathbb{E}[x_1 x_2] & \mathbb{E}[x_1^2 x_2] & \mathbb{E}[x_1 x_2^2] & \mathbb{E}[x_1^2 x_2^2] \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Therefore, bias will be driven by $\mathbb{E}[X^T y]$ only.

Second, under our assumptions about (x_1, x_2, z), $\mathbb{E}[X^T y]$ reads as follows:

$$\mathbb{E}[X^T y] = \begin{bmatrix} \alpha + \delta_2 \mathbb{E}[x_1 z] + \delta_3 \mathbb{E}[x_2 z] \\ \beta_1 + \delta_1 \mathbb{E}[x_1 z] \\ \beta_2 + \delta_1 \mathbb{E}[x_2 z] \\ \beta_3 + \delta_2 \mathbb{E}[x_1^2 x_2 z] + \delta_3 \mathbb{E}[x_1 x_2^2 z] \end{bmatrix} = \begin{bmatrix} \alpha + \delta_2 \sigma^2_{x_1 z} + \delta_3 \sigma^2_{x_2 z} \\ \beta_1 + \delta_1 \sigma^2_{x_1 z} \\ \beta_2 + \delta_1 \sigma^2_{x_2 z} \\ \beta_3 + \delta_2 \sigma^2_{x_2 z} + \delta_3 \sigma^2_{x_1 z} \end{bmatrix}.$$

Therefore, bias in the gene-environment interaction is captured by:

$$\hat{\tilde{\beta}}_3 - \beta_3 = \delta_2 \sigma^2_{x_2 z} + \delta_3 \sigma^2_{x_1 z}.$$

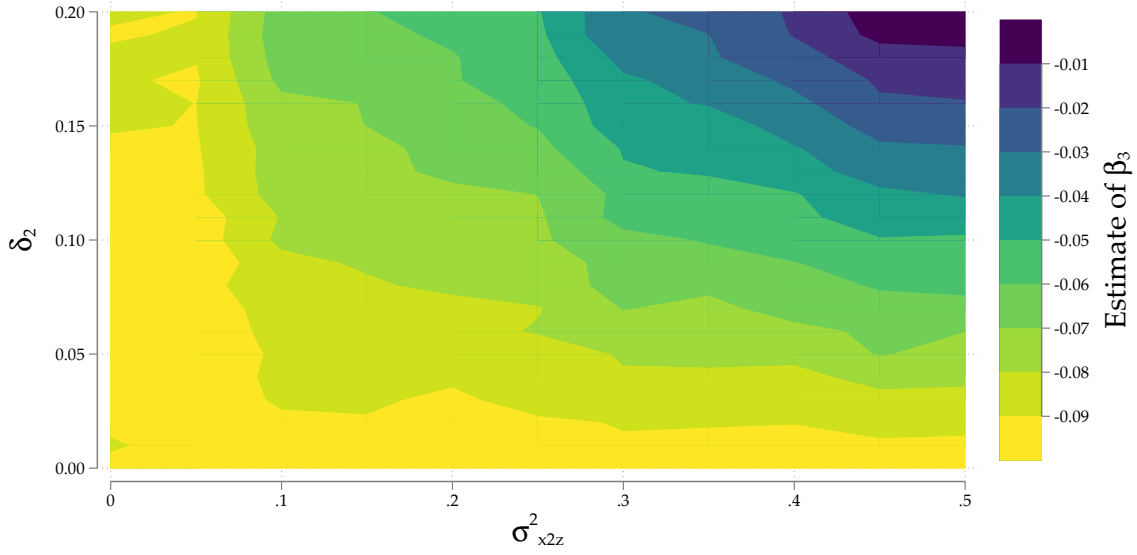
Note our assumptions $\sigma^2_{x_1 z} \geq 0, \sigma^2_{x_2 z} \geq 0$ are insufficient to sign the bias since the signs of δ_2 and δ_3 are unknown. However, we can sign the bias by two additional assumptions that are corroborated by our analysis and existing empirical literature:

$\sigma^2_{x_1 z} = 0 \rightarrow$ We assume the absence of confounding through genetic nurture. This assumption is supported by the similarity of results from between- and within-family models presented in Table 2.

$\delta_2 \geq 0 \rightarrow$ We assume (weak) complementarity between genetic endowments and (unobserved) family background characteristics. This assumption is consistent with the evidence presented in Figure 4 and existing literature confirming (weakly) positive Scarr-Rowe interactions (Figlio et al., 2017; Ronda et al., 2022; Turkheimer et al., 2003).

Under these additional assumptions $\hat{\tilde{\beta}}_3 - \beta_3 \geq 0$. Hence, we would estimate a lower bound of the substitutability between genetic endowments and school quality.

FIGURE S.10 – Simulation for $\beta_3 = -0.10$



Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This figure shows a heatmap of estimates for β_3 under different assumptions about $\sigma^2_{x_2 z}$ and δ_2. The simulation is based on the following data generating process: $y_i = 0.4x_{1i} + 0.2x_{2i} - 0.1(x_{1i} \cdot x_{2i}) + 0.2z_i + \delta_2(x_{1i} \cdot z_i) + 0.2(x_{2i} \cdot z_i) + \epsilon_i$, where $\sigma^2_{x_1 x_2} = \sigma^2_{x_1 z} = 0$, $\sigma^2_{x_2 z} \in [0.0(0.05)0.50]$, and $\delta_2 \in [0.0(0.01)0.20]$. We run 100 iterations for each combination of $\sigma^2_{x_2 z}$ and δ_2.

We confirm this result empirically through a simulation where we estimate β_3 under different assumptions about $\sigma^2_{x_2 z}$ and δ_2. Figure S.10 shows that β_3 increases, i.e., substitutability decreases, with increasing $\sigma^2_{x_2 z}$ and δ_2. This suggests that residual selection into schools that is not accounted for by our control variables would attenuate our estimated gene-environment interaction towards zero.

D DATA APPENDIX

D.1 Outcome Variables

Educational attainment. We measure educational attainment by total *years of education*. In each wave, respondents were asked about their highest level of education at the time of the interview. For each respondent, we use the most recent information and transform education levels into years of education following the mapping suggested by Domingue et al. (2015). Numeric values in parentheses: eighth grade or less (8), some high school (10), high school graduate (12), GED (12), some vocational/technical training (13), some community college (14), some college (14), completed vocational/technical training (14), associate or junior college degree (14), completed college (16), some graduate school (17), completed a Master degree (18), some post-baccalaureate professional education (18), some graduate training beyond a master's degree (19), completed post-baccalaureate professional education (19), completed a doctoral degree (20).

We use the most recent available information to construct the following measures for educational degrees: *High School* (including GED), *2-year College*, *4-year College*, and *Post-Graduate*. Two-year college degrees include associate and junior college degrees as well as vocational and technical training after high school. Four-year college degrees include bachelor's degrees. Post-graduate degrees include master's degrees, doctoral degrees, and post-baccalaureate professional degrees. If available, information is taken from wave 5; otherwise, we take it from wave 4 or 3, respectively. We only include respondents for which we observe educational degrees when they are at least 27 years old at the time of observation. We assume an ordinal ranking of degrees (high school < 2-year college < 4-year college < post-graduate) and assign the possession of a lower-ranked degree if a respondent obtained a higher-ranked degree. For example, we assume that a respondent has finished high school if he or she has obtained a college degree, even if we do not have explicit information about high school graduation status.

Health. We proxy *subjective health* by quality-adjusted life years (QALY) that we derive from self-assessed health (SAH) measures. We use information from waves 3 and 4, where participants were asked "in general, how is your health?" We convert their (categorical) responses into a continuous measure using a mapping proposed by Van Doorslaer and Jones (2003). Using information about objective health—the Health Utility Index Mark III—Van Doorslaer and Jones (2003) scale the intervals of the SAH categories. This approach yields "quality weights" for health between 0 and 1. The values for each health status category are as follows (quality weights in parentheses): "excellent" (0.9833), "very good" (0.9311), "good" (0.841), "fair" (0.707), and "poor" (0.401).[1] We average resulting QALY measures across waves 3 and 4.

[1] See Table 4 in Van Doorslaer and Jones (2003).

We construct an index of *objective health* based on information from wave 4. Specifically, we sum the standardized values about whether a respondent (i) is obese, (ii) has stage one hypertension, and (iii) has high cholesterol (as indicated by the respondent). Each item was answered with either "yes" (= 1) or "no" (= 0). We reverse-code our measure of objective health such that higher values indicate better health.

Cognitive skills. The *Picture Vocabulary Test* (PVT) is a test for receptive hearing vocabulary and is a widely-used proxy for verbal ability and scholastic aptitude. To administer the PVT, an examiner presents a series of pictures to the respondent. There are four pictures per page, and the examiner speaks a word describing one of the pictures. The respondent then has to indicate the picture that the word describes. In our analysis, we use age-adjusted PVT percentile ranks from wave 3 (Harris, 2020).

Preferences. We construct two measures of preferences: *risk aversion* and *patience*. In waves 3 and 4, participants were asked (i) whether they like to take risks, and (ii) whether they live their life without much thought for the future. Questions were answered on a five-point Likert scale ranging from "strongly agree" to "strongly disagree." We reverse-code both measures and use averages from waves 3 and 4 in our analysis.

Personality. The Big Five personality traits are openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (Almlund et al., 2011). We use information from wave 4 to construct personality measures. Participants were asked a set of questions that each relate to one of the five personality traits. Questions were answered on a five-point Likert scale ranging from "strongly agree" to "strongly disagree." We use averages of the following questions in our analysis. *Openness*: (i) "I have a vivid imagination," (ii) "I have difficulty understanding abstract ideas" (reverse-coded), (iii) "I am not interested in abstract ideas" (reverse-coded), (iv) "I do not have a good imagination" (reverse-coded). *Conscientiousness*: (i) "I get chores done right away," (ii) "I like order," (iii) "I often forget to put things back in their proper place" (reverse-coded), (iv) "I make a mess of things" (reverse-coded). *Extraversion*: (i) "I am the life of the party," (ii) "I talk to a lot of different people at parties," (iii) "I don't talk a lot" (reverse-coded), (iv) "I keep in the background" (reverse-coded). Agreeableness: (i) "I sympathize with others' feelings," (ii) "I feel others' emotions," (iii) "I am not interested in other people's problems" (reverse-coded), (iv) "I am not really interested in others" (reverse-coded). *Neuroticism*: (i) "I have frequent mood swings," (ii) "I get upset easily," (iii) "I am relaxed most of the time" (reverse-coded), (iv) "I seldom feel blue" (reverse-coded).

TABLE S.7 – Summary Statistics (Outcomes)

	Obs.	Mean	SD	Min	Max
Educational Attainment					
Years of Education	4,036	14.68	2.27	8.00	20.00
High School Degree	4,034	0.96	0.20	0.00	1.00
2-year College Degree	4,036	0.50	0.50	0.00	1.00
4-year College Degree	4,036	0.39	0.49	0.00	1.00
Post-Graduate Degree	4,036	0.14	0.35	0.00	1.00
Health					
Subjective	3,473	0.91	0.07	0.40	0.98
Objective	4,036	0.05	1.94	-6.46	1.62
Cognitive Skills					
Picture Vocabulary Test	3,382	59.26	26.34	0.00	100.00
Preferences					
Risk Aversion	3,467	2.82	0.86	1.00	5.00
Patience	3,467	3.91	0.73	1.00	5.00
Personality					
Openness	3,998	3.63	0.63	1.00	5.00
Conscientousness	4,030	3.64	0.69	1.25	5.00
Extraversion	4,029	3.33	0.77	1.00	5.00
Agreeableness	4,028	3.86	0.59	1.00	5.00
Neuroticism	4,024	2.57	0.70	1.00	5.00

Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This table shows summary statistics for outcome variables in our core analysis sample. The sample is restricted to genotyped individuals who (i) are of European descent, and (ii) attended an Add Health high school or an associated feeder school in wave 1.

D.2 Variables of Interest

Polygenic Indexes. Add Health obtained saliva samples from consenting participants in wave 4. After quality control procedures, genotyped data is available for 9,974 individuals and 609,130 SNPs. Add Health uses this data and calculates a set of different PGI using summary statistics from existing GWAS. Our baseline measure PGI^{EA} is based on statistics from Lee et al. (2018).

School characteristics. In waves 1 and 2, Add Health administered questionnaires to headmasters of Add Health schools. We use this information to construct an indicator for high school quality using principal components analysis (PCA). We extract the first component from a PCA that includes the following school-level information (component loadings in parentheses): (i) the average class size (-0.31), (ii) the share of teachers with a master's degree (+0.63), (iii) the share of new teachers in the current school year (-0.36), (iv) the share of teachers with

school-specific tenure of more than five years (+0.62). We apply a factor rotation for interpretability reasons (oblique oblimin rotation of the Kaiser normalized matrix with $\gamma = 0$; see Gorsuch, 1983). The calculated factor is standardized to $\mu = 0$ and $\sigma = 1$ on the full sample of Add Health respondents in wave 1.

Alternatively, we use the same information and aggregate across dimensions using the procedure suggested in Anderson (2008) and Kling et al. (2007). The calculated factor is standardized to $\mu = 0$ and $\sigma = 1$ on the full sample of Add Health respondents in wave 1.

Family socioeconomic status. We use the *social origins factor score* constructed by Belsky et al. (2018). Their measure uses information about parental education, parental occupation, household income, and household receipt of public assistance in wave 1. The score is standardized to $\mu = 0$ and $\sigma = 1$ on the full sample of Add Health respondents in wave 1.

TABLE S.8 – Summary Statistics (Variables of Interest)

	Obs.	Mean	SD	Min	Max
Polygenic Scores					
PGI^{EA}	4,036	0.02	0.99	-4.13	3.39
School Characteristics					
Q^S (PCA)	4,036	0.26	1.23	-3.17	2.49
Q^S (Anderson, 2008)	4,036	0.25	0.90	-2.59	1.98
Teacher w/ MA (%)	4,036	51.20	24.10	0.00	95.00
Experienced Teacher (%)	4,036	66.66	23.43	0.00	98.00
New Teacher (%)	4,036	7.87	7.28	0.00	47.00
Class Size	4,036	24.40	4.50	12.00	38.00
Family SES					
Social Origins Factor Score	3,960	0.33	1.14	-4.51	3.51

Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This table shows summary statistics for variables of interest in our core analysis sample. The sample is restricted to genotyped individuals who (i) are of European descent, and (ii) attended an Add Health high school or an associated feeder school in wave 1.

D.3 Control Variables

Child characteristics. The child's *gender* (female or male, as indicated by the interviewer) is taken from the in-home questionnaire in wave 1.

We calculate the child's *age* (in months) at each wave by subtracting the child's birth date from the date of the interview. Because birth dates have minor inconsistencies across waves, we take

averages across waves 1 to 4.

We use the first 20 *principal components* of the full matrix of the genetic data. The components are obtained from a principal components analysis on the matrix of SNPs in Add Health (see Braudt and Harris, 2020, for a discussion). The principal components are standardized to $\mu = 0$ and $\sigma = 1$ on the full sample of genotyped Add Health respondents.

Family socioeconomic status. We use information from wave 1 to construct measures of *parents' education*. We transform parents' highest degree into years of education following the mapping suggested by Domingue et al. (2015). Numeric values in parentheses: never went to school (0), eighth grade or less (8), some high school (10), completed vocational/technical training instead of high school (10), went to school but level unknown (12), respondent doesn't know (12), high school graduate (12), GED (12), completed vocational/technical training after high school (14), some college (14), completed college (16), professional training beyond a master's degree (19). Where available, mothers' and fathers' education refers to the resident parent. If this information is not available, we use the biological parents' education instead.

Information about *mother's age at birth* (in years) is obtained from wave 1 if available, and wave 2 otherwise. To calculate age at birth, we take information about the mother's age (as indicated by the child) and subtract the age of the child at the respective wave.

Information about religion (*Christian* or not) is obtained from wave 1 (as indicated by the child).

We calculate *potential wages* for population group g in time period t according to the following formula (Shenhav, 2021):

$$\hat{w}_{gt} = \sum_{j} \frac{E_{jg,1970}}{E_{g,1970}} \times \sum_{o} \frac{E_{ojg,1970}}{E_{jg,1970}} \left(\pi_{ojt,-r}\right) \times w_{ojt,-r},$$

where $\frac{E_{jg,1970}}{E_{g,1970}}$ describes the group-specific employment share of industry j in 1970, $\frac{E_{ojg,1970}}{E_{jg,1970}}$ describes the group- and industry-specific employment share of occupation o in 1970, $\pi_{ojt,-r}$ describes the leave-region-out industry-specific employment growth in occupation o for the year t relative to 1970 (scaled by the overall employment growth in occupation o for the year t relative to 1970), and $w_{ojt,-s}$ describes the leave-region-out average hourly wage paid in year t for each occupation/industry/region cell. We define groups g by individuals that are homogeneous in gender (male, female), educational attainment ($<$ High School, High School, $>$ High School), and ethnicity (Non-Hispanic White, Hispanic, Non-Hispanic Black). We define regions r by census regions (North-East, Midwest, South, West). Employment shares in 1970 are taken from the 1970 decennial census. Employment shares and wages in periods t are taken from the March Supplements of the Current Population Survey (CPS) over the period 1975-2000. We match the time series of $\hat{w}_{gt}$ to the parents of respondents in Add Health based on information about g. Then we calculate (i) mean potential wages across respondents ages 0–14, and (ii) the

standard deviation in potential wages across respondents ages 0–14.

Control function We construct the control function variables from the in-school questionnaires in wave 1. Specifically, for each school, we calculate the (i) *share of white students*, (ii) the *share of students in single-parent households*, (iii) *average years of education of students' mothers* and school averages for (iv) *GPAs in Math* and (v) *GPAs in English*. We transform mothers' highest degree into years of education following the mapping suggested by Domingue et al. (2015). Numeric values in parentheses: never went to school (0), eighth grade or less (8), some high school (10), went to school but level unknown (12), respondent doesn't know (12), high school graduate (12), GED (12), completed vocational/technical training after high school (14), some college (14), completed college (16), professional training beyond a four-year college (19). Student GPAs are standardized within grade times federal state cells. To prevent mechanical correlation, we calculate all control function variables while excluding the cohort of the respondent (leave-cohort-out).

Other school characteristics. We use information from the school administrator questionnaire in wave 1 to construct for other school policies. Specifically, we construct a binary indicator that assumes a value of 1 if the school uses *ability groups* based on English ability, proxy for the strictness of *retention policy* by calculating the school average of retained students in 1993, and construct a *strictness index* for sanction policies. The strictness index is constructed as follows. School administrators were asked what happens to first-time offenders in the following domains: (i) cheating, (ii) fighting with another student, (iii) injuring another student, (iv) possessing alcohol, (v) possessing an illegal drug, (vi) possessing a weapon, (vii) drinking alcohol at school, (viii) using an illegal drug at school, (ix) smoking at school, (x) verbally abusing a teacher, (xi) physically injuring a teacher, and (xii) stealing school property. Responses are "minor action", "in-school suspension", "out-of-school suspension", and "expulsion." We standardize answers to $\mu = 0$ and $\sigma = 1$ on the full sample of Add Health respondents and then sum across dimensions (Anderson, 2008; Kling et al., 2007).

We also use information from the school administrator questionnaire in wave 1 to construct measures for teacher composition. Specifically, we calculate the schools' share of full-time classroom teachers that the school administrator identifies as (i) *White*, (ii) *Hispanic*, and (iii) *Female*.

Other PGI. We use the PGI for body mass index (*BMI*) (Yengo et al., 2018), attention deficit hyperactivity disorder (*ADHD*) (Demontis et al., 2019), *depressive symptoms* (Howard et al., 2019), *intelligence* (Savage et al., 2018), *smoking* (Liu et al., 2019), and *sleep duration* (Jansen et al., 2019). All polygenic indexes are standardized to $\mu = 0$ and $\sigma = 1$ on the full sample of genotyped Add Health respondents.

Endogenous controls. To measure *parental time investments*, we use the information on a series of activities that children have done with their mother or father in the last four weeks. Specifically, the child is asked whether he or she has (i) gone shopping, (ii) played a sport, (iii) gone to a religious service or church-related event, (iv) talked about someone he or she is dating, or a party he or she went to, (v) gone to a movie, play, museum, concert, or sports event, (vi) had a talk about a personal problem he or she was having, (vii) talked about his or her school work or grades, (viii) worked on a project for school, (ix) talked about other things he or she is doing in school. Questions were answered with "yes" (= 1) or "no" (= 0). We standardize answers to $\mu = 0$ and $\sigma = 1$ on the full sample of Add Health respondents and then sum across dimensions (Anderson, 2008; Kling et al., 2007).

Information about *breastfeeding* and *family income* is taken from wave 1. Parents were asked about whether the child was breastfed (yes or no) and about their income. We use the logarithm of the latter (replacing zero incomes with a 1 to prevent a loss of observations).

TABLE S.9 – Summary Statistics (Controls)

	Obs.	Mean	SD	Min	Max
Child Controls					
Female	4,036	0.54	0.50	0.00	1.00
Age in Months (Wave 1)	4,036	192.41	19.62	144.00	256.00
Firstborn	4,036	0.48	0.50	0.00	1.00
Family Controls					
Education Mother (in Years)	4,036	13.54	2.48	0.00	19.00
Education Father (in Years)	4,036	13.56	2.68	0.00	19.00
Maternal Age at Birth	4,036	25.33	4.84	16.00	46.08
Christian	4,036	0.82	0.38	0.00	1.00
Foreign-born Father	4,036	0.03	0.16	0.00	1.00
Foreign-born Mother	4,036	0.03	0.17	0.00	1.00
Potential Wage/Hour Mother (Mean)	4,036	12.57	1.39	9.40	14.27
Potential Wage/Hour Father (Mean)	4,036	15.40	1.32	11.14	17.11
Potential Wage/Hour Mother (SD)	4,036	0.35	0.11	0.12	0.51
Potential Wage/Hour Father (SD)	4,036	0.39	0.08	0.20	0.65
Control Function					
White Student (School Share)	4,036	0.81	0.18	0.05	1.00
Single Parents (School Share)	4,036	0.24	0.08	0.03	0.58
Educ. Mother (School Average)	4,036	13.40	0.68	12.03	16.25
GPA English (School Average)	4,036	0.02	0.13	-0.45	0.51
GPA Math (School Average)	4,036	0.02	0.15	-0.45	0.54
Other School Characteristics					
Strictness Index	4,036	0.19	0.45	-3.07	0.77
Ability Groups	4,036	0.66	0.47	0.00	1.00
Retention Policy (%)	3,971	3.55	3.72	0.00	20.75
White Teacher (%)	4,036	93.23	11.49	18.00	100.00
Hispanic Teacher (%)	4,036	1.60	4.25	0.00	52.00
Female Teacher (%)	4,036	57.08	13.85	25.00	96.00
Polygenic Scores					
BMI	4,036	-0.02	1.00	-3.42	3.56
ADHD	4,036	-0.04	1.00	-3.82	3.48
Depressive Symptoms	4,036	-0.02	1.00	-3.79	3.55
Intelligence	4,036	0.02	0.99	-3.57	4.64
Ever Smoker	4,036	-0.02	1.00	-4.25	4.25
Sleep Duration	4,036	0.01	0.99	-3.82	2.97
Endogenous Controls					
Breastfed	3,731	0.49	0.50	0.00	1.00
Time Investment Mother	4,035	0.44	4.17	-7.80	13.48
Family income (log)	3,377	3.74	0.74	0.00	6.91

Data: National Longitudinal Study of Adolescent to Adult Health.
Note: Own calculations. This table shows summary statistics for control variables in our core analysis sample. The sample is restricted to genotyped individuals who (i) are of European descent, and (ii) attended an Add Health high school or an associated feeder school in wave 1.

References

ALMLUND, M., A. L. DUCKWORTH, J. HECKMAN, and T. KAUTZ (2011). "Personality Psychology and Economics". *Handbook of the Economics of Education*. Ed. by E. A. HANUSHEK, S. MACHIN, and L. WOESSMANN. Vol. 4. Elsevier, pp. 1–181.

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

BECKER, J. et al. (2021). "Resource profile and user guide of the Polygenic Index Repository". *Nature Human Behaviour* 5 (12), pp. 1744–1758.

BELSKY, D. W. et al. (2018). "Genetic Analysis of Social-class Mobility in Five Longitudinal Studies". *Proceedings of the National Academy of Sciences* 115 (31), E7275–E7284.

BRAUDT, D. and K. M. HARRIS (2020). "Polygenic Scores (PGSs) in the National Longitudinal Study of Adolescent to Adult Health (Add Health)—Release 2".

CINELLI, C. and C. HAZLETT (2020). "Making Sense of Sensitivity: Extending Omitted Variable Bias". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82 (1), pp. 39–67.

DEMONTIS, D., R. K. WALTERS, J. MARTIN, M. MATTHEISEN, T. D. ALS, E. AGERBO, G. BALDURSSON, R. BELLIVEAU, J. BYBJERG-GRAUHOLM, M. BÆKVAD-HANSEN, et al. (2019). "Discovery of the First Genome-wide Significant Risk Loci for Attention Deficit/Hyperactivity Disorder". *Nature Genetics* 51 (1), pp. 63–75.

DOMINGUE, B. W., D. BELSKY, D. CONLEY, K. M. HARRIS, and J. BOARDMAN (2015). "Polygenic Influence on Educational Attainment: New Evidence From the National Longitudinal Study of Adolescent to Adult Health". *AERA Open* 1 (3), pp. 1–13.

FIGLIO, D. N., J. FREESE, K. KARBOWNIK, and J. ROTH (2017). "Socioeconomic status and genetic influences on cognitive development". *Proceedings of the National Academy of Sciences* 114 (53), pp. 13441–13446.

FLOOD, S., M. KING, R. RODGERS, S. RUGGLES, and J. R. WARREN (2020). "Integrated Public Use Microdata Series, Current Population Survey: Version 8.0 [dataset]". *Minneapolis, MN: IPUMS*.

GORSUCH, R. (1983). *Factor Analysis*. 2nd ed. Hillsdale, NJ: Psychology Press.

HARRIS, K. M. (2020). "National Longitudinal Study of Adolescent to Adult Health, Public Use Peabody Picture Vocabulary Test (PVT) Score Data, Wave III [dataset]". *National Longitudinal Study of Adolescent to Adult Health*.

HOWARD, D. M., M. J. ADAMS, T.-K. CLARKE, J. D. HAFFERTY, J. GIBSON, M. SHIRALI, J. R. COLEMAN, S. P. HAGENAARS, J. WARD, E. M. WIGMORE, et al. (2019). "Genome-wide Meta-analysis of Depression Identifies 102 Independent Variants and Highlights the Importance of the Prefrontal Brain Regions". *Nature Neuroscience* 22 (3), pp. 343–352.

JANSEN, P., K. WATANABE, S. STRINGER, N. SKENE, J. BRYOIS, A. HAMMERSCHLAG, C. DE LEEUW, J. BENJAMINS, A. MUÑOZ-MANCHADO, M. NAGEL, et al. (2019). "Genome-wide Analysis of Insomnia in 1,331,010 Individuals Identifies New Risk Loci and Functional Pathways". *Nature Genetics* 51 (3), pp. 394–403.

KLING, J. R., J. B. LIEBMAN, and L. F. KATZ (2007). "Experimental Analysis of Neighborhood Effects". *Econometrica* 75 (1), pp. 83–119.

LEE, J. J. et al. (2018). "Gene Discovery and Polygenic Prediction from a Genome-wide Association Study of Educational Attainment in 1.1 Million Individuals". *Nature Genetics* 50 (8), pp. 1112–1121.

LIU, M., Y. JIANG, R. WEDOW, Y. LI, D. BRAZEL, F. CHEN, G. DATTA, J. DAVILA-VELDERRAIN, D. MCGUIRE, C. TIAN, et al. (2019). "Association Studies of up to 1.2 Million Individuals Yield New Insights into the Genetic Etiology of Tobacco and Alcohol Use". *Nature Genetics* 51 (2), pp. 237–244.

RONDA, V., E. AGERBO, D. BLESES, P. B. MORTENSEN, A. BØRGLUM, D. M. HOUGAARD, O. MORS, M. NORDENTOFT, T. WERGE, and M. ROSHOLM (2022). "Family Disadvantage, Gender and the Returns to Genetic Human Capital". *Scandinavian Journal of Economics* 124 (2), pp. 550–578.

RUGGLES, S., S. FLOOD, R. GOEKEN, J. GROVER, E. MEYER, J. PACAS, and M. SOBEK (2020). "Integrated Public Use Microdata Series USA: Version 10.0 [dataset]". *Minneapolis, MN: IPUMS*.

SANZ-DE-GALDEANO, A. and A. TERSKAYA (forthcoming). "Sibling Differences in Educational Polygenic Scores: How Do Parents React?" *Review of Economics and Statistics*.

SAVAGE, J. E. et al. (2018). "Genome-wide Association Meta-analysis in 269,867 Individuals Identifies New Genetic and Functional Links to Intelligence". *Nature Genetics* 50 (7), pp. 912–919.

SHENHAV, N. (2021). "Lowering Standards to Wed? Spouse Quality, Marriage, and Labor Market Responses to the Gender Wage Gap". *Review of Economics and Statistics* 103 (2), pp. 265–279.

TURKHEIMER, E., A. HALEY, M. WALDRON, B. D'ONOFRIO, and I. I. GOTTESMAN (2003). "Socioeconomic Status Modifies Heritability of IQ in Young Children". *Psychological Science* 14 (6), pp. 623–628.

VAN DOORSLAER, E. and A. JONES (2003). "Inequalities in Self-reported Health: Validation of a New Approach to Measurement". *Journal of Health Economics* 22 (1), pp. 61–87.

YENGO, L., J. SIDORENKO, K. E. KEMPER, Z. ZHENG, A. R. WOOD, M. N. WEEDON, T. M. FRAYLING, J. HIRSCHHORN, J. YANG, P. M. VISSCHER, et al. (2018). "Meta-analysis of Genome-wide Association Studies for Height and Body Mass Index in 700000 Individuals of European Ancestry". *Human Molecular Genetics* 27 (20), pp. 3641–3649.