Educational Attainment and Intergenerational Mobility: A Polygenic Score Analysis

Aldo Rustichini

University of Minnesota

William G. Iacono

University of Minnesota

James J. Lee

University of Minnesota

Matt McGue

University of Minnesota

We extend a standard model of parental investment and intergenerational mobility to include a fully specified genetic analysis of skill transmission. The model's predictions differ substantially from the standard model's. The coefficient of intergenerational income elasticity (IGE) may be larger than that in the standard model and depends on the distribution of the genotype. The distribution of genetic endowments may be stratified according to income. The model is tested on data, including genetic information, of twins and their parents, estimating how IGE is affected by genetic factors and how environment and genes interact. The effect of intelligence is substantially stronger than that of other traits.

We thank three anonymous referees and the editor, who gave advice that led to a complete reorganization of the paper, the analysis as well as the exposition. We are in particular thankful

Electronically published August 30, 2023

Journal of Political Economy, volume 131, number 10, October 2023.

https://doi.org/10.1086/724860

I. Introduction

In recent research on heritability of phenotypes based on genome-wide association studies (GWASs), a number of markers have been identified. A GWAS is a study of common genetic variants spanning the entire genome (typically 1 million single-nucleotide polymorphisms [SNPs] or more) in a typically large set of individuals to determine whether and how much any variant is associated with a trait. The markers that achieve significance at the conventional GWAS threshold[1] are still limited in number, and together they explain a limited fraction of the variability of the phenotype. In spite of this, a considerable fraction of phenotypic variation can be explained by a larger set of genetic markers that includes variants that are not significant by GWAS standards.

A way to take into account the information available in markers, including, perhaps, those with significance lower than the GWAS threshold, is to compute a polygenic score (PGS). A PGS is an individual-specific score, obtained as sum of the value of the markers in a selected set, each value weighted by a coefficient that has been estimated separately on an independent training sample (Dudbridge 2013). Our analysis here is based on the large GWAS of educational attainment reported by Lee et al. (2018; see also Rietveld et al. 2013 and Okbay et al. 2016). An illuminating discussion of the analysis of educational attainment in the modern GWAS era is in Cesarini and Visscher (2017).

Theoretical framework.—We set up the investigation in a fully specified model of parental investment in education of children. Some classical papers establishing this tradition are Becker and Tomes (1979, 1986) and Loury (1981). Important developments of the early model are in, among many, Solon (1992, 2004), Mulligan (1997, 1999), Black and Devereux (2011), and Black et al. (2017). Our model differs from the existing ones in the field in two respects, both introduced because we need to take into account information on genotype and its transmission. First, we introduce explicitly the fact that children are the outcome of a joint process involving a father and a mother; so we need to include in the model a theory of mating (similarly to Aiyagari, Greenwood, and Guner 2000 and Greenwood,

to one of the referees, who urged us to clarify the relationship between the research reported here and earlier publications. We thank Aysu Okbay for generously running the meta-analysis on the data, Peter Visscher for a clarification on Robinson et al. (2017), and Philippe Köllinger for the help in the process. We also thank Tom Holmes, Andrea Ichino, Chris Phelan, Joel Waldfogel, and Giulio Zanella for very useful observations, criticisms, and suggestions, and audiences in many seminars for very lively, illuminating discussions. This work was supported in part by grants from the National Science Foundation to Rustichini (SES1728056), the National Institute on Alcohol Abuse and Alcoholism (AA09367), and the National Institute of Drug Abuse (DA05147). This paper was edited by James J. Heckman.

[1] The threshold is 5×10^{-8}; the factor 10^{-8} corrects (Bonferroni) for multiple comparisons.

Guner, and Knowles 2003).[2] The importance of assortative mating has been well documented in the past. For instance, Greenwood et al. (2016) document that assortative mating along educational characteristics has increased in the United States. We build here on research like that by Fernández and Rogerson (2001) and Fernández, Guner, and Knowles (2005), who study models where assortative mating directly affects intergenerational mobility. Second, we model the process of skill formation consistently with the transmission of genotype from parents to children, along well-known lines in genetics (see, e.g., Nagylaki 1992). From our vantage point, after so much research, we can revisit the classical debate between Goldberger (1989) and Becker (1989) and realize that both models were, in some important measure, imprecise. We take this opportunity to illustrate the implications of our work.[3] Becker had in mind the autoregressive process assumed in his earlier work (Becker and Tomes 1979, 1986), which we discuss more in detail below (sec. III.C). In his thought-provoking 1985 Woytinsky lecture, Goldberger (1989, 505) suggests a modification of Galton's (1886) "Regression towards Mediocrity," presenting Galton's argument that the characteristic of the individual is some weighted average of the characteristics of the entire history of ancestors. But the expectation of the child's phenotype conditional on the entire history of genotypes of the ancestors is equal to the expectation conditional on the parents' genotype only. It also has a precise form,[4] which is neither the one in Becker and Tomes (1979) nor the one in Goldberger (1989).

Empirical questions.—Within our theoretical framework we address two basic sets of questions. First, how much of the variance in income and educational achievement is explained by the PGS, and how does family structure affect the transmission? Similar questions have been investigated with the same data by McGue, Rustichini, and Iacono (2017) and McGue et al. (2020), but they simply examined correlational results, rather than tests of a well-specified model of parental investment. Tests of the effectiveness of the PGS in predicting a variety of variables are presented in existing literature: for educational attainment, see Rietveld et al. (2013), Okbay et al. (2016), Kong et al. (2018), Lee et al. (2018), and Willoughby et al. (2021); for wealth, see Barth, Papageorge, and Thom (2020); for social mobility of children compared to parents, see Belsky et al. (2018); and

[2] In this paper, two terms, "matching" and "mating," are used interchangeably, as synonymous for partnership among parents. The reason for the coexistence of the two terms is that the "matching" is used more frequently in the economics literature and "mating" in behavioral genetics. In each instance we use the term more appropriate in the context.

[3] The discussion between Goldberger and Becker centered on two main points: whether adopting a utility-maximization framework makes a difference for the predictions of the theory, and what is the stochastic process of skill. For the first, we adopt here the utility-maximization setup, but as long as the comparison of policies is not explicitly modeled, choosing one or the other makes little difference. We focus here on the second point.

[4] See, e.g., eq. 10.104 in Nagylaki (1992).

for health outcomes, see Barcellos, Carvalho, and Turley (2018). Earlier contributions on the issue of health conditions and academic performance using genetic markers are in Ding et al. (2009) and Fletcher and Lehrer (2011). This estimate would give us a lower bound on how much of the variance of success in education can be attributed to the individual's genotype. How is this effect modified by assortative mating among parents and the correlation among their genotypes? And finally, how is the effect of genes mediated by the direct effect on the genotype of the children, and how much is mediated by the indirect effect on the environment provided to them, as well as parental investment?

Second, what are the channels through which the effect of genotype, as summarized by the PGS, operates in each individual? Recall that the score is built on a simple statistical association between genotype and the phenotype of interest, in our case success in education, and that no mechanism underlying the association is identified. A natural first channel to consider is intelligence: the score likely summarizes a set of highly polygenic effects on intelligence, and in turn intelligence improves the chances of success in education. But intelligence is not the only plausible channel; personality traits are an important additional way. We use the term *personality* to indicate a set of individual characteristics possessed by a person that together determine a consistent pattern of cognition, emotions, motivations, and behaviors in various situations. A substantial fraction of success in education might be traced back to motivation, self-control, ambition: in general, personality traits distinct from pure cognitive skills. A gene affecting these traits would also appear as a contribution to the PGS, even if unrelated to intelligence. These are all natural channels. The effect of genes on education could operate, however, along completely different pathways, involving individual characteristics that have no bearing on the technology of educational attainment, for example, discrimination. Clearly, understanding which of these pathways operates, and in what measure, is essential, particularly for policy guidance. We now review our answers to these questions.

Outline of main results.—We develop (sec. II) a model of intergenerational mobility, building on classical parental investment models but replacing their ad hoc skill-transmission equation with a precise and correct model of genetic inheritance from the two parents. In the model, the coefficient of persistence of skills is endogenous, depending on the distribution of the genotypes in the populations; thus, most of the conclusions of the classical model are now invalid. We provide the correct predictions.

An important component of the theory is the model of assortative matching among parents according to characteristics, some endowed with a natural order (such as income and skills) and some not (such as personality traits and physical appearance); we show how this affects the distribution of the genotypes at the invariant distribution of the system. A state of the system is described by a joint probability on genotypes and endogenous

variables, such as income and education. With assortative matching, the transition function is nonlinear, so existence of a stationary distribution is not simple. We prove its existence and some basic properties.

At the stationary distribution, within each class of matching, alleles are in Hardy-Weinberg equilibrium.[5] More notably, the frequency of alleles with positive effect on educational attainment, and thus on income, positively correlates with income. The correlation is stronger, the stronger the effect of the allele. These results identify a powerful force producing lasting inequality, which has been ignored so far and is absent by assumption in standard models.[6]

The model leads to a natural empirical test, using data described in section IX. Information on the genotype of individuals is summarized into a PGS obtained from a large GWAS on educational attainment. All the predictions of the model are tested with a unique set of data in which we have complete genetic information on parents and children, in addition to information on education, personality traits, intelligence, family environment, and income. We estimate equation (5), the intergenerational elasticity coefficient of income, which is at the lower end of existing estimates for the overall United States. We compare it to the effect size of genetic factors measured by the PGS; we find that the latter is approximately half of that of income. In section VI, we identify the pathways of the effect on income through human capital formation.

In section VII, we use the twin structure of our data to check for the robustness of the results and investigate passive gene-environment correlation, that is, how the genetic endowment of the parents affects the phenotype of the children through the family environment. Natural and significant channels of this effect are the education of parents and their income, and we prove that this channel is significant. However, there is no additional residual channel through family environment in addition to these two. When we study the pathways of the genetic effect measured by the PGS, we find that after correction for measurement errors, the effect from genotype to educational and economic success is mediated mostly by intelligence and only weakly by noncognitive skills. Conclusions are presented in section VIII.

II. Genetic Skill Transmission and Parental Investment

We begin with the conceptual and theoretical structure for our empirical analysis, introducing a model and an equilibrium concept. The complete

[5] The definition of Hardy-Weinberg equilibrium is recalled in sec. III.

[6] We use "standard model" in a broad sense here, which includes Goldberger (1989)'s model.

model to be tested is presented in section II.G. Our first aim is to show that the standard analysis of parental investment in education, and intergenerational mobility (as pioneered in Becker and Tomes 1979, where the skill transmission follows a simple AR(1) [autoregressive] process), should be modified—if one wants to avoid significant misunderstandings—to take into account a fully specified genetic mechanism of skill transmission. A core feature of the model we propose is the combination of the theory of marriage (Becker 1973) to predict mating, with a model of genetic transmission. A comparison of the prediction of the two models is provided in section III.C; there we show that they differ substantially on key predictions about, for instance, intergenerational mobility.

Our model has several components. After defining the basic environment (sec. II.A), in section II.B we describe how the skills of the children are affected by the genetic endowment inherited by parents, family environment, and random events. Then, in section II.D, we describe the decision of parents to invest in education of the children.

A. *Setup*

A population of individuals, constant in number over time, is organized into households. A household maximizes a utility function of own consumption and future income of two children, which in turn is affected by the genetic endowment of the children, parental investment in education, and environment. The restriction to two children is consistent with the assumption that population size is constant. In our data, the two children also happen to be twins: this detail has little importance when we study parental investment,[7] but it becomes important when we study the correlation of skill and income across siblings. We denote y the natural log of the income (so this value ranges in the real line), E consumption expenditure, I parental investment in education of children, and h human capital measured by the education level. The terms ϵ^e and ϵ^y denote the random shocks to education and income, respectively: each one is i.i.d. (independently and identically distributed) across periods, and the two are independent within periods. Subscripted α's denote productivity parameters of the variable in the index; so α_{I}, α_{h} denote positive real numbers associated with parental investment and human capital; $\delta \in (0, 1)$ is the discount factor. A vector of real numbers $\theta = (\theta^1, \ldots, \theta^{n_{\mathrm{I}}}, \theta^{n_{\mathrm{I}}+1}, \ldots, \theta^n)$ describes the n skills, where the index from 1 to n_{I} refers to hard or cognitive skills and that from $n_{\mathrm{I}} + 1$ to n to soft or noncognitive skills (Heckman and Kautz 2012; Heckman, Pinto, and Savelyev 2013). Skills enter linearly into the production of the

[7] Note, however, that two children who are also twins have the same age, so the parental investment in this case does not concern two individuals of different age, as instead is typical for siblings.

education level though an n-dimensional vector of coefficients α_θ. The superscript i refers to the family and the subscript $j = 1, 2$ to the siblings, so a sibling is uniquely identified by the pair ij. Household log income y^i is some combination of the log incomes of the father y^i_f and the mother, y^i_m.[8] The precise form of the combination is specified below. We denote **E** the expectation of a random variable.

We emphasize that the model is not a two-period model but an overlapping-generations model, so each individual appears in the model as a child and then as a parent, and the model describes behavior in both stages of life. So when we model, for instance, how genetic factors affect skill formation, human capital accumulation, and income of children, we also model how the same variables have been determined for the parents of these children. We make full use of this in some crucial step—for example, in section E, where we describe how genetic factors affect behavior both as children and as parents.

B. Skill Transmission

We replace the standard AR(1) mechanism of skill transmission (discussed more extensively in sec. III.C) with a detailed model where the skill vector θ results from genetic factors, parental investment in education, family environment common to all children, and idiosyncratic random events for each individual.

We examine these components separately, beginning with the genetic component.[9] If K is the number of loci, a genotype is a $g \in G \equiv \{0, 1, 2\}^K$, so $g = (g(k) : k = 1, \ldots, K)$. Here "0, 1, 2" refers to the count of one of the alleles in a biallelic system (a GWAS typically deals with variants, SNPs, that are biallelic in the analysis). The joint distribution of genotypes of the two children, given the genotype of the two parents, depends on the twin type, which may be monozygotic, MZ, or dizygotic, DZ. To describe how the distribution is determined, we start with the function from parents' genotype to the probability over genotypes of an individual offspring, given by a function H from $G \times G$ to $\Delta(G)$:

[8] The use of letters "f" and "m" avoids confusion with the family index.

[9] In the context of the twin-studies model, the integration of parental investment models with a more realistic model of skill transmission has been explored in Rustichini, Iacono, and McGue (2017), where the realistic model of genetic transmission is used to provide a justification for the standard ACE models (see sec. II.G for a definition) in twin studies in the context of economic analysis of parental investment. However, in Rustichini, Iacono, and McGue (2017), there is no analysis of the invariant measure produced by assortative mating according to characteristics (which we consider one of the main contributions of this paper), nor is there a comparison of the predictions of the standard skill transmission model in the tradition of Becker and Tomes (1979). Finally, the empirical analysis does not make any use of the genetic information used in this paper.

$$H : (g_m, g_f) \mapsto H(g_m, g_f). \tag{1}$$

We write $H(\cdot \mid g_m, g_f)$ when we want to indicate explicitly the set on which this measure operates. The function H follows well-known rules of Mendelian inheritance (see, e.g., Crow and Kimura 1970 or Nagylaki 1992); for instance, if $K = 1$, so that $G = \{0, 1, 2\}$, then $H(\cdot \mid 1, 1)$ is (0.25, 0.5, 0.25), and so $H(2 \mid 1, 1) = 0.25$. Similarly, $H(\cdot \mid 0, 2)$ is (0, 1, 0).

The map in equation (1) is well defined only under the assumption, which we make, that the distribution across loci is independent. Simple examples show that we may have two different haplotype pairs that induce the same genotype profile (g_m, g_f) for the parents but, without this assumption, induce different elements in $\Delta(G)$ for the children.

C. PGSs

Let w denote an n-valued function determining skills as a function of the genotype g. The PGSs are denoted by $w(g)$. They are computed under the assumption of additivity across loci and within each locus, so that

$$w(g) = \sum_{k=1}^{K} \alpha(k) g(k), \tag{2}$$

where α is a vector of parameters. The values w are latent variables, and they would be of little use if we did not have an estimate. We rely on estimates, called *estimated PGSs*, of the true effect $w(g)$ of the genotype g:

$$\mathrm{PGS}(g) = \sum_{k=1}^{K} \beta(k) g(k), \tag{3}$$

where β's are weights derived from a GWAS. We should note that the weights obtained in a GWAS do not give a full account of the variability in educational attainment. There may be rare variants (Yengo et al. 2020), as well as structural variants (Chiang et al. 2017), that are not well captured by a GWAS.[10] We let X_j^i denote a vector of variables associated with twin ij. These variables may be observable or not and may include, for instance, the parents' education, personality traits of the child, and the family's social status. Also, let Π denote a matrix with n rows, F a family-specific n-dimensional vector (common to both twins in family i, either MZ or DZ), and ϵ^θ an individual-specific n-dimensional environmental zero-mean shock on the skill. We specifically denote the effect of family income, which is assumed to be linear with coefficient π.

[10] We ignore the possible measurement error of PGSs here, since we are not primarily interested in heritability per se. A possible extension of our research would reduce this attenuation using, e.g., methods described in Becker et al. (2021).

The skill of twin ij is thus given by[11]

$$\theta_j^i = w(g_j^i) + \pi y^i + \Pi X_j^i + F^i + \epsilon_j^{\theta,i}. \tag{4}$$

We assume the no-correlation,

$$\forall\ i, j,\ \ \forall\ k \in \{\mathrm{h}, \mathrm{y}\} : \mathbf{E}\epsilon_j^{k,i}\epsilon_j^{\theta,i} = 0, \tag{5}$$

and zero-mean conditions:

$$\forall\ i, j : \mathbf{E}F^i = 0, \mathbf{E}(\epsilon_j^{\theta,i}) = 0.$$

D. *Parental Investment*

The ith household solves in the variables E expenditure in consumption and I^i pair of investment in the two children:

$$\max_{(E^i, I_1^i, I_2^i)} \mathbf{E}_{(\theta_1^i, \theta_2^i)} \left((1 - \delta) \ln E^i + \delta \sum_{j=1,2} y_j^i \right), \tag{6}$$

subject to the budget constraint given by the household's income (recall that y is the natural log of income):

$$E^i + \sum_{j=1,2} I_j^i = \exp(y^i). \tag{7}$$

The choice on consumption and educational investment is taken with the knowledge of the skills (θ_1^i, θ_2^i) of the children; hence the subscript in the expectation of equation (6), which refers to the random shocks ϵ^{h} and ϵ^{y}. Human capital accumulation is described by[12]

$$h_j^i = \alpha_{\mathrm{I}} \ln I_j^i + \alpha_\theta \theta_j^i + \epsilon_j^{\mathrm{h},i}, \quad j = 1, 2, \tag{8}$$

and income is given by

$$y_j^i = \alpha_{\mathrm{h}} h_j^i + \epsilon_j^{\mathrm{y},i}, \quad j = 1, 2. \tag{9}$$

[11] The effect of family income on skill in eq. (4) is taken here as given. One can easily set up a more complex model in which parents also decide on an investment in skill formation, in addition to human capital accumulation as described in sec. II.D. This more complex model is described in sec. S-0.1, where we show that it yields a skill equation just like eq. (4) and where the income term is produced by a household optimization problem, just as it is in eq. (8).

[12] In both eqq. (8) and (9), we could add on the right-hand side a term $w(g)$, multiplied by some additional parameter, to allow direct influence of genetic component on the variable. However, since this term already appears in the right-hand side of eq. (4), this genetic component will, even in the simple version presented in eqq. (8) and (9), be considered in empirical estimates, and this addition would make the model more complex with no substantial gain.

We assume a zero mean for shocks to human capital and income,

$$\forall\ i, j,\quad \forall\ k \in \{\mathrm{h}, \mathrm{y}\} : \mathbf{E}\epsilon_j^{k,i} = 0, \tag{10}$$

and assume that the shocks to human capital and income are not correlated:

$$\forall\ i, j : \mathbf{E}(\epsilon_j^{\mathrm{h},i}\epsilon_j^{\mathrm{y},i}) = 0.$$

At the optimal solution of the problem in equations (6)–(10), optimal parental investment is equal for the two siblings ($\hat{I}_1^i = \hat{I}_2^i \equiv \hat{I}^i$) and is a constant fraction of household income:

$$\hat{I}^i = \frac{\delta\alpha_{\mathrm{Ih}}}{1 - \delta + 2\delta\alpha_{\mathrm{Ih}}}\exp(y^i) \equiv \psi\exp(y^i), \tag{11}$$

where $\alpha_{\mathrm{Ih}} \equiv \alpha_{\mathrm{I}}\alpha_{\mathrm{h}}$. Equal investment in education for the two children is, of course, a very special feature due to the preferences we have adopted.

E. Income of the Children

In the analysis below, we also use this more general model to control for education of parents, college degree of parents, and work status of the father. Substituting the optimal investment reported in equation (11) into the human capital equation (8) and substituting the result into the equation for income (eq. [9]), we get the reduced equation for income:

$$y_j^i = a + \alpha_{\mathrm{Ih}}y^i + \alpha_{\theta\mathrm{h}}\theta_j^i + \alpha_{\mathrm{h}}\epsilon_j^{\mathrm{h},i} + \epsilon_j^{\mathrm{y},i}, \tag{12}$$

where $a = \alpha_{\mathrm{Ih}}\ln\psi$ and $\alpha_{\theta\mathrm{h}} = \alpha_\theta\alpha_{\mathrm{h}}$.

To complete the model, we need to specify how the pairs of parents are selected. To this we turn now.

F. Matching Processes

To complete the system described by equations (4), (8), (12), (18), and (19), we need to specify the matching process for parents. We assume that this process depends not only on the individual characteristics that we have described so far, namely, skill and income, which are relevant for economic outcomes, but also on characteristics in a set C that are important for matching but not for economic activity (such as the personality traits, different from cognitive or noncognitive skills, that are recorded in our data). Recall that Y is the set of log incomes; let $Z \equiv G \times Y \times \Theta \times C$ and the observable characteristics $Z_{\mathrm{O}} \equiv Y \times \Theta \times C$, with generic element z_{O}; for convenience, we indicate with a subscript (as in $\Delta_{\mathrm{m}}(Z)$) whether the element in $\Delta(Z)$ refers to the mother or the father.

A *matching* associates with a pair of distributions $(\mu_m, \mu_f) \in \Delta_m(Z) \times \Delta_f(Z)$ an element denoted $M(\mu_m, \mu_f) \equiv \nu \in \Delta(Z \times Z)$, describing the distribution of pairs of genotypes, skills, income, and characteristics of the two parents. The matching process is required to:

1. be feasible: the marginal of each type of parent distribution is the same as the original distribution for that type:
$$M(\mu_m, \mu_f)_{\Delta_i(Z)} = \mu_i, \quad i \in \{m, f\};$$
2. have conditional independence of genotype: the random variables g_m and g_f (genotype of mother and father, respectively) are independent, conditional on the information of observable characteristics.

The conditional-independence assumption requires that matching depends only on the observable characteristics $z_O \in Y \times \Theta \times C$; in other words, matches are made on the basis of observable characteristic and not on the genotype. Thus, matching of genotypes is not random within the population, but it is random within the set of individuals with given observable characteristics. The assumption is very weak, at least as long as individuals choose their partners without taking into account the results of genetic tests, which is typically not yet the case.

Random matching within the entire population is a special example of matching: in this case, a mother of type z_{mO} is selected and independently a father of type z_{fO}, according to μ_m and μ_f, respectively. This model is convenient for its simplicity, but it is not entirely supported by the data, which show instead substantial positive correlation between several characteristics of the parents. Thus, a model induced by preferences over matchings is desirable and will provide a better approximation. A detailed analysis of the equilibrium concept is presented in section A.

G. *Matching According to Worth*

The analysis of the invariant distribution is simpler if matching is dependent only on the income and skill of the spouse. So we set

$$\Pi = 0, \; F^i = 0, \; \epsilon^\theta = 0, \; \epsilon^h = 0, \text{ and } \epsilon^y \sim N(0, \sigma^2_{\epsilon^y}). \tag{13}$$

We call the set of individuals with the same worth a *worth class*. In this model, in each generation children are born of spouses of same worth (not necessarily income: higher skill may compensate for a lower income).

A pair of genotype and income (g, y) has a worth $w(g) + w_y y$. Mating is random within each worth class. To define these classes, we consider partitions of the set. A possible partition is the *discrete partition*, in which mating occurs only within pairs of exactly the same worth; we use this partition

as a simple but not very realistic example. A more realistic model has a "countable partition." To define it, we take a countable set of values, indexed by the integers

$$\mathcal{V} \equiv \{v_i : i \in Z\}. \tag{14}$$

We assume that these values are increasing in the index and that the distance between successive terms is uniformly bounded above and below:

$$\exists\ \underline{M}, \overline{M},\ \ \forall\ i : 0 < \underline{M} \ \leq\ v_{i+1} - v_i \ \leq\ \overline{M}. \tag{15}$$

The class of genotype and income pairs of worth v_i is defined as

$$C(v_i) \equiv \big\{(g, y) : w(g) + w_y y \in [v_i, v_{i+1})\big\}. \tag{16}$$

The worth function $W : G \times Y \to \mathcal{V}$ is defined as

$$W(g, y) \equiv v_i \quad \text{if } (g, y) \in C(v_i). \tag{17}$$

We consider a probability measure $\mu \in \Delta(G \times Y, \mathcal{B}(G \times Y))$, where $\mathcal{B}$ are the Borel subsets, as the description of the current distribution in the population of pairs of genotype and income. Genotype G is finite, so the Borel σ-field is the power set; using the Borel definition and notation for both G and Y simplifies the exposition.

As we mentioned above, children in our sample are all twins. The genetic transmission function in equation (1) is obviously true in particular for each individual twin. In addition to that equation, we have two additional conditions restricting the joint transmission to the pair of twins. These conditions depend on the twin type, an element on the set {DZ, MZ}, and are defined as

$$H_{\mathrm{DZ}}(g_{\mathrm{m}}, g_{\mathrm{f}})(g^1, g^2) \ =\ H(g_{\mathrm{m}}, g_{\mathrm{f}})(g^1) H(g_{\mathrm{m}}, g_{\mathrm{f}})(g^2) \tag{18}$$

for the genotype pair (g^1, g^2) of DZ twins and

$$H_{\mathrm{MZ}}(g_{\mathrm{m}}, g_{\mathrm{f}})(g^1, g^2) \ =\ \begin{cases} H(g_{\mathrm{m}}, g_{\mathrm{f}})(g^1) & \text{if } g^1 = g^2, \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

for that of MZ twins.

For the given μ, we describe the next-period measure as follows. Each worth class is chosen with the probability induced by μ on the worth space, denoted by μ_V. Two parents (that is, two pairs of genotype and income) are chosen according to the probability on that class of genotypes and income. Within the class, mating is random. The genotype of parents then determines the genotype of the child, and parents' income and education, together with the child's genotype, determine the child's

income. This entire process yields the new measure.[13] The complete model of the process on genotype, income, education, and skill is given by equations (4) for skill and (8) for education, the reduced equation (12) for income, and equations (18) and (19) for the genotype transmission. Together with the mating process presented in section II.F, these equations completely determine a nonlinear (because of the function H in eq. [1]) transition on measures on the space of genotypes and income, $\Delta(G \times Y)$. An invariant distribution is a fixed point of this transition function.

If an invariant distribution exists, we can then subtract from the variables $(y_j^i, \theta_j^i, h_j^i, w(g_j^i))$ their expected value with respect to the invariant distribution; so the constants are eliminated (e.g., the a term in the reduced equation for income is eliminated). Since no confusion is possible, we keep the same names for these variables, which now have a zero mean. We write equations (18) and (19) in the compact form:

$$g_j^i \text{ is distributed as } H_k(g_{\mathrm{m}}^i, g_{\mathrm{f}}^i), \quad k \in \{\mathrm{MZ}, \mathrm{DZ}\}. \tag{20}$$

If we substitute equation (4) into the reduced equation for income (eq. [12]), we get the twin's income y_j^i as a linear function of genetic endowment g_j^i, family income y^i, and environment F^i, and a weighted sum of idiosyncratic (j-dependent) variables:

$$\begin{aligned} y_j^i = {} & \alpha_{\theta\mathrm{h}} w(g_j^i) + (\alpha_{\mathrm{Ih}} + \alpha_{\theta\mathrm{h}}\pi) y^i + \alpha_{\theta\mathrm{h}} F^i \\ & + \alpha_{\theta\mathrm{h}} \Pi X_j^i + \alpha_{\theta\mathrm{h}} \epsilon_j^{\theta,i} + \alpha_{\mathrm{h}} \epsilon_j^{\mathrm{h},i} + \epsilon_j^{\mathrm{y},i}. \end{aligned} \tag{21}$$

The decomposition in equation (21) is a more detailed version of the standard ACE decomposition in behavioral genetics (see, e.g., Knopik et al. 2017, 358), where the phenotype is income; the "A" term is the additive contribution of genotype, $\alpha_{\theta\mathrm{h}} w(g_j^i)$; the common, or shared-environment, component "C" is the sum of the two terms $(\alpha_{\mathrm{Ih}} + \alpha_{\theta\mathrm{h}}\pi) y^i$ and $\alpha_{\theta\mathrm{h}} F^i$; and the sum of the last four terms is the "E" component.

We assume that

$$\alpha_{\mathrm{Ih}} + \alpha_{\theta\mathrm{h}}\pi < 1 \text{ and } \alpha_{\theta\mathrm{h}} > 0, \tag{22}$$

to ensure that (the first inequality) the income process is bounded and an invariant measure exists and that (the second inequality) skill has a nontrivial effect on income. The equation describing human capital accumulation is similar, up to the constant multiplier α_{h}; we report it here for convenience because we cite it below in the empirical analysis:

[13] For a precise definition of the transition from one period's measure to the next, we refer to sec. D; here, the income of the child is described by eq. (A9) and the genotype of the child by eq. (A11).

$$h_j^i = \alpha_\theta w(g_j^i) + (\alpha_1 + \alpha_\theta \pi) y^i + \alpha_\theta \Pi X_j^i + \alpha_\theta F^i + \alpha_\theta \epsilon_j^{\theta,i} + \epsilon_j^{h,i}, \quad (23)$$

and it is obtained by substituting equation (11) into equation (8) and subtracting the constant term.

Different further specifications of the model are possible, depending on how we model the variables X_j^i and F^i in equation (23) and therefore in equation (21). We explore these possibilities in detail in the rest of the paper. In particular, the equation modeling the variable F^i is examined in the section on passive gene-environment correlation (sec. III.A), and the model for the variable X_j^i is analyzed in the section on measurement error (sec. IV.B), where we discuss how we plan to estimate equations (21) and (23), thus providing a link between theory and empirical analysis.

III. Invariant Measures

We now show that an invariant measure exists and has some interesting properties. Existence of the invariant measure is far from immediate, because the process on distributions of skills and income in our model is nonlinear. The nonlinearity follows from the matching process: in every period, the two distributions (for potential mothers and fathers) are shuffled by the matching to produce a measure on the product space of spousal pairs.

A few preliminaries are necessary for a good understanding of the statement. We call the *skill allele* the allele at some locus that yields a higher value of the skill (more precisely, it has a higher genic value).[14] We find that, at equilibrium, matching is random within each worth class; thus, alleles are in Hardy-Weinberg equilibrium at all loci, but the frequency may differ across classes. We recall that a population is in Hardy-Weinberg equilibrium at a biallelic locus (with alleles denoted A and a and the frequency of A equal to p) if the frequencies of the three combinations (aa, aA, AA) are, respectively, $(1 - p)^2$, $2p(1 - p)$, and p^2; these are the combinations obtained by independent combination of two gametes carrying A or a (one from the father and one from the mother) with probabilities p and $1 - p$, respectively. Under some assumptions (described in detail in, e.g., sec. 3.1 of Nagylaki 1992 or sec. 2.2 of Crow and Kimura 1970), in particular the assumption that mating among male and female is random, a Hardy-Weinberg equilibrium is reached in one generation and maintained in all following generations. Finally, recall that we assume (eq. [22]) that skill affects income but that the total coefficient of household's income on children's is less than 1. We can state the following:

[14] The genic value is a measure of the contribution of the allele to the phenotype of interest, the skill in our case (see, e.g., Crow and Kimura 1970, 117).

THEOREM 3.1. Assume equation (22) and that the worth of an individual depends linearly on income and skill. Then, for any vector of allele frequencies:

1. an invariant measure exists, which induces that allele frequency;
2. within each worth class, alleles at each locus are in Hardy-Weinberg equilibrium;
3. within each worth class of the discrete partition, a higher income of both parents implies a lower expected PGS of the child; and
4. the allele frequencies are invariant across periods.

Some remarks may help to clarify the statement. An invariant measure exists in spite of the process being nonlinear and for any initial allele frequency. The proof relies on the order structure of the genotype and income space. The Hardy-Weinberg equilibrium holds, but only within worth classes. One can thus compute the fixation index, which is a measure of population differentiation due to genetic structure across populations (in our model, populations are income and skill classes). The deviations from Hardy-Weinberg equilibrium in the population may be small, since the phenotype is highly polygenic and the size of the GWAS coefficient declines quickly. Still, as we see in section III.B.2 below, the model predicts a stratification across populations of the alleles with stronger effect. Higher income of both parents is compensated for by the lower skill implicit in the genotype (the third statement). The last statement shows that frequency in the population of each allele does not change from one period to the next. So there may be many invariant measures, depending on the initial condition (at least 2^K; see proposition A4). The intuitive reason for the invariance property is that, as long as income does not affect the relative fertility for different genotypes and incomes, the specific features of the mating process may affect the association of genotype and income but can only reshuffle the existing alleles. The lack of differential effects on fertility is a strong assumption, particularly when we are interested in secular development, and examining the implications of relaxing it is an essential next step in research.

A. Gene-Environment Correlation

In our estimation (presented in sec. V) of the two equations (21) and (23), we consider, among the independent variables, the PGSs of the parents. We justify here the reason for this choice. Clearly, all the information on the genotypes of the parents that could be potentially relevant for the determination of the genotype of the twins is rendered irrelevant by the direct information that we have on the genotype of the twins. However, the genotypes of the parents can very well have an additional indirect effect

on the phenotype of interest of the offspring (educational achievement, in our case) through the effect of the environment on the phenotype (passive gene-environment correlation; Plomin, DeFries, and Loehlin 1977; Scarr and McCartney 1983; Jaffee and Price 2007).

The idea of gene-environment correlation (usually denoted rGE) rejects the assumption that environment and genes are uncorrelated.[15] The correlation may arise in three main ways. The most important for our purposes is the "passive rGE effect."[16] Genes of the parents affect directly the genes of the children, but they also affect the environment in which the child grows; hence the potential for correlation between G and E. For example, higher intelligence of parents, due in part to the genes of the parents, may be transferred directly through genes to children but also through the family environment created by parents. A related concept, "genetic nurture," has been extensively explored by Kong et al. (2018) and Okbay et al. (2022), and we discuss it below.[17]

We now discuss how rGE can be analyzed within our model and how we can then estimate it in our data analysis. First, the household income (y^j in eq. [23]) is already an example of an rGE path: the income of the parents is determined in part by their genes (this follows applying the income eq. [21] to the parents) and also by the grandparents' genes (iterating the process), and so on. Similarly, if we include among the variables in the vector X_j^i the human capital h^i of the parents, then applying the human capital equation (23) to the parents and iterating, we see that parents' genes, grandparents' genes, and so on are relevant. Since the entire ancestry of the individual enters into the determination of the family income and parents' education, we refer to this as *ancestral rGE*. Models of parental investment, as in Becker and Tomes (1979), are a special, very simplified,

[15] The rGE is different from the gene-environment interaction (usually denoted $G \times E$). The latter describes the idea that even if genes and environment are independent, the way in which each of the two operates on personality and behavior may depend on the value of the other; i.e., genes and environment do not operate additively. For example, genes may determine the motivation of an individual (as a personality trait, measured, e.g., by tasks or survey questions), and environment may offer opportunities (measured, for instance, by schooling available in the place of origin), but the resulting success of the individual (measured by education or income) may be different from a linear combination of the two. For example, in a poor environment where opportunities are severely constrained, a person with high motivation and intelligence may fail just as one with low values, and the difference may emerge only when adequate opportunities are offered.

[16] The other two effects are evocative and active. The *evocative* effect refers to the difference in response that different genotypes induce in the environment; for instance, more active children are more likely to induce stronger social stimulation from the environment and hence richer learning. The *active* effect is produced by the selection, perhaps purposeful, of different environments in which to operate by different genetic types. These two effects are harder to estimate in our data.

[17] Genetic nurture in Kong et al. (2018) is defined to operate through those genes that are not transmitted from parents to children. The role of family environment in considered in detail in Willoughby et al. (2021), which is discussed in detail in sec. VII.

case of ancestral rGE. We have information on family income and parents' education in our data, and so we can control for its effects. But passive rGE may arise in a different, more subtle way, which we model by considering the case in which, in equation (23), the variable F has the special form

$$F^i = \alpha^C_{\mathrm{m}} g^i_{\mathrm{m}} + \alpha^C_{\mathrm{f}} g^i_{\mathrm{f}}, \tag{24}$$

that is, the family environment depends on the genetic profile of the parents though some k-dimensional vectors that may differ for father and mother. The additive form is the same we make for the genes affecting directly educational attainment. In section E, we provide a detailed analysis of this case.

We emphasize that the weights α^C in equation (24) may be very different from those estimated by β in section II.C. In particular, very different genes (more precisely, SNPs) can be relevant in equations (3) and (24). We provide an example of this difference below, where the two sets of genes are disjoint. We also emphasize that "parents" in equation (24) should be interpreted in the more precise meaning of individuals providing parents' role. For example, if the child is adopted, then the genotypes $(g^i_{\mathrm{m}}, g^i_{\mathrm{f}})$ in equation (24) are those of the adopting parents, not the biological ones (and the same holds for y^i and h^i). With minor changes, the proof of theorem 3.1 holds, and thus in particular existence of an invariant measure holds. We refer to this component of rGE as "parental rGE."

B. Numerical Computation

The main properties of the process and equilibrium distribution of the model in section III can be illustrated with a numerical computation of the equilibrium distribution.[18] We study the distribution in $\Delta(G \times Y)$ in successive generations of a constant-size population where each household has two children. The sex of each child is determined independently (from each other and from the other variables), with probability 1/2 on each sex.

1. Speed of Convergence

Convergence to the invariant distribution is fast and approximately achieved in our model within five generations. The value of the ratio of the norm of the difference between current and past μ to the norm of the current μ is within 10% after five generations and within 2.26% after 10 generations.

[18] Coding is in Matlab (release R2022b). The Matlab code is available upon request.

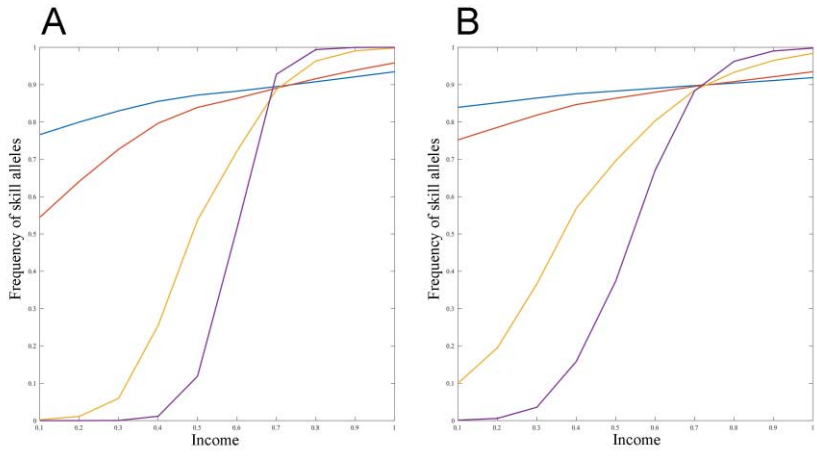

FIG. 1.—Population stratification and rGE. Both panels display the frequency of alleles by income. The flattest line, with smallest difference across income, describes the frequency of the allele with smallest genic value; the others are in increasing order, with purple line for the highest-effect allele. *A*, Only child's genotype affects income (no passive rGE). *B*, Only parents' genotypes affect income (full rGE). The figure illustrates how two very different economies may have very similar statistical properties.

2. Endogenous Population Stratification

The skill alleles have at equilibrium a frequency that is increasing with worth, education, and income. As we mentioned in theorem 3.1, society is stratified. The effect is strong, and it is stronger the higher the genic value of the allele. Both facts are illustrated in figure 1*A*.

3. Parental rGE

An intuitive reason for the next result is provided by a simplified example. Consider the case in which the set of genes (or more precisely the SNPs) that are relevant in equation (3) and the other set of those relevant for equation (24) have an empty intersection. We refer to the first set as EA (for educational attainment) SNPs and to the second as PC (for parental care) SNPs. SNPs improving parental care also affect positively the education and income of the children. Obviously, children's PC SNPs are correlated with those of the biological parents, by 50% or more (because of assortative mating); and since parents' PC SNPs affect educational attainment of the children, these SNPs will be correlated to educational attainment and thus will appear to influence educational attainment directly even if they are not. This is illustrated by a comparison of the two panels of figure 1. The panels report the main features of two economies that have the same underlying preferences and technology but completely different

pathways from genes to traits; that is, they differ only in the functions w and F. We refer to the economy in panel A as the EA economy and to that in panel B as the PC economy for short.[19]

The figure illustrates the following results. First, just as in the case in which the effect on educational attainment is direct, there is also population stratification when the genetic effect occurs only through parental care, with higher frequency of the alleles with positive effect in the richer, more educated population (see fig. 1*A*). Second, for each allele k, the distribution of income for the three subgroups of the population with $g(k) = 0, 1, 2$ is different, even in the economy where there is no direct genetic effect on education (see fig. 1*B*). Third, as a consequence of the second point, the estimated GWAS coefficients for educational attainment are significant and positive even in this latter economy, where there is no direct effect of genes on educational attainment.

Obviously, in principle parental rGE affects children's phenotype. The real question is, Once we control for ancestral rGE, is parental rGE quantitatively important? In section VII, we show that the answer is negative.

C. *Intergenerational Mobility: Standard and Genetic Models*

In this section, we compare the predictions of the model we have presented with those of the standard model of parental investment. The model with autoregressive transmission of skill (as introduced in Becker and Tomes 1979) has (adopting our notation to this case) the following equations: for income in generation t,

$$y_{t+1} = \alpha_{\text{Ih}} y_t + \alpha_{\theta\text{h}} \theta_{t+1} + \epsilon^{y}_{t+1}, \tag{25}$$

and for skill,

$$\theta_{t+1} = \eta \theta_t + \epsilon^{\theta}_{t+1}, \tag{26}$$

where $\eta \in (0, 1)$ is a fixed "heritability" parameter. Note that there is only one type of skill. At the stationary distribution, we can compute, using the Yule-Walker equations, the intergenerational income elasticity ρ_{PM} (the subscript "PM" stands for perfect matching; the reason for this will be clear in the comments following eq. [32]) to be

$$\rho_{\text{PM}} = \alpha_{\text{Ih}} + \alpha_{\theta\text{h}} \frac{\eta \mathbf{E}(\theta y)}{\mathbf{V}(y)}, \tag{27}$$

[19] The figure relies on the analysis developed in sec. E. In the notation of that section, there is a K-dimensional vector α such that, in panel A, $\alpha^{\text{A}} = \alpha$ and $\alpha^{\text{C}}_{\text{m}} = \alpha^{\text{C}}_{\text{m}} = 0$, and, in panel B, $\alpha^{\text{A}} = 0$, $\alpha^{\text{C}}_{\text{m}} = \alpha^{\text{C}}_{\text{m}} = (1/2)\alpha$. In simple words, panel A describes an economy where all alleles are EA and no passive rGE exists; in panel B, no allele affects educational attainment, and the effect is only through the environment provided by the parents.

where $\mathbf{V}$ denotes the variance of a random variable and $\mathbf{E}(\theta y)$ and $\mathbf{V}(y)$ have an explicit expression in terms of the primitive parameters.[20] When $\sigma_{\epsilon^y} = 0$, the intergenerational persistence formula (27) becomes the well-known formula (see, e.g., Solon 2004) in which persistence is a simple weighted average of the income and skill transmission:

$$\rho_{\mathrm{PM}} = \frac{\alpha_{\mathrm{Ih}} + \eta}{1 + \alpha_{\mathrm{Ih}}\eta}. \tag{31}$$

A direct comparison of the standard model (eqq. [25]–[26]) with a genetic model like equations (20)–(21), where sex is an essential component of reproduction, is meaningless, since, apart from the genes, there are not even two parents in the standard model. So we must first build a more general model that includes the standard one as a special case of the general class of models (with gametic reproduction, as is the case for human population) in sections A and II.G. We assume income and skill to be the weighted average of the income and skill of the two parents, as in equations (A3) and (A4). Thus, the income of the child follows the equation

$$y_{t+1} = \alpha_{\mathrm{Ih}} \sum_{i=\mathrm{m,f}} w_i^{\mathrm{y}} y_{it} + \alpha_{\theta\mathrm{h}}\theta_{t+1} + \epsilon_{t+1}^{\mathrm{y}}, \tag{32}$$

and the skill transmission follows

$$\theta_{t+1} = \eta \sum_{i=\mathrm{m,f}} w_i^{\theta}\theta_{it} + \epsilon_{t+1}^{\theta}. \tag{33}$$

The matching between parents that decides the pairing of $(\theta_{\mathrm{m}t}, y_{\mathrm{m}t})$ with $(\theta_{\mathrm{f}t}, y_{\mathrm{f}t})$ is determined by preferences and stable matching, as in section A. The standard model (eqq. [25]–[26]) becomes a special case of equations (32)–(33) when we assume that preferences of mothers and fathers are lexicographic (with any order on θ and y) and that $\mu_{\mathrm{m}} = \mu_{\mathrm{f}}$, so matching occurs only among identical types (perfect matching, hence the PM subscript).

We now show that the formulas for intergenerational income elasticity (eq. [27] or [31]) of the standard model are an upper bound on the persistence within the class of models requiring equations (25), (32), and (33). The reason is that, as we have just seen, the standard model maximizes

[20] The explicit expressions are

$$\mathbf{V}(\theta) = \frac{\sigma_{\epsilon^\theta}^2}{1 - \eta^2}, \tag{28}$$

$$\mathbf{E}(\theta y) = \frac{\alpha_{\theta\mathrm{h}}\mathbf{V}(\theta)}{1 - \alpha_{\mathrm{Ih}}\eta}, \tag{29}$$

$$\mathbf{V}(y) = \frac{1}{1 - \alpha_{\mathrm{Ih}}^2}(\alpha_{\theta\mathrm{h}}^2\mathbf{V}(\theta) + \sigma_{\epsilon^y}^2 + 2\alpha_{\mathrm{Ih}}\alpha_{\theta\mathrm{h}}\eta\mathbf{E}(\theta y)). \tag{30}$$

the similarity among parents, forcing their income and skill to be identical. For example, consider the case where parents match only on income but may differ in skill. This happens when preferences are linearly ordered by the income of the spouse. In this case, the corresponding intergenerational elasticity, call it ρ_{MY}, can be shown to satisfy

$$\rho_{\text{MY}} < \rho_{\text{PM}}. \tag{34}$$

The proof is in section B. We can now discuss the relation between predictions of the standard and genetic models on the important issue of the size of intergenerational mobility. The standard model with autoregressive transmission of skill assumes a fixed η (in eq. [26]). Such a fixed parameter, however, has no correspondent in reality: the genetic model shows that the persistence represented by that η is endogenous and depends on the distribution of the genotype. Therefore, the corresponding elasticity, call it ρ_{G}, also does depend on the distribution, which is different in different populations. So persistence may differ among populations independently of preferences, technology, and institutions in the economy, depending only on the distribution of the genotype in that population.

An important implication of the differences we have highlighted so far is that the persistence in a model with genetic transmission of skill can be higher than that in the standard model, even higher than the highest possible value in the class of standard models with sexual reproduction (presented in eqq. [32]–[33]). That is, it may be the case that $\rho_{\text{G}} > \rho_{\text{PM}}$. It follows, in particular, that the adoption of the amended model with AR(1) transmission and sexual reproduction (eqq. [32]–[33]) might make predictions worse, by further underestimating the persistence.

We illustrate this possibility in a simple but clarifying example. Take $K = 1$ (a single locus with alleles $\{A, a\}$), with frequency $p(A)$ of A, determining a one-dimensional skill $\theta \in \{\theta_0, \theta_1, \theta_2\}$, ordered as the index. Preferences are determined by the household maximization problem and hence are described by equation (A6); and to ease comparison with the simple form (eq. [31]), we assume $\sigma_{\epsilon^y} = 0, \Pi = 0, F = 0$, and $\epsilon^\theta = 0$.

This economy has a stationary distribution at two values:

$$(0, y_0, \theta_0) \text{ with probability } 1 - p(A) \text{ and } (2, y_2, \theta_2) \text{ with probability } p(A), \tag{35}$$

where

$$y_i = \frac{\alpha_{\theta \text{h}} \theta_i}{1 - \alpha_{\text{Ih}}}.$$

The persistence here is 1, and this can never occur in an autoregressive model with $\eta < 1$.

The example is obviously artificial in the assumption that a skill phenotype is determined by a single locus, whereas the skills of interest for

economic applications are highly polygenic. The force highlighted by the example, however, is not at all artificial, and it points to the effect that assortative mating has on increasing the variance of genetic values and magnifying the heritability and the resemblance between relatives.[21] This effect is absent by assumption in the autoregressive model, even in the amended version in which two partners are introduced, given by equations (25), (32), and (33).

IV. Estimation Strategy

Our empirical analysis will estimate equations (21) and (23). In the next two subsections, we discuss the introduction into the analysis of the genotype of the parents among the explanatory variables (X_j^i) and the additional information we can derive from a special subset of our data, the DZ twins. We recall that the joint distribution of genotypes is described by equations (18) and (19).

A. *Correlation among Twins*

In the fixed-effects analysis below, we rely on the fact that DZ twins share important environmental characteristics but do not entirely share the genotype. The degree of the sharing depends on the nature and strength of the assortative matching between parents. Genetic correlation among parents may occur for two different types of reasons. Correlation may exist because matching is directly on the relevant phenotype (e.g., the correlation on genes affecting intelligence among parents occurs because parents match according to intelligence), or it may occur indirectly, when matching occurs along dimensions unrelated to the phenotype (e.g., matching occurs along the characteristics in the set C of physical appearance), but because of population stratification a correlation between genes affecting variables in C and Θ exists.[22]

Whatever the cause, the correlation for DZ twins is a simple function of the correlation between the PGS of the parents. We use the subscripts 1

[21] This force is well recognized in population genetics: see chap. 4 of Crow and Kimura (1970), in particular sec. 4.6 for our single-locus example and sec. 4.7 for a multivariable example. The analysis in population genetics is very different from the one we present here because the assortative mating in our model is endogenous and determined at equilibrium in the marriage market.

[22] We can illustrate this second possibility, considering the extreme case in which there is no overlap between loci affecting the θ skills and the characteristics in C and matching along C characteristics is perfect. In this case, the stationary distribution has segregated populations with different frequencies on the alleles determining θ and thus different distributions on the θ skills. This equilibrium is not robust, of course: with a small imperfection in the C-matching the frequency of the θ alleles converges exponentially in the long run to a value independent of the C characteristics; however, the transition is slow when the imperfection is small, and in the transition the correlation may be substantial.

and 2 to indicate that the variable refers to first and second siblings and the subscripts "m" and "f," for mother and father, respectively. Then:

LEMMA 4.1. The correlation between the standardized PGSs of nonidentical full siblings, hence in particular of DZ twins, is equal to 1/2 plus half of the correlation between the standardized PGSs of the parents, that is,

$$\mathbf{E}(\mathrm{PGS}_1\mathrm{PGS}_2) = \frac{1}{2} + \frac{1}{2}\mathbf{E}(\mathrm{PGS}_\mathrm{m}\mathrm{PGS}_\mathrm{f}).$$

The proof is in section C.[23] Lemma 4.1 gives the predicted correlation among DZ twins as a function of the correlation among parents. In section VII, we present the correlation among parents' PGSs and find that data are consistent with the prediction of the lemma.

In the next sections, we test and estimate the parameters of the two equations (21) for income and (23) for human capital. The data we use are described in detail in section IX.

B. Measurement Error and Structural Equation Models

Reliable estimates—for example, that of the path from genetic factors to educational outcomes—must take into account errors in measurement of the variables. This is obviously important if we want to minimize downward biases of single coefficients, but it is even more important if we want to compare the relative sizes of the effect operating through cognitive and noncognitive skills, since the error in measurement might be different for the two groups of variables. For example, it might be natural to expect a larger error in measures of noncognitive skills, based on surveys, than in those of cognitive skills, based on tests. We model explicitly and estimate errors in measurement using a structural equation model (SEM).

The SEM we consider is of the usual form (see, e.g., Bollen 1989):

$$\mathbf{Y} = B\mathbf{Y} + \Gamma\mathbf{X} + \alpha_Y + \zeta, \tag{36}$$

with $\mathbf{Y}$ an m-vector of m_Y endogenous observed variables y and m_η endogenous unobserved variables η; $\mathbf{X}$ an n-vector of n_X exogenous observed variables x and n_ξ exogenous unobserved variables ξ; α_Y a vector of means; and ζ a vector of errors. Entries of B are denoted by β's, entries of Γ by γ's. In this section, we adopt the notation convention that variables with a capital first letter are endogenous and those with a lowercase first letter are exogenous.[24]

[23] On the related, but different, issue of segregation variance (i.e., the variance of the offspring about the mid–parent value), see Rogers (1983).

[24] This carries the modest price of changing PGS to pGS.

We set up our analysis by adopting a general form (eq. [36]) to test the basic equations of the model, with basic equations (21) and (23). Specific examples are the system of equations (39)–(40) and that of equations (43)–(45). We recall that the variables in the vector X_j^i for ij (introduced in sec. II.C) are not necessarily observed, so we add equations providing measurements of these latent variables. They may also be endogenous, so we add equations describing how they are determined. Examples of variables that are components of X_j^i are the latent endogenous variables C and NC (cognitive and noncognitive skills, in eqq. [37]–[38]; these are the η-variables); observed endogenous variables e_h and y_h in equations (43)–(44), (y-variables); and finally pGS_m and pGS_f, exogenous observable x-variables in equations (43)–(45).

V. Income and Human Capital Determination

We first estimate the parameters of the model presented in section II.G. Table 1 reports the panel regression of the log income at the age-29 assessment over family income, PGS, and other control variables. Estimates reported in the table control for the difference in the age of the individuals (parent or child) at which the information on income was collected. Since wage increases with age at a rate that may be heterogeneous (Rupert and Zanella 2015; Lagakos et al. 2018), this difference may introduce a bias in the estimated coefficient if the slope depends on characteristics, such as education, that are correlated with wage. We use a specification of the Mincer (1974) equation that has that of Lagakos et al. (2018),[25] as special case, allowing the slope to depend on education. If the slope increases with education, we expect the estimated elasticity coefficient to overestimate the true value; thus, we control for the time difference, the education of the parents, and an interaction term.

The estimated unconditional intergenerational elasticity (IGE) is 0.134 (SE = 0.027); the table (col. 1) reports the values after sex and the interaction between sex and family income are controlled for. Age does not have a significant effect, as might be expected, since individuals in the sample are approximately the same age. Sex of the individual has a strong and significant effect: income for male individuals has a substantially larger intercept (27.7%) but a smaller (by 6%) dependence on the family income. The fraction of males in the twins population is 48%; thus, the standardized male variable is approximately equal to 1 for male and −1 for female.

In column 2, the coefficient of the individual PGS is 7.8% (SE = 0.025, p-value = .002). Its size is approximately half of that of family income (12.8%). Considering that the PGS we are using is estimated from

[25] Specifically, the formulation given in sec. VI.A of Lagakos et al. (2018).

TABLE 1
INCOME AT THE AGE-29 ASSESSMENT, FAMILY INCOME, PGS, AND PERSONALITY

	COEFFICIENTS (SE)		
VARIABLE	(1)	(2)	(3)
Family income	.134***	.128***	.078**
	(.027)	(.027)	(.032)
Male	.277***	.276***	.313***
	(.025)	(.025)	(.029)
Male × family income	−.060**	−.060**	−.050*
	(.025)	(.025)	(.030)
PGS		.078***	.021
		(.025)	(.028)
Education in years			.256***
			(.035)
IQ			.008
			(.029)
MPQ PA			.061**
			(.026)
MPQ NA			−.024
			(.027)
MPQ CN			.034
			(.032)
Externalizing			−.072*
			(.037)
Academic effort			.057
			(.038)
Academic problems			−.017
			(.034)
Observations	2,100	2,100	1,485

NOTE.—All variables, including parent college and male, are standardized to mean zero and SD 1. The signs of the MPQ variable NA, externalizing, and academic problems are reversed. Estimates control for principal components and the parent-child time difference in age at income data collection.

* $p < .1$.
** $p < .5$.
*** $p < .01$.

coefficients from a GWAS for education, it is likely that the weight of genetic factors affecting income is higher.

Column 3 in the table presents controls for some of the variables that are likely to mediate the effect of the PGS. Education years is the most natural variable to capture the effect of the PGS in education, and in fact the estimated coefficient is large (25.6% [SE = 0.035], p-value < .001) and significant.[26]

The controls for principal components and the difference in the age of parents and children at the collection of data on income produce no significant coefficient; the IGE falls after the control for difference in age, as expected, but in small measure (on the order of 10%). Controls

[26] Note that the sample size is smaller because several variables are missing for some subjects.

for additional variables (in particular education of parents and PGSs of father and mother) produces elasticity coefficients that are small and nonsignificant, with no effect on the coefficients of the variables of more significant interest.[27]

The values of IGE are on the lower side of the currently available estimates for developed countries, which vary between a minimum of 0.2 and 0.4 (see, e.g., Solon 1992, Zimmerman 1992, Mazumder 2005, Lee and Solon 2009, and surveys in Blanden 2011 and Björklund, Roine, and Waldenström 2012). The coefficient reaches higher values in some studies: see, for example, Palomino, Marrero, and Rodríguez (2018), who in a finer analysis (taking into account quartiles of the distribution) show that it can take higher values for the highest and lowest levels of income. There are some possible explanations for this difference. One is measurement error in our income data. Another is that in some developed countries with European populations the IGE coefficient is lower. For example, in Sweden (a country that is more relevant, given the demographic composition of Minnesota at the time in which the data were collected), values are lower (see, e.g., Österberg 2000, 427, where values are around 0.125), although they can be substantially higher at higher values of income (see Björklund, Roine, and Waldenström 2012) that are less relevant for our sample.[28]

VI. Identifying the Path from PGS to Education

In this section, we identify how much of the effect of PGS on educational achievement can be attributed to factors such as cognitive or noncognitive skills.

In our estimates below, the vector $\mathbf{Y}$ has a vector η of endogenous latent variables equal to (C, NC), denoting cognitive and noncognitive skills, respectively. The structural observed component of the $\mathbf{Y}$ vector is number of education years of the twins, e: we focus on this measure (rather than college or GPA [grade point average]) because it is the most relevant for economic consequences. The measurement variables in $\mathbf{Y}$ are a vector $((\mathrm{ct}_i)_{i=1,\ldots,I_{\mathrm{ct}}}, (\mathrm{nct}_j)_{j=1,\ldots,J_{\mathrm{nct}}})$ of measurements of cognitive and noncognitive skills. Turning to the vector $\mathbf{X}$, in our case $x \equiv ((x_k)_{k=1,\ldots,K})$ is a vector of control variables, such as household income variables, education and PGS of parents, the principal components, age, and sex. The system we estimate is

[27] The coefficients are 0.004 (SE = 0.166) for parents' education, −0.03 (SE = 0.037) for mother's PGS, and 0.04 (SE = 0.038) for father's PGS.

[28] See Björklund and Jäntti (1997) for a comparison of Sweden and United States on intergenerational mobility; they find that mobility is higher in Sweden.

$$C = \beta_{\mathrm{C}}\mathrm{pGS} + \zeta_{\mathrm{C}}, \tag{37}$$

$$\mathrm{NC} = \beta_{\mathrm{NC}}\mathrm{pGS} + \zeta_{\mathrm{NC}}, \tag{38}$$

$$\mathrm{ct}_i = \alpha_{\mathrm{ct}_i} + \gamma^{\mathrm{C}}_{\mathrm{ct}_i} C + \zeta_{\mathrm{ct}_i},\ i = 1, \ldots, I_{\mathrm{ct}}, \tag{39}$$

$$\mathrm{nct}_j = \alpha_{\mathrm{nct}_j} + \gamma^{\mathrm{NC}}_{\mathrm{nct}_j}\mathrm{NC} + \zeta_{\mathrm{nct}_i},\ j = 1, \ldots, J_{\mathrm{nct}}, \tag{40}$$

$$e = \alpha_{\mathrm{e}} + \gamma^{\mathrm{C}}_{\mathrm{e}} C + \gamma^{\mathrm{NC}}_{\mathrm{e}}\mathrm{NC} + \sum_k \gamma^{x_k}_{\mathrm{e}} x_k + \zeta_{\mathrm{e}}, \tag{41}$$

$$\gamma^{\mathrm{C}}_{\mathrm{ct}_1} = \gamma^{\mathrm{NC}}_{\mathrm{nct}_1} = 1. \tag{42}$$

The PGS may be added to the right-hand side of equation (41) with little consequence. The normalization condition (42) is necessary because any multiplication of the variables β_{C} and ζ_{C} by a positive constant, and corresponding division by the same constant of the vector $(\gamma_{\mathrm{ct}_i} : i = 1, \ldots, I_{\mathrm{ct}})$, gives a new vector of parameters, with the corresponding random variables still satisfying the system of identification equations; a similar rescaling of β_{NC}, ζ_{NC} and $(\gamma_{\mathrm{nct}_j} : j = 1, \ldots, J_{\mathrm{nct}})$ would have the same effect. Hence the two normalization conditions (42). With this normalization, the model is identified, if there are at least two cognitive and two noncognitive tests. More precisely:

PROPOSITION 6.1. Assume that $I_{\mathrm{ct}} \geq 2$ and $J_{\mathrm{nct}} \geq 2$; then the system (37)–(42) is identified.

Proof. Substituting equations (37)–(38) into equations (39)–(41) reduces the system to a system of observed variables. We indicate by σ^2_X the variance of a variable X. The simpler system in observed variables can be solved recursively for the parameters in the following order: σ^2_{pGS}, β_{C}, β_{NC}, $\gamma^{\mathrm{C}}_{\mathrm{ct}_i}$, $\gamma^{\mathrm{NC}}_{\mathrm{nct}_i}$, $\sigma^2_{\zeta_{\mathrm{C}}}$, $\sigma^2_{\zeta_{\mathrm{NC}}}$, $\sigma^2_{\zeta_{\mathrm{ct}_i}}$ for all $i \neq 1$, $\sigma^2_{\zeta_{\mathrm{nct}_j}}$ for all $j \neq 1$, $\gamma^{\mathrm{C}}_{\mathrm{e}}$, $\gamma^{\mathrm{NC}}_{\mathrm{e}}$, and finally $\sigma^2_{\zeta_{\mathrm{e}}}$. QED

The structural component of the SEM estimation is reported in table 2. In the equation for education years, the coefficients for both C and NC are significant. We can compute with the delta method the product of the coefficient for the link from the PGS to the variable C, times the coefficient from C to education years. The value of the product is 0.082 (SE = 0.018, $z = 4.53$, p-value < .001), with confidence interval [0.046, 0.117]. The corresponding product for the path passing through NC has a value of 0.034 (SE = 0.019, $z = 1.8$, p-value = .071), with confidence interval [−0.003, 0.072].

Once we control for C and NC, the coefficient of the PGS is not significant (p-value = .725). For comparison, we note that in the regression restricted to twins, controlling only for sex, the coefficient is 18.7% (SE = 0.022, $z = 8.37$, p-value < .001). The coefficients of education of parents and family income are both significant and of the same order

TABLE 2
SEM OF PATHWAYS FROM PGS TO EDUCATION YEARS ($N = 852$)

Equation, Variable	Coefficient	z	p-Value	Confidence Interval
Education years:				
C	.285	4.87	<.001	[.171, .401]
	(.058)			
NC	.856	3.11	.002	[.315, 1.4397]
	(.276)			
PGS	.014	.35	.725	[−.066, .94]
	(.041)			
PGS mother	.033	.71	.282	[−.027, .093]
	(.030)			
PGS father	.019	.66	.512	[−.039, .078]
	(.030)			
Education of parents	.136	4.58	<.001	[.078, .194]
	(.29)			
Family income	.075	2.38	.017	[.013, .137]
	(.031)			
Male	−.151	−2.77	.007	[−.260, −.041]
	(.055)			
Constant	.376	9.85	<.001	[.301, .450]
	(.027)			
C:				
PGS	.287	9.21	<.001	[.226, .349]
	(.031)			
NC:				
PGS	.040	1.95	.051	[−.0002, .081]
	(.025)			

NOTE.—The model estimated is described in eqq. (37)–(41). All observed variables are standardized to mean zero and SD 1. Cognitive-skills test scores (C) are verbal and performance IQ; the noncognitive-skills test scores (NC) are the three broad MPQ dimensions. Standard errors are estimated by bootstrapping. Model vs. saturated: $\Pr > \chi^2 < 0.0001$.

of magnitude, but that of education of parents (13.6%, SE = 0.029, $z = 4.58$, p-value < .001) is approximately twice that of family income (7.5%, SE = 0.031, $z = 2.38$, p-value = .017). The PGS of parents is not significant.

VII. Fixed-Effects Analysis and Parental rGE

In this section, we estimate the equations for income and human capital, using three important additional pieces of information: the fact that children are twins, both DZ and MZ, the overlapping-generations structure of the model, and the information, including genetic information, on parents. We begin with the analysis based on DZ twins.

A. *Fixed-Effects Analysis with DZ Twins*

DZ twins offer a uniquely informative way for the analysis of the effect of genetic variables on educational achievement. DZ twins share many

significant variables: date and condition of birth, family background, and very similar family environment in the following years. Therefore, a fixed-effects analysis of measures of educational achievements regressed on PGS, once restricted to DZ twins, will control for the effect of environmental factors common to the two twins.

We have seen in section IV.A the theoretical estimate of the correlation among DZ twins, depending on the degree of assortative mating of the parents. The difference in PGS correlation and the predicted correlation with random assortative matching (which is 1/2) is 0.083, and it must be due to the assortative matching among parents. In our case, we are considering not the genome-wide correlation[29] but that between the PGSs of parents. The correlation coefficient between the PGSs of the two parents is $r = 0.152$. As discussed recently in the literature (see Abdellaoui, Verweij, and Zietsch 2014; Robinson et al. 2017), the estimate of genetic assortative mating can be influenced by population stratification, which may produce spurious correlation. For example, the genetic assortative mating estimated in Domingue et al. (2014) becomes insignificant when a control with principal components is performed.[30] In table S-10 (sec. S-0.5) we report the controls for principal components in our data. The table shows that the estimated correlation in PGSs of spouses is robust to such control. This correlation is to be expected, given the strong correlation between education years of the two parents: for education years, the correlation coefficient is $r = 0.522$, and for IQ it is 0.37.

The fixed-effects analysis is presented in tables S-2 to S-5 for education years, GPA, college, and IQ score, respectively. All the regressions show that the coefficient of the PGS is significant in the fixed-effects regression. In the case of GPA, the coefficient is large and is approximately equal in the two regressions.

B. The Explanatory Power of Parents' PGS

A different way to control for the effect of genetic endowment of parents on family environment is to control directly for their PGS; in this case, we can use the information on both types of twins, including MZ.

The system we estimate is presented in equations (43)–(45) and is a special case of the general SEM structure in the general system (eq. [36]), with the same interpretation for the parameters α_Y, β, γ, and σ as in section IV.B. As usual, the superscript i refers to the family and the subscript j to the twin. The variables e_h and y_h denote, respectively, education of parents (average

[29] See p. 12 of the Supplementary Information in Robinson et al. (2017).

[30] See sec. S2 ("Principal Components") of the Supporting Information for Domingue et al. (2014), in particular table S1. These are the same tests we use in table S-10.

of the education years of the two parents) and family income. The y-variables are e_{h}, y_{h}, and e; the x-variables are $\mathrm{pGS_m}$, $\mathrm{pGS_f}$, and pGS. There are no exogenous latent ξ-variables. The equations of the model (note that, unlike the estimate reported in table 2, we are not controlling for C and NC variables) are

$$e_{\mathrm{h}}^{i} = \alpha_{e_{\mathrm{h}}} + \gamma_{e_{\mathrm{h}}}^{\mathrm{pGS_m}}\mathrm{pGS}_{\mathrm{m}}^{i} + \gamma_{e_{\mathrm{h}}}^{\mathrm{pGS_f}}\mathrm{pGS}_{\mathrm{f}}^{i} + \zeta_{e_{\mathrm{h}}}, \tag{43}$$

$$y_{\mathrm{h}}^{i} = \alpha_{y_{\mathrm{h}}} + \gamma_{y_{\mathrm{h}}}^{\mathrm{pGS_m}}\mathrm{pGS}_{\mathrm{m}}^{i} + \gamma_{y_{\mathrm{h}}}^{\mathrm{pGS_f}}\mathrm{pGS}_{\mathrm{f}}^{i} + \zeta_{y_{\mathrm{h}}}, \tag{44}$$

$$e_{j}^{i} = \alpha_{\mathrm{e}} + \gamma_{\mathrm{e}}^{e_{\mathrm{h}}}e_{\mathrm{h}}^{i} + \gamma_{\mathrm{e}}^{y_{\mathrm{h}}}y_{\mathrm{h}}^{i} + \gamma_{\mathrm{e}}^{\mathrm{pGS}}\mathrm{pGS}_{j}^{i} + \gamma_{\mathrm{e}}^{\mathrm{pGS_m}}\mathrm{pGS}_{\mathrm{m}}^{i} + \gamma_{\mathrm{e}}^{\mathrm{pGS_f}}\mathrm{pGS}_{\mathrm{f}}^{i} + \zeta_{\mathrm{e}}. \tag{45}$$

With this formulation we can take into account the difference in the PGSs of the DZ twins and still use the information on MZ twins (see Okbay et al. 2022 for a justification of this method). The estimate of the SEM is presented in table 3.

Education of parents and family income have a strong and significant influence on educational attainment of the twins; thus, they exert their influence though this channel in addition to the direct one of the genotype of the twins. However, the coefficients of the two parental PGSs, which could potentially report additional unobserved channels from genotype of parents to education years, are not significant, although they are of course large and significant in the equations for both family income and parents' education. This finding is consistent with the result reported in Willoughby et al. (2021): conditioning on parental IQ and socioeconomic status substantially reduces the effect of parental genotype. Within our model, this result is an implication of the identification of family income and parental education[31] as the pathways of the effect of family background.[32]

The results are similar if we introduce explicitly a latent variable F of family environment, affected by the PGS of the parents, and modify the education-years equation as

$$e_{j}^{i} = \alpha_{e} + \gamma_{e}^{e_{\mathrm{h}}}e_{\mathrm{h}}^{i} + \gamma_{e}^{y_{\mathrm{h}}}y_{\mathrm{h}}^{i} + \gamma_{e}^{F}F^{i} + \gamma_{e}^{\mathrm{pGS}}\mathrm{pGS}_{j}^{i} + \zeta_{e}\,. \tag{46}$$

The PGSs of parents significantly affect the education of parents and income of the family, and in turn the education of parents and income of the family affect education years of children, but F has little residual influence.

We find similar results if we consider different measures of educational attainment. For example, if we take the variable e_j^i to be a binary variable indicating whether the twin has a college degree or not (and estimate the correspondent of eq. [45] with a probit model), we find the coefficients of e_{h} to be 0.35 (SE = 0.05, $z = 6.98$, p-value $< .001$; marginal effect 12%);

[31] These variables were not considered by Willoughby et al. (2021).

[32] For the record, the coefficient of the score of the mother is significant at the 10% level.

TABLE 3
SEM OF PATHWAYS FROM PGS TO EDUCATION YEARS ($N = 802$)

Equation, Variable	Coefficient (SE)	z	p-Value	Confidence Interval
Education of parents:				
PGS mother	.182 (.032)	5.62	<.001	[.118, .245]
PGS father	.301 (.033)	8.96	<.001	[.235, .367]
Constant	.066 (.033)	2.00	.045	[.001, .132]
Family income:				
PGS mother	.091 (.029)	3.12	<.001	[.034, .149]
PGS father	.154 (.030)	5.05	<.001	[.094, .213]
Constant	.131 (.030)	4.28	<.001	[.070, .198]
Education years:				
Education of parents	.183 (.021)	8.76	<.001	[.142, .224]
Family Income	.112 (.023)	4.84	<.001	[.066, .157]
PGS	.103 (.032)	4.84	.002	[.038, .167]
PGS mother	.052 (.023)	2.26	.094	[−.006, .084]
PGS father	−.003 (.024)	−.13	.899	[−.051, .044]
Male	−.139 (.048)	−2.85	.004	[−.235, −.043]
Constant	.345 (.025)	13.43	<.001	[.284, .395]

NOTE.—The model estimated is described in eqq. (43)–(45). All observed variables are standardized to mean zero and SD 1. Standard errors are estimated by bootstrapping. Model vs. saturated $Pr > \chi^2 < 0.0001$.

for y_h the coefficient is 0.21 (SE = 0.057, $z = 3.76$, p-value < .001; marginal effect 6.7%). The estimated coefficient for the PGS of the twin is 0.16 (SE = 0.043, $z = 3.67$, p-value < .001; marginal effect 5.4%).

C. *Regression on Parents' PGS*

In this section, we see that if we regress variables of interest on the PGS of the children and we include that of the parents, we typically find the coefficient of the parents' score to be significant and positive. This finding provides evidence that the genes of parents affect the success of children, in addition to the direct effect on the genes of the children. After we control for education of parents and family income, the coefficient of the parents' PGS is insignificant, while the coefficient of the PGS of the twin stays significant. This second finding suggests that income and education of parents channel most of the additional effect of parents' genes.

We present the results in section S-0.4 for education years (table S-6), GPA (table S-7), college (table S-8), and IQ (table S-9). These results are consistent with earlier findings of passive rGE[33] but add insight into the mechanism from genetic profile of parents to children's outcomes: most of this effect is channeled by parents' education and income, with parents' education typically having the largest and most significant role.

When we control for education of parents and family income (see col. 4 in table S-6 [app. S-0.4], the coefficients of the PGS of the parents are substantially reduced and not significant. In this model, the fraction explained by the education of parents is large (coefficient is 0.116, [SE = 0.025]), and so is the case for family income (coefficient is 0.083 [SE = 0.028]). Interestingly, the coefficient of the mother's PGS shows some modest effect in columns 2 and 3, that is, even after we condition for IQ and soft skills. The same result of the decline in significance of the PGS of parents holds for other indicators of educational attainment, such as college and the GPA index, reported in tables S-7 and S-8, respectively.

In conclusion, we add two findings to the analysis in Kong et al. (2018) and Willoughby et al. (2021), where evidence of a passive rGE is reported. First, we identify, consistently with the model we developed in section II and with the more general theory of parental investment, two paths through which genetic factors of the parents operate indirectly, namely, family income and education of parents; education of parents has a larger coefficient than family income. Second, we show that once these two factors are taken into account, there is no significant residual indirect effect.[34]

VIII. Conclusions

Our analysis has been set up as a natural extension of theories of parental investment and intergenerational mobility (as in Becker and Tomes 1979 and in the large literature building on that model), but it replaces the ad hoc assumption of an asexual AR(1) process with a fully specified formulation of genetic transmission of skills from a pair of parents in a stable, nonrandom mating equilibrium. Our model provides the basis for an economic analysis of genetic factors in education and intergenerational mobility; it is more realistic than the existing models, and it is still analytically manageable, so that it can be tested in the data. Our data analysis provides a proof of concept of this statement.

[33] See Kong et al. (2018); see also the analysis in Willoughby et al. (2021).

[34] Within the model defined precisely here, there is little evidence of genetic nurture, as defined in recent literature (see, for an in depth discussion, Wang et al. 2021 and Okbay et al. 2022).

Realism of the assumptions would matter little, perhaps, if the predictions of the alternative models were similar. We have shown instead that the predictions of our model of intergenerational mobility differ substantially from those of the standard model. Most notably, there is no constant heritability coefficient as in the standard model; instead, heritability is determined endogenously and depends on the probability distribution of the genotype and on the features of the assortative mating, hence ultimately on the mating preferences of the agents. We have concluded in our analysis that the standard model is likely to underestimate the intergenerational elasticity of income. Our model also allows a precise test of important features affecting intergenerational mobility, such as assortative mating and passive gene-environment correlation, which is the effect of genes of parents operating (over and above the direct effect on genes) through the environment provided by parents to children. If we want to analyze precisely the relative weights of nature and nurture, an issue that is crucial for a variety of public policies, economic theory will have to adopt models that incorporate this information explicitly. The difference between standard and fully specified genetic models will become even more consequential as more precise estimates of the link between genes and phenotypes of economic interest, as well as richer information on the genetic profile of individuals, become available.

In our empirical analysis, we confirm earlier results that genetic factors measured by the PGS have a large effect on educational achievement, for example, raising the fraction achieving college from about 20% in the low decile of the score to about 60% in the top decile. Very different pathways of the effect of PGS could be consistent with this finding: for example, the effect might be entirely due to discrimination operating on individual characteristics that are genetically based but irrelevant for the technology of educational achievement. These discrimination effects are less likely for components that operate through intelligence and personality; any fraction of the explanatory power of the PGS that can be attributed to the mediation of these individual characteristics is less likely to operate through discrimination. Regression analysis shows that the pathways occur in significant part through intelligence and personality and that the size of the effect of intelligence is stronger overall.

Our data include information on the genetic profile of the parents, so we can test directly the size and significance of the effect of the genotype of parents on the environment of children (passive rGE). Our analysis decomposes this effect into two different paths: one operates through genes that directly affect educational attainment of the parents but influence the environment of the children indirectly through the effects on income and education. This first is the path that economists have analyzed with models of parental investment. A second path operates through genes that affect directly the environment of the children without affecting educational

attainment of the parents and thus their income and education. Our analysis of the data suggests that most passive rGE operates though the first channel; within this channel, education matters more than income.

Fixed-effects analysis on DZ twins is performed, exploiting the information we have on the genotype, summarized by the PGS, which is identical for MZ twins and differs for DZ twins, in a measure that depends on chance and the degree of assortative mating between the child's parents. Our results shows a significant effect of PGS on a measure of academic performance at school (the GPA score) and on intelligence, as well as in educational achievement, in particular college degree. This final result provides an important support for our conclusion, since DZ twins share very similar environments in their formative years but are significantly different in genotype, in spite of assortative mating. The analysis of the pathways operating from genes associated with educational attainment though cognitive and noncognitive skills shows that the largest effect is through cognitive skills.[35]

IX. Description of the Data

Individuals in the sample we use here are twin participants in the Minnesota Twin Family Study (MTFS; Disney et al. 1999; Iacono et al. 1999), which includes two cohorts of twins, one assessed initially at a target age of 11 ($N = 1{,}512$) and a second assessed initially at a target age of 17 ($N = 1{,}252$), and subsequent follow-up assessments undertaken at target ages of 20, 24, and 29 for the older cohort and 14, 17, 20, 24, and 29 for the younger cohort. The participation rates in the follow-ups of MTFS have generally been above 90% (see McGue, Irons, and Iacono 2014).

A. *Measures of Income*

Data on income of parents and twins were collected at different points in time. The age of parents at the moment at which the data on income were collected is higher than the age of the children by approximately 10 years. We control for this difference in the estimation (see the discussion preceding table 1). The measure of parents' income was collected on a 13-point, self-reported scale that ranged from less than $10,000 to over $80,000.[36]

[35] This conclusion is different from the one reached in McGue, Rustichini, and Iacono (2017) and McGue et al. (2020), using the same data. The reason for the discrepancy is ex post clear. Neither of these papers sets up the analysis as a test of a fully specified model of parental investment, and both ignore key variables in the analysis, such as household income.

[36] The precise bands were (1) less than $10,000, (2) $10,001–$15,000, (3) $15,001–$20,000, (4) $20,001–$25,000, (5) $25,001–$30,000, (6) $30,001–$35,000, (7) $35,001–$40,000, (8) $40,001–$45,000, (9) $45,001–$50,000, (10) $50,001–$60,000, (11) $60,001–$70,000, (12) $70,001–$80,000, and (13) more than $80,000.

A first assessment of the income of the twins was collected at the age-29 assessment and was the answer to the question What is your annual income before taxes (in thousands of dollars? No specific band of income was suggested. In the analysis, the data on income are translated into dollar amounts, then log transformed and standardized.

B. Measures of Human Capital

Information on educational achievement in the sample is provided by a classification of the individual into one of seven classes, described in table 4. Data on academic performance of the twins in school were collected in a dedicated academic history interview, given to both mother and child. Four scores were calculated: GPA, behavior problems, academic problems, and academic motivation.

The GPA score used here is a GPA-like index, not the actual GPA. Five questions in the academic history survey asked separately both the mother and the child about grades the child was getting in school. The questions provided a 5-point letter scale, from A to F, for the answer. The questions asked about grades in (*a*) reading/English, (*b*) arithmetic/math, (*c*) science, (*d*) social studies/history, and (*e*) overall. The GPA score was then calculated to represent an average of items *a*–*d* transformed to a four-point scale. In a validation sample (Johnson, McGue, and Iacono 2004), the correlation between reported grades and actual GPA from school transcripts exceeded 0.8.

C. Explanatory Variables

A specific strength of our data is the availability of information on variables that are natural candidates to provide an explanation of the way in which the genetic profile of individuals, summarized by the PGS, can affect educational achievement. We describe these data here.

TABLE 4
EDUCATION YEARS VARIABLE

Education Level	Class	Years
Less than high school	1	10
GED	1	11
High school	2	13
High school + vocation	3	14
Community college	3	15
College	4	19
Professional degree	5	22

NOTE.—The variable "Class" is a coarser classification used in the analysis.

Computation of PGSs.—We constructed the PGSs predicting years of education from the summary statistics released by Lee et al. (2018), with the cohorts 23andMe and MCTFR (Minnesota Center for Twin and Family Research) removed. The weights of the SNPs in the score were then calculated with the software tool LDpred (Vilhjálmsson et al. 2015), which uses an external sample to estimate the correlations between SNPs in order to convert the univariate regression coefficients in GWAS summary statistics to partial regression coefficients. We used the data in MCTFR for parents of European ancestry to estimate the correlations between SNPs and calculated the partial regression coefficients of the 450,000 SNPs that were originally genotyped in MCTFR and survived all default software filters. We set the LDpred shrinkage parameter equal to unity—the highest possible value and the one leading to the least shrinkage of the PGS weights. This choice, sometimes regarded as the most conservative, was followed by Lee et al. (2018). Our experience has shown that varying this parameter over a tenfold range scarcely influences the prediction R^2 (e.g., Willoughby et al. 2021).

Cognitive ability.—Cognitive ability was assessed at intake for both MTFS cohorts by means of four subtests from the age-appropriate Wechsler Intelligence Scale. Twins in the younger cohort were assessed with the Wechsler Intelligence Scale for Children–Revised (WISC-R), and twins in the older cohort were assessed with the Wechsler Adult Intelligence Scale–Revised (WAIS-R). The short forms consisted of two performance subtests (block design and picture arrangement) and two verbal subtests (information and vocabulary), and the scaled scores from these subtests were prorated to determine overall IQ. IQ from this short form has been shown to correlate ($r = 0.94$) with IQ from the complete test (Sattler 1974).

Noncognitive skills: personality measures.—Six measures of noncognitive skills derived from the age-17 assessment of both cohorts were used. First, we used three higher-order scales from the Multidimensional Personality Questionnaire (MPQ; Tellegen and Waller 2008). The MPQ has 11 primary trait scales (absorption, well-being, social potency, achievement, social closeness, stress reaction, aggression, alienation, control, harm avoidance, and traditionalism). Each is assessed with 18 self-reported items. The three higher-order MPQ scales (positive emotionality of affectivity [PA, associated with well-being, social potency, achievement, and social closeness], negative emotionality or affectivity [NA, associated with stress reaction, alienation, and aggression], and constraint (CN, associated with control, harm avoidance, and traditionalism]) are computed as linear functions of the 11 primary scales.[37]

[37] For details, see https://www.upress.umn.edu/test-division/mpq/copy_of_mpq_BF-overview.

High constraint is associated with tendencies to inhibit and constrain impulsive as well as risk-taking behavior. Individuals with higher NA scores are more prone to experience anxiety, anger, and, in general, negative engagement. Positive emotionality is associated with search for rewarding behavior and experience, while low PA may be associated with loss of interest, depressive engagement, and fatigue. In our sample, the three higher-order dimensions, as well as IQ, are approximately normally distributed.

Additional noncognitive skills.—Three additional measures of soft skills were derived from answers to questionnaires.

1. *Externalizing* was the total number of *DSM-IV* (*Diagnostic and Statistical Manual of Mental Disorders*, fourth edition) symptoms of oppositional defiant disorder, conduct disorder, and adult antisocial behavior (i.e., the adult symptoms used in diagnosing antisocial personality disorder) obtained by interviewing the twin with the Diagnostic Interview for Children and Adolescents (DICA-R; Welner et al. 1987; Reich 2000) and the Structured Clinical Interview for *DSM-III-R* (SCID; Spitzer et al. 1992). The interviews were modified to ensure complete coverage of *DSM-IV*, and symptoms were reported over the lifetime of the adolescent. In the analysis reported here, the "externalizing" scale was log-transformed (after adding 1) to minimize positive skew.
2. The *academic effort* scale consisted of eight items answered by the twins' mother on a four-point scale (definitely false, probably false, probably true, definitely true). Items on this scale (with $\alpha = 0.91$)[38] cover academic effort (e.g., "Turns in homework on time") and motivation ("Wants to earn good grades").
3. Finally, the *academic problems* scale consisted of three items ($\alpha = 0.77$) answered on the same four-point format by the mother and covering behavioral problems in a school setting (e.g., "Easily distracted in class").

Family background.—Three indicators of family background assessed at intake were analyzed here. First, parent occupational status was based on mothers' and fathers' reports and coded on the Hollingshead scale (Hollingshead 1957). We inverted the seven-point Hollingshead scale so that higher scores represented higher occupational status. Individuals were coded as missing if they did not work full-time, were disabled or institutionalized, or reported their occupation as homemaker. The occupation status of the home was taken as the maximum of the two parent

[38] Cronbach's (1951) alpha is a good lower bound on the reliability when the scale measures only one common factor.

reports. Parent college was the number of parents having completed a 4-year college degree.

Data Availability

The data and codes necessary to replicate the empirical results in the paper are available in Rustichini et al. (2023), in the Harvard Dataverse: https://doi.org/10.7910/DVN/OYHSJL. The folder includes the Stata code (Stata17) and data file (dta format) to reproduce the tables in the paper.

Appendix

Proofs and Additional Material

A. Preferences and Stable Matchings

We assume a preference order over matchings; consistent with our assumption on matchings, the order is defined on the observable vector z_O of each of the two mates. It is also monotonic in the $\Theta \times Y$ component and is homophilic in the C component. More precisely, recall that $\Theta \equiv \times_{l=1}^{n}\Theta_k$: each component has a natural order (such as "taller," "more intelligent," "lower Neuroticism score," and so on), and Y has the natural order over the real numbers, so Θ and $\Theta \times Y$ have an induced partial order. An *individual* in the marriage market is a type $z_O \in \Theta \times Y \times C$. Preferences over mates of the individual z_O of sex $s \in \{m, f\}$ (recall that "m" is for mother, assumed to be female) are represented by a weak order $\succcurlyeq_{z_O}$ that is monotonic,

$$\forall\ z''_M, z'_M : z''_M \geq z'_M \text{ implies } \forall\ c \in C, (z''_M, c) \succcurlyeq_{z_O} (z'_M, c), \tag{A1}$$

and homophilic,

$$\forall\ z_M, c, e, f : d(f, c) \leq d(e, c) \text{ implies } \forall\ z'_M (z'_M, f) \succcurlyeq_{(z_M, c),} (z'_M, e). \tag{A2}$$

The household maximization problem described in equations (6)–(10), which depends only on the $\Theta \times Y$ components, defines a preference over matches. In the maximization problem, an individual (θ_m, y_m) evaluates the utility $U(\theta_m, y_m, \theta_f, y_f)$ from a match with an individual (θ_f, y_f) anticipating the household income and the skill of the two children; so her preferences (if the preferences are completely described by the household maximization problem) are represented by $U(\theta_m, y_m, \cdot)$. The same holds for the "f" potential spouse. We assume that household log income y^h is a linear combination of the income of the two spouses, with weights w^y adding to 1, and that the expected (by the parents) skill of each child θ^c is a linear combination of the skills of the parents with weight w_i^θ, $i \in \{m, f\}$ also adding to 1. In summary, we assume

$$y^h = w_m^y y_m + w_f^y y_f \tag{A3}$$

and

$$\theta^c = w_{\mathrm{m}}^{\theta}\theta_{\mathrm{m}} + w_{\mathrm{f}}^{\theta}\theta_{\mathrm{f}}; \tag{A4}$$

Substituting the optimal investment (eq. [11]) into the budget constraint, education, and income equations ([7], [8], and [9], respectively), we find that, up to a constant independent of θ and y, the *worth* in the marriage market of a type (θ, y) of sex $i \in \{\mathrm{m}, \mathrm{f}\}$ is

$$W_i(\theta, y) \equiv (1 - \delta + 2\delta\alpha_{\mathrm{Ih}})w_i^{y}y + 2\delta\alpha_{\theta\mathrm{h}}w_i^{\theta}\theta, \tag{A5}$$

and the utility of a household is the sum of the worth of the spouses,

$$U(\theta_{\mathrm{m}}, y_{\mathrm{m}}, \theta_{\mathrm{f}}, y_{\mathrm{f}}) = W_{\mathrm{m}}(\theta_{\mathrm{m}}, y_{\mathrm{m}}) + W_{\mathrm{f}}(\theta_{\mathrm{f}}, y_{\mathrm{f}}), \tag{A6}$$

so the household utility from the household maximization problem is linear and monotonically increasing in the parents' types and income, and hence the overall utility is (if we assume that any additional components are monotonically increasing) monotonically increasing.

A *stable matching* is defined as usual: a matching that cannot be blocked by individuals or pairs of mates.[39] By the properties we have derived, we conclude, using standard arguments:

PROPOSITION A1. A stable matching exists. There is complete segregation over C. Parents' genotypes (the random variables g_{m} and g_{f}) are conditionally independent for any vector of observable characteristics.

B. *Proof of Inequality (34)*

The inequality follows because when parents match on income and only on income the system (28)–(30) is as follows. Equation (28) becomes

$$\mathbf{V}(\theta) = \frac{\sigma_{\epsilon^{\theta}}^2 + (\eta^2/2)\mathbf{E}(\theta_{\mathrm{m}}, \theta_{\mathrm{f}})}{1 - (\eta^2/2)}. \tag{A7}$$

Equations (29) and (30) are unchanged. Rearranging one obtains the inequality (34). QED

C. *Proof of Lemma 4.1*

We denote PGS_i the PGS of twin i and $\mathrm{PGS}_{\mathrm{m}}$ and $\mathrm{PGS}_{\mathrm{f}}$, as indicated by the subscript, those of the mother and father, respectively. We use similar notation for g, the genotype of various individuals. The proof uses the fact that the genotype of the child is (after meiotic recombination) the sum of one haplotype of the mother and one of the father, each chosen with equal probability. Recall that we are considering an additive model, as stated in equation (2). Given these premises, we have

[39] More precisely, a matching ν is stable if and only if for all—except possibly a zero-measure set (with respect to the product measure $\nu \otimes \nu$)—pairs $(z_{\mathrm{m}}, z_{\mathrm{f}}, z'_{\mathrm{m}}, z'_{\mathrm{f}})$,

$$z_{\mathrm{f}} \succ_{z_{\mathrm{m}}} z'_{\mathrm{f}} \text{ or } z'_{\mathrm{m}} \succ_{z'_{\mathrm{f}}} z_{\mathrm{m}} \text{ or } (z_{\mathrm{f}} \succcurlyeq_{z_{\mathrm{m}}} z'_{\mathrm{f}} \text{ and } z'_{\mathrm{m}} \succcurlyeq_{z'_{\mathrm{f}}} z_{\mathrm{m}}).$$

$$\mathbf{E}(\mathrm{PGS}_i|g_m, g_f) = \frac{\mathrm{PGS}_m + \mathrm{PGS}_f}{2}. \tag{A8}$$

We then have

$$\begin{aligned}
\mathbf{E}(\mathrm{PGS}_1\mathrm{PGS}_2) &= \mathbf{E}(\mathbf{E}(\mathrm{PGS}_1\mathrm{PGS}_2)|g_m, g_f) \\
&= \mathbf{E}(\mathbf{E}(\mathrm{PGS}_1|g_m, g_f)\mathbf{E}(\mathrm{PGS}_2|g_m, g_f)) \\
&= \mathbf{E}\left(\mathbf{E}\left(\frac{1}{2}(\mathrm{PGS}_m + \mathrm{PGS}_f)\middle|g_m, g_f\right)\mathbf{E}\left(\frac{1}{2}(\mathrm{PGS}_m + \mathrm{PGS}_f)\middle|g_m, g_f\right)\right) \\
&= \frac{1}{2}\mathbf{E}(\mathbf{E}((\mathrm{PGS}_m)^2 + \mathrm{PGS}_m\mathrm{PGS}_f)|g_m, g_f) \\
&= \frac{1}{2} + \frac{1}{2}\mathbf{E}(\mathrm{PGS}_m\mathrm{PGS}_f),
\end{aligned}$$

where the first equality follows from elementary property of expectation, the second from the conditional independence of PGS with respect to parents' genotype, the third from additivity of PGS of each offspring (eq. [A8]), the fourth from symmetry between PGS_m and PGS_f, and the fifth again from elementary properties of expectation. QED

D. *Proof of Theorem 3.1*

We recall the equations describing the process on income and genetic profile, simplifying the notation for clarity in exposition.

We write the income equation in the compact form:

$$y_c = \beta C(y_m, y_f) + w(g_c) + \sigma Z, \tag{A9}$$

where $\beta < 1$, $\sigma > 0$, Z is a standard normal, and C is a composition map giving household income as a function of the income of the two parents. We assume that C is continuous and satisfies

$$\min\{y_m, y_f\} \leq C(y_m, y_f) \leq \max\{y_m, y_f\}; C(y, y) = y. \tag{A10}$$

We call the value of this composition *household income* and denote it y_h. This form includes the special cases in which the household income is the average of the two parents' income, possibly with different weights.

The genotype of the child, given the pair of parents' genotypes (g_m, g_f), is a random variable with distribution, conditional on (g_m, g_f),

$$H(\cdot|g_m, g_f) \in \Delta(G). \tag{A11}$$

We can now be more precise. We first assign to μ its disintegration according to the partition, $W^{-1}(\mathcal{V})$, that is, the vector of pairs of probability of the class v_i and the conditional probability, given v_i,

$$((\mu_{\mathcal{V}}(v_i), \mu(\cdot|v_i)) : i \in Z). \tag{A12}$$

By Rohlin's (1952) theorem, such a disintegration exists, and in addition, (i) $\mu_{\mathcal{V}}$ is a probability measure on $\mathcal{V}$, (*ii*) for every i, $\mu(\cdot|v_i)$ is a probability measure on $G \times Y$ that satisfies $\mu(C(v_i|v_i)) = 1$, and (iii) $\mu(\cdot) = \Sigma_{i\in Z}\mu_{\mathcal{V}}(v_i)\mu(\cdot|v_i)$.

We now describe the function giving the next-period measure, examining each component of this object separately. First, there is a Markov kernel assigning to a parents' profile (g_m, y_m, g_f, y_f) a probability on $G \times Y$, interpreted as the child's genotype and household income, assigning to $O_G \times O_Y \in \mathcal{B}(G) \times \mathcal{B}(Y)$

$$K_F((g_m, y_m, g_f, y_f), O_G \times O_Y) = H(O_G|g_m, g_f)\delta_{C(y_m,y_f)}(O_Y). \tag{A13}$$

Next is the Markov kernel assigning to a pair (g_c, y_h) of child's genotype and household income a distribution on child's income, as by equation (A9), assigning to a Borel subset of Y, O_Y

$$K_I((g_c, y_h), O_Y) \equiv \Phi\left[\frac{1}{\sigma}(O_Y - \beta y_h - w(g_c))\right], \tag{A14}$$

where Φ is the measure induced by the standard normal. We define the function $\Psi : \Delta(G \times Y) \to \Delta(G \times Y)$ as

$$\Psi(\rho) \equiv (\rho \otimes \rho)K_F K_I, \tag{A15}$$

where $\rho \otimes \rho$ is the independent product of the two measures, and $K_F K_I$ is the composition of the two kernels.

LEMMA A2. The map Ψ is continuous in the weak topology.

Proof. The map $\rho \to \rho \otimes \rho$ is continuous (see lemma 1.1 of Parthasarathy 1967, chap. 3). The rest follows from the continuity assumption on the combination function C and the fact that the topology on G is discrete; thus, H is continuous. QED

The next-period measure is defined by the function

$$T : \Delta(G \times Y, \mathcal{B}(G \times Y)) \to \Delta(G \times Y, \mathcal{B}(G \times Y)),$$

where for every set $O \in \mathcal{B}(G \times Y))$,

$$(T\mu)(O) \equiv \sum_{i \in \mathbb{N}} \mu_V(v_i)\Psi(\mu(\cdot|v_i))(O). \tag{A16}$$

As is standard in economics, we study the distribution on population characteristics (genotype and income), considering the invariant distributions.

D1. Invariant Measures

This section illustrates a reason why the model with a fully specified genetic transmission is different from the standard model.

The following invariance property is true for any function T' (including the function T we defined above) on $\Delta(G \times Y)$ that has two basic properties. The first is the *mating property*: the mating process operates though a mating function $M : (\Delta(G \times Y))^2 \to \Delta((G \times Y)^2)$ that preserves marginals. In the case of T defined in equation (A16), the mating function is

$$dM(\mu, \mu) = \sum_{v_i} \mu_V(v_i)(\mu(\cdot|v_i) \otimes \mu(\cdot|v_i)).$$

The second is the *factor property*: the distribution of child's genotype and income factor through the H function in equation (A11) and a kernel

$S:(G\times Y)^2\to\Delta(Y)$ denoted $S(\cdot;(g_m,y_m,g_f,y_f))$ (in the case of T, this is the Markov kernel K_rK_l).

LEMMA A3. The set of measures with the same minor-allele frequency is invariant under any T' that satisfies the mating and factor properties.

Proof. Let $C:G\to\{0,0.5,1\}^K$, defined by $C(g,k)\equiv g(k)/2$, and the push-forward mapping $\mu\in\Delta(G)$ to $C_*\mu$; the expectation with respect to $C_*\mu$ at k gives the frequency of the allele at locus k as

$$\mathrm{AF}(k)=\int_G d\mu_G(g)C(g,k).$$

Then, denoting $X\equiv(G\times Y)^2$, with generic element $x\equiv(g_m,y_m,g_f,y_f)$, the next-period allele frequency at k is

$$\begin{aligned}
&\int_{G\times Y}\int_X dM(\mu,\mu)(x)H(g_c;g_m,g_f)S(dy_c;x)C(g_c,k)=\\
&\int_G\int_X dM(\mu,\mu)(x)H(g_c;g_m,g_f)\int_Y S(dy_c;x)C(g_c,k)=\\
&\int_G\int_X dM(\mu,\mu)(x)H(g_c;g_m,g_f)C(g_c,k)=\\
&\int_G\int_{G^2} dM(\mu,\mu)_{G^2}(g_m,g_f)H(g_c;g_m,g_f)C(g_c,k)=\\
&\int_G d\mu_G(g)C(g,k),
\end{aligned}$$

where the first equality follows from Fubini's theorem; for the second, we have used the obvious fact that, for all x,

$$\int_Y S(dy_c;x)=1;$$

for the third, we have defined, for $O\in\mathcal{B}(G^2)$,

$$M(\mu,\mu)_{G^2}(O)=M(\mu,\mu)(O\times Y^2);$$

and the last follows from the basic properties of the function H. QED

The following proposition examines a case that is uninteresting from a substantial point of view (because it excludes heterogeneity) but is very useful for illustration of the differences between our model and the standard model of parental skill transmission. Let us define the set of genotypes that are homozygotes at all loci:

$$\mathrm{Hom}\equiv\{g\in G:\forall\ k,g(k)\in\{0,2\}\},\tag{A17}$$

a set of 2^K elements. If the marginal of the initial measure is concentrated on a single element in Hom, then all the iterates have the same property.

PROPOSITION A4. The map T has at least 2^K fixed points.

Proof. Take the initial measure to be concentrated on a single genotype $g\in\mathrm{Hom}$. We consider for illustration the case in which the partition is fine. In the general case, the result follows as a corollary of our results below. With the fine partition mating takes place among individuals with the same income

and genotype. There is no dynamics involving G, so there is a unique invariant measure, distributed as $N(w(g)/(1-\beta), \sigma^2/(1-\beta^2))$. QED

Note that the dynamic is entirely in the set $\Delta(Y)$; restricted to this set, the iterates of T are weakly asymptotically stable. Of course, the initial condition is not, in the interesting case, concentrated on an element in Hom.

D2. Estimates of T

As we mentioned in the main text, the specific difficulty in analyzing T derives from the fact that, because of the product of measures in the definition of Ψ, T is not linear. Thus, standard theorems on existence of invariant measures, such as the Krylov-Bogoliuvov, which is based on averaging, are not available.

To address this difficulty, we first endow $G \times Y$ with a partial order. We say that $g' \succcurlyeq_G g$ if $w(g') \geq w(g)$, and we define the partial order on $G \times Y$, denoted $\succcurlyeq$, as the one induced by the $\succcurlyeq_G$ and the natural order over the real numbers. The order $\succcurlyeq$ allows us to define the set of increasing functions on $G \times Y$ as

$$\mathcal{I} \equiv \{f : G \times Y \to \mathbb{R}, (g', y') \succcurlyeq (g, y) \Rightarrow f(g', y') \geq f(g, y)\}. \tag{A18}$$

In turn, we can now define the first-order stochastic dominance order on probability measures on $G \times Y$ as the stochastic order induced by the cone $\mathcal{I}$.

We can now construct our estimates of the function T. In simple terms, the idea is to construct a function that is defined by the same process on income and genotype as T is but gives the best possible income and the best possible genotype to the child. This will give us a control from above, and a similar procedure will give the control from below. Since the construction for the lower bound is completely symmetric to that of the upper bound, we develop in detail only the first.

Our control from above will operate on the subset of measures that have support on the best possible genotype, which we now define. We let g^* and g_* be any choice of g providing the maximum and minimum values, respectively, of the function w on the finite set G, arbitrarily selecting one of the optimal values if necessary; that is,

$$\forall\ g \in G : w(g^*) \geq w(g) \geq w(g_*).$$

We refer to g^* (g_*) as the selected best (worst) genotype. The first step is to define the largest class to which an income can belong, for some genotype:

$$V(y) \equiv \max\{v_i : G \times \{y\} \cap C(v_i) \neq \varnothing\}. \tag{A19}$$

Conditions (15) and (16) insure that the function V is well defined, that is, that the supremum is finite and it is achieved. Also note that by definitions (16) and (17), $V(y) = W(g^*, y)$; definition (A19) is more convenient for future use. Next, we define the supremum over the incomes in a class:

$$\bar{Y}(v_i) \equiv \sup\{y : G \times \{y\} \cap C(v_i) \neq \varnothing\}. \tag{A20}$$

Note that $W(g_*, \bar{Y}(v_i)) = v_{i+1}$.

Lemma A5. The function V is piecewise constant, increasing, and right-continuous. The function $y \to \bar{Y}(V(y))$

1. is piecewise constant, increasing, and right-continuous; and
2. is such that, for all $y \in Y$ such that $V(y) = v_i$,

$$\frac{w(g^*) - w(g_*)}{w_y} \leq \bar{Y}(V(y)) - y \leq \frac{v_{i+1} - v_i + w(g^*) - w(g_*)}{w_y} \equiv y_{\bar{Q}}.$$

Proof. Let $B^* : \mathcal{V} \to Y$ be defined by

$$B^*(v_i) = \frac{v_i - w(g^*)}{w_y}.$$

Note that

$$(\{g^*\} \times Y) \cap C(v_i) = \{g^*\} \times [B^*(v_i), B^*(v_{i+1})).$$

The function V is constant and equal to v_i on the interval $[B^*(v_i), B^*(v_{i+1}))$; hence, the statement concerning V follows. The function $\bar{Y}(V(\cdot))$ inherits the properties of V and so is piecewise constant and right continuous. The function $y \to \bar{Y}(V(y))$ has, on the interval $[B^*(v_i), B^*(v_{i+1}))$, the minimum at $B^*(v_{i+1})$ and the maximum at $B^*(v_i)$, and the values in the statement follow from simple computations, with the difference $B^*(v_{i+1}) - B^*(v_i)$ providing the additional term $(v_{i+1} - v_i)/w_y$ in the upper estimate. QED

We denote the subset of measures with full support on the selected best genotype

$$\Delta^*(G \times Y) \equiv \{\nu \in \Delta(G \times Y) : \nu(\{g^*\} \times Y) = 1\}. \tag{A21}$$

Lemma A6. For $\mu \in \Delta(G \times Y)$ and $\nu \in \Delta^*(G \times Y)$,

$$\nu \succcurlyeq \mu \text{ if and only if } \nu_Y \succcurlyeq \mu_Y.$$

Proof. If $\nu \succcurlyeq \mu$, then considering functions that are constant with respect to G proves that $\nu_Y \succcurlyeq \mu_Y$. If $\nu_Y \succcurlyeq \mu_Y$, then for any $h \in \mathcal{I}$,

$$\begin{aligned}(\nu, h) &= \int_Y d\nu(g^*, y)h(g^*, y) \\ &\geq \int_Y d\mu(g, y)h(g^*, y) \\ &\geq \int_Y d\mu(g, y)h(g, y)) \\ &= (\mu, h),\end{aligned}$$

where the first equality is the definition, the second is the hypothesis we made, the third follows because $h \in \mathcal{I}$, and the last is the definition. QED

We now introduce the function on measures that will provide the upper bound for T; it is denoted $\bar{Q}$, and we provide first a description of its definition. Take any y, and assign to both parents the income $y + y_Q$, so that $y_h = y + y_Q$, and genotype g^*. Then apply the same transition from the pair of parents' genotype and income, as we do for T. The induced function on measures $\bar{Q}$ is linear.

Definition A7. The Markov kernel $S^{\bar{Q}}$ is defined as, for any $O_Y \in \mathcal{B}(Y)$,

$$S^{\bar{Q}}(y, O_Y) \equiv \Phi\left\{\frac{1}{\sigma}\left[O_Y - \beta(y + y_{\bar{Q}}) - w(g^*)\right]\right\}.$$

The function $\bar{Q}$ from $\Delta^*(G \times Y)$ to itself is defined as

$$(\bar{Q}\nu)(\{g^*\} \times O_Y) = \int_Y d\nu(\{g^*\}, y) S^{\bar{Q}}(y, O_Y). \tag{A22}$$

Lemma A5 implies that any household income obtained by a match in the class $V(y)$ is less than $y + y_{\bar{Q}}$; since any genotype g_c obtained by that match has $w(g_c) \leq w(g^*)$, the next-period income obtained by this process dominates in first-order stochastic dominance induced by the process underlying T. Thus, for every y and μ in the order interval,

$$S^{\bar{Q}}(y, \cdot) \succcurlyeq S^T_\mu(y, \cdot),$$

where $\succcurlyeq$ is the order on operators (see the second part of definition 5.2.1 in Müller and Stoyan 2002, chap. 5; see also O'Brien 1975 and Kamae, Krengel, and O'Brien 1977).

We also recall that the sequence of iterates P^n, $n \in \mathbb{N}$, of a Markov operator on a metric space X is called weakly asymptotically stable if P has a unique invariant distribution μ^* and

$$\forall\ \mu \in \Delta(X, \mathcal{B}) : P^n \text{ converges weakly to } \mu^*.$$

Lemma A8. The function $\bar{Q}$ has a unique fixed point, $\bar{\nu}^\infty$, given by

$$\bar{\nu}^\infty(\{g^*\}, \cdot) \sim N\left(\frac{\beta y_{\bar{Q}} + w(g^*)}{1 - \beta}, \frac{\sigma^2}{1 - \beta^2}\right). \tag{A23}$$

The sequence of its iterates is weakly asymptotically stable.

Proof. Take the moment-generating function of the nth iterate of the function defining the next-period income random variable,

$$y' = \beta(y + y_{\bar{Q}}) + w(g^*) + \sigma Z,$$

and consider the limit. QED

To allow the comparison between T and $\bar{Q}$, we represent the action of T in a form similar to equation (A22) for $\bar{Q}$. Since T is not linear, the Markov kernel corresponding to $S^{\bar{Q}}$ must depend on the current measure and is written $S^T_\mu(y, O_Y)$ as the probability of a Borel set O_Y at the point y and population measure μ.

We first provide an informal description of the process underlying this special Markov kernel. The income of the parent m is chosen (this is chosen according to the measure μ_Y). The genotype g_m is then chosen according to a version of the conditional measure $\mu(\cdot|y_m)$. The parent belongs to the class of worth $v_i = W(g_m, y_m)$, and a mate is chosen randomly in that class, with probability $\mu(\cdot|v_i)$. The parents' profile (g_m, y_m, g_f, y_f) gives the probability on child's pair (g_c, y_c).

The precise definition is given next:

Definition A9. For $\mu \in \Delta(G \times Y)$, $S^T_\mu : Y \to \Delta(Y, \mathcal{B}(Y))$ is defined as

$$\begin{aligned} S^T_\mu(y, O_Y) \equiv & \int_{G^2 \times (G \times Y)} d\mu(g_m|y) \sum_{v_i} \delta_{v_i}(W(g_m, y)) d\mu(g_f, y_f|v_i) \\ & \times H(g_c|g_m, g_f) \Phi\left[\frac{1}{\sigma}(O_Y - \beta C(y_m, y_f) - w(g_c))\right] \end{aligned} \tag{A24}$$

for any $O_Y \in \mathcal{B}(Y)$.

The Y-marginal of T_μ is an average of $S_\mu^T(y, \cdot)$:

LEMMA A10. For all $\mu \in \Delta(G \times Y)$ and $O_Y \in \mathcal{B}(Y)$,

$$(T\mu)(G \times O_Y) = \int_Y d\mu_Y(y) S_\mu^T(y, O_Y). \tag{A25}$$

Proof. We first observe that for any $\mu \in \Delta(G \times Y)$,

$$\begin{aligned} &\sum_{v_i} \mu_V(v_i)(\mu(\cdot|v_i) \otimes \mu(\cdot|v_i))(g_{\mathrm{m}}, y_{\mathrm{m}}, g_{\mathrm{f}}, y_{\mathrm{f}}) = \\ &\mu_Y(y_{\mathrm{m}}) d\mu(g_{\mathrm{m}}|y_{\mathrm{m}}) \sum_{v_i} d\mu(g_{\mathrm{f}}, y_{\mathrm{f}}|v_i) \delta_{v_i}(W(g_{\mathrm{m}}, y)). \end{aligned} \tag{A26}$$

Take now any real-valued bounded continuous function f on Y:

$$\begin{aligned} (T\mu, f) &= \int_{G\times Y} d(T\mu)(g, y) f(y_{\mathrm{c}}) \\ &= \int_Y f(y_{\mathrm{c}}) \int_G d(T\mu)(g, y) \\ &= \int_Y f(y_{\mathrm{c}}) \int_G \int_{(G\times Y)^2} \sum_{v_i} \mu_V(v_i)(\mu(\cdot|v_i) \otimes \mu(\cdot|v_i))(g_{\mathrm{m}}, y_{\mathrm{m}}, g_{\mathrm{f}}, y_{\mathrm{f}}) H(g_{\mathrm{c}}|g_{\mathrm{m}}, g_{\mathrm{f}}) \Pr(y_{\mathrm{c}}|y_{\mathrm{m}}, y_{\mathrm{f}}, g_{\mathrm{c}}) \\ &= \int_Y f(y_{\mathrm{c}}) \int_{(G\times Y)^2} \int_G \mu_Y(y_{\mathrm{m}})\, d\mu(g_{\mathrm{m}}|y_{\mathrm{m}}) \sum_{v_i} d\mu(g_{\mathrm{f}}, y_{\mathrm{f}})|v_i) \delta_{v_i}(W(g_{\mathrm{m}}, y)) H(g_{\mathrm{c}}|g_{\mathrm{m}}, g_{\mathrm{f}}) \Pr(y_{\mathrm{c}}|y_{\mathrm{m}}, y_{\mathrm{f}}, g_{\mathrm{c}}) \\ &= \int_Y d\mu_Y(y) \int_Y S_\mu^T(y, dy_{\mathrm{c}}) f(y_{\mathrm{c}}), \end{aligned}$$

where in the fourth equality we have used the initial observation (A26), the second follows because f depends only on y, the third is the definition of T, and the last is the definition of $S_\mu^T(y, \cdot)$. QED

We define the function $\underline{Q}$, the set $\Delta_*(G \times Y)$, the kernel $S^{\underline{Q}}$, and the measure $\underline{\nu}^\infty$, in a manner similar to $\bar{Q}$, $\Delta^*(G \times Y)$, $S^{\bar{Q}}$, and $\bar{\nu}^\infty$, respectively.

We can now define the order interval

$$[\underline{\nu}^\infty, \bar{\nu}^\infty] \equiv \{\mu : \underline{\nu}^\infty \preceq \mu \preceq \bar{\nu}^\infty\}. \tag{A27}$$

LEMMA A11. For every $\mu \in [\underline{\nu}^\infty, \bar{\nu}^\infty]$, $T\mu \in [\underline{\nu}^\infty, \bar{\nu}^\infty]$.

Proof. For μ in the order interval,

$$\begin{aligned} T\mu &\preceq \bar{Q}\mu \\ &\preceq \bar{Q}\bar{\nu}^\infty \\ &= \bar{\nu}^\infty, \end{aligned}$$

where the first relation follows from $T \preceq \bar{Q}$, the second from the monotonicity of $\bar{Q}$ (first part of definition 5.2.1 in Müller and Stoyan 2002), and the last because $\bar{\nu}^\infty$ is a fixed point of $\bar{Q}$. QED

The order interval has a key property, proved in the next lemma:

LEMMA A12. The set $[\underline{\nu}^\infty, \bar{\nu}^\infty]$ is weakly compact and is convex.

Proof. Convexity is clear. We first prove that the set is relatively compact in the weak topology. By Prohorov's theorem (Parthasarathy 1967), it suffices to show that it is uniformly tight. Let $\epsilon > 0$ be given: we claim that there exists a compact set $K \subseteq G \times Y$ such that for any μ in the set, $\mu(K) \geq 1 - \epsilon$. We will find

a set $K = G \times [-M, M]$ for some M. For such a K, $\mu(K) = \mu_Y([-M, M])$. By lemma A6, we derive that

$$\underline{\nu}_Y^{\infty} \preceq \mu_Y \preceq \bar{\nu}_Y^{\infty}.$$

Find M large enough that

$$\max\{\underline{\nu}_Y^{\infty}(-\infty, -M], \bar{\nu}_Y^{\infty}[M, +\infty)\} < \frac{\epsilon}{2},$$

so that

$$\mu_Y([-M, M]^c) < \epsilon,$$

as required.

Finally, the order interval is weakly closed (see, e.g., proposition 3 of Kamae, Krengel, and O'Brien 1977). QED

LEMMA A13. The function T on $[\underline{\nu}^{\infty}, \bar{\nu}^{\infty}]$ is continuous in the weak topology.

Proof. The measures in the set are uniformly absolutely continuous with respect to the Lebesgue measure by equation (A9); note that the variance σ is independent of the income. Recall now that a sequence μ^n converges weakly to μ if and only if

$$\lim_{n \to \infty} \mu^n(A) = \mu(A)$$

for any Borel set A whose topological boundary ∂A has μ-measure zero. Now the statement follows from the fact that for any $i \in \mathcal{Z}$,

$$\partial C(v_i) = \cup_g \left\{ \left(g, \frac{v_i - w(g)}{w_y} \right), \left(g, \frac{v_{i+1} - w(g)}{w_y} \right) \right\},$$

which is a set of finite points in $G \times Y$. QED

A simple example shows that continuity may fail when the uniform absolute continuity with respect to the Lebesgue measure fails.

EXAMPLE A14. Let $K = 1$, $G \equiv \{aa, aA, AA\}$, $w(aa) = 0$, $w(aA) = 1$, $w(AA) = 2$, and $w_y = 1$. Let $v_1 = 0$, and $\mathcal{V} \equiv \{v_1\}$. Denote $(G \times Y) \setminus C(v_1) \equiv C(v_0)$, and denote the conditioning on the set $C(v_0)$ as conditioning on v_0.

Consider the sequence in $\Delta(G \times Y)$:

$$\mu^n = \frac{1}{2}(p^n \delta_{(aa,1/n)} + (1 - p^n)\delta_{(AA,-2+(1/n))}) + \frac{1}{2}((1 - p^n)\delta_{(aa,-1/n)} + p^n \delta_{(AA,-2-(1/n))}), \tag{A28}$$

with $p^n = 2/3$ if n is even and $1/3$ when odd. If we also let

$$\mu = \frac{1}{2}\delta_{(aa,0)} + \frac{1}{2}\delta_{(AA,-2)},$$

then μ^n converges weakly to μ.

We now consider the disintegration of the measures. For any n,

$$\mu_V^n(v_1) \equiv \mu^n(C(v_1)) = \mu^n(C(v_0)) = \frac{1}{2},$$

but

$$\mu(C(v_1)) = 1 \text{ and } \mu(C(v_0)) = 0.$$

Also,

$$\mu^n(\cdot|v_1) = p^n\delta_{(aa,1/n)} + (1 - p^n)\delta_{(AA,-2+(1/n))}$$

and

$$\mu^n(\cdot|v_0) = (1 - p^n)\delta_{(aa,-1/n)} + p^n\delta_{(AA,-2-(1/n))}.$$

On the other hand, $\mu(\cdot|v_0)$ is undefined, and

$$\mu(\cdot|v_1) = \frac{1}{2}\delta_{(aa,0)} + \frac{1}{2}\delta_{(AA,-2)}.$$

Thus, the sequence of conditional expectations at a given worth oscillates with no limit, and the limit of any subsequence (when it exists) is different from the conditional value of the limit measure.

Also, the function $\mu \to \mu_V(v_i)$ is not continuous.

We can now summarize the analysis developed so far, recalling the statement of theorem 3.1:

THEOREM A15. Assume equation (22) and that the worth of an individual depends linearly on income and skill. Then, for any vector of allele frequencies:

1. an invariant measure exists, which induces that allele frequency;
2. within each worth class, alleles at each locus are in Hardy-Weinberg equilibrium;
3. within each worth class of the discrete partition, a higher income of both parents implies a lower expected PGS of the child; and
4. the allele frequencies are invariant across periods.

Proof. The first part follows, given the previous analysis, from Himmelberg's (1972) theorem.

The second part follows, applying Hardy-Weinberg's theorem to the population within the worth class and using the fact that equilibrium is reached in one period.

For the third part of the theorem, consider in the discrete partition case two families, indexed by $i = 1, 2$ with $y_j^2 > y_j^1$, $j \in \{\text{m}, \text{f}\}$, so the genotype worth, denoted w_j^i, is such that $w_j^2 < w_j^1$, $j \in \{\text{m}, \text{f}\}$. The proof is very simple when the function w is injective. For any pair (w_m, w_f),

$$\begin{aligned}
E(w(g_c)|w_\text{m}), w_\text{f}) &= \sum_k \beta(k) E(g_c(k)|w_\text{m}, w_\text{f}) \\
&= \sum_k \beta(k) E(g_c(k)|g_\text{m}, g_\text{f}) \\
&= \sum_k \beta(k) E(g_c(k)|g_\text{m}(k), g_\text{f}(k)) \\
&= \sum_k \beta(k) \frac{1}{2}(g_\text{m}(k) + g_\text{f}(k)) \\
&= \frac{1}{2}(w(g_\text{m})) + w(g_\text{f})) \\
&= \frac{1}{2}(w_\text{m} + w_\text{f}).
\end{aligned}$$

Injectivity is used in the second equality. The third equality uses the absence of linkage disequilibrium among the SNPs in the PGS. In the general case in which $w^{-1}(w_j)$ is not a singleton, it suffices to take averages. Note that the probability on the finite set $w^{-1}(w_{\mathrm{m}}) \times w^{-1}(w_{\mathrm{f}})$ is uniform.

The fourth part follows from lemma A3. QED

E. *Passive Gene-Environment Correlation*

We focus on triples of a child, mother, and father. Let $g_l^s \in \{0, 1, 2\}^K$ be the genotype of $l \in \{\mathrm{c}, \mathrm{m}, \mathrm{f}\}$, and with $s \in \{\mathrm{t}, \mathrm{nt}\}$, let $g_l^s \in \{0, 1\}$ the transmitted ($s = \mathrm{t}$) and nontransmitted parts of the genotype of l; $g_l(k)$ and $g_l^s(k)$ are the values at the kth locus. Note that

$$g_{\mathrm{c}} = g_{\mathrm{m}}^{\mathrm{t}} + g_{\mathrm{f}}^{\mathrm{t}},\ g_{\mathrm{f}} = g_{\mathrm{f}}^{\mathrm{t}} + g_{\mathrm{f}}^{\mathrm{nt}},\ \text{and}\ g_{\mathrm{m}} = g_{\mathrm{m}}^{\mathrm{t}} + g_{\mathrm{m}}^{\mathrm{nt}}. \tag{A29}$$

We take α^{A} as the 3^K-dimensional vector of true genic values of the genes as they affect directly the phenotype of interest (here, superscript A refers to the additive part in the standard ACE decomposition); α_l^{C} is the vector for the effect on the environment provided to the child by the parent of type l.

Recalling the form of the family environment variable in equation (24) and using equation (23), if we set $\Pi = 0$ to focus on the issue of interest and take the value α_θ to be part of the genic values,

$$h_j^i = \alpha^{\mathrm{A}} g_{\mathrm{c}} + \rho y^i + \alpha_{\mathrm{m}}^{\mathrm{C}} g_{\mathrm{m}} + \alpha_{\mathrm{f}}^{\mathrm{C}} g_{\mathrm{f}} + \zeta_j^{\mathrm{h},i}, \tag{A30}$$

where we have denoted, to lighten notation:

$$\rho \equiv \alpha_I + \alpha_\theta \pi, \alpha_\theta;\ \zeta_j^{\mathrm{h},i} \equiv \epsilon_j^{\theta,i} + \epsilon_j^{\mathrm{h},i}.$$

Equation (A30) clarifies the different ways in which passive gene-environment interaction occurs. The first way is described by terms of the form $\alpha_l^{\mathrm{C}} g_l$, which express the direct effect of the parents on the child's environment, through pathways that are possibly completely unrelated to the phenotype of interest (which is human capital in our case).

The second way operates through the term ρy^i, which contains implicitly terms of the form $\alpha_l^{\mathrm{A}} g_l$, relative to parents, grandparents, and so on, that affected the child's household income. Unlike the first, this pathway involves genes that are relevant for the phenotype of interest.

E1. Fully Genetic Decomposition of Income

Recall that income is in our model a linear function of human capital with coefficient α_{h}. In the following, we assume that we have rescaled the index of human capital so that $\alpha_{\mathrm{h}} = 1$. We can now express the income of an individual as the discounted series of all past genetic contributions of ancestors, plus a random, zero-mean term. To denote in a simple way the ancestors of an individual i, we use the following notation. For any $n \in \{0, 1, 2, \ldots\}$, a list of possible ancestors of depth n is an element s in the set $\{\mathrm{m}, \mathrm{f}\}^n$. For instance, $\mathrm{m}i$ is the mother of i, $\mathrm{fm}i$ is the father of the mother of i, and so on. We adopt the convention that at $n = 0$, the only element s in $\{\mathrm{m}, \mathrm{f}\}^n$ is the identity, so for such s, $si = i$, $s\mathrm{m}i = \mathrm{m}i$, and so on. We denote as $h(i)$ the family of individual i.

LEMMA A16. For every individual i,

$$y_i = \sum_{n=0}^{\infty} \left(\frac{\rho}{2}\right)^n \left(\sum_{s \in \{\mathrm{m,f}\}^n} (\alpha^{\mathrm{A}} g_{si} + \alpha^{\mathrm{C}}_{\mathrm{m}} g_{\mathrm{m}si} + \alpha^{\mathrm{C}}_{\mathrm{f}} g_{\mathrm{f}si}) + \sum_{s \in \{\mathrm{m,f}\}^n} \zeta_{si} \right). \tag{A31}$$

Proof. Using equation (A30) and recalling that $\alpha_{\mathrm{h}} = 1$, we get, for every individual i,

$$y_i = \rho y_{\mathrm{h}(i)} + \alpha^{\mathrm{A}} g_i + \alpha^{\mathrm{C}}_{\mathrm{m}} g_{\mathrm{m}i} + \alpha^{\mathrm{C}}_{\mathrm{f}} g_{\mathrm{f}i} + \zeta_i, \tag{A32}$$

where

$$y_{\mathrm{h}(i)} = \frac{1}{2}(y_{\mathrm{m}i} + y_{\mathrm{f}i}). \tag{A33}$$

Substituting equation (A33) formulated for each ancestor repeatedly into equation (A32) yields equation (A31). The series converges under our assumption that $\rho < 1$. QED

E2. Estimation

Lemma A16 has some useful implications for our estimations.

E2.1. GWAS Coefficients

The estimated GWAS coefficients $\beta(k)$ of the kth SNPs are obtained as a linear univariate regression of the h^i_j values (or, given our normalization $\alpha_{\mathrm{h}} = 1$, of y^i_j) on the $g_c(k)$ values. They are a biased estimate of the α^{C} values, for three reasons. The first reason is due to the term y^i in equation (A30), because y^i is obviously correlated with g_c, since they are both affected by the parents' and other ancestors' genotype. The second factor is the term introduced by the environmental value F, given by the parents' genotypes, again correlated with g_c. The third factor is the linkage disequilibrium (LD) correlation between different loci.

We standardize the genotype variables to have mean zero and variance equal to 1 (for the $g(k)$ variable) and 1/2 (for the $g^s(k)$ variables), obtaining the new variables $\mathrm{Sg}(k)$ and $\mathrm{Sg}^s(k)$.[40] Using the formula in lemma A16, if we ignore the LD correlation we find:

[40] That is, we call $p(k)$ the frequency of the allele with value 1, and define

$$\mathrm{Sg}(k) \equiv \frac{g(k) - 2p(k)}{\sqrt{2p(k)(1-p(k))}};\ \mathrm{Sg}^l(k) \equiv \frac{g^l(k) - p(k)}{\sqrt{2p(k)(1-p(k))}}, \quad \text{for } l = \mathrm{t, nt}. \tag{A34}$$

Of course, at Hardy-Weinberg equilibrium,

$$\mathbf{E}\,\mathrm{Sg}(k) = \mathbf{E}\,\mathrm{Sg}^l(k) = 0,\ \mathrm{Var}\,\mathrm{Sg}(k) = 1,\ \mathrm{Var}\,\mathrm{Sg}^l(k) = 1/2,\ l = \mathrm{t, nt},$$

and

$$\mathrm{Sg}(k) = \mathrm{Sg}^{\mathrm{t}}(k) + \mathrm{Sg}^{\mathrm{nt}}(k). \tag{A35}$$

LEMMA A17. For every k,

$$\mathbf{E}\,\beta(k) = \alpha^{\mathrm{A}}(k) + \frac{1}{2}(\alpha^{C}_{\mathrm{m}}(k) + \alpha^{C}_{\mathrm{f}}(k)) + \rho C, \tag{A36}$$

where C is a constant.

The term multiplied by ρ takes into account the effect occurring through grandparents and previous generations. As we have seen, ρ is between 0.2 and 0.4; thus, terms with ρ or higher order are small. Eliminating the potential bias introduced by the terms of the form α^{C} is possible using information of the genotype of parents, direct or imputed (see Kong et al. 2018 and Young et al. 2022). Complete elimination of the bias would require information on the infinite sequence of ancestors, although the complete formula shows that the effects decays exponentially; thus, effects of generations beyond parents is small.

We emphasize that, even if $\alpha^{\mathrm{C}}_{\mathrm{m}} = \alpha^{\mathrm{C}}_{\mathrm{f}} = 0$, a passive rGE effect persists through the influence on the environment of the children that genes influencing educational attainment produce on family income and parents' education. This effect may be substantial, and in our data it is. This what we consider next.

E2.2. Controlling for Parental PGS

We consider first the case with no effect of parents' genotype on environment, that is,

$$\alpha^{\mathrm{C}}_{\mathrm{m}} = \alpha^{\mathrm{C}}_{\mathrm{f}} = 0. \tag{A37}$$

In this case, substituting equation (A33) into equation (A32), we obtain

$$y_i = \alpha^{\mathrm{A}} g_i + \frac{\rho}{2}\alpha^{\mathrm{A}}(g_{\mathrm{m}i} + g_{\mathrm{f}i}) + \frac{\rho^2}{2}(y_{\mathrm{h(m}i)} + y_{\mathrm{h(f}i)}) + \zeta_i.$$

The PGS of the child is an unbiased measure of the term $\alpha^{\mathrm{A}} g_i$, and so are the parental scores for $\alpha^{\mathrm{A}} g_{si}$, $s \in \{\mathrm{m}, \mathrm{f}\}$. Since ρ is relatively small, the larger part of the effect on income is produced by terms measured by the PGS of the child and the PGSs of the parents. This is the model we estimate in section S-0.4.

When the assumption in equation (A37) does not hold, we have the bias described in equation (A36), and at the current state of knowledge one has to accept it. However, the estimates presented in section S-0.4 suggest that adding the terms modeling the environmental effect changes little of the results.

References

Abdellaoui, Abdel, Karin J. H. Verweij, and Brendan P. Zietsch. 2014. "No Evidence for Genetic Assortative Mating beyond That Due to Population Stratification." *Proc. Nat. Acad. Sci. USA* 111 (40): E4137.

Aiyagari, S. Rao, Jeremy Greenwood, and Nezih Guner. 2000. "On the State of the Union." *J.P.E.* 108 (2): 213–44.

Barcellos, Silvia H., Leandro S. Carvalho, and Patrick Turley. 2018. "Education Can Reduce Health Differences Related to Genetic Risk of Obesity." *Proc. Nat. Acad. Sci. USA* 115 (42): E9765–E9772.

Barth, Daniel, Nicholas W. Papageorge, and Kevin Thom. 2020. "Genetic Endowments and Wealth Inequality." *J.P.E.* 128 (4): 1474–522.
Becker, Gary S. 1973. "A Theory of Marriage: Part I." *J.P.E.* 81 (4): 813–47.
———. 1989. "On the Economics of the Family: Reply to a Skeptic." *A.E.R.* 79 (3): 514–18.
Becker, Gary S., and Nigel Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *J.P.E.* 87 (6): 1153–89.
———. 1986. "Human Capital and the Rise and Fall of Families." *J. Labor Econ.* 4 (3, pt. 2): S1–S39.
Becker, Joel, Casper A. P. Burik, Grant Goldman, Nancy Wang, Hariharan Jayashankar, Michael Bennett, Daniel W. Belsky, et al. 2021. "Resource Profile and User Guide of the Polygenic Index Repository." *Nature Human Behaviour* 5:1744–58.
Belsky, Daniel W., Benjamin W. Domingue, Robbee Wedow, Louise Arseneault, Jason D. Boardman, Avshalom Caspi, Dalton Conley, et al. 2018. "Genetic Analysis of Social-Class Mobility in Five Longitudinal Studies." *Proc. Nat. Acad. Sci. USA* 115 (31): E7275–E7284.
Björklund, Anders, and Markus Jäntti. 1997. "Intergenerational Income Mobility in Sweden Compared to the United States." *A.E.R.* 87 (5): 1009–18.
Björklund, Anders, Jesper Roine, and Daniel Waldenström. 2012. "Intergenerational Top Income Mobility in Sweden: Capitalist Dynasties in the Land of Equal Opportunity?" *J. Public Econ.* 96 (5–6): 474–84.
Black, Sandra E., and Paul J. Devereux. 2011. "Recent Developments in Intergenerational Mobility." In *Handbook of Labor Economics*, vol. 4B, edited by David Card and Orley Ashenfelter, 1487–541. Amsterdam: North-Holland.
Black, Sandra E., Paul J. Devereux, Petter Lundborg, and Kaveh Majlesi. 2017. "On the Origins of Risk-Taking in Financial Markets." *J. Finance* 72 (5): 2229–78.
Blanden, Jo. 2011. "Cross-Country Rankings in Intergenerational Mobility: A Comparison of Approaches from Economics and Sociology." *J. Econ. Surveys* 27 (1): 38–73.
Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables*. New York: Wiley.
Cesarini, David, and Peter M. Visscher. 2017. "Genetics and Educational Attainment." *npj Sci. Learning* 2:4.
Chiang, Colby, Alexandra J. Scott, Joe R. Davis, Emily K. Tsang, Xin Li, Yungil Kim, Tarik Hadzic, et al. 2017. "The Impact of Structural Variation on Human Gene Expression." *Nature Genetics* 49:692–99.
Cronbach, Lee J. 1951. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika* 16 (3): 297–334.
Crow, James F., and Motoo Kimura. 1970. *An Introduction to Population Genetics Theory*. New York: Harper & Row.
Ding, Weili, Steven F. Lehrer, J. Niels Rosenquist, and Janet Audrain-McGovern. 2009. "The Impact of Poor Health on Academic Performance: New Evidence Using Genetic Markers." *J. Health Econ.* 28 (3): 578–97.
Disney, Elizabeth R., Irene J. Elkins, Matt McGue, and William G. Iacono. 1999. "Effects of ADHD, Conduct Disorder, and Gender on Substance Use and Abuse in Adolescence." *American J. Psychiatry* 156 (10): 1515–21.
Domingue, Benjamin W., Jason Fletcher, Dalton Conley, and Jason D. Boardman. 2014. "Genetic and Educational Assortative Mating among US Adults." *Proc. Nat. Acad. Sci. USA* 111 (22): 7996–8000.
Dudbridge, Frank. 2013. "Power and Predictive Accuracy of Polygenic Risk Scores." *PLoS Genetics* 9 (4). https://doi.org/10.1371/journal.pgen.1003348.

Fernández, Raquel, Nezih Guner, and John Knowles. 2005. "Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality." *Q.J.E.* 120 (1): 273–344.

Fernández, Raquel, and Richard Rogerson. 2001. "Sorting and Long-Run Inequality." *Q.J.E.* 116 (4): 1305–41.

Fletcher, Jason M., and Steven F. Lehrer. 2011. "Genetic Lotteries within Families." *J. Health Econ.* 30 (4): 647–59.

Galton, Francis. 1886. "Regression towards Mediocrity in Hereditary Stature." *J. Anthropological Inst. Great Britain and Ireland* 15:246–63.

Goldberger, Arthur S. 1989. "Economic and Mechanical Models of Intergenerational Transmission." *A.E.R.* 79 (3): 504–13.

Greenwood, Jeremy, Nezih Guner, and John A. Knowles. 2003. "More on Marriage, Fertility, and the Distribution of Income." *Internat. Econ. Rev.* 44 (3): 827–62.

Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos. 2016. "Technology and the Changing Family: A Unified Model of Marriage, Divorce, Educational Attainment, and Married Female Labor-Force Participation." *American Econ. J. Macroeconomics* 8 (1): 1–41.

Heckman, James J., and Tim Kautz. 2012. "Hard Evidence on Soft Skills." *Labour Econ.* 19 (4): 451–64.

Heckman, James J., Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *A.E.R.* 103 (6): 2052–86.

Himmelberg, C. J. 1972. "Fixed Points of Compact Multifunctions." *J. Math. Analysis and Applications* 38 (1): 205–7.

Hollingshead, August. 1957. *Two Factor Index of Social Position.* New Haven, CT: privately printed.

Iacono, William G., Scott R. Carlson, Jeanette Taylor, Irene J. Elkins, and Matt McGue. 1999. "Behavioral Disinhibition and the Development of Substance-Use Disorders: Findings from the Minnesota Twin Family Study." *Development and Psychopathology* 11 (4): 869–900.

Jaffee, S. R., and T. S. Price. 2007. "Gene-Environment Correlations: A Review of the Evidence and Implications for Prevention of Mental Illness." *Molecular Psychiatry* 12:432–42.

Johnson, Wendy, Matt McGue, and William G. Iacono. 2004. "Genetic and Environmental Influences on Academic Achievement Trajectories during Adolescence." *Developmental Psychology* 42 (3): 514–32.

Kamae, T., U. Krengel, and G. L. O'Brien. 1977. "Stochastic Inequalities on Partially Ordered Spaces." *Ann. Probability* 5 (6): 899–912.

Knopik, Valerie S., Jenae M. Neiderhiser, John C. DeFries, and Robert Plomin. 2017. *Behavioral Genetics*, 7th ed. New York: Worth.

Kong, Augustine, Gudmar Thorleifsson, Michael L. Frigge, Bjarni J. Vilhjálmsson, Alexander I. Young, Thorgeir E. Thorgeirsson, Stefania Benonisdottir, et al. 2018. "The Nature of Nurture: Effects of Parental Genotypes." *Science* 359 (6374): 424–28.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. 2018. "Life Cycle Wage Growth across Countries." *J.P.E.* 126 (2): 797–849.

Lee, Chul-In, and Gary Solon. 2009. "Trends in Intergenerational Income Mobility." *Rev. Econ. and Statis.* 91 (4): 766–72.

Lee, James J., Robbee Wedow, Aysu Okbay, Edward Kong, Omeed Maghzian, Meghan Zacher, Tuan Anh Nguyen-Viet, et al. 2018. "Gene Discovery and Polygenic Prediction from a Genome-Wide Association Study of Educational Attainment in 1.1 Million Individuals." *Nature Genetics* 50:1112–21.

Loury, Glenn C. 1981. "Intergenerational Transfers and the Distribution of Earnings." *Econometrica* 49 (4): 843–67.

Mazumder, Bhashkar. 2005. "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data." *Rev. Econ. and Statis.* 87 (2): 235–55.

McGue, Matt, Dan Irons, and William Iacono. 2014. "The Adolescent Origins of Substance Use Disorders: A Behavioral Genetic Perspective." In *Genes and the Motivation to Use Substances*, edited by Scott F. Stoltenberg, 31–50. New York: Springer.

McGue, Matt, Aldo Rustichini, and William G. Iacono. 2017. "Cognitive, Noncognitive, and Family Background Contributions to College Attainment: A Behavioral Genetic Perspective." *J. Personality* 85 (1): 65–78.

McGue, Matt, Emily A. Willoughby, Aldo Rustichini, Wendy Johnson, William G. Iacono, and James J. Lee. 2020. "The Contribution of Cognitive and Noncognitive Skills to Intergenerational Social Mobility." *Psychological Sci.* 31 (7): 835–47.

Mincer, Jacob A. 1974. *Schooling, Experience, and Earnings.* New York: NBER.

Müller, Alfred, and Dietrich Stoyan. 2002. *Comparison Methods for Stochastic Models and Risks.* Chichester: Wiley.

Mulligan, Casey B. 1997. *Parental Priorities and Economic Inequality.* Chicago: Univ. Chicago Press.

———. 1999. "Galton versus the Human Capital Approach to Inheritance." *J.P.E.* 107 (S6): S184–S224.

Nagylaki, Thomas. 1992. *Introduction to Theoretical Population Genetics.* Berlin: Springer.

O'Brien, G. L. 1975. "The Comparison Method for Stochastic Processes." *Ann. Probability* 3 (1): 80–88.

Okbay, Aysu, Jonathan P. Beauchamp, Mark Alan Fontana, James J. Lee, Tune H. Pers, Cornelius A. Rietveld, Patrick Turley, et al. 2016. "Genome-Wide Association Study Identifies 74 Loci Associated with Educational Attainment." *Nature* 533:539–42.

Okbay, Aysu, Yeda Wu, Nancy Wang, Hariharan Jayashankar, Michael Bennett, Seyed Moeen Nehzati, Julia Sidorenko, et al. 2022. "Polygenic Prediction of Educational Attainment within and between Families from Genome-Wide Association Analyses in 3 Million Individuals." *Nature Genetics* 54:437–49.

Österberg, Torun. 2000. "Inter-generational Income Mobility in Sweden: What Do Tax-Data Show?" *Rev. Income and Wealth* 46 (4): 421–36.

Palomino, Juan C., Gustavo A. Marrero, and Juan G. Rodríguez. 2018. "One Size Doesn't Fit All: A Quantile Analysis of Intergenerational Income Mobility in the U.S. (1980–2010)." *J. Econ. Inequality* 16 (3): 347–67.

Parthasarathy, K. R. 1967. *Probability Measures on Metric Spaces.* New York: Academic Press.

Plomin, R., J. C. DeFries, and J. C. Loehlin. 1977. "Genotype-Environment Interaction and Correlation in the Analysis of Human Behavior." *Psychological Bull.* 84 (2): 309–22.

Reich, Wendy. 2000. "Diagnostic Interview for Children and Adolescents (DICA)." *J. American Acad. Child and Adolescent Psychiatry* 39 (1): 59–66.

Rietveld, Cornelius A., Sarah E. Medland, Jaime Derringer, Jian Yang, Tõnu Esko, Nicolas W. Martin, Harm-Jan Westra, et al. 2013. "Individuals Identifies Genetic Variants Associated with Educational Attainment." *Science* 340 (6139): 1467–71.

Robinson, Matthew R., Aaron Kleinman, Mariaelisa Graff, Anna A. E. Vinkhuyzen, David Couper, Michael B. Miller, Wouter J. Peyrot, et al. 2017. "Genetic Evidence of Assortative Mating in Humans." *Nature Human Behaviour* 1:0016.

Rogers, Alan. 1983. "Assortative Mating and the Segregation Variance." *Theoretical Population Biology* 23 (1): 110–13.

Rohlin, V. A. 1952. "On the Fundamental Ideas of Measure Theory." *American Math. Soc. Translations*, series 1, vol. 71. Providence, RI: American Math. Soc.

Rupert, Peter, and Giulio Zanella. 2015. "Revisiting Wage, Earnings, and Hours Profiles." *J. Monetary Econ.* 72:114–30.

Rustichini, Aldo, William G. Iacono, James J. Lee, and Matt McGue. 2023. Replication Data for: "Educational Attainment and Intergenerational Mobility: A Polygenic Score Analysis." Harvard Dataverse. https://doi.org/10.7910/DVN/OYHSJL.

Rustichini, Aldo, William G. Iacono, and Matt McGue. 2017. "The Contribution of Skills and Family Background to Educational Mobility." *Scandinavian J. Econ.* 119 (1): 148–77.

Sattler, Jerome M. 1974. *Assessment of Children's Intelligence.* Philadelphia: Saunders.

Scarr, Sandra, and Kathleen McCartney. 1983. "How People Make Their Own Environments: A Theory of Genotype → Environment Effects." *Child Development* 54 (2): 424–35.

Solon, Gary. 1992. "Intergenerational Income Mobility in the United States." *A.E.R.* 82 (3): 393–408.

———. 2004. "A Model of Intergenerational Mobility Variation over Time and Place." In *Generational Income Mobility in North America and Europe*, edited by Miles Corak, 38–47. Cambridge: Cambridge Univ. Press.

Spitzer, Robert L., Janet B. Williams, Miriam Gibbon, and Michael B. First. 1992. "The Structured Clinical Interview for *DSM-III-R* (SCID). I: History, Rationale, and Description." *Archives General Psychiatry* 49:624–29.

Tellegen, Auke, and Niels G. Waller. 2008. "Exploring Personality through Test Construction: Development of the Multidimensional Personality Questionnaire." In *The SAGE Handbook of Personality Theory and Assessment.* Vol. 2, *Personality Measurement and Testing*, edited by Gregory J. Boyle, Gerald Matthews, and Donald H. Saklofske, 261–92. London: Sage.

Vilhjálmsson, Bjarni J., Jian Yang, Hilary K. Finucane, Alexander Gusev, Sara Lindström, Stephan Ripke, Giulio Genovese, et al. 2015. "Modeling Linkage Disequilibrium Increases the Accuracy of Polygenic Risk Scores." *American J. Human Genetics* 97 (4): 576–92.

Wang, Biyao, Jessie R. Baldwin, Tabea Schoeler, Rosa Cheesman, Wikus Barkhuizen, Frank Dudbridge, David Bann, Tim T. Morris, and Jean-Baptiste Pingault. 2021. "Genetic Nurture Effects on Education: A Systematic Review and Meta-analysis." bioRχiv preprint. https://doi.org/10.1101/2021.01.15.426782.

Welner, Zila, Wendy Reich, Barbara Herjanic, Kenneth G. Jung, and Henry Amado. 1987. "Reliability, Validity, and Parent-Child Agreement Studies of the Diagnostic Interview for Children and Adolescents (DICA)." *J. American Acad. Child and Adolescent Psychiatry* 26 (5): 649–53.

Willoughby, Emily A., Matt McGue, William G. Iacono, Aldo Rustichini, and James J. Lee. 2021. "The Role of Parental Genotype in Predicting Offspring Years of Education: Evidence for Genetic Nurture." *Molecular Psychiatry* 26 (8): 3896–904.

Yengo, L., A. R. Wood, S. Vedantam, E. Marouli, J. Sidorenko, S. Sakaue, S. Raghavan, et al. 2020. "A Meta-analysis of Height in 4.1 Million European-Ancestry Individuals Identifies ~10,000 SNPs Accounting for Nearly All Heritability Attributable to Common Variants." Manuscript, Univ. Queensland, Brisbane.

Young, Alexander I., Seyed Moeen Nehzati, Stefania Benonisdottir, Aysu Okbay, Hariharan Jayashankar, Chanwook Lee, David Cesarini, Daniel J. Benjamin, Patrick Turley, and Augustine Kong. 2022. "Mendelian Imputation of Parental Genotypes Improves Estimates of Direct Genetic Effects." *Nature Genetics* 54:897–905.

Zimmerman, David J. 1992. "Regression toward Mediocrity in Economic Stature." *A.E.R.* 82 (3): 409–29.