A Chip Off The Old Block? Genetics and The Intergenerational Transmission of Socioeconomic Status

Sjoerd van Alten[1,3], Leandro Carvalho[3], Marina Aguiar Palma[1], Silvia Barcellos[3], and Titus Galama[1,3]

[1] *Vrije Universiteit Amsterdam*

[2] *Tinbergen Institute*

[3] *University of Southern California*

February 28, 2025

PRELIMINARY DRAFT: PLEASE DON'T CITE OR CIRCULATE

Abstract

We study how one generation's genetics affects the next generation's economic prospects. Genetic data from the Dutch Lifelines Biobank are merged to annual tax records for the period 2006-2021. Our identification strategy exploits the randomness in the genetic transmission from parents to offspring. A parent's genetics influences their offspring's education, income, and wealth. These effects are relatively large when compared to the impacts on parent's own outcomes, and cannot be explained by the transmission of genetic makeup alone. We study the channels through which parent's genetics influences their offspring's socioeconomic status. Genetic transmission accounts for at least 2/3 of effects on education and income but for only 1/3 of the effect on net wealth.

1 Introduction

Increasing economic inequality in many developed countries has made promoting mobility one of the priorities of policy makers. Designing effective policies to tackle this issue requires a better understanding of how socioeconomic advantage is transmitted across generations. Economists and other social scientists have studied extensively the advantages conferred by being born into advantageous environmental circumstances: e.g., having more educated parents (Tsai, Liu, Chou, & Grossman, 2011; Chevalier, Harmon, O'Sullivan, & Walker, 2013; Dickson, Gregg, & Robinson, 2016), access to better schools, or growing up in a neighborhood with lower poverty rates (Chetty, Hendren, & Katz, 2016). We know, however, much less about the role played by genetics.

How does the genetics of one generation affect the economic prospects of the next? There are two challenges to answering this question, concerning measurement and identification. The first challenge is that many datasets containing molecular genetic data were designed to study health and have little or no socioeconomic data. We overcome this challenge by merging genetic data from the Dutch Lifelines Biobank with administrative longitudinal tax records. The tax records include detailed data on wealth, income, savings, and investments from 2006 to 2021 of Lifelines participants and their children. As an individual-level measure of one's genetics, we use the polygenic index (hereafter, PGI) for educational attainment, the best-established measure in the social genomics literature. In our data, this linear index constructed from millions of genetic markers explains about 6.2% of the cross-sectional variation in educational attainment.

The identification challenge requires isolating exogenous variation in genetics, holding constant one's socioeconomic background. Human biology provides a source of exogenous variation. Each one of us have two copies of every genetic marker. During conception, only one of a parent's two copies is transmitted to the offspring. It is random which copy is transmitted. The practical implication is that, once we condition on the PGIs of one's parents, the variation left in one's PGI is random. To verify, we show that, while one's PGI

is associated with the socioeconomic status (SES) of one's parents, the two are orthogonal to each other once we condition on the PGIs of one's parents.

In our empirical analysis, we use data from three generations of a family. We refer to the oldest and youngest generations as the first and third generations, respectively. We estimate the effects of the PGI of the second generation ("the parent") on the SES of the third generation ("the offspring"), holding the PGIs of the first generation ("the grandparents") constant. Our main sample consists of 11,788 third-generation individuals who were 30 years of age or older in 2021. On average, we observe these individuals for a period longer than 8 years.

One generation's genetics does impact the next generation's economic prospects. We find that a one standard deviation increase in a parent's PGI increases her offspring's education by 0.31 years of schooling and the offspring's likelihood of graduating from college by 7.5 percentage points. It also moves the offspring 1.7 percentiles up in the within-cohort distribution of individual income and 1.7 percentiles up in the within-cohortdistribution of net wealth.

These effect sizes are relatively large when compared to the effects that the parent's PGI has on her own outcomes. A one standard deviation increase in the parent's PGI increases the parent's education by 0.41 years of schooling and her likelihood of graduating from college by 7.5 percentage points. It also moves the parent 2.4 percentiles up in the distribution of individual income and 2.2 percentiles up in the distribution of net wealth. These results indicate that the effects of the shock to the parent's genetics have a high persistence across the two generations.

We proceed to study the channels through which a parent's genetics affects her offspring's SES. We distinguish between two broadly-defined channels. Because high-PGI parents tend to be more educated, have higher incomes, and be wealthier than their peers, their offspring will benefit from growing up in higher-SES family environments. We refer to this as the "family background channel." At the same time, children of high-PGI parents will "inherit

higher PGIs" themselves. We refer to this as the "genetic transmission channel." These two channels may be moderated by assortative mating. A parent's PGI may affect her choice of (biological) co-parent. The co-parent in turn would not only transmit half of his genetics to their offspring but could also influence the family environment in which this offspring is raised.

The genetic transmission channel explains about 2/3 of the effect of a parent's genetics on her offspring's education and an even larger fraction of its effects on the offspring's income. By contrast, this channel explains just about 1/3 of the effect of a parent's genetics on her offspring's net wealth. This is in line with evidence from adoption studies that suggest that wealth is subject to a larger share of environmental transmission, than education and income (Black, Devereux, Lundborg, & Majlesi, 2020).

Assortative mating seems to be more important for the family background channel than for the genetic transmission channel. A one standard deviation increase in a parent's PGI leads her to choose as a co-parent someone with 0.23 years more of schooling, who is 4.8 percentage points more likely to have a college degree, and who is 0.9 percentiles higher in the distribution of individual income. In contrast, we cannot reject the hypothesis that there is no "genotypic assortative mating": an exogenous increase in a parent's PGI does not lead her to choose as a co-parent someone who himself has a higher PGI (or a lower PGI).

We illustrate the importance of genetics for the intergenerational transmission of SES by benchmarking our estimates against existing estimates of the causal effects of family environment – in particular, estimates from Black et al. (2020) and from Fagereng et al. (2021) identified using the quasi-random assignment of adopted children into adoptive families. A parent's genetics seems to affect her offspring's education and income at least as much as the family environment. In contrast, the offspring's net wealth seems to depend more on the family environment than on the parent's genetics – consistent with Black et al. (2020)'s findings. Moving a parent 10 percentiles up in the PGI distribution increases the offspring's education by 0.15 years of schooling, their income by 0.9 percentiles, and their net wealth

by 0.9 percentiles. Moving the adoptive parent 10 percentiles up in the distribution of net wealth would increase the adopted child's education by 0.1 years of schooling, their income by 0.5 percentiles, and their new wealth by 1.7 or 2.7 percentiles (depending on the study).

This paper contributes to the large literature in economics interested in understanding how socioeconomic advantage is transmitted across generations[1]. Our contribution is to shed light on how genetics contributes to this transmission.

To our knowledge, our paper is the first to use molecular genetic data to study how the genetics of one generation affects the economic prospects of the next. Previous studies had to infer genetics' role indirectly by combining model assumptions with data on the similarity between the outcomes of different family members (e.g., adopted child vs. biological parent; dizygotic vs monozygotic twins; or adopted children vs. biological children). One other contribution of this study is to quantify the relative importance of the two channels through which a parent's genetics can affect the SES of her offspring. It also speaks to a growing literature in social genomics that have exploited the exogenous variation in one's genetics to study its effects on one's own outcomes.[2]. This paper makes two contributions to this literature. One is to estimate these effects using longitudinal tax records containing data on both income and wealth. The other is to advance this literature by studying how one's genetics affects not only one's own outcomes but also how it affects the outcomes of the next generation.

The rest of the paper is structured as follows. Section 2 describes our data. Section 3 introduces the genetic measure used in the empirical analysis. Section 4 outlines our empirical strategy. Section 5 shows our results. Section 6 explores the mechanisms that drive our results. Section 7 discusses the size of our estimates in context with the literature on intergenerational transmission of SES. Section 8 concludes.

[1]see (Bowles & Gintis, 2002; Björklund, Lindahl, & Plug, 2006; Sacerdote, 2007; Black et al., 2020; Fagereng, Mogstad, & Rønning, 2021; Beauchamp, Schmitz, McGue, & Lee, 2023)

[2]see (Belsky et al., 2018; Muslimova, Van Kippersluis, Rietveld, Von Hinke, & Meddens, 2020; Burik, Kweon, & Koellinger, 2021; Okbay et al., 2022; Barcellos, Carvalho, & Turley, 2021; Houmark, Ronda, & Rosholm, 2024; Carvalho, 2024; Buser, Ahlskog, Johannesson, Koellinger, & Oskarsson, 2024)

2 Data

We merge genetic data from the Dutch Lifelines Biobank with administrative tax records made available by the Dutch tax authorities (Belastingdienst) and Statistics Netherlands (CBS). The tax records include detailed data on both income and wealth for the entire population of the Netherlands. We supplement these data with information on completed education from both surveys and administrative data.

2.1 Dutch tax administrative records

We use data on annual income (before taxes) and wealth for the period between 2006 and 2021.[3] Data on income are available both at the individual and the household level.[4] We also use data on annual individual earnings.[5] There is information on various components of net wealth, all measured at the household level. These include total assets, total debt, financial wealth, housing wealth, business assets, other assets, mortgage debt, and other debt. Financial wealth is further divided into (i) checking and savings balances and (ii) bonds and shares.[6] In some of our analyses, we measure income and wealth in terms of percentile ranks, which are calculated within calendar year and year of birth. When estimating the effects on levels, we winsorize the top and bottom 1% of the data.[7]

2.2 The Dutch Lifelines Biobank

Lifelines is a prospective cohort of 167,729 individuals who, at the time of recruitment, lived in Northern Netherlands (i.e., in the provinces of Drenthe, Friesland and Groningen). The sample size corresponds to about 10% of this region's population. Recruitment occurred

[3]Wealth is measured as of January 1st of each year.

[4]Household income is calculated by aggregating the individual incomes of everyone living in the household. CBS provides a household-specific identifier.

[5]Earnings include both labor earnings and self-employed income.

[6]Our measures of wealth exclude pension funds. In the Netherlands, pension funds are not freely transferable and are therefore not considered assets.

[7]Data were deflated using the consumer price index, provided by CBS, with 2015 as the base year.

between 2006 and 2013 using a two-step procedure (Scholtens et al., 2015). First, general practitioners recruited individuals of ages 25 to 49.[8] These original participants were then encouraged to invite their family members to also enroll in the study (no age restrictions applied to these family members). Because of this recruitment design, Lifelines contains multiple members of the same family, which will be instrumental for our empirical strategy. To date, about 78,700 Lifelines participants have been genotyped. We subjected the genetic data to standard quality control steps, in order to ensure that only well-measured genetic variants were included in the analysis (see appendix A1). These quality control checks imply that, in total, we have genetic data of 78,038 Lifelines respondents at our disposal.

2.3 Merging Tax Records and Genetic Data

To study how one generation's genetics affects the SES of the next generation, we need to merge the genetic data of a Lifelines participant with the administrative data on the income and wealth of this respondent's adult children. We do that using a parent-child linkage file ("KINDOUDERTAB") which identifies the legal parents of each person living in the Netherlands. In the Netherlands, legal parenthood is a very reliable predictor for biological parenthood, as adoption is sufficiently rare (<0.5% of the Dutch population (CBS, 2008, 2010)). In any case, any misclassification would bias our estimates towards zero because adoptive parents and their adopted children do not share their genetics.

2.4 Educational attainment

Lifelines collected data on participants' completed education. In some of our analyses, we will also need data on the completed education of family members of Lifelines participants, especially their children. For this reason, we complement the Lifelines self-reported data on education with administrative individual-level data on highest degree completed made

[8]GPs were asked to exclude the following people from participation: those with terminal illness or severe mental illness, those unable to visit the GP, those unable to fill in questionnaires, or those unable to understand Dutch.

available by CBS.[9] The information of CBS is not complete for the full Dutch population, with missingness occurring more frequently in the older cohorts. Education is available for 10,412 out of 11,788 children (88%) within our main sample. The information on highest degree is used to construct a measure of years of schooling and an indicator for college graduates.[10]

3 Genetic Measure

Humans are identical in over 99% of their genomes. They do differ, however, in loci called *Single nucleotide polymorphisms* (SNPs). Although these are not the only type of genetic variation in humans, they are the most common and the most studied. There are about 20 million SNPs spread out across the human genome. At most SNPs, people can have one of two possible genetic variants. In genetic data, one of these two possible genetic variants is arbitrarily chosen as the "reference allele". Because a person has two copies of each chromosome, they may have 0, 1, or 2 copies of the reference allele.

Our genetic measure will be a polygenic index ("PGI"). To maximize predictive power, a PGI combines information from millions of genetic markers, each with a small influence on the outcome of interest. PGIs are widely used in the social sciences genetics literature (Becker et al., 2021). In practice, a PGI is a weighted average of the number of reference alleles an individual has across various SNPs:

$$PGI_i = \sum_{h}^{H} x_{ih} w_h \tag{1}$$

where i indexes an individual and h indexes a SNP. PGI_i is individual i's PGI, x_{ih} is the number of reference alleles (0, 1 or 2) they have at SNP h, and w_h is the weight assigned to

[9] The CBS information is sourced from various administrative sources and population surveys.

[10] Our measure of having a university degree equals one when someone holds a bachelor's degree (or higher) from a higher educational institution (i.e. $\geq$ 16 years of schooling). Note that in the Dutch educational system, such degrees are rewarded by a university ("wetenschappelijk onderwijs") or university of applied sciences ("hoger beroepsonderwijs"). The latter institutions typically reward a bachelor of arts to those completing their degree.

SNP h. The weights are specifically constructed to maximize the PGI's power to predict a given outcome of interest.

In this paper, we use a PGI constructed to predict educational attainment ("EA"). To construct it, we used the publicly-available weights of Okbay et al. (2022). In practice, this "Genome-Wide Association Study" (GWAS) consisted of a series of regressions – one for each SNP – of years of schooling on the individual's number of reference alleles at that SNP.[11] Roughly speaking, the GWAS weights are the coefficients from these regressions. Appendix A2 provides more detail on how these weights are constructed.

Because effect sizes of each individual SNP are typically very small, a large sample is critical for constructing a PGI with high predictive power. Okbay et al.'s GWAS (2022) was originally conducted in a sample of about 3 million individuals. However, their publicly-available GWAS results are based on a smaller subsample of 765,283 individuals that excluded over 2 million participants of 23andMe. This is nevertheless the largest publicly-available GWAS for EA available to date. Following the standard practice in social genomics, the sample is restricted to participants with European genetic ancestries.[12] We standardize the PGI using its cross-sectional standard deviation.

The PGI we construct explains 6.2% of the cross-sectional variation in years of schooling in our data. A one standard deviation increase in one's PGI is associated with an increase in education of about 0.6 years of schooling and a 10 percentage-point increase in the likelihood of graduating from college (see tables A1 and A2.

As we will discuss in the next section, our identification strategy requires having PGIs

[11]The regressions controlled for sex, age, and the first principal components of the genetic data. The principal components are used to control for population stratification: subgroups in a European ancestry population will have different values of $E[x_{ik}]$ due to different ancestral backgrounds. This could lead to spurious associations between x_{ik} and EA if these different subgroups differ in their EA for environmental reasons (Price et al., 2006).

[12]The data providers flagged 249 genotyped respondents as being of non-European ancestry, as based on a combination of variables reported by the respondents and the genetic data, assessing ancestry using a principal components analysis. We do not include non-European ancestry individuals because of the *problem of portability:* due to differing *linkage disequlibrium* patterns among groups of different ancestry, constructing a PGI based on GWAS results estimated in a different ancestry group will result in large measurement error of this PGI and lower predictiveness as a result (Bitarello & Mathieson, 2020; Privé et al., 2022)

of two generations: the PGI of a parent in the second generation, and the PGI of this participant's parents, i.e., the grandparents. Lifelines includes only a few trios where a participant and their two biological parents were all genotyped. Nevertheless, it is possible to impute the genotypes of an individual's parents using *Mendellian imputation*, a methodology recently developed by Young et al. (2022). This method imputes the sum of parental SNPs of each participant, using the SNP value of *either* a genotyped sibling *or* a genotyped parent. Intuitively, the method leverages that, if sibling B has a reference allele at a given SNP that sibling A does not have, then this must be an allele that was transmitted from their parents to sibling B, but not to A. If there is any remaining information about the parents' genotypes that is missing, the population average is imputed. Crucially, while the imputed parental genotype is measured with error, the estimates of the causal effects of the offspring's PGI will remain unbiased in our empirical specification. Intuitively, the reason for unbiasedness is that the parents' genotypes that cannot be imputed are the ones which were not transmitted to the child and are therefore unrelated to the child's genetic variants. See Young et al. (2022) for more details.

4 Empirical Strategy

4.1 Identification

A longstanding challenge in understanding how genetics influences SES is disentangling genetic effects from environmental influences. To address this, we leverage the random genetic variation generated by Mendelian inheritance. Humans have two copies of each genetic marker. During conception, each parent transmits only one of their two copies to their offspring, and the selection of which copy is transmitted occurs at random. This randomness ensures that, once we condition on the sum of the parents' PGIs, the remaining variation in an individual's PGI is effectively random.

Figure 1 provides evidence to support this claim. The top panel shows the association

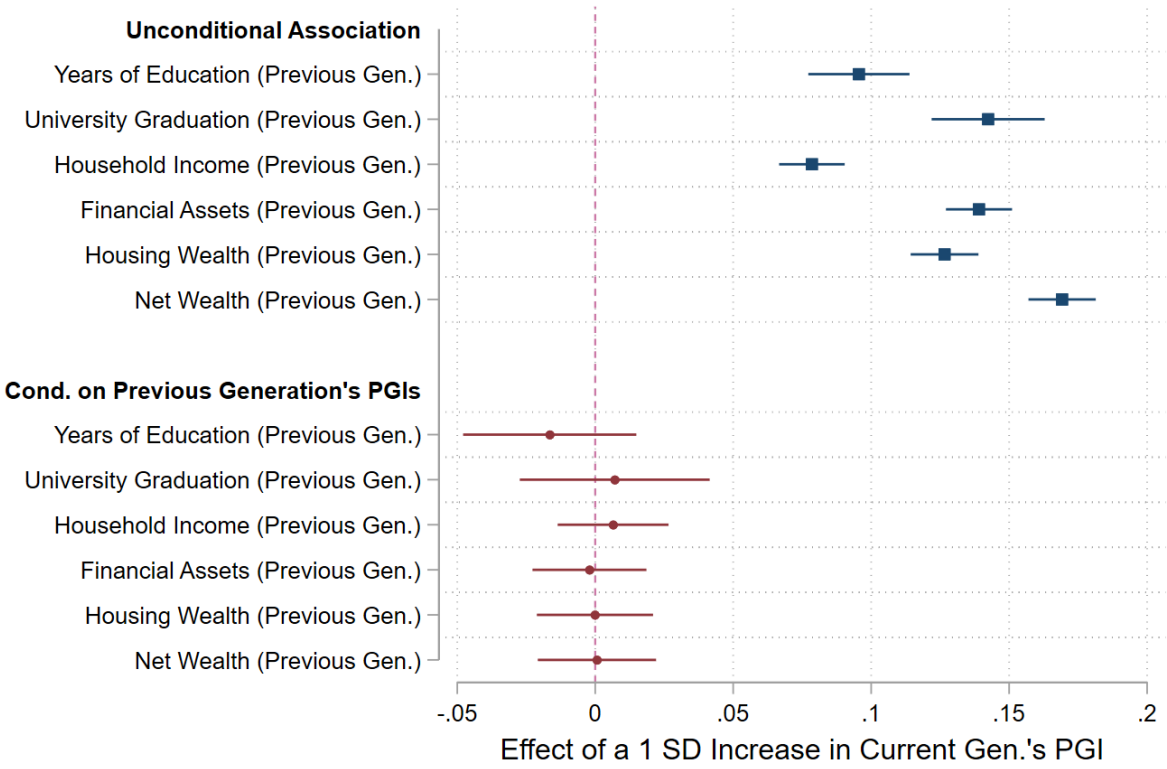

Figure 1: The effect of a PGI for educational attainment on various measures of mean SES of the previous generation (i.e., one's parents). The upper panel of the figure shows the coefficient of the PGI as estimated in a model in which each measure of previous generation's SES is regressed on the PGI for EA. Here, the PGI significantly predicts previous generation's SES, which indicates that the effect of the PGI is biased by environmental confounders. The bottom panel shows the same coefficient when the previous generation's PGI, based on the sum of imputed alleles of one's parents, is included as a control variable. 95% confidence intervals around each point estimate are included. Standard errors are clustered to account for serial correlation in annually measured outcomes. In the second panel, all effects are centered around zero. This indicates that the bias in the upper panel is adequately captured by this control variable. All regressions were estimated within the sample for which we had (1) PGIs of a Lifelines respondent, (2) the sum of imputed PGIs of the Lifelines' respondents parents and (3) the relevant measure of SES for at least one of these Lifelines respondents' parents.

of one's PGI with different measures of the SES of one's parents (i.e., the association of one generation's PGI with the the previous generation's SES). Each row corresponds to a separate regression of one of these measures on one's PGI (the row label identifies the specific SES measure). The marker displays the coefficient on one's PGI. The brackets show 95% confidence intervals. The top panel shows that the parents of high-PGI individuals tend to be more educated, to have higher incomes, and to have accumulated more wealth. As alluded above, one's PGI is confounded with one's socioeconomic background.

The bottom panel of Figure 1 shows what happens to the coefficient on one's PGI when we now control for the (sum of the) PGIs of one's parents. It is clear that, once we control for the PGIs of one's parents, the variation left in one's PGI is orthogonal to the environmental circumstances into which one was born.

4.2 Regression analyses

We can exploit the random recombination of genetic makeup at conception of the Lifelines participant to study how one's genetics affects one's own SES. In particular, we can estimate the causal effect of one's PGI on one's own SES by controlling for the PGIs of one's parents:

$$Y_{i,j}(t) = \alpha_0 + \alpha_1 G_{i,j} + \alpha_2 \bar{G}_i + X_{i,j}(t)\alpha_3 + \epsilon_{i,j}(t) \tag{2}$$

where i indexes this individual's parents and i, j indexes the individual. $Y_{i,j}(t)$ is the individual's outcome in year t and $G_{i,j}$ is their PGI. $\bar{G}_i^1$ is the sum of the PGIs of their parents (the bar in $\bar{G}_i$ is used to remind that it aggregates the PGIs of the individual's father and

mother).[13] The vector $X_{i,j}(t)$ contains a set of control variables.[14]

Moreover, we can leverage the same exogenous variation to study how one's genetics affects the SES of one's offspring by estimating

$$Y_{i,j,k}(t) = \beta_0 + \beta_1 G_{i,j} + \beta_2 \bar{G}_i + X_{i,j,k}(t)\beta_3 + \varepsilon_{i,j,k}(t), \tag{3}$$

where we have replaced the measure of one's own SES on the left-hand side with a measure of the SES of one's offspring. The subscript i, j, k indexes one of the individual's children and $Y_{i,j,k}(t)$ is the socioeconomic outcome of this child. We will distinguish between equations (2) and (3), by saying that the former estimates "same-generation" genetic effects while the latter estimates "next-generation" genetic effects.

To estimate equation (3), we need data from 3 generations of a family: (1) The sum of the PGIs of the oldest generation, the parents of individual i, j; (2) the PGI of individual i, j; and (3) the socioeconomic outcome of child i, j, k in the youngest generation. Hereafter, we will refer to the oldest generation as "the first generation" or as "the grandparents"; the intermediate generation as "the second generation" or as "the parent"; and the youngest generation as "the third generation" or "the offspring". Although $\bar{G}_i$ is the sum of the PGIs of the grandfather and grandmother in the first generation, of the side of parent i, j, for simplicity we will refer to it as the "grandparents' PGI" or as the "first generation's PGI".

Our main object of interest is β_1, the coefficient on the second generation's PGI in

[13] Prior studies have used this model to estimate the effect of the PGI on one's own socioeconomic outcomes (Okbay et al., 2022; Sanz-de Galdeano & Terskaya, 2023). Other papers have relied on within-family estimation to estimate the same type of treatment effect (Domingue, Belsky, Conley, Harris, & Boardman, 2015; Kweon et al., 2020; Buser, Ahlskog, Johannesson, Koellinger, & Oskarsson, 2021). The advantages of this Mendellian imputation vis à vis within-family estimation are that in our specification (1) the outcome Y need not be observed in the sibling, (2) individuals who do not have a genotyped sibling (because their sibling did not participate in Lifelines or because they are an only child), but that do have a genotyped parent, can be included, (3) within-family estimation is biased when the PGI of sibling $j' \neq j$ affects the outcome of individual j, and (4) this specification does not require the inclusion of within-family fixed effects, increasing degrees of freedom and thus power.

[14] The inclusion of these controls is not necessary for achieving unbiasedness. Indeed, conditional on $\bar{G}_i$, the controls are orthogonal to $G_{i,j}$. Rather, we include them to absorb any variance in the error term, increasing the precision of our estimates. Control variables that we typically include are sex, year of birth, and year of measurement of the outcome variable.

equation (3). β_1 describes a clearly defined counterfactual, namely, the increase in the expected value of Y for offspring i, j, k that occurs if the parent would have inherited, at her conception, a different set of genetic variants such that her PGI would be one standard deviation higher. These genes will encompass a large variety of biological effects, which are mostly unknown. Namely, these genes code for proteins that may, in turn, influence various cognitive and non-cognitive skills of the individual. The skills influenced by these genes are likely to influence the human capital production function of both the parent and the offspring.

However, it should be noted that our estimates of β_1 should be considered a lower bound of the true next-generation effect, for various reasons. First, we use the EA PGI as a proxy for the overall genetic propensity of the parent towards an SES-related outcome of the child. The EA PGI was designed to predict years of schooling of parents, and therefore weighs certain genetic markers higher than others. It is certainly possible that genetic markers exist that are relevant for the SES of a child, but that are not too predictive of parents' years of education. These genes will not be captured by the EA PGI, resulting in a lower estimate of β_1 compared to the effect as captured by full parental genetic makeup. However, we expect such bias to be quite limited, as evidence from GWAS suggests that PGIs for SES-related traits other than EA have little to no predictive power, conditional on EA[15]. Further, the EA PGI is subject to measurement error, as its weights are estimated in a GWAS with final sample size. This ensures that the EA PGI that we constructed is only a proxy of the "true" EA PGI, which This results in dilution bias in our estimates 3. Methods to correct for such measurement error exist, and we apply these in Appendix A3 (Becker et al., 2021; Sanz-de Galdeano & Terskaya, 2023; van Kippersluis et al., 2023).

It is important to stress that, even after controlling for the PGIs of the first generation, there is substantial variation left in the second generation's PGI to identify the effects of

[15]Precisely, this statement is rooted in the fact that *genetic correlations*, which measure the overlap between genetic associations predictive for EA and other SES-related outcomes are very high. For example, the genetic correlation between years of education and income is 0.92 (Koellinger et al., 2024) and between years of education and occupational status is 0.99 (Akimova, Wolfram, Ding, Tropf, & Mills, 2023)

interest. In particular, this variation is about 35% of the cross-sectional variation in the PGI.[16] The imputation of the PGIs of the first generation does not bias our estimates of β_1 because the part of the first generation's PGI that cannot be imputed is, by definition, orthogonal to the second generation's PGI (Young et al., 2022). Our balance tests in Figure 1 further illustrate this point.

4.3 Sample and Summary statistics

Our main sample consists of 11,788 individuals in the the third generation of their families – i.e., those indexed by i, j, k in equation (3). For each of these individuals, we have data, not only on their socioeconomic outcomes, but also data on the PGIs of the two previous generations on one side of this individual's family.[17]

We restrict our sample to individuals who were at least 30 years old as of 2021. We impose this age restriction to ensure that they have completed their education and that their income and wealth is indicative of their socioeconomic status.

Table 1 provides summary statistics of the education, income, wealth, and other related characteristics of these individuals. They were born between 1964 and 1991 and were on average 39 years old in 2021. They are highly educated with 15 years of schooling on average. More than 60% of them have a college degree.[18] The average individual earnings and the average net wealth are €41,006 and €92,114, respectively. The medians are €38,978 and €30,607. The portfolio of the typical participant is mostly in checking and savings. The median home equity is only €1,551; its average is €33,322.

Many of the individuals in our sample are siblings. While there are 11,788 offspring in the third generation of our sample, there are only 5,736 distinct parents in the second generation. When we estimate equation (3), we cluster the standard errors at the level of

[16] According to the laws of genetics, this fraction should be around 0.5.(Wang & Xu, 2019) It is lower in our sample because the PGIs of the first generation are imputed.

[17] For about 11% of our sample (1,278 of 11,788 cases), we have PGIs of the two previous generations on both sides of the individual's family. In these cases, we randomly include only the maternal or the paternal family.

[18] As anticipated in Section 2.3, the information on education is missing for about 12% of our sample.

Variable	Mean	Median	SD	P1	P99	N
Time-Invariant Characteristics						
Male	0.5	1	0.5	0	1	11788
Birth year	1982.38	1984	6.93	1964	1991	11788
Years of Schooling	15.02	16	1.92	10	17	10412
University Graduate	0.63	1	0.48	0	1	10412
Time-Variant Characteristics						
Positive Earnings	0.93	1	0.25	0	1	101303
Individual Earnings	41006	38978	26528	0	112108	101303
Individual Income	43425	40398	25762	0	116452	101303
Household Income	82662	77250	42934	10033	219082	101425
Net wealth, levels	92114	30607	208647	-150156	1222568	100180
Assets	256101	210853	257940	0	1553482	100180
Checkings & Savings	26752	13471	39330	0	215341	100180
Stocks & Bonds	5798	0	47252	0	100144	100180
Home equity	33322	1551	91922	-180999	366506	100180
Debt	162888	157350	129394	0	623892	100180
Mortgage	151924	152085	120275	0	531255	100180

Table 1: Summary statistics of the third generation (offsprig). To be included in our main sample, individuals needed to be (1) at least age 30 in 2021, (2) have at least one parent who participated and was genotyped in Lifelines, and (3) this parent needed to have at least a parent *or* a sibling who participated in Lifelines and was also genotyped.

parent i, j – this clustering also corrects for the possibility of serial correlation in the error term.

Summary statistics for the second generation are presented in Appendix Table A3. The second generation is less educated and has lower incomes than the third generation. However, the second generation has accumulated more wealth – that is because the second generation is observed at a much later stage of their life cycles than the third generation: the parents in our sample were born between 1939 and 1968.

Figure 1 used the largest sample possible we could use to conduct the balance test (in order to maximize the statistical power to fail the test). Appendix Figure A1 reproduces the balance test for the sample that we will use in the analyses that follow. The conclusion is the same as in Figure 1, as it shows that the PGI of the second generation is orthogonal to the SES of the first generation, once we condition on the first generation's PGIs.

	In Levels		Percentile Rank	
	β	S.E.	β	S.E.
Education				
Years of Schooling	0.41	0.05		
University	0.075	0.008		
Income & Earnings				
Any Individual Earnings	0.008	0.006		
Individual Earnings	1.6	0.4	1.6	0.5
Individual Income	2.2	0.4	2.4	0.5
Household Income	2.9	0.7	2	0.5
Household Wealth				
Net Wealth	19.2	5.7	2.2	0.5
Assets	23.7	6.5	2.1	0.5
Financial Wealth	4.7	1.7	1.8	0.5
Checkings and Savings	3.7	1.1	1.8	0.5
Stocks and Bonds	3.6	3.1	1.9	0.7
Real estate	13.1	3.2	2.2	0.5
Debt	3.8	2.2	1.1	0.6
Mortgage	1.2	1.9	0.6	0.6

Table 2: Effects of second generation's PGI on second generation's outcomes (estimates of equation 2 for various outcomes $Y_{i,j}$).

5 Results

5.1 Same-Generation Genetic Effects

We first examine how a second-generation individual's genetics influences their SES outcomes. Estimates from equation (2) are presented in Table 2. The odd-numbered columns report estimates of α_1, the coefficient on the second generation's PGI, while the even-numbered columns show robust standard errors. All regressions control for the first generation's PGI, as well as the second-generation's year of birth and sex, and year-of-measurement fixed effects for the time-variant variables (coefficients not reported in the table) .

Second-generation individuals with higher PGIs tend to achieve higher SES in adulthood. On average, they attain more education, earn higher incomes, and accumulate greater wealth than their peers. A one standard deviation increase in a second-generation individual's PGI raises their education by 0.4 years of schooling and increases their likelihood of college graduation by 7 percentage points. It also boosts annual individual and household incomes by €2,200 and €2,900, respectively, corresponding to increases of 2.2 and 2.4 percentiles up in their income distributions. Additionally, it raises net wealth by €19,200, moving them 2.2 percentiles up in the wealth distribution.

These results preview one of the pathways through which second-generation genetics may influence the SES of the third generation. Second-generation individuals with higher PGIs tend to be more educated, earn higher incomes, and accumulate more wealth than their peers. As a result, their children are more likely to grow up in higher-SES family environments, which may, in turn, enhance the child's own SES outcomes. We refer to this as the "family background channel." Simultaneously, the children of high-PGI individuals will inherit higher PGIs themselves, a mechanism we term the "genetic transmission channel." We proceed now to estimate the "total next-generation effect', without distinguishing between these two channels. Later in the paper, we will quantify the relative contributions.

5.2 Next-Generation Genetic Effects

Next, we examine how the genetics of the second generation affects the SES of the third generation. Table 3 estimates equation (3). The odd-numbered columns report estimates of β_1, the coefficient on the second generation's PGI, while the even-numbered columns show clustered standard errors . All regressions control for the first generation's PGI (not reported in the table), as well as the second generation's year of birth and sex, the third generation's sex, and year fixed effects. The last column presents the mean of the dependent variable.

The genetics of one generation causally influence the SES of the next. Just like their parents, the children of high-PGI parents attain more education, earn higher incomes, and accumulate greater wealth than their peers. A one standard deviation increase in a parent's PGI raises the offspring's education by a third of a year of schooling and increases their likelihood of college graduation by 7 percentage points. It also boosts the offspring's average annual individual income by €2,010 and household income by €3,033, moving them up by approximately 1.8 percentiles in these income distributions.

The table also studies the effects on the offspring's non-labor individual income (i.e., individual income minus individual earnings) and on the total income of all other members of the offspring's household (i.e., household income minus the offspring's individual income).

	In Levels		Percentile Rank	
	β	S.E.	β	S.E.
Education				
Years of Schooling	0.3	0.04		
University	0.075	0.009		
Income & Earnings				
Any Individual Earnings	0.003	0.004		
Individual Earnings	1.7	0.4	1.6	0.5
Individual Income	1.7	0.4	1.7	0.4
Household Income	2.9	0.7	1.8	0.5
Non-labor Individual Income	0.02	0.08	0.02	0.43
Income, other household members	1	0.51	0.5	0.42

Table 3: Estimates of the effect of the parent's PGI on SES-related outcomes of their children (estimates of equation 3 for various outcomes $Y_{i,j,k}$). All outcome variables are measured annually except years of schooling and being a university graduate, which are measured only once per child. Additional control variables include parent year of birth fixed effects, the gender of the parent of the child, and year fixed effects for variables that are measured annually. Standard errors are clustered at the level of the parent. In the columns "Percentile Rank", all income and wealth-related outcomes are measured as their within-year-of-birth percentiles.

There is no effect on the former, indicating that the offspring of high-PGI parents have higher individual incomes exclusively because they have higher earnings. There is a marginally significant effect on the income of other household members, which increases by €1,013. One potential explanation for this result is if the offspring of high-PGI parents had spouses with higher incomes on average – something we will return to in section 6.3.

Table 4 studies how the parent's genetics affects the offspring's household wealth. A one standard deviation increase in the parent's PGI increases average financial wealth by €3,100, average total assets by €9,400, and average net wealth by €6,900. In terms of percentile ranks, the increase in the parent's PGI moves the offspring up in the distributions of financial wealth, total assets, and net wealth by about 2.2, 1.7 and 1.7 percentiles, respectively. Interestingly, the point estimates of the effects on financial wealth and on its components are generally a little bit larger than the point estimates of the effects on income (when all these outcomes are measured in percentile ranks). In contrast, the parent's PGI has a relatively smaller effect on housing wealth (1.1 percentiles). As we will discuss in section 5.3, this is partly due to the age composition of our sample.

The effects of parental genes on child's outcomes as shown in table 3 and 4 are reflective of a sizable role that the parental genome plays in shaping the SES-related outcomes of

children, despite the OLS estimates using the EA PGI providing a lower-bound of the true effect of an exogenous shock in parental genetic makeup. In Appendix A3, we correct our estimates for measurement error in the EA PGI, which boosts effect sizes by roughly one third to one half. For example, the effect of a one-standard deviation increase in the EA PGI of the parent on child's years of schooling goes from 0.3 to 0.42 after measurement error correction (see Table A5) and the effect on net wealth percentile rank increases from 1.7 to 2.7.

The next section compares these estimates to the effects of the parental PGI on the parent's own outcomes, and shows that these estimates are reflective of a remarkably strong intergenerational persistence of an exogenous shock to the parental PGI, with persistence being at least 66%, depending on the outcome. Further, section 7 transforms our estimates to rank-rank effects and compares these estimates to existing estimates of the causal effect of *family environment* on children's outcome. This shows that the effect of being born to a parent with a one percentile-rank higher PGI exhibits a larger effect on years of schooling and individual income, keeping constant the parent's environment, than being born to a parent with one-percentile rank higher wealth, keeping constant the parent's genetics.

	In Levels		**Percentile Rank**	
	β	S.E.	β	S.E.
Household Wealth				
Net Wealth	6.9	3.6	1.7	0.5
Assets	9.4	4.6	1.7	0.5
Financial Wealth	3.1	0.9	2.2	0.5
Checkings and Savings	2.9	0.7	2.1	0.5
Stocks and Bonds	-0.1	0.5	1.8	0.6
Real estate	6.0	2.6	1.1	0.5
Debt	2.2	2.2	0.4	0.5
Mortgage	1.8	2.1	0.1	0.5

Table 4: Estimates of the effect of parent's PGI on wealth-related outcomes in their children. All outcomes are measured annually. Additional control variables include parent year of birth fixed effects, the gender of the parent and the child, and year fixed effects. Standard errors are clustered at the level of the parent. In the columns "Percentile Rank", all income and wealth-related outcomes are measured as their within-year-of-birth percentiles.

5.3 Persistence of Genetic Effects across Generations

Figure 2 contrasts the same-generation (reported in table 2) and next-generation genetic effects (reported in table 3 and 4), for selected outcomes. The light blue bars show the same-generation effects, i.e., the effects of the second generation's genetics on the second generation's SES. The dark blue bars show the next-generation effects, i.e., the effects of the second generation's genetics on the SES of the third generation. The brackets show 95% confidence intervals. The left y-axis shows the effects on education. The right y-axis shows the effects on income and wealth, both measured in terms of percentile ranks.

The figure suggests a high persistence of the shock's effects across both generations. Keeping the the confidence intervals into account, we cannot reject the hypothesis that the effects are the same across the generations. If we take the point estimates at face value, they indicate that the effects on the third generation are at least two-thirds as large as the effects on the second generation. One caveat about this exercise is that the two generations are at different points of their life cycles when we observe them. That said, Table A4 shows that, for wealth-related outcomes, the effects on the third generation seem to increase as the individuals in this generation age, suggesting that the persistence may be even higher than shown in Figure 2. Such an age-dependent effect is not observed for education or income-related outcomes. As we will discuss next, this high level of persistence is informative about the mechanisms underlying the next-generation genetic effects: a persistence of 50% is to be expected if the sole reason that the parental PGI affects offspring's SES is because the parent transmits half of their genetic material to the offspring at the offspring's conception. Hence, the pattern of intergenerational persistence above 50% suggests that other mechanisms than mere genetic transmission must mediate the effect of parental PGI on child's SES.

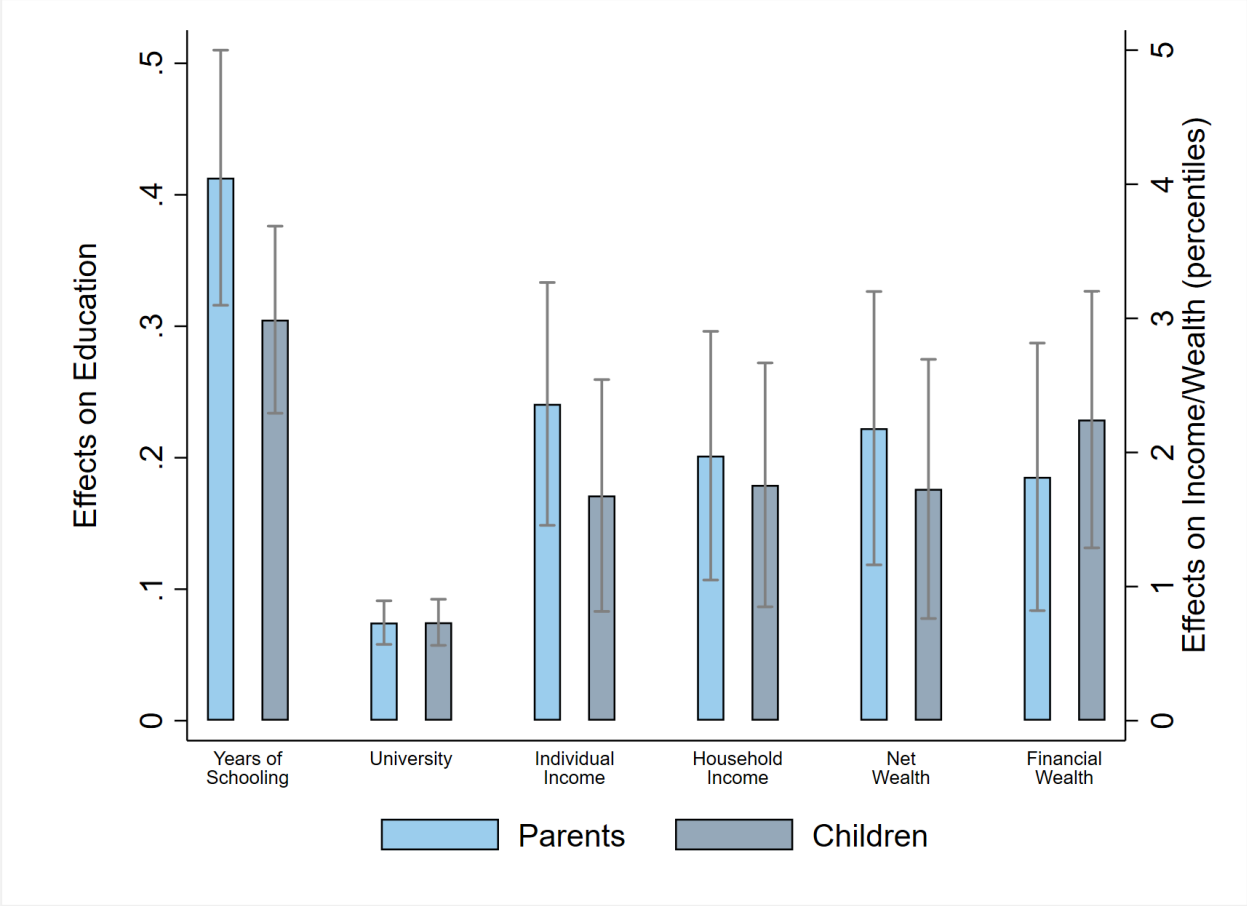

Figure 2: This figure contains the effects of the parental PGI (estimated using specification 3) on the socioeconomic outcomes of (1) the parents themselves and (2) their children. Each bar represents the point estimate of the parental PGI on the parent's outcome (light blue bar) or the child's outcome (dark blue bar). 95% confidence intervals are included. All standard errors are clustered at the level of the parent. Additional control variables include parent-year-of-birth fixed effects, parent's sex, sex of the person whose outcome is tested, and, for annually measured outcomes, year-of-observation fixed effects. Income and wealth-related outcomes are measured in within-year-of-birth percentiles.

6 Mechanisms

In this section, we study the channels through which the genetics of one generation may affect the economic prospects of the next. Figure 3 illustrates these channels. In this figure, we imagine that an individual in the second generation inherits by chance a set of genetic variants that lead them to get more education, resulting also in higher income and more wealth. These are the same-generation genetic effects for the second generation shown in Table 2. We call this individual parent #1. We consider the effects of such an exogenous shock to the genetics of parent #1 in the second generation on the SES-related outcome of this parent's offspring.

Broadly, we can categorize these effects into two conceptually distinct mechanisms: those driven by "genetic transmission" (the left-hand side of the diagram) and those driven by a "family background channel" (the right-hand side of the diagram).

We refer to the effects of the genetic variants transmitted from parent #1 to the offspring (illustrated by the path on the left-hand side of the diagram formed by solid arrows) as the "genetic transmission channel". On average, the offspring should inherit about half of the shock to the genetics of parent #1 (illustrated by the solid red arrow). This genetic inheritance will in turn impact the offspring's SES outcomes through genetic effects (illustrated by the solid light green arrow) – these effects correspond to same-generation genetic effects for the third generation.The genetic transmission channel is equal to the same-generation genetic effect for the third generation times 0.5.

We refer to the effects of parent #1's genetics channeled through her characteristics and behaviors as the "family background channel" (illustrated by the path on the right-hand side of the diagram formed by solid arrows). Independent of genetic transmission, the exogenous shock to parent #1's genetics affects the characteristics and behaviors of parent #1 (illustrated by the solid dark green arrow) – these effects correspond to the same-generation genetic effects for the second generation, which were shown in Table 2. The characteristics and behaviors of parent #1 will in turn shape the family environment in

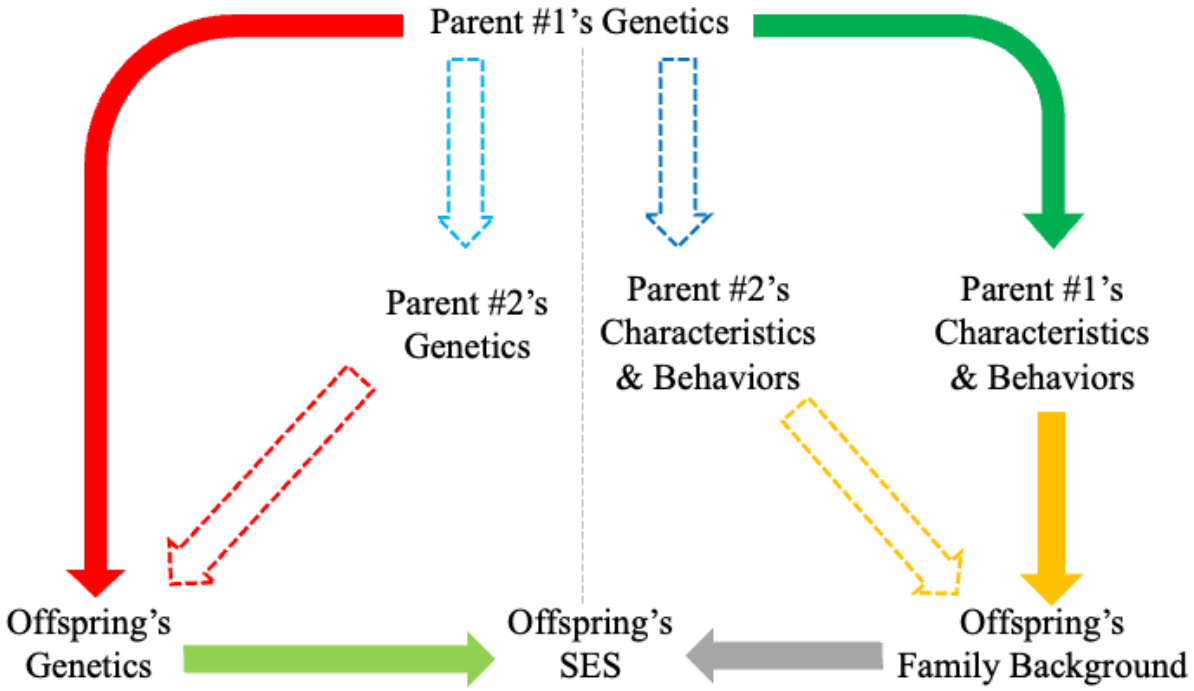

Figure 3: The figure illustrates the channels through which one's genetics (parent #1) may affect the economic prospects of one's offspring. The path on the left-hand side of the diagram formed by solid arrows illustrates the effects of the genetic variants transmitted from parent #1 to the offspring: the solid red arrow illustrates these genetic variants and the solid light green arrow illustrates the effects these variants have on the offspring's SES. The path on the right-hand side of the diagram formed by solid arrows illustrates effects working through the characteristics and behaviors of parent #1. Parent #1's genetics will also influence her characteristics and behaviors (solid dark green arrow on the right-hand side of the diagram), shaping the environment in which her offspring grows up (solid orange arrow). The gray arrow represents environmental effects. The paths formed by dashed arrows capture the contribution of assortative mating, since parent #1's genetics may influence her choice of reproductive partner.

which the offspring grows (illustrated by the solid orange arrow), impacting the offspring's SES outcomes through environmental effects (illustrated by the solid gray arrow).

In the economic literature, widespread evidence is found that family background influences the SES of children. Such evidence is found by looking at the intergenerational spillovers of natural experiments or by estimating adoptive parent-adoptee associations in wealth and other SES-related variables, since the parental genome and the genome of the adopted child are uncorrelated. Such evidence motivates the existence of the solid gray arrow in Figure 3. In our decomposition, the strength of the "family background channel" is driven by the effect of parent #1's genetics on parent #1's characteristics and behaviors, and by the subsequent effect of these characteristics and behaviors on the offspring's SES (the solid green, orange and gray arrows, combined). As such, it is important to clarify that even relatively low estimates of the family background channel do *not* rule out the importance of family background for shaping offspring's outcomes overall. After all, our decomposition only considers changes to offspring's SES due to family background that find their origin in genetic variation of parent #1. Variation in offspring's family background that is orthogonal to parent #1's genetics may have separate effects that our mechanism decomposition does not include.

Finally, assortative mating may contribute towards the next-generation genetic effects (illustrated by the paths formed by the hollow arrows). The shock to parent #1's genetics may influence her choice of biological co-parent (i.e., parent #2). If a one standard deviation increase in parent #1's PGI leads her to have a child with someone who has himself a PGI that is η standard deviation units higher (where η quantifies the hollow light blue arrow), then the offspring will not only inherit 0.5 standard deviation units more from parent #1 but also inherit $\eta/2$ standard deviation units more from parent #2. We will refer to this pathway as "genetic assortative mating". Similarly, if the genetics of parent #1 leads her to have children with someone with particular characteristics and who exhibits particular behaviors (hollow dark blue arrow), this will shape the environment in which their offspring

grows. We will refer to this pathway as "environmental assortative mating".

To quantify the various mechanisms shown in figure 3 that mediate the next-generation estimates previously shown in tables 3 and 4, we proceed in three steps. First, we quantify the genetic transmission channel and its relative contribution for the next-generation genetic effects (the path on the left-hand side of the diagram formed by solid arrows). Here, identification comes from the fact that the natural experiment of randomization of genetic material repeats itself when the offspring is conceived. Second, we investigate the contribution of the genetics of parent #2 (the path on the left-hand side of the diagram formed by hollow arrows). Finally, we study how much assortative mating contributes to shaping the offspring's family background, comparing the relative importance of the two paths illustrated on the right-side of the diagram, the one formed by solid arrows and the one formed by hollow arrows.

6.1 Quantifying Genetic Transmission Channel

To quantify the genetic transmission channel, we need to estimate the same-generation effect for the third generation:

$$Y_{i,j,k}(t) = \gamma_0 + \gamma_1 G_{i,j,k} + \gamma_2 \bar{G}_{i,j} + X_{i,j,k}(t)\gamma_3 + \epsilon_{i,j,k}(t) \tag{4}$$

where $G_{i,j,k}$ is the PGI of the third-generation individual and $\bar{G}_{i,j}$ is the sum of the PGI of parent i, j with the PGI of her biological co-parent.

While the left-hand side of equations (3) and (4) is the same, the requirements in terms of the genetic data needed to estimate these two equations are different. Estimating how one's SES is impacted by the genetics of one's parents (i.e., next-generation effects) requires genetic data of one of this individual's parents ($G_{i,j}$) and genetic data of one's grandparents on this side of the family ($\bar{G}_i$). In contrast, to estimate how one's SES is impacted by one's own genetics (i.e., same-generation effects) requires one's own genetic data ($G_{i,j,k}$) and the

genetic data of one's parents ($\bar{G}_{i,j}$). We would be underpowered to estimate equations (3) and (4) if we confined ourselves to a sample that satisfied all of the genetic data requirements shown in the table.

For this reason, we will use a different sample to estimate equation (4) from the one we used to estimate equation (3). We will refer to the latter as the "next-generation sample" and to the former as the "same-generation sample". The next-generation sample was born between 1964 and 1991. The same-generation sample comprises individuals born between 1964 and 1991 who satisfy the genetic data requirements for estimating same-generation genetic effects (i.e., equation 4). To make the same-generation sample as similar as possible to the next-generation sample, we re-weight the former so it matches the latter in terms of the distributions of year of birth and gender. Appendix A5 details how these weights were constructed. Naturally, there is a concern as to whether the two samples are different in terms of unobservables. Appendix Table A9 shows that, while there are some differences in the unweighted summary statistics of income and wealth of the two samples, these differences are greatly reduced when the summary statistics of the same-generation sample are re-weighted. As such, we conclude that the weights based on year of birth and gender are sufficient such for the weighted same-generation sample to uncover the direct genetic effects of the children used to estimate next-generation effects.

Table 5 shows estimates of the genetic transmission channel. Column (1) and (2) reproduce estimates of next-generation effects from Table 3. Column (3) shows estimates of the genetic transmission channel, namely estimates of same-generation genetic effects for the third generation multiplied by 0.5. Column (4) reports the standard errors of these estimates. Column (5) estimates the fraction of the next-generation genetic effects that is explained by the genetic transmission channel. It corresponds to column (3) divided by column (1). The last columns of the table show the bounds of bootstrapped 95% confidence intervals for the estimates in column (5). Notice that these estimates do not take into account genotypic assortative mating.

	Total Effect		Genetic Transmission		Ratio	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Education						
Years of Schooling	0.30	0.04	0.20	0.01	0.65	0.09
University	0.07	0.01	0.05	0.00	0.65	0.09
Income & Earnings						
Individual Earnings	1.58	0.52	1.49	0.14	0.94	0.85
Individual Income	1.68	0.54	1.50	0.14	0.89	0.72
Household Income	1.76	0.50	1.11	0.12	0.63	0.35
Household Wealth						
Net Wealth	1.73	0.43	0.54	0.18	0.31	0.15
Assets	1.69	0.34	0.72	0.16	0.43	0.14
Financial Wealth	2.24	0.41	1.23	0.15	0.55	0.13
Checkings and Savings	2.12	0.40	1.16	0.14	0.55	0.13
Stocks and Bonds	1.84	0.52	0.97	0.20	0.53	0.56
Real Estate	1.07	0.40	0.51	0.15	0.48	0.28

Table 5: This table shows the extent to which the effect of parental PGI on socioeconomic outcomes of children occurs through a genetic transmission channel: the transmission of genetic material from parent to child, at conception of the child. The left column shows the effect of the parental PGI on the socioeconomic outcomes of the child (as previously shown in table 3). The middle columns show the part of this total effect that can be attributed to the genetic transmission channel. To estimate it, we restrict the effect of assortative mating to be zero, and estimate κ by estimating the effect of one's PGI on one's *own* outcome in the direct effect sample, with inverse probability weights applied to make this sample comparable to the third generation of the main sample. Finally, the right column shows the ratio of the genetic transmission channel to the total effect. When it is lower than one, this implies the remainder of the total effect operates through the family background of the child.

The point estimates imply that the genetic transmission channel explains about 2/3 of the next-generation effects on education. It explains an even larger share of the effects on individual earnings and on individual income. The genetic transmission channel seems to play a smaller role in the next-generation effects on wealth. It explains about a third of the effect on net wealth, suggesting that the remaining two-thirds may be explained by the family background channel.[19]

6.2 Quantifying Genetic Assortative Mating

We study genotypic assortative mating by estimating how an exogenous variation in one's PGI affects the PGI of one's children:

$$G_{i,j,k} = \theta_0 + \theta_1 G_{i,j} + \theta_2 \bar{G}_i + X_{i,j,k}\theta_3 + \varepsilon_{i,j,k}. \tag{5}$$

[19]While genotypic assortative mating could also explain part of the two-thirds, we will show below that the amount of genotypic assortative mating seems to be small.

If there was random mating, then θ_1 should be equal to 0.5. If there was genotypic assortative mating, then θ_1 should be equal to 0.5 $(1 + \eta)$, where η is the parameter quantifying the amount of genotypic assortative mating. If η were larger than zero, it would imply that our estimates of the "genetic transmission channel" as shown in the previous section understate the importance of this channel in explaining the overall next-generation effect. In this section, we will estimate η.

Notice that equation (8) has one more genetic data requirement than equation (3). It also requires genetic data of the third generation, of the offspring. We refer to the sample that satisfies all these genetic data requirements as the "3-generation sample" (N = 2,257). To compensate for the loss in sample size resulting from this additional sample restriction, we include in the 3-generation offspring of all ages, including those under the age of 30 (since the outcome now is time-invariant).

Figure 4 illustrates the causal relationship between the parent's PGI and the offspring's PGI. First, we ran a regression of parent #1's PGI on the PGIs of grandparents #1. We refer to the residuals from such a regression as the parent's residualized PGI. Second, we ran a regression of the offspring's PGI on the PGIs of grandparents #1. We refer to the residuals from such regression as the offspring's residualized PGI. Figure 4 shows a scatterplot of the offspring's residualized PGI against the parent's residualized PGI. The figure also plots the corresponding fitted regression (solid line) and the 22.5 degree line (dashed line). The difference between the two lines quantifies the amount of genotypic assortative mating. If there was random mating, then the two lines would be on top of each other. If the solid line is steeper than the 22.5 degree line indicates that high-PGI individuals have children with people who also have higher PGIs themselves.

The slope of the fitted regression line is 0.48 (s.e. 0.034). We cannot reject the hypothesis of random mating (i.e., $\theta_1 = 0.5$). The p-value of such test is 0.56.[20] This finding contrasts

[20]In appendix A4, we show additional specifications, combining information on the PGIs of parent #1, parent #2, and their children. In our best-powered specification, we find $\hat{\eta} = 0.0137$ $(s.e. = 0.0352, P = 0.698)$.

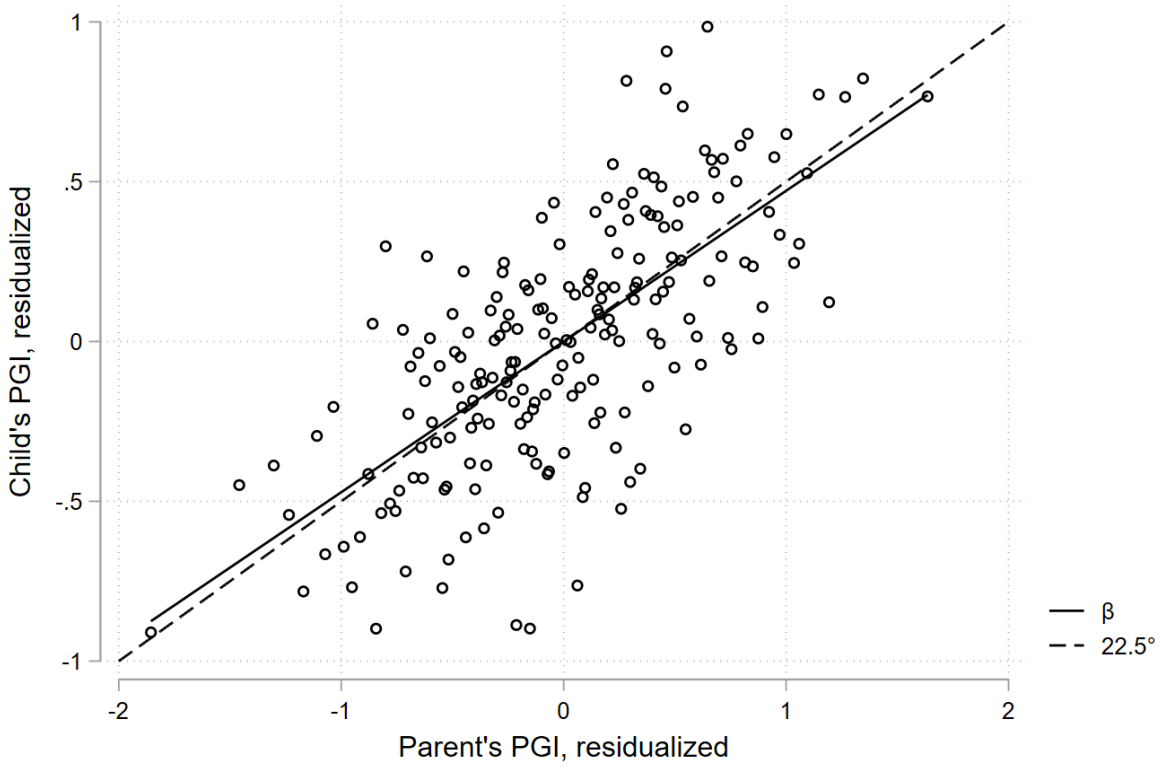

Figure 4: This figure shows evidence that the inheritance of genetic material of the third-generation individuals in our data is not impacted by assortative mating. The y-axis shows the PGI of 2,257 children in our data who participated in Lifelines, residualized from the PGI of their grandparents (from the side of the parent for whom we observe the PGI). The x-axis displays the PGI of the parents, residualized from the same grandparental PGI. The dotted line shows the 22.5-degree line. The solid line has slope equal to the regression coefficient of the Child's residualized PGI on the parental residualized PGI.

with the rest of the literature. For example, Okbay et al. study genotypic assortative mating in terms of the PGI for educational attainment among 862 couples in the UK Biobank and 1,603 couples in Generation Scotland. They document a mate-pair correlation of 0.175 ($s.e. = 0.020$). However, their results do not condition on the genetic makeup of the previous generation. We are not aware of any other study that has tried to quantify genotypic assortative mating using exogenous variation in the PGI of one of the partners as we do.

6.3 Quantifying The Family Background Effect and Environmental Assortative Mating

Figure 2 showed a remarkable persistence of the effect of parental genetics on the SES of the next generation. Table 5 explained this persistence. It showed that genetic transmission only accounts for part of the effect that having a high-PGI parent has on child's SES. The remaining part of the effect cannot be explained through genetic assortative mating. As such, we must conclude that the residual effect of parent #1's PGI operates through family background (the left-hand side of Figure 3: having a parent with a high PGI positively impacts the family background in which the child grows up, which translates in higher SES in adulthood. The exact mechanisms through which this family background effect operates are likely multifaceted and difficult to unpack.

Table 2 already illustrates that high-PGI parents have higher education, income, and wealth themselves. They could transmit these outcomes to the child in various ways. For example, parents with higher education may help the child succeed in the educational system themselves, or make it possible for the parents to help the child with their homework. Income and wealth can be used to buy better educational opportunities, or having a parent with a high-income job can open doors for high-paying job opportunities for the children as well. Finally, wealth can be transmitted to children directly.

Alternatively, the second parent may have a strong impact on the child through environmental assortative mating, the hollow dark blue arrow in Figure 3. In section 6.2, we argued

that the PGI of parent #1 does not strongly impact the genetic makeup of parent #2. However, this does not exclude that the PGI of parent #1 may impact the choice of spouse more generally, and hence influences parent #2's characteristics and behaviors that contextualize the environment in which the offspring goes up. In table A10 we show that the effects of parent #1's PGI on the *individual* SES-related outcomes of the other parent are relatively large. Having a parent with a higher PGI makes it more likely that the other parent of the child is higher educated, and able to earn a higher income. Hence, the choice of one's spouse is an important mechanism through which the family background effect operates.

7 Comparison of Estimates of Parental Genetics to Effects of Family Environment

Some readers may find hard to get a feel for the effect sizes of the next-generation effects shown in Table 3 and in Table 4, and consequently struggle to assess how important the genetics of one generation is for the economic prospects of the next. To this end, we propose to benchmark our estimates against existing estimates of how those prospects are impacted by one's family environment. Our goal is not to perpetuate the false dichotomy between nature and nurture, but rather to drive home that the genetic effects we estimate are not small: they are as large as estimates of the importance of family environment.

Studies in Economics have used the quasi-random assignment of adopted children into adoptive families to estimate the causal effects of family environment (Björklund et al., 2006; Sacerdote, 2007; Black et al., 2020; Fagereng et al., 2021; Beauchamp et al., 2023). Because adoptive families and their adopted children do not share their genetics, any effects of the adoptive family on the adopted child can be attributed to the effects of the environmental circumstances under which adoptive parents raise their adopted children.

We focus on the estimates of Black et al. (2020); Fagereng et al. (2021) because they provide as close to an apples-to-apples comparison as possible. In particular, we focus on

specifications in which they measured family environment characteristics in percentile ranks.

For the purposes of comparison, we make two changes to our estimates. First, we measure the second generation's PGI in percentile ranks rather than in standard deviation units. Second, we correct our estimates for measurement error in the PGI (see appendix A3) .

Figure 5 compares our estimates of the causal impacts of a parent's PGI on the SES outcomes of her offspring to the existing estimates of the causal effects of the family environment on the same outcomes. The pink bars show the effect moving a parent up in the PGI distribution by one percentile rank. The other bars show estimates of the effect of moving adoptive parents up in the distribution of either income or wealth by one percentile rank. The labels in the x-axis identify the offspring's SES outcome being impacted. Impacts on the offspring's years of schooling are shown in the left y-axis. Impacts on the offspring's income and wealth are shown in the right y-axis.

The effects of the parents' genetics on the education and the income of the adult children seem to be at least as large as the effects of family environment. For example, moving a parent up 10 percentiles in the PGI distribution would increase the offspring's education by 0.15 years of schooling and her income by 0.9 percentiles. Moving the adoptive parent up 10 percentiles in the distribution of net wealth would increase the adopted child's education by 0.1 years of schooling and her income by 0.5 percentiles.

On the other hand, the estimated effect of parents' genetics on the offspring's net wealth is smaller than the estimates of how it is impacted by the family's environment. Black et al. (2020) and Fagereng et al. (2021) estimate that moving the adoptive parent up 10 percentiles in the distribution of net wealth would increase the net wealth of the adopted child by 2.7 and by 1.7 percentiles, respectively. In contrast, we estimate that moving a parent up 10 percentiles in the PGI distribution would increase the offspring's net wealth by 0.9 percentiles. This is in line with Black et al.'s (2020) observation that wealth is different: an environment conducive to the generation of wealth seems to be of larger importance to the intergenerational transmission of wealth, than to other measures of SES such as years of

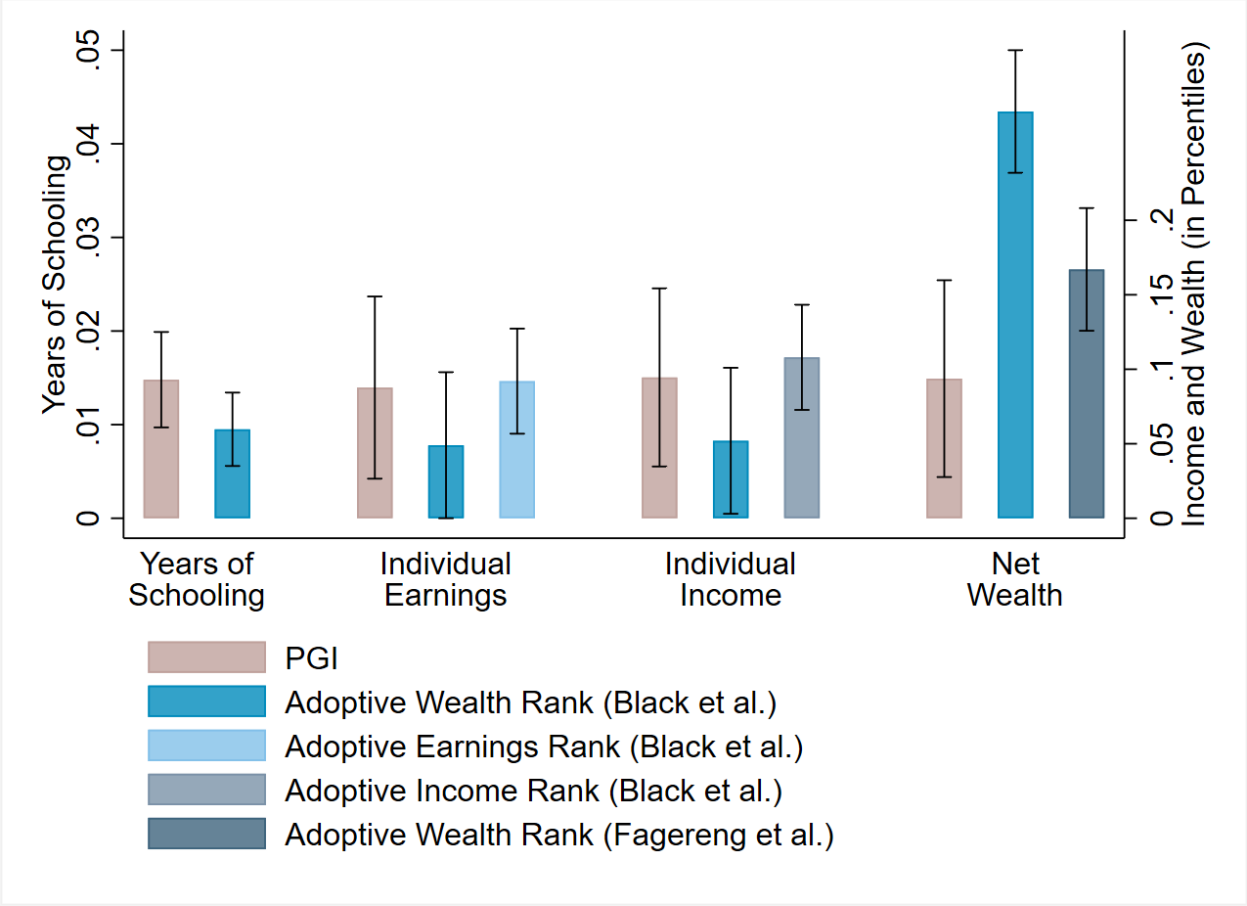

Figure 5: This figure compares the effects of the parental PGI on outcomes of children, as identified by us, to other estimates obtained from previously published adoption studies. The pink bars show the point estimates of the effect of the percentile rank of the parental PGI on the outcomes of children, correct for measurement error due to sampling error in GWAS. The blue bars show the effect of the wealth, earnings, or income rank of adoptive parents on the same outcomes in adoptees, as estimated in previous studies.

schooling and income-related variables. This is because wealth is not only obtained through skills of the children (for which genetic transmission plays a substantial role), but, additionally, wealth can be directly transferred across generations, resulting in a large environmental component (Black et al., 2020).

8 Conclusion

We have shown that the genetic makeup of parents does not only affect their own socioeconomic outcomes but also those of their children once these children have reached adulthood. Parental genetics, as measured through a polygenic index for educational attainment, have a substantial influence on the educational attainment, income, and wealth of their children. Our findings imply that genetic transmission does, at least in part, play a role in the formation of intergenerational correlations of socioeconomic outcomes.

We have further shown that parental genes do not only impact the outcomes of their children because they are transmitted genetically at conception of these children, but that they separately influence children's socioeconomic outcomes because the same genetic index impacts the family environment, which in turn affects the socioeconomic outcomes of these children. In the sociogenetic literature, the idea of genetic makeup affecting offspring through non-genetically transmitted pathways is termed "genetic nurture". Evidence for its existence for educational attainment has been mixed (Kong et al., 2018; Nivard et al., 2022). We show that such "genetic nurture" plays a role in educational attainment, explaining about 1/3rd of the total effect of genetic transmission. We also show, for the first time, that this family background channel plays an even more important role in the intergenerational transmission of wealth.

For our wealth-related estimates, the effect of the parental PGI operating through family background trumps the genetically transmitted effect. This reiterates the point made by Black et al. (2020) that the intergenerational transmission of wealth is different than that

of education and income, and thus needs to be studied separately. Namely, education and income are likely to a larger extent influenced by skills, for which genetic transmission may play a substantial role. Although these same skills may be beneficial to wealth creation as well, wealth has the additional characteristic that it can be directly transferred across generations: an additional environmental channel. Unfortunately, we were not able to study consumption as an overall measure of welfare. Black et al. (2020) suggest that, like wealth, it is more determined by environment than by biology.

Overall, our results imply that the effects of genetics should not be studied in a vacuum and that strong genetic effects are compatible with strong environmental mediation. Our results imply that the intergenerational transmission of SES is partially formed because parents transfer their genetic makeup conducive to obtaining certain values of SES to their children at conception. This genetic makeup, in turn, affects the SES of children throughout their life cycle. It is important to note that such direct effects driven by the child's genetic makeup should *not* be interpreted as evidence for genetic determinism. Studies have shown that PGI effects can interact with the environment, and are thus modifiable by, for example, policy. For example, in the UK, the effect of the EA PGI on wages increased after the school leaving age was raised (Barcellos et al., 2021). The genes included in the EA PGI may operate through a wide variety of behaviors and environments, and, for example, induce people to move to higher educated areas (Abdellaoui et al., 2019). The PGI may also influence various traits that are irrelevant to an individual's educational merit, but that still affect one's education due to social structure or discrimination (Jencks et al., 1972; Coop & Przeworski, 2022). Although we estimate our results in individuals of White European background only, and control for gender, effects of the PGI could still arise due to various sorts of discrimination.

Importantly, the extent to which intergerational persistence of SES is genetically driven puts no upper bound on the extent to which policy can attenuate (or even eradicate) this persistence (Goldberger, 1979). Instead, our findings are useful to understand how a specific

measure of genetics affects socioeconomic inequality given how the society of the Netherlands is structured. Our results are thus of relevance to policy makers: knowing the extent of genetic transmission of socioeconomic outcomes is vital to knowing *which* policy is most effective in decreasing socioeconomic inequality of subsequent generations.

References

Abdellaoui, A., Hugh-Jones, D., Yengo, L., Kemper, K. E., Nivard, M. G., Veul, L., ... others (2019). Genetic correlates of social stratification in great britain. *Nature human behaviour*, *3*(12), 1332–1342.

Akimova, E. T., Wolfram, T., Ding, X., Tropf, F. C., & Mills, M. C. (2023). Polygenic predictions of occupational status gwas elucidate genetic and environmental interplay for intergenerational status transmission, careers, and health. *bioRxiv*, 2023–03.

Barcellos, S. H., Carvalho, L., & Turley, P. (2021). *The effect of education on the relationship between genetics, early-life disadvantages, and later-life ses* (Tech. Rep.). National Bureau of Economic Research.

Beauchamp, J., Schmitz, L., McGue, M., & Lee, J. (2023). Nature-nurture interplay: Evidence from molecular genetic and pedigree data in korean american adoptees. *Available at SSRN 4491976*.

Becker, J., Burik, C. A., Goldman, G., Wang, N., Jayashankar, H., Bennett, M., ... others (2021). Resource profile and user guide of the polygenic index repository. *Nature human behaviour*, *5*(12), 1744–1758.

Belsky, D. W., Domingue, B. W., Wedow, R., Arseneault, L., Boardman, J. D., Caspi, A., ... others (2018). Genetic analysis of social-class mobility in five longitudinal studies. *Proceedings of the National Academy of Sciences*, *115*(31), E7275–E7284.

Bitarello, B. D., & Mathieson, I. (2020). Polygenic scores for height in admixed populations. *G3: Genes, Genomes, Genetics*, *10*(11), 4027–4036.

Björklund, A., Lindahl, M., & Plug, E. (2006). The origins of intergenerational associations: Lessons from swedish adoption data. *The Quarterly Journal of Economics*, *121*(3), 999–1028.

Black, S. E., Devereux, P. J., Lundborg, P., & Majlesi, K. (2020). Poor little rich kids? the role of nature versus nurture in wealth and other economic outcomes and behaviours. *The Review of Economic Studies*, *87*(4), 1683–1725.

Bowles, S., & Gintis, H. (2002). The inheritance of inequality. *Journal of economic Perspectives*, *16*(3), 3–30.

Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., of the Psychiatric Genomics Consortium, S. W. G., ... Neale, B. M. (2015). Ld score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics*, *47*(3), 291–295.

Burik, C., Kweon, H., & Koellinger, P. (2021). Disparities in socio-economic status and bmi in the uk are partly due to genetic and environmental luck.

Buser, T., Ahlskog, R., Johannesson, M., Koellinger, P., & Oskarsson, S. (2021). Using genes to explore the effects of cognitive and non-cognitive skills on education and labor market outcomes.

Buser, T., Ahlskog, R., Johannesson, M., Koellinger, P., & Oskarsson, S. (2024). The causal effect of genetic variants linked to cognitive and non-cognitive skills on education and labor market outcomes. *Labour Economics*, *90*, 102544.

Carvalho, L. (2024). Genetics and socioeconomic status. *Forthcoming in the Journal of Political Economy Micro*.

CBS. (2008). *Stiefouderadopties naar leeftijd kind bij indiening van het adoptieverzoek.* Retrieved from `https://www.cbs.nl/nl-nl/cijfers/detail/71395ned#shortTableDescription` (February 20th, 2024)

CBS. (2010). *Sinds 1956 meer dan 55 duizend kinderen geadopteerd.* Retrieved from `https://www.cbs.nl/nl-nl/nieuws/2010/51/sinds-1956-meer-dan`

-55-duizend-kinderen-geadopteerd (February 20th, 2024)

Cheesman, R., Selzam, S., Ronald, A., Dale, P. S., McAdams, T. A., Eley, T. C., & Plomin, R. (2017). Childhood behaviour problems show the greatest gap between dna-based and twin heritability. *Translational psychiatry*, *7*(12), 1284.

Chetty, R., Hendren, N., & Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, *106*(4), 855–902.

Chevalier, A., Harmon, C., O'Sullivan, V., & Walker, I. (2013). The impact of parental income and education on the schooling of their children. *IZA Journal of Labor Economics*, *2*(1), 1–22.

Coop, G., & Przeworski, M. (2022). Lottery, luck, or legacy. a review of "the genetic lottery: Why dna matters for social equality". *Evolution; International Journal of Organic Evolution*, *76*(4), 846.

Dickson, M., Gregg, P., & Robinson, H. (2016). Early, late or never? when does parental education impact child outcomes? *The Economic Journal*, *126*(596), F184–F231.

Domingue, B. W., Belsky, D. W., Conley, D., Harris, K. M., & Boardman, J. D. (2015). Polygenic influence on educational attainment: New evidence from the national longitudinal study of adolescent to adult health. *AERA open*, *1*(3), 2332858415599972.

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS genetics*, *9*(3), e1003348.

Fagereng, A., Mogstad, M., & Rønning, M. (2021). Why do wealthy parents have wealthy children? *Journal of Political Economy*, *129*(3), 703–756.

Goldberger, A. S. (1979). Heritability. *Economica*, *46*(184), 327–347.

Houmark, M. A., Ronda, V., & Rosholm, M. (2024). The nurture of nature and the nature of nurture: How genes and investments interact in the formation of skills. *American Economic Review*, *114*(2), 385–425.

Jencks, C., et al. (1972). Inequality: A reassessment of the effect of family and schooling in

america.

Koellinger, P. D., Kweon, H., Burik, C., Ning, Y., Ahlskog, R., Xia, C., ... others (2024). Associations between common genetic variants and income provide insights about the socioeconomic health gradient.

Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., ... others (2018). The nature of nurture: Effects of parental genotypes. *Science*, *359*(6374), 424–428.

Kweon, H., Burik, C., Karlsson Linnér, R., De Vlaming, R., Okbay, A., Martschenko, D., ... Koellinger, P. (2020). Genetic fortune: Winning or losing education, income, and health.

Lloyd-Jones, L. R., Zeng, J., Sidorenko, J., Yengo, L., Moser, G., Kemper, K. E., ... others (2019). Improved polygenic prediction by bayesian multiple regression on summary statistics. *Nature communications*, *10*(1), 5086.

Marees, A. T., de Kluiver, H., Stringer, S., Vorspan, F., Curis, E., Marie-Claire, C., & Derks, E. M. (2018). A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *International journal of methods in psychiatric research*, *27*(2), e1608.

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

Nivard, M. G., Belsky, D., Harden, K. P., Baier, T., Ystrom, E., van Bergen, E., & Lyngstad, T. H. (2022). Neither nature nor nurture: Using extended pedigree data to elucidate the origins of indirect genetic effects on offspring educational outcomes.

Nolte, I. M. (2020). Metasubtract: an r-package to analytically produce leave-one-out meta-analysis gwas summary statistics. *Bioinformatics*, *36*(16), 4521–4522.

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., ... others (2022). Polygenic prediction of educational attainment within and between families

from genome-wide association analyses in 3 million individuals. *Nature genetics*, *54*(4), 437–449.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*, *38*(8), 904–909.

Privé, F., Aschard, H., Carmi, S., Folkersen, L., Hoggart, C., O'Reilly, P. F., & Vilhjálmsson, B. J. (2022). Portability of 245 polygenic scores when derived from the uk biobank and applied to 9 ancestry groups from the same cohort. *The American Journal of Human Genetics*, *109*(1), 12–23.

Sacerdote, B. (2007). How large are the effects from changes in family environment? a study of korean american adoptees. *The Quarterly Journal of Economics*, *122*(1), 119–157.

Sanz-de Galdeano, A., & Terskaya, A. (2023). Sibling differences in genetic propensity for education: How do parents react? *Review of Economics and Statistics*, 1–44.

Scholtens, S., Smidt, N., Swertz, M. A., Bakker, S. J., Dotinga, A., Vonk, J. M., ... others (2015). Cohort profile: Lifelines, a three-generation cohort study and biobank. *International journal of epidemiology*, *44*(4), 1172–1180.

Trejo, S., & Domingue, B. W. (2018). Genetic nature or genetic nurture? introducing social genetic parameters to quantify bias in polygenic score analyses. *Biodemography and Social Biology*, *64*(3-4), 187–215.

Tsai, W.-J., Liu, J.-T., Chou, S.-Y., & Grossman, M. (2011). *Intergeneration transfer of human capital: Results from a natural experiment in taiwan* (Tech. Rep.). National Bureau of Economic Research.

van Kippersluis, H., Biroli, P., Dias Pereira, R., Galama, T. J., von Hinke, S., Meddens, S. F. W., ... Rietveld, C. A. (2023). Overcoming attenuation bias in regressions using polygenic indices. *Nature communications*, *14*(1), 4473.

Wang, M., & Xu, S. (2019). Statistics of mendelian segregation—a mixture model. *Journal of Animal Breeding and Genetics*, *136*(5), 341–350.

Young, A. I., Nehzati, S. M., Benonisdottir, S., Okbay, A., Jayashankar, H., Lee, C., ... Kong, A. (2022). Mendelian imputation of parental genotypes improves estimates of direct genetic effects. *Nature genetics*, *54*(6), 897–905.

Acknowledgements

We thank Hans van Kippersluis and Dilnoza Mislimova for providing us with split-sample summary statistics of the UK Biobank to create split-sample IV PGIs for educational attainment. We thank Dalton Conley for his valuable comments and feedback. We further thank participants at the European Social Science Genetics Conference 2023, the Behavior Genetics Association Meeting 2023, the Integrating Genetics and Social Science Conference 2023, and the NBER Social Mobility conference for valuable comments and insights.

A1 Quality control procedures of genetic data

In Lifelines, Study participants were genotyped using blood samples drawn during the first assessment visit. Lifelines genotyped data was released for two subsamples. The first is the GWAS CytoSNP cohort which consists of 15,400 unrelated respondents that were 18 years or older. The second is the UGLI cohort which consists of 64,589 respondents aged 8 or older.

Genotypes of the CytoSNP cohort were measured using the Illumina CytoSNP-12v2 array, measuring $\sim$ 300,000 SNPs. Genotypes of the UGLI cohort were measured using the Infinium Global Screening Array®(GSA) MultiEthnic Disease Version, measuring 691,072 SNPs. Both genetic data cohorts were subject to strict quality control procedures prior to release.[21] Further, missing SNPs not measured by the genotyping arrays were imputed using the dense reference panels Genome of the Netherlands and 1000 Genomes. As a result, $\sim$40,000,000 loci are assessed in both subcohorts.[22]

To construct well-estimated PGIs, we performed various quality control (QC) procedures: ensuring only well-estimated SNPs and respondents with reliable genetic data were included. Most of these are recommended by Marees et al. (2018). We only use data on the first 22 chromosomes, ignoring the sex chromosome, which is the smallest. Restricting to the first 22 chromosomes ensures that the EA PGI will not be artificially higher in any of the sexes.

First, we treated the CytoSNP and UGLI cohorts as separate cohorts. Within each cohort, we dropped multiallelic SNPs and loci with a minor allele frequency of $< 1\%$ ($\sim$33.5 million in CytoSNP, $\sim$27.2 million in UGLI). We further dropped SNPs with low imputation quality, as determined by an info score of < 0.8 (744,661 SNPs in CytoSNP in 333,409 in

[21]The quality control reports for CytoSNP and UGLI are available from `http://wiki.lifelines.nl/doku.php?id=gwas` and `http://wiki.lifelines.nl/lib/exe/fetch.php?media=qc_report_ugli_r1.pdf`, respectively.

[22]Reliable imputation is feasible because SNPs are inheritated in chunks (called haplotype blocks). This implies that SNPs that are closely located to one another in the genome are highly correlated (R^2 values > 0.99 are not uncommon). To save costs, genotyping arrays are designed to only measure a subset of SNPs in a given genomic region, knowing that reliable imputation can be used to map out the non-measured nearby SNPs.

UGLI). Last, we dropped individuals with excess homozygosity rates (3 standard deviations above or below the average), removing 126 respondents in CytoSNP and 529 in UGLI.

Further, we dropped SNPs that were not in Hardy-Weinberg Equilibrium (with p-value threshold 10^{-6}) (1,163 SNPs in CytoSNP, 22,549 in UGLI)[23])

1,289 respondents in the CytoSNP cohort were also part of the UGLI cohort. We removed these respondents from the CytoSNP cohort to avoid double counting. After all QC steps, there were 6,789,250 SNPs present in CytoSNP, and 7,000,369 in UGLI. We restricted both data sets to the 6,408,251 SNPs that they both had in common, and combined them using the –bmerge command in PLINK.[24]

A2 Details on estimation of PGI weights

We calculated polygenic indices on the genetic data in the Lifelines as follows. First, we used MetaSubtract to correct the publicly available summary statistics in the most recent GWAS on years of education for the inclusion of the Lifelines CytoSNP cohort in GWAS discovery (Nolte, 2020; Okbay et al., 2022). To implement Metasubtract, we first replicated Okbay et al.'s (2022) GWAS on years of education in the CytoSNP cohort: we first residualized years of education in this cohort from the first 10 principal components of the genetic data, a cube in age, a sex dummy, and an interaction of this sex dummy with the cube in age. We next used plink2 to perform a GWAS on this residualized years of education variable. We verified the validity of this GWAS analysis by checking the genetic correlation between Okbay et al.'s GWAS summary statistics and our own summary statistics conducted within the Lifelines CytoSNP cohort, and found no significant difference from one, as expected.[25]

[23]In both data cohorts, the data providers already performed this quality check prior to releasing the data, but in CytoSNP, they used a more lenient threshold of $P < 0.0001$

[24]The QC report of the first release of the UGLI cohort includes 606 respondents that were also genotyped using the CytoSNP genotyping array. The UGLI Quality Control report assessed that the concordance within individuals of the genotypes assessed using both arrays was extremely high, being 99.82% in the respondent with lowest concordance

[25]Genetic correlations were calculated using LD-score regression (Bulik-Sullivan et al., 2015): $r_g = 1.07$ $(s.e. = 0.10)$.

Finally, we used Metasubtract to process the summary statistics of Okbay et al. (2022) prior to PGI construction. Metasubtract anallytically subtracts the GWAS summary statistics of CytoSNP from Okbay et al.'s (2022) summary statistics, using inverted versions of the formulas used to meta-analyze GWAS summary statistics. The resulting processed versions of Okbay et al.'s (2022) summary statistics are therefore independent from the CytoSNP cohort.

Using these processed EA GWAS summary statistics, we next constructed PGIs for EA in the Lifelines genetic data, using SbayesR (Lloyd-Jones et al., 2019). This algorithm uses Bayesian regression to correct GWAS summary statistics for linkage disequilibrium, using a mixture of normal distributions as the prior. We use linkage disequilibrium scores included in the SBayesR software, estimated on respondents of the UK Biobank.

A3 Correcting estimates for Measurement Error in the EA PGI

The EA PGI is measured with error because of the measurement error in the estimated GWAS weights used to construct the PGI (Dudbridge, 2013). This measurement error arises because the weights are estimated in GWAS that use a large, but finite, sample size. Hence, these weights are subject to classic sampling error, which causes dilution bias in our OLS regressions that test for the effect of the EA PGI on same-generation and next-generation outcomes. This diluation errors would not occur if one could observe the "true PGI", which is defined as the EA PGI as constructed using weights that are estimated in a hypothetical GWAS of infinite sample size.

To correct our estimates for this type of measurement error, we use the measurement error correction adjustment proposed by van Kippersluis et al. (2023). To implement it, we need two different noisy proxies of "the true PGI". We can then use one proxy to instrument for the other using Obviously Related Instrumental Variables (ORIV) To construct two different

PGIs, a GWAS sample is randomly split into two and two separate independent GWAS are conducted. The two sets of GWAS weights are then used to construct two different PGIs. van Kippersluis et al. (2023) show that this estimator yields a consistent estimate irrespective of the sample size used in GWAS discovery. This results in the following specification:

$$\begin{bmatrix} Y_{i,j,k}(t) \\ Y_{i,j,k}(t) \end{bmatrix} = \begin{bmatrix} \beta_{01} + \beta_1 G_{1i,j} + \beta_2 \bar{G1}_i + X_{i,j}(t)\beta_3 + \varepsilon_{1i,j,k}(t) \\ \beta_{02} + \beta_1 G_{2i,j} + \beta_2 \bar{G2}_i + X_{i,j}(t)\beta_3 + \varepsilon_{2i,j,k}(t) \end{bmatrix} \tag{6}$$

Tables A5 and A6 show how the estimates in Tables 3 4 change once we correct for measurement error.

where $\begin{bmatrix} G_{1i,j} \\ G_{2i,j} \end{bmatrix}$ and $\begin{bmatrix} \bar{G1}_i \\ \bar{G2}_i \end{bmatrix}$ are instrumented by $\begin{bmatrix} G_{2i,j} \\ G_{1i,j} \end{bmatrix}$ and $\begin{bmatrix} \bar{G2}_i \\ \bar{G1}_i \end{bmatrix}$.

The estimates in tables A5 and A6 corrrect for measurement error in the EA PGI that arises due to classis sampling error in the GWAS weights. However, other sources of measurement error in the PGI remain. Other possible sources of measurement error in the PGI are (1) the PGI does not capture sources of genetic variation beyond SNPs, such as rare genetic variants (Cheesman et al., 2017); (2) The GWAS coefficients used to construct the EA PGI are consistent for the *cross-sectional* association between the SNPs and years of schooling, and not its causal effect (Trejo & Domingue, 2018); and (3) there may exist SNPs that predict SES-related traits of the child such as net wealth and income, but that do *not* predict years of schooling, and that do therefore not show up in the GWAS for educational attainment that we used for PGI construction. Because of these various sources of measurement error, our coefficients in tables A5 and A6 remain conservative: they are downward biased estimates of the effects of a hypothetical genetic index that could perfectly rank individuals based on their genetic likelihood to receive high educational attainment.

In Figure 5, we further transform the effect of the EA PGI to a rank effect. That, we estimate

$$\begin{bmatrix} Y_{i,j,k}(t) \\ Y_{i,j,k}(t) \end{bmatrix} = \begin{bmatrix} \beta_{01} + \beta_1 Rank(G_{1i,j}) + \beta_2 \bar{G}1_i + \varepsilon_{1i,j,k}(t) \\ \beta_{02} + \beta_1 Rank(G_{2i,j}) + \beta_2 \bar{G}2_i + \varepsilon_{2i,j,k}(t) \end{bmatrix} \tag{7}$$

where $\begin{bmatrix} Rank(G_{1i,j}) \\ Rank(G_{2i,j}) \end{bmatrix}$ and $\begin{bmatrix} \bar{G}1_i \\ \bar{G}2_i \end{bmatrix}$ are instrumented by $\begin{bmatrix} G_{2i,j} \\ G_{1i,j} \end{bmatrix}$ and $\begin{bmatrix} \bar{G}2_i \\ \bar{G}1_i \end{bmatrix}$.

A4 Genotypic assortative mating (estimation of parameter η)

We study genotypic assortative mating by estimating how an exogenous variation in one's PGI affects the PGI of one's children:

$$G_{i,j,k} = \theta_0 + \theta_1 G_{i,j} + \theta_2 \bar{G}_i + X_{i,j,k}\theta_3 + \varepsilon_{i,j,k}. \tag{8}$$

If mating in our data is random, then θ_1 should be equal to 0.5. If there is genotypic assortative mating, then θ_1 should be equal to $0.5(1+\eta)$, where η is the parameter quantifying the amount of genotypic assortative mating (see figure 3).

Notice that equation (8) has one more genetic data requirement than equation (3). It also requires genetic data of the third generation, the children. We refer to the sample that satisfies all these genetic data requirements as the "3-generation sample" (N = 2,257).

Figure 4 illustrates the causal relationship between the parent's PGI and the offspring's PGI. First, we ran a regression of parent #1's PGI on the PGIs of grandparents #1. We refer to the residuals from such a regression as the parent's residualized PGI. Second, we ran a regression of the child's PGI on the PGIs of grandparents #1. We refer to the residuals from such regression as the child's residualized PGI. Figure 4 shows a scatterplot of the offspring's residualized PGI against the parent's residualized PGI. The figure also plots the corresponding fitted regression (solid line) and the 22.5 degree line (dashed line). The difference between the two lines quantifies the amount of genotypic assortative mating. If there was random mating, then the two lines would be on top of each other. If the solid line is steeper than the 22.5 degree line indicates that high-PGI individuals have children with people who also have higher PGIs themselves. The slope of the fitted regression line is 0.48 (s.e. 0.034). We cannot reject the hypothesis of random mating (i.e., $\theta_1 = 0.5$). The p-value of such test is 0.56.

To ensure that our failure to reject random mating is not due to low power, we consid-

ered alternative strategies to estimate the assortative mating parameter η. All estimation strategies are consistent with η being zero. A first alternative strategy regresses the PGI of parent #2 no the PGI of parent #1, conditional on the sum of PGIs of the grandparents from the side of parent #1, as follows:

$$G_{ij\#2} = \alpha + \eta G_{ij\#1} + \xi_1 \bar{G}_{i\#1} + X_{i,j}\xi_2 + \varepsilon_{i,j} \tag{9}$$

Here, $G_{ij\#2}$ is the PGI of the spouse of parent #1, in other words, the second parent of child j. Note that we can only estimate this regression in the subsample of parents for whom spouses were also genotyped. This is the case for 1547 parent pairs. The averages of $G_{ij\#2}$ and $G_{ij\#1}$ in this subsample are 0.0417 ($s.e. = 0.025$) and 0.0340 ($s.e. = 0.025$) respectively, insignificantly different from zero. This indicates that there is no selection based on the PGI into this subsample.

Note that this regression again conditions on the PGI of the first generation, as defined by the sum of grandparental alleles *from the side of parent i*. Hence, η reflects the causal effect from an increase in the EA PGI of parent #1 on the PGI of the spouse, which decides the expected value of the PGI of their child. Table A8 shows the results of these regression, in the second column. Using this strategy, we estimate $\hat{\eta} = 0.0333 (s.e. = 0.0419)$. This value is insignificant from zero ($p = 0.427$), which is similar to our other procedure that estimates η by regressing the PGIs of children on the PGIs of their parents.

As the final (and best-powered) alternative estimation strategy to estimate η, we combine both specifications in a single maximum likelihood function. That is, we estimate η jointly on a stacked data set of all observations that have either the PGI of the child (G_{ijk}) or the PGI of parent #2 ($G_{ij\#2}$) as the outcome variable. We maximize the following likelihood function:

$$log(L) = \sum_{i}^{I} -\mathbb{1}(\text{Spouse}_{ij}) \left[log(\sigma_1) + \frac{1}{2} \left(\frac{G_{ij\#2} - \widehat{G}_{ij\#2}}{\sigma_1} \right)^2 \right] - \mathbb{1}(\text{child}_{ijk}) \left[log(\sigma_2) + \left(\frac{G_{ijk} - \widehat{G}_{ijk}}{\sigma_2} \right)^2 \right] \tag{10}$$

,

where the indicator function indicates whether the PGI of the spouse of parent $ij\#1$, or the kth child of parent $ij\#1$ is observed in our data, and $G_{ij\#2} - \widehat{G}_{ij\#2}$ and $G_{ijk} - \widehat{G}_{ijk}$ the residuals as defined by equation 9 and 8, respectively. Our ML estimates of the parameter η (reported in the third column of table A8) again indicate no significant evidence for assortative mating ($\hat{\eta} = -0.014$ ($s.e. = 0.035$)).

This rejection of genetic assortative mating on the PGI for educational attainment contrasts with the rest of the literature. For example, Okbay et al. (2022) study genotypic assortative mating in terms of the PGI for educational attainment among 862 couples in the UK Biobank and 1,603 couples in Generation Scotland. They document a mate-pair correlation of 0.175 ($s.e. = 0.020$). However, their results do not condition on the genetic makeup of the previous generation. We are not aware of any other study that has tried to quantify genotypic assortative mating using exogenous variation in the PGI of one of the partners, as we do.

A5 Construction of inverse probability weights

We calculated weights inversely proportional to the probability of being included in the direct effect sample, by comparing the frequency counts of sex-year-of-birth cells in the main sample and direct effect sample. The middle panel of table A9 shows the summary statistics of the direct effect sample after applying these inverse probability weights. As expected, the probability of being male is now 0.5, and the average year of birth is now 1982.5, very similar to the 3rd generation of the main sample. Looking at the wealth-related outcomes, the differences between the direct effect sample and the main sample are now also much

reduced.

A6 Additional Tables

	OLS	OLS	OLS	OLS	OLS	OLS	FE	FE
EA PGI	0.614	0.647	0.574	0.610	0.446	0.445	0.464	0.466
	(0.010)	(0.009)	(0.016)	(0.015)	(0.027)	(0.026)	(0.028)	(0.028)
EA PGI (Parental)					0.088	0.115		
					(0.015)	(0.014)		
N	59093	59093	20294	20294	20294	20294	17053	17053
R squared	0.0621	0.1441	0.0602	0.1479	0.0618	0.1506	0.0301	0.0375
GWAS Controls	NO	YES	NO	YES	NO	YES	NO	YES

Table A1: Various models that test the performance of the EA PGI in Lifelines data on Years of Education. The outcome is years of education as self-reported by Lifelines respondents that were 30 years or older at their assessment. The first 2 columns include all respondents. Columns (3)-(6) include the subsample for which, additionally, the imputed parental EA PGI is available. Columns (7) and (8) is estimated using within-family fixed effect estimator among families with multiple genotyped full siblings. The coefficient on EA PGI is possibly confounded upwards in specifications (1)-(4) due to environmental correlates with the EA PGI. In specifications (5)-(8), assignment of the EA PGI is plausibly exogenous (due to conditioning on the imputed sum of parental PGIs or sibling fixed effects). Some specifications control for additional controls that are typical for GWAS analyses. These are age, sex, and the first ten principal components of the genetic data to control for population stratification.

	OLS	OLS	OLS	OLS	OLS	OLS	FE	FE
EA PGI	0.104	0.107	0.103	0.107	0.091	0.091	0.089	0.089
	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.005)	(0.006)	(0.006)
EA PGI (Parental)					0.008	0.011		
					(0.003)	(0.003)		
N	59093	59093	20294	20294	20294	20294	17053	17053
R squared	0.0517	0.0780	0.0487	0.0843	0.0490	0.0850	0.0249	0.0265
GWAS Controls	NO	YES	NO	YES	NO	YES	NO	YES

Table A2: Various models that test the performance of the EA PGI in Lifelines data on the probability of having a university degree. The outcome is having at least a bachelor's degree as self-reported by Lifelines respondents that were 30 years or older at their assessment. The first 2 columns include all respondents. Columns (3)-(6) include the subsample for which, additionally, the imputed parental EA PGI is available. Columns (7) and (8) is estimated using within-family fixed effect estimator among families with multiple genotyped full siblings. The coefficient on EA PGI is possibly confounded upwards in specifications (1)-(4) due to environmental correlates with the EA PGI. In specifications (5)-(8), assignment of the EA PGI is plausibly exogenous (due to conditioning on the imputed sum of parental PGIs or sibling fixed effects). Some specifications control for additional controls that are typical for GWAS analyses. These are age, sex, and the first ten principal components of the genetic data to control for population stratification.

Variable	Mean	Median	SD	P1	P99	N
Time-Invariant Characteristics						
Male	0.33	0	0.47	0	1	5736
Birth Year	1956.29	1957	6.64	1939	1968	5736
Years of Schooling	12.34	13	2.31	7	17	5609
EA4 PGI	0.01	0.01	0.99	-2.36	2.3	5736
EA4 PGI, First Generation	0.02	0	1.89	-4.59	4.57	5736
University Graduate	0.18	0	0.39	0	1	5609
Time-Variant Characteristics						
Positive Earnings	0.69	1	0.46	0	1	81071
Individual Earnings	23956	16081	26509	0	106116	81071
Individual Income	32065	26554	25486	0	112206	81071
Household Income	75259	68810	42034	12878	209854	81102
Net wealth, levels	228664	157555	281260	-107324	1319719	79956
Assets	337393	260374	314910	1	1650384	79956
Checkings & Savings	42843	23321	56070	0	284289	79956
Stocks & Bonds	19969	0	170933	0	322902	79956
Home equity	120507	109481	119063	-119234	460700	79956
Debt	104996	79533	114505	0	593028	79956
Mortgage	93468	73118	98136	0	452372	79956

Table A3: Summary statistics of genotyped parents in the main estimation sample. To be included, parents had to (1) be a genotyped respondent in Lifelines, (2) have either a full sibling or a biological parent that was also a genotyped Lifelines respondent, and (3) have at least one child who was 30 years or older in 2021.

	$\mathbf{PGI^{2nd} \times Age^{3rd}}$		$\mathbf{PGI^{2nd} \times Male^{3rd}}$		$\mathbf{PGI^{2nd}}$	
	β	S.E.	β	S.E.	β	S.E.
Education						
Years of Schooling	0.0098	0.0054	0.110	0.06	0.25	0.05
University	0.0019	0.0012	0.031	0.02	0.06	0.01
Income & Earnings						
Any Individual Earnings	0.0000	0.0006	-0.006	0.007	0.006	0.006
Individual Earnings	-0.02	0.06	-1.16	0.84	2.09	0.63
Individual Income	-0.01	0.06	-1.36	0.82	2.29	0.64
Household Income	0.06	0.06	-0.16	0.85	1.77	0.63
Household Wealth						
Net Wealth	0.13	0.06	1.28	0.87	1.02	0.66
Assets	0.20	0.06	1.54	0.80	0.81	0.63
Financial Wealth	0.08	0.06	0.30	0.83	1.98	0.64
Checkings and Savings	0.08	0.06	0.18	0.82	1.92	0.63
Stocks and Bonds	0.00	0.09	1.09	1.14	1.33	0.82
Real Estate	0.25	0.07	1.21	0.91	0.42	0.70
Debt	0.09	0.07	0.78	0.83	-0.07	0.62
Mortgage	0.15	0.08	0.62	0.94	-0.22	0.70

Table A4: Interactions of the effect of the parental PGI (PGI^{2nd}) on various socioeconomic outcomes of the children. The estimates use a specification similar to equation 3, but with the following additional variables: the interaction between the parent's PGI (PGI^{2nd}) and the age of the child (Age^{3rd}), the interaction between parent's PGI (PGI^{2nd}) and an indicator for the child being male ($Male^{3rd}$), and a linear control in age, with age demeaned, so that the effect of the parental PGI (PGI^{2nd}) can be interpreted as the effect for a daughter of average age in the sample (39 years in 2021). As in our main specification, additional control variables include parent year of birth fixed effects, a dummy for the sex of the parent and of the child, and, for time-varying variables, year-of-measurement fixed effects. Additionally, each variable that is interacted with the parental PGI was also interacted with the grandparental PGI. All standard errors are clustered at the level of the parent.

	In Levels		Percentile Rank	
	β	S.E.	β	S.E.
Education				
Years of Schooling	0.42	0.07		
University	0.101	0.018		
Income & Earnings				
Any Individual Earnings	0.009	0.008		
Individual Earnings	2.8	0.9	2.5	0.9
Individual Income	2.9	0.9	2.7	0.9
Household Income	4.5	1.5	2.7	0.9
Non-labor Individual Income	0.14	0.15	0.92	0.83
Income, other household members	1.7	1.03	0.8	0.85

Table A5: Estimates of the effect of the parent's PGI on education- and income-related outcomes of their children that correct for measurement error in the PGI due to sampling error in GWAS. In each specification, β shows the effect of the parental PGI on the outcome as measured in the third generation. In these specifications, the EA PGI is constructed using GWAS weights that are estimated in one half of the UK Biobank, and instrumented with an EA PGI that is constructed using weights estimated in the other half. All outcome variables are measured annually except years of schooling and being a university graduate, which are measured only once per child. Additional control variables include parent year of birth fixed effects, the gender of the parent and the child, and year fixed effects for variables that are measured annually. Standard errors are clustered at the level of the parent. In the columns "Percentile Rank", all are measured as their within-year-of-birth percentiles. Monetary variables in levels are measured in thousands of euros, according to 2015 price levels

	In Levels		Percentile Rank	
	β	S.E.	β	S.E.
Household Wealth				
Net Wealth	4.9	7.1	2.7	1.0
Assets	5.4	9.3	2.0	0.9
Financial Wealth	4.4	2.0	3.6	1.0
Checkings and Savings	5.0	1.5	3.5	1.0
Stocks and Bonds	-1.4	1.4	1.7	1.2
Real estate	6.5	5.1	1.2	1.0
Debt	0.1	4.6	0.0	0.9
Mortgage	0.0	4.2	-0.3	1.1

Table A6: Estimates of the effect of the parent's PGI on wealth-related outcomes of their children that correct for measurement error in the PGI due to sampling error in GWAS. In each specification, β shows the effect of the parental PGI on the outcome as measured in the third generation. In these specifications, the EA PGI is constructed using GWAS weights that are estimated in one half of the UK Biobank, and instrumented with an EA PGI that is constructed using weights estimated in the other half. All outcome variables are measured annually. Additional control variables include parent year of birth fixed effects, the gender of the parent and the child, and year fixed effects. Standard errors are clustered at the level of the parent. In the columns "Percentile Rank", all are measured as their within-year-of-birth percentiles. Variables in levels are measured in thousands of euros, according to 2015 price levels

	$\mathbf{PGI^{2nd} \times Wealth^{1st}}$	
	β	S.E.
Education		
Years of Schooling	-0.0006	0.0013
University	-0.0002	0.0003
Income & Earnings		
Any Individual Earnings	-0.0000	0.000
Individual Earnings	0.006	0.014
Individual Income	0.004	0.013
Household Income	0.007	0.014
Household Wealth		
Net Wealth	0.003	0.014
Assets	0.005	0.013
Financial Wealth	0.002	0.014
Checkings and Savings	0.006	0.014
Stocks and Bonds	0.001	0.018
Real Estate	0.008	0.015
Debt	0.009	0.015
Mortgage	0.012	0.016

Table A7: Interactions of the effect of the parental PGI^{2nd} on various socioeconomic outcomes of the children with long-run family wealth. The estimates use a specification similar to equation 3, but with the following additional variables: the interaction between the parent's PGI (PGI^{2nd}) and the net wealth as averaged over this parent's siblings and the same measure of sibling's net wealth as an additional control variable. Sibling's net wealth serves as a proxy of the SES of the first generation (i.e., the grandparents). As in our main specification, additional control variables include parent year of birth fixed effects, a dummy for the sex of the parent and of the child, and, for time-varying variables, year-of-measurement fixed effects. All standard errors are clustered at the level of the parent.

	OLS		ML
	G_{ijk}	$G_{ij\#2}$	
$\hat{\eta}$	-0.03	0.03	0.014
s.e.	0.03	0.04	0.035
N	2,257	1,547	3,378

Table A8: Estimates of η using different estimation strategies. The first column estimates η by regressing the PGI of parent #1 ($G_{ij\#1}$ on the PGI of the child G_{ijk}). η is next constructed as the coefficient minus 0.5. The second column estimates η by regressing the PGI of parent #1 on the PGI of parent #2 (($G_{ij\#2}$. Last, the third column combines both types of regressions in a single maximum likelihood estimator (see equation 10). In column 1, N refers to the amount of child-parent#1 pairs for which the PGIs were available. In column 2, N similarly refers to the amount of available parent#1-parent#2 pairs. In the third column, N refers to the amount of parents #1 for which the ML-function could be esitmated.

Variable	Same-generation Sample N	Mean	SD	Same-generation Sample, weighted Mean	SD	Next-generation Sample (children) N	Mean	SD
Time-Invariant Characteristics								
Male	19431	0.4	0.49	0.5	0.5	11788	0.5	0.5
Birth Year	19431	1976.51	7.99	1982.54	6.65	11788	1982.38	6.93
EA4 PGI	19431	-0.01	0.99	0.02	0.99			
EA4 PGI, First Generation	19431	-0.07	1.68	-0.02	1.64			
Positive Earnings	236970	0.94	0.23	0.96	0.2	101303	0.93	0.25
Time-Variant Characteristics								
Individual Earnings	236970	39588	25623	42473	24920	101303	41006	26528
Individual Income	236970	41999	25357	44416	24406	101303	43425	25762
Household Income	237118	84835	40292	83663	38877	101425	82662	42934
Net wealth, levels	233550	132345	242488	100231	215713	100180	92114	208647
Assets	233550	296523	276120	267117	251067	100180	256101	257940
Checkings & Savings	233550	32398	44386	28463	38711	100180	26752	39330
Stocks & Bonds	233550	8146	87577	5331	63516	100180	5798	47252
Home equity	233550	52213	102225	35631	90468	100180	33322	91922
Debt	233550	161932	122337	165181	118405	100180	162888	129394
Mortgage	233550	152309	111593	155507	109775	100180	151924	120275

Table A9: Summary statistics of the same-generation sample. The same-generation sample consists of all Lifelines respondents who (1) were genotyped in Lifelines, (2) have either a full sibling or biological parent who was also a genotyped Lifelines respondent, (3) were born between 1964 and 1991. The left panel shows the summary statistics for this sample. The middle panel shows the same sample, but after applying inverse probability weighting (computed using frequencies of sex and year of birth) to make the sample comparable to the children of the next-generation sample. The right panel shows the summary statistics for these children in the next-generation sample summary statistics (as in table 1), and is included for comparison.

	In Levels β	S.E.	Percentile Rank β	S.E.
Education				
Years of Schooling	0.22	0.07		
University	0.051	0.013		
Income & Earnings				
Any Individual Earnings	-0.002	0.007		
Individual Earnings	0.5	0.5	0.4	0.6
Individual Income	0.9	0.5	0.9	0.5

Table A10: Estimates of the effect of parent's PGI on the individual socioeconomic outcomes of the other parent, i.e., the spouse of the parent whose PGI was measured. The variables, "any individual earnings", "individual earnings" and "individual income" are measured annually. "Years of schooling" and "university" are measured once for each co-parent. Additional control variables include parent year of birth fixed effects, the first parent's gender, and year fixed effects for variables that are measured annually. Standard errors are clustered at the level of the parent. In the columns "Percentile Rank", all coutcomes are measured as their within-year-of-birth percentiles. Monetary varibles in levels are measured in thousands of euros, according to 2015 price levels.

A7 Additional Figures

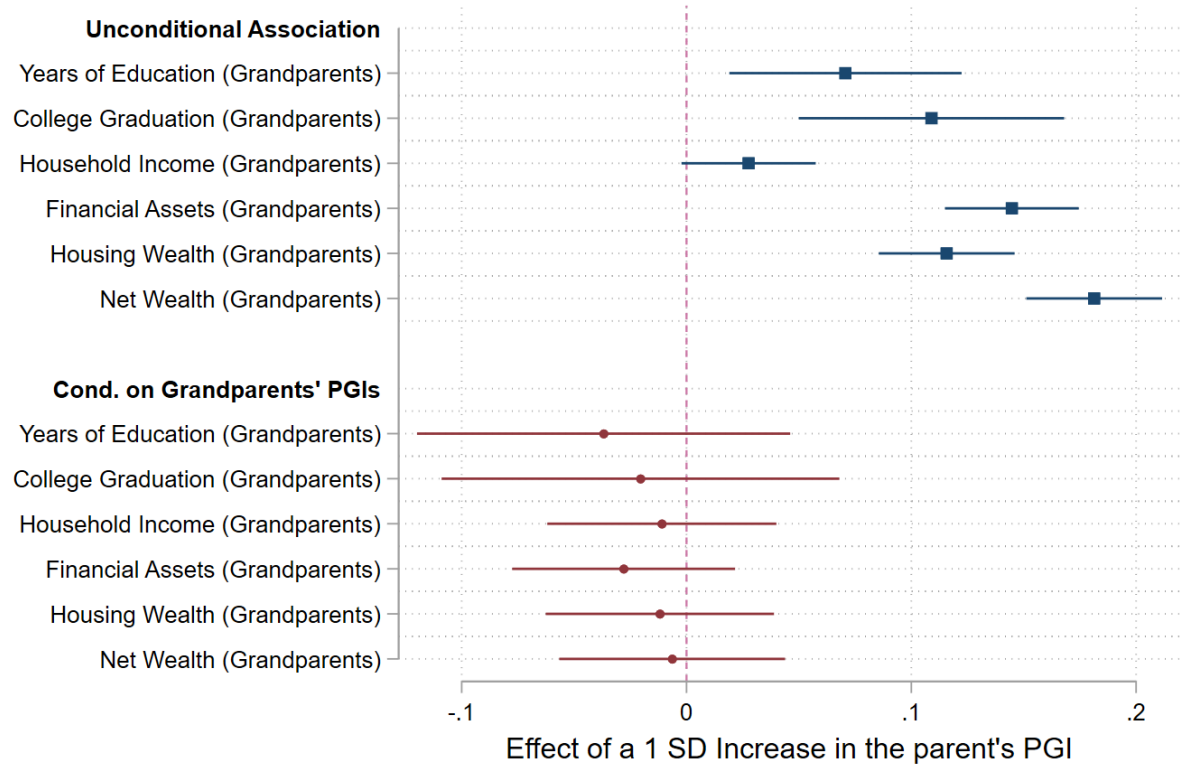

Figure A1: The effect of a PGI for educational attainment on various measures of mean SES of the previous generation (i.e., one's parents). The upper panel of the figure shows the coefficient of the PGI as estimated in a model in which each measure of previous generation's SES is regressed on the PGI for EA. Here, the PGI significantly predicts previous generation's SES, which indicates that the effect of the PGI is biased by environmental confounders. The bottom panel shows the same coefficient when the previous generation's PGI, based on the sum of alleles of one's parents, is included as a control variable. In this panel, all effects are centered around zero. This indicates that the bias in the upper panel is adequately captured by this control variable. All regressions were estimated for the parents in our main sample for which we had the relevant measure of SES for at least one of these Lifelines respondents' parents