Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Sibling Differences in Genetic Propensity for Education: How do Parents React?*

Anna Sanz-de-Galdeano† Anastasia Terskaya‡

March 16, 2023

Abstract

We take advantage of recent advances in genomics to revisit a classic question in economics: how do parents respond to children's endowments

*We thank the Editor (Brian A. Jacob), referee Dan Benjamin, and two anonymous referees for their feedback. We are grateful to Alejandra Abufhele, Tiziano Arduini, Jere Behrman, Pietro Biroli, Antonio Cabrales, David Cesarini, Damian Clarke, Lola Collado, Dalton Conley, Jorge García-Hombrados, Daniel Hamermesh, Aysu Okbay, Daniel Rees, Peter Savelyev, Alba Sóñora, Patrick Turley, and Alexander Young, as well to numerous seminars participants for their comments. We acknowledge 23andMe, Inc. employees and participants as the source of the data in this work.

†FAE, Universidad de Alicante, Carretera de San Vicente s/n, 03080 San Vicente - Alicante, Spain. Also affiliated with IZA. Acknowledges financial support from Project PID2021-124237NB-I00 financed by MCIN/ AEI /10.13039/501100011033/ and by FEDER, Una manera de hacer Europa; and from Project CIPROM/2021/068 financed by Consellería de Innovación, Universidades, Ciencia y Sociedad Digital. Email: anna.sanzdegaldeano@gmail.com.

‡Corresponding author. Department of Economics and Institut d'Economia de Barcelona (IEB), University of Barcelona, Carrer de John Maynard Keynes, 1, 11, 08034, Barcelona Spain. Acknowledges financial support from the Spanish Ministry of Economy and Competitiveness Grant PID2020-120589RA-I00. Email: a.terskaya@ub.edu.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

and to sibling differences in endowments? We use an index based on DNA, which is fixed at conception and assigned randomly across siblings, as a proxy for educational endowments. We find that parents of non-twins display inequality aversion: given the absolute endowment level of one child, they invest less in him/her if his/her sibling has a lower genetic predisposition to education. In contrast, we find no evidence that parents of dizygotic twins react to endowment differences between children.

JEL classification: D13, D64, J13, J24.

Keywords: Intra-household allocation of resources, educational polygenic index, parental investments, Add Health.

1 Introduction

Is the family an equalising agent? Do parents exacerbate or mitigate differences in children's endowments by reallocating resources within the family? These are crucial questions for both academics and policy makers, as parental responses should be considered when designing policies aimed at fostering human capital and reducing inequalities between children.

The literature analysing how parental investments are related to children's endowments is vast and it has been continuously expanding since the seminal contributions of Becker and Tomes (1976) and Behrman et al. (1982). Becker and Tomes (1976) proposed a model of resource allocation within the family and analysed how parental investments are affected by differences in their children's abilities or other aspects of their endowments. They show that, if the cost of increasing children's human capital is negatively related to their endowments (that is, if such cost is higher for less able children), parents may reinforce differences in children's endowments by investing more in higher-endowed children. By contrast, Behrman et al. (1982) develop a general preference model that introduces parental aversion to inequality in the distribution of their children's human capital. In their framework, the degree of parental inequality aversion plays a central role in determining whether parents will follow a compensating strategy (devoting more resources to children with lower endowments)

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

or a reinforcing strategy (devoting more resources to their better-endowed siblings).

The subsequent empirical literature has so far reached mixed conclusions on whether parents compensate or reinforce differences in their children's endowments.[1] Some studies have found evidence of parental compensatory behavior (Behrman et al., 1982; Pitt et al., 1990; Bharadwaj et al., 2018; Terskaya, 2023; Savelyev et al., 2022), while others have found that parents follow a reinforcing strategy (Behrman et al., 1994; Datar et al., 2010; Aizer and Cunha, 2012; Hsin, 2012; Frijters et al., 2013; Rosales-Rueda, 2014). Some authors have also uncovered different patterns of parental behavior across contexts or socio-demographic groups (Behrman, 1988b,a; Datar et al., 2010; Hsin, 2012; Restrepo, 2016; Abufhele et al., 2017; Bhalotra and Clarke, 2019). Interestingly, Yi et al. (2015) provide evidence that, when faced with differences in health endowments between their children, parents react by compensating in terms of health investments while, on the other hand, they reinforce inequalities through their educational investment decisions. This lack of consensus is to be expected because different authors use different measures of children's endowments and/or parental investments in different contexts, and, perhaps more importantly, this literature presents several identification challenges that are dealt with in various ways.

This article combines the traditional literature on intra-household resource allocation with recent advances in behavioral genetics to study how parents respond to adolescents' educational genetic endowments and to differences in educational genetic endowments between siblings. We extend the previous empirical literature on intra-household resource allocation in three ways.

First, we take into account that parental investment decisions depend both on parental preferences regarding inequality in the distribution of their children's human capital (Behrman et al., 1982), as well as on how costly it is for parents to increase their children's human capital —also known as the price effect (Becker and Tomes, 1976). We motivate our empirical analysis by means of a general parental preference model that incorporates both mechanisms, as previously proposed by Terskaya (2023). Importantly, evidence based on family fixed effects models —which compare

[1] See Almond and Mazumder (2013) for a review.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

parental investments across children within the family— is not fully informative on whether parents are inequality averse or not, as even inequality averse parents may follow a reinforcing strategy if the cost of investing in lower-endowed children is sufficiently higher than the cost of investing in higher-endowed children (Terskaya, 2023). To address this issue, we estimate how parental investments in one child are affected by the divergence between his/her genetic endowment and that of his/her sibling, while holding constant the child's own genetic endowment, which serves as a proxy for the costs to adding to his/her quality faced by the parents.

We pose the following questions: do parents react to adolescents' endowment levels? Do they invest more or less in children who are more or less able than their siblings but who are otherwise comparable in terms of their own endowment, and hence in the costs of investing in them? Distinguishing parental preferences for equality *versus* efficiency from the price effect is important for policy design and for assessing the effectiveness of compensatory interventions designed to help disadvantaged children because parents will reinforce or attenuate (depending on their preferences) the impact of such interventions by reallocating resources within the family.

Second, we use an educational attainment polygenic index —a summary measure of genetic variants that predicts educational attainment— as an indicator of children's educational genetic endowments.[2] Not only do these indexes explain about 12-15% of variation in educational attainment (Okbay et al., 2016; Lee et al., 2018; Becker et al., 2021; Okbay et al., 2022), but they also allow endogeneity issues to be alleviated. In particular, endowment indicators measured during childhood may be the result of prior parental (both post- and pre-natal) investments, while endowment indicators measured at birth (*e.g.*, birth weight) for singleton siblings may reflect differences in pre-natal investment decisions (Almond and Currie, 2011; Currie, 2011; Del Bono et al., 2012, among others). By contrast, genes are fixed at conception and hence cannot be the consequence of parental investment choices. Another important advantage of using an endowment measure based on DNA is that it is randomly assigned across siblings conditional on parental genotype. Our empirical model relies on this ran-

[2]In our data up to 14% of the variation in educational attainment is explained by the educational attainment polygenic index.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

domization because we control for parental polygenic indexes. Specifically, since we do not observe parental genetic variants, we conduct Mendelian imputation of parental genotypes, a technique proposed by Young et al. (2020, 2022).

Third, we focus on parental responses to differences in adolescents' educational genetic endowments rather than health endowments or shocks, while, with some exceptions, most previous studies have focused on the latter.[3] This may be due to the fact that at-birth measures of endowments (other than birth weight), which are less likely to suffer from reverse causality than indicators measured later on in life, are not readily available. Be that as it may, we know much less about parental responses to differences in their children's educational endowments than about how parents react when they face differences in their children's health endowments.

We use data from the sibling sample of the National Longitudinal Survey of Adolescent to Adult Health, a survey of youth in the 7^{th} through 12^{th} grades in the United States in 1995. We find evidence that American parents of non-twin siblings of European-ancestry display inequality aversion because, given a child's absolute level of educational genetic endowment, they invest less in him/her if his/her sibling is less-endowed, while also holding constant other characteristics, including parental genetic propensity for education. However, for parents of twins, the effect is not statistically distinguishable from zero. To explain this result, we incorporate a public good component when modelling parental investments. Bharadwaj et al. (2018) estimate that the public good component of parental investments is especially important for parents of twins, potentially because it is more difficult to invest differently across siblings who are closer together in age than among siblings who are farther apart. Our theoretical results suggest that the public good component attenuates the effect of parental inequality aversion on parental investments, which is consistent with the difference in our estimated effects for non-twins and twins.

Our study speaks to a broad and emerging literature that aims at integrating genetics and the social sciences (Beauchamp et al., 2011; Benjamin et al., 2008, 2012; Conley and Fletcher, 2017; Lehrer and Ding, 2017). For example, recent contributions

[3]See, for example, Rosales-Rueda (2014), Halla and Zweimueller (2014), Yi et al. (2015).

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

have studied the association between educational polygenic indexes or other genetic markers and human capital accumulation (Ding et al., 2006; Domingue et al., 2015; Papageorge and Thom, 2020; Ronda et al., 2020), labor market outcomes (Papageorge and Thom, 2020), and wealth at retirement (Barth et al., 2020, 2022). However, there is still much to learn regarding the mechanisms through which genetic endowments affect socioeconomic outcomes, and whether their impact is reinforced or mitigated by environmental factors in different contexts. For instance, Breinholt and Conley (2020) and Houmark et al. (2020) find that early parental investments increase with young children's genetic endowments even when accounting for parental genes.[4] We complement these studies by separately assessing the relevance for later parental investment decisions of adolescents' absolute and relative (with respect to their siblings) genetic propensity for education, while controlling for parental genes. We also contribute to studies that show that genetically influenced characteristics of close relatives can affect individuals' outcomes (Kong et al., 2018; Young et al., 2020, 2022; Houmark et al., 2020; Wertz et al., 2019).

We also take into account empirically and analytically several potential limitations of polygenic indexes. For instance, as argued in Becker et al. (2021), since the weights used to compute polygenic indexes are estimated in samples, they measure the additive genetic component of educational attainment with noise. Following Becker et al. (2021), we adjust our estimated effects for measurement error in polygenic indexes. We extend the Becker et al. (2021) measurement-error correction method for regression models that include not only individuals' polygenic indexes, but also the polygenic indexes of their relatives. We show that even the corrected estimated effects and standard errors are likely to be conservative.

The remainder of the paper is organized as follows. The next section lays out a model that guides our empirical estimation. In Section 3, we discuss our empirical strategy, focusing on how we exploit the availability of genetic data to address the empirical challenges involved in disentangling the price effect from the impact of

[4]Behrman et al. (1994) and Savelyev et al. (2022) exploit differences between allocations for identical *versus* nonidentical twins to estimate the effects of relative endowments on parental investment decisions.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

parental preferences for equality *versus* efficiency. In Section 4 we describe the Add Health data used as well as our measure of genetic predisposition to educational attainment, and we show that it correlates with several education-related indicators in our sample. Section 5 discusses our estimation results, Section 6 presents robustness checks, and Section 7 concludes.

2 Theoretical Model

We propose a theoretical framework that builds upon the classical intra-household allocation models of Becker and Tomes (1976) and Behrman et al. (1982) in order to illustrate how parental investment decisions depend on children's endowments. This framework is similar to the one presented in Terskaya (2023) to study the schooling gap between disabled and non-disabled individuals in Mexico.

We assume that parental preferences in family f can be represented by the utility function $U_{pf} = U_{pf}(c, V_{1f}, .., V_{nf})$, where c denotes parental consumption and V_{if} is the human capital of child i in family f. For simplicity of notation, in the remainder of this section we omit the family subscript, f, and we consider allocations within a particular two-child family. However, note that parental preferences might vary across families depending on the age gap between children (as discussed at the end of this section), children's sex, parental socio-economic status, institutions, etc. Such heterogeneity is ultimately an empirical question. We also assume that parental preferences are separable in consumption, which allows one to analyse the allocation of resources among children regardless of parental consumption. In particular, we specify parental preferences using a CES utility function as in Behrman et al. (1982):

$$U = \left\{ V_1^{\rho} + V_2^{\rho} \right\}^{\frac{1}{\rho}} \tag{1}$$

The main advantage of this utility form is that ρ measures the degree of parental inequality aversion across children. When $0 < \rho < 1$ parents do not dislike inequality in the distribution of human capital of their children and, instead, they care about efficiency. In this case, parents follow a "reinforcement strategy". In the extreme case of linear preferences, parents are indifferent about inequality between children, and

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

they maximize the sum of the expected human capital of their children. When $\rho < 0$, parents are more concerned about equality than efficiency, that is, they are inequality averse. A special case of inequality aversion is the Rawlsian case, in which parents are unwilling to accept any difference between children's human capital. When $\rho = 0$, parents trade-off equality and efficiency.

Following Behrman et al. (1982), we assume that a child i's human capital function has the following form:

$$V(e_i, PI_i) = e_i^{\alpha_e} PI_i^{\alpha_p} \tag{2}$$

, where e_i denotes the genetic endowments of child i. By genetic endowments we mean a set of genetic characteristics that directly affect a child's human capital through biological channels (*e.g.*, brain function, neuronal development). If human capital is determined by cognitive and non-cognitive ability, then e_i will incorporate genes that are associated with greater cognitive and non-cognitive ability. PI_i denotes parental inputs devoted to child i. Positive and diminishing returns to parental inputs require $0 < \alpha_p < 1$, and positive returns to genetic endowments imply that $\alpha_e > 0$.

With this human capital function, marginal returns to parental inputs are positively related to genetic endowments. That is, endowments and parental inputs are complements in the production of human capital. The complementarity between parental inputs and children's endowments is important in our context because it introduces a trade-off between efficiency and equality in parental investment decisions. While some may question this assumption at early ages of childhood, there is strong empirical evidence of complementarities between skills and investments at later stages of childhood (Cameron and Heckman, 2001; Cunha et al., 2006; Cunha and Heckman, 2008; Cunha et al., 2010).[5] In our analysis we focus on parental investments in adolescents, for whom skills and parental inputs are likely to be complements as in (2).[6] However, we acknowledge that our results might not be generaliz-

[5]See Heckman and Mosso (2014) for an extensive review.

[6]One may argue that parental investments in teenagers depend on endowments (or the skills stock) in adolescence rather than on genetic endowments (which are fixed at conception). Our conclusions would be the same if we used endowment

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

able to early parental investments because: (1) there might be no equality-efficiency trade-off when children are younger, (2) it may be harder for parents to observe the manifestations of children's genetic endowments at earlier ages, which may in turn attenuate their responses to such endowments, and (3) parents might be less or more inequality averse when they invest in young *versus* older children.

Finally, the parental budget constraint has the following form:

$$p_1 PI_1 + p_2 PI_2 = I \tag{3}$$

, where p_i is the cost of parental inputs for child i and I denotes total parental investments in children.[7] Furthermore, following Becker and Tomes (1976), we allow the cost of parental inputs to differ with children's initial endowments e, assuming in addition that $p_i = p(e_i)$ is not increasing in e and that $p(e_i)$ is homogeneous of degree one.

In Appendix A we solve the utility maximization problem (1) subject to (2) and (3), which yields the following expression for parental investments in child 1:

$$log(PI_1) = log(I) + G(e_1) + F\left(\frac{e_1}{e_2}\right) \tag{4}$$

, where $G(e_1) = -log(p(e_1))$, and $F(\frac{e_1}{e_2})$ is a function of the parameters of the model and $\frac{e_1}{e_2}$.

This equation shows that parental inputs in child 1 positively depend on total parental investments in children, while depending negatively on the price of parental

indicators measured in adolescence instead of genetic endowments under the assumption that differences in endowments due to differences in genetic endowments have not been completely eliminated by the time individuals reach adolescence. To our knowledge, there is no empirical evidence that differences in capability due to disparities in genetic endowments decrease throughout development stages. On the contrary, one of the best documented and most replicated findings in behavioral genetics is that the influence of genes on intelligence increases throughout development (Plomin et al., 2016; Houmark et al., 2020).

[7]Note that both the cost of parental inputs and total investments include monetary and non-pecuniary costs, such as time.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

inputs in child 1 ("the price effect"). Since p_1 may be negatively associated with child 1's endowment level, parental inputs devoted to child 1 will positively depend on his/her endowments (holding endowment differences across siblings constant). Furthermore, (4) indicates that parental inputs in child 1 depend on the relative endowments of child 1 with respect to child 2 ($\frac{e_1}{e_2}$).

Additionally, it can be shown that the following holds (see Appendix A for the derivations):

- $\frac{\partial log(PI_1|e_1)}{\partial\left(\frac{e_1}{e_2}\right)} < 0$ if and only if $\rho < 0$
- $\frac{\partial log(PI_1|e_1)}{\partial\left(\frac{e_1}{e_2}\right)} > 0$ if and only if $0 < \rho < 1$
- $\frac{\partial log(PI_1|e_1)}{\partial\left(\frac{e_1}{e_2}\right)} = 0$ if and only if $\rho = 0$

Hence, an increase in relative genetic endowments of a child with respect to his/her sibling's genetic endowments (holding constant the child's own absolute level of endowments) will decrease parental investments in this child if and only if parents are inequality averse, while it will increase parental investments in this child if and only if parents care more about efficiency than equality. The intuition for this result is that inequality averse parents will try to compensate for sibling differences in initial endowments through their investments, while parents who care more about efficiency will reinforce these differences because higher initial endowments are associated with higher returns to parental investments.

In Section A.1 of the Appendix we follow Terskaya (2023) and consider a scenario in which parents cannot completely separate the inputs devoted to each child. That is, we assume that there is a public good component of parental investments. As shown in Bharadwaj et al. (2018), this public good component is especially important when siblings are close in age and also in the case of twins. Consistent with Bharadwaj et al. (2018) and Terskaya (2023), we show that the effect of parental inequality aversion is attenuated when there are externalities in parental investments across siblings. This is because, when the public good dimension is important, parents cannot fully compensate for the differences between their children.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

3 Empirical Strategy

Our goal is to distinguish the two mechanisms considered in the theoretical model that may induce parents to follow a "reinforcing strategy" (that is, to invest more in better-endowed children than in their less well-endowed siblings) *versus* a "compensating strategy" (that is, to invest more in children with lower relative endowments). Importantly, when using sibling or twin fixed effects models to compare parental investments devoted to children with different initial endowments, one can only identify the composite impact of parental preferences regarding equality *versus* efficiency and the price effect.

3.1 Parental Preferences Regarding Equality *versus* the Price Effect

Our empirical strategy involves identifying the impact on parental investment decisions of children's relative (with respect to their siblings) educational genetic endowments, while holding children's own endowments constant (that is, by holding parental costs of increasing their children's human capital constant), as in equation (4). We consider the following empirical specification:

$$PI_{if} = \beta_0 + \beta_1(\hat{g}_{if} - \hat{g}_{jf}) + \beta_2\hat{g}_{if} + \beta_3\hat{g}_{pf} + X'_{if}\alpha + S'_{jf}\delta + F'_f\gamma + u_{if} \qquad (5)$$

, where PI_{if} is a parental investment indicator for child i in family f, $\hat{g}_{if}$ stands for a measure of child i's educational genetic endowment based on i's DNA, and $\hat{g}_{jf}$ stands for the same measure of educational genetic endowment for child j, with subscript j denoting child i's sibling. The educational genetic endowment of parents in family f are denoted by $\hat{g}_{pf}$. Individual-level characteristics of children i and j in family f are denoted by X'_{if} and S'_{jf} respectively, and F'_f denotes family-level characteristics (shared by children i and j) that may influence parental investment choices. Note that $(\hat{g}_{if} - \hat{g}_{jf})$ is our measure of child i's relative genetic educational endowment, as it is the difference between i's endowment and his/her sibling's endowment.

As we are controlling for child i's own endowment ($\hat{g}_{if}$), β_1 measures the effect of parental preferences regarding equality in the distribution of children's human capital on parental investment decisions. For any given level of child i's endowment,

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

$\beta_1 < 0$ is consistent with parental inequality aversion ($\rho < 0$), as it indicates that parents will invest less (more) in child i if he/she is higher-endowed (lower-endowed) than his/her sibling j. By contrast, $\beta_1 > 0$ is consistent with parents valuing efficiency more than equality ($0 > \rho > 1$), as it indicates that parents will invest more (less) in child i if his/her endowment is higher (lower) than that of his/her sibling j. Finally, $\beta_1 = 0$ is consistent with parents having neutral preferences regarding equality in the distribution of their children's human capital ($\rho = 0$).

As for β_2 in equation (5), this parameter is informative about the price effect. In particular, $\beta_2 > 0$ implies that, for any given level of inequality in siblings' endowments $(\hat{g}_{if} - \hat{g}_{jf})$, parents will invest more in children whose level of endowments $(\hat{g}_{if})$ is higher because the cost of investing in them is lower.[8]

3.2 Genetic Lotteries

Our empirical strategy relies on the fact that genetic information from parents is randomly assigned to children, a phenomenon that is sometimes referred to as genetic lotteries (Fletcher and Lehrer, 2011). This implies that, conditional on parental genetic endowments, the genetic endowment of children is random. Given that we control for parental genetic endowments in model (5), our identification relies on random variation of $\hat{g}_{if}$ and $\hat{g}_{jf}$. However, we do not give a causal interpretation to the effect of parental genetic endowments estimated in equation (5) because parental genetic endowments might be correlated with parental characteristics, as well as with genetic endowments of earlier ancestors, which may also influence parental behavior.

[8]Equation (5) can be rewritten as equations (4) or (5) in Behrman et al. (1994):

$$PI_{if} = \beta_0 + \beta_1(g_{if} - \hat{g}_{jf}) + \beta_2\hat{g}_{if} + \beta_3\hat{g}_{pf} + X'_{if}\alpha + S'_{jf}\delta + F'_f\gamma + u_{if} =$$
$$\beta_0 + (\beta_1 + \beta_2)\hat{g}_{if} - \beta_1\hat{g}_{jf} + \beta_3\hat{g}_{pf} + X'_{if}\alpha + S'_{jf}\delta + F'_f\gamma + u_{if}$$

However, in this case, the coefficient for $\hat{g}_{if}$ measures the combination of the price effect and the effect of parental preferences for equality *versus* efficiency.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

3.3 Endogenous Fertility

An additional issue that studies analysing parental responses to children's endowment differences must confront is the potential endogeneity of fertility. If fertility decisions were exogenously fixed, one could compare the parental investments in equally endowed children with differently endowed siblings regardless of birth order. However, parental decisions to have another child may depend on the endowments of previously born children. While this is not an issue when analyzing twins, for whom we estimate equation (5) as it is, it may be a problem for analyses based on non-twin siblings.

Ejrnæs and Pörtner (2004) provide a theoretical model of intra-household resource allocation where fertility decisions are treated as endogenous and parental investments and fertility decisions depend on children's genetic endowments. Importantly, they show that, if parents care more about efficiency (or have strong preferences for highly endowed children), parents who have a highly endowed child will stop having children earlier than parents who have a less well-endowed child. This however does not hold for inequality-averse parents.[9] In Table 1 we illustrate that highly-endowed children with highly-endowed older siblings are born to parents who are indifferent towards their children's endowments (indifferent parents, for short). In contrast, highly-endowed children with less well-endowed older siblings could have been born both to parents with strong preferences for highly-endowed children or to indifferent parents. Therefore, the comparison of children with the same absolute level of endowments but who differ in terms of their older sibling's endowments is complicated by the fact that these children are born to parents with systematically different preferences regarding the endowments of their offspring.

On the bright side, Table 1 also illustrates that children with the same absolute endowment levels but who differ in terms of their younger sibling's endowments are

[9]Ejrnæs and Pörtner (2004) test their model using longitudinal data from the Philippines to examine the effect of birth order on the number of hours in school and completed education. Consistently with the predictions of their model, they find that children with a higher birth order have an advantage over siblings with a lower birth order.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

born to parents with similar preferences. As a consequence, one can circumvent the endogenous fertility issues that affect the analysis of non-twin siblings by focusing on the genetic endowments of older children with respect to their younger siblings.

[Table 1 HERE]

We estimate equation (5) in the pooled sample of twins and non-twins, as well as for twins and non-twins separately. As we discuss in Section 2 and in Appendix A.1, it may be harder for parents of twins to compensate or reinforce differences between them since the public good component of parental investments is more important for twins than for non-twins (Bharadwaj et al., 2018). This may imply that $\hat{\beta}_1$ is likely lower for twins than for non-twins.

4 Data and Descriptive Statistics

4.1 The Add Health Dataset

We use data from the National Longitudinal Study of Adolescent to Adult Health (Add Health in what follows), a nationally representative longitudinal survey of U.S. 7^{th} to 12^{th} graders during the school year 1994/95 drawn from a stratified sample of 80 high schools and 52 middle schools. Within each school and grade, a random sample of approximately 17 males and 17 females, as well as an oversample of siblings and specific minorities were selected for interviews (the so-called *in-home sample*). The interviews were conducted in 1994/95 (Wave I, ages 12-20 years), 1996 (Wave II), 2001/02 (Wave III), 2008 (Wave IV), and 2016-18 (Wave V). Information on all our variables of interest was collected in Wave I except for the genetic information, as genotyping was performed in Wave IV, and completed years of education measured in Wave IV.

In Wave IV, respondents were asked for consent for the collection of saliva samples. Approximately 80% of respondents consented to long-term archiving of their samples and were eligible for genotyping. After quality control procedures, genetic information was available for approximately 65% of Wave IV respondents and poly-

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

genic indexes were constructed. We provide a detailed description of genotyping procedure and genetic information in Add Health in Appendix C.[10]

Another crucial advantage of Add Health for our purposes is that its *in-home* survey collects information on respondents' relationship with their parents, which allows us to construct indicators of parental investments. Importantly, the Add Health *in-home* and parental questionnaires also provide detailed information on individual and family background characteristics.

4.2 Genome-Wide Data to Measure Genetic Endowments

There are several ways to study the impact of genetic endowments on outcomes. Early studies estimated the genetic components of different traits by comparing the correlation of outcomes between dizygotic and monozygotic twin pairs (Taubman, 1976; Behrman et al., 1994; Savelyev et al., 2022). These studies showed that many socioeconomic outcomes have a strong genetic component. The main advantage of this methodology is that it does not require genetic data. However, it can only be used to infer which proportion of the outcome's variance is explained by the genetic component, but it cannot quantify genetic endowments.

As DNA sequencing has become cheaper, genome-wide association studies (GWAS) have been implemented to study the genetic determinants of different traits (*e.g*, depression, educational attainment, body mass, cognition). The human genome consists of a large set of DNA molecules, and in most locations in the genome (approximately 99%) there is no variation among humans. However, in some locations there is some variation and these genetic variants are used in GWAS to study the impact of genes on different outcomes.

GWAS analyse associations between an outcome of interest (a phenotype) and hundreds of thousands of genetic variants through a data-mining approach. Therefore, a GWAS requires large samples of genotyped individuals with available phenotypic information in order to provide precise estimates of the genetic effects. The

[10]For more information, see the Add Health documentation available at: `https://addhealth.cpc.unc.edu/wp-content/uploads/docs/user_guides/SSGAC-PGS_UsersGuide.pdf`.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

most powerful GWAS are usually conducted for easily available outcomes, such as educational attainment or body mass index, while less available outcomes, such as cognitive performance, are usually analyzed in smaller GWAS, and therefore the genetic effects on such outcomes are estimated with less precision.[11]

The effects of genetic variants on a phenotype estimated in a GWAS can be used as weights for construction of a summary index that measures individuals' genetic predisposition to this phenotype. These summary indexes are called polygenic indexes. In our analysis we use an educational attainment polygenic index as a measure of educational genetic endowments. In the sensitivity analysis, we also use a cognitive performance polygenic index.

To date, all GWAS of educational attainment rely on samples of individuals of European descent, so polygenic indexes are likely to be much less predictive and subject to more measurement error for other groups (Lee et al., 2018). Therefore, we conduct our analysis on a sample of European descent, as done in most previous studies using educational polygenic indexes.

4.2.1 Imputation of Parental Genotypes

Our strategy for identifying β_1 and β_2 in equation (5) relies on random assignment of genes across sibling and on random assignment of siblings' genes conditional on parental genes, respectively. However, there is no genetic information for parents in Add Health. To circumvent this issue, we rely on Mendelian imputation of parental genotypes, a technique proposed by Young et al. (2020, 2022). This method allows one to impute missing parental genotypes using available information on children's genotypes and the fact that, at each genome location, children inherit one molecule from each parent at random.[12]

[11] We provide additional details about GWAS and how genetic information is coded in Appendix B.

[12] For instance, suppose that the mother has the $++$ variant and the father has the $+-$ at genome location l. Then, the child will inherit a $+$ from the mother, and a $+$ or a $-$ from the father with a 50% probability, so the child's variant will be $++$ or $+-$.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Specifically, the method uses information on whether siblings inherited the same or different genetic variants from their parents. For instance, suppose that sibling 1 has the $++$ variant at genome location l and sibling 2 has the $--$ variant at the same location. Hence, sibling 1 has inherited a $+$ from the mother and a $+$ from the father, and sibling 2 has inherited a $-$ from the mother and a $-$ from the father, which implies that the mother and the father have the $-+$ variant at location l. When siblings' genetic variants are the same, it is not possible to be sure of which variants the parents have, so in this case the method uses the frequency with which a certain genetic variant occurs in the population.

By comparing imputed and observed parental genotypes, Young et al. (2020, 2022) demonstrate that this imputation method provides approximately unbiased estimates of parental genotypes. Additionally, Young et al. (2020, 2022) show that controlling for the imputed parental polygenic index enables unbiased and consistent estimation of the coefficient of the child's polygenic index. We provide additional details about the imputation of parental genetic data in Appendix D.

4.2.2 Polygenic Indexes

Polygenic indexes for Add Health are constructed as a standardized weighted sum of approximately 1.2 million genetic variants:

$$\hat{g}_{ti} = \frac{\sum_{l=1}^{L} x_{il}\hat{\gamma}_{tl}}{sd\left(\sum_{l=1}^{L} x_{il}\hat{\gamma}_{tl}\right)} \tag{6}$$

, where x_{il} is genetic variant l of individual i (measured with respect to a reference genotype and demeaned), $\hat{\gamma}_{tl}$ is a weight for genetic variant l associated with trait t (*e.g.*, educational attainment or cognitive performance),[13] and *sd* stands for standard

[13]For example, suppose that i has the $++$ variant at location l and $+$ is a reference molecule at this location, so i has two reference molecules at location l. Then, $x_{il} = 2 - \bar{x}_l$. That is, x_{il} is measured as the demeaned number of reference molecules that the individual has at genome location l. See Appendix B for further details on how genetic markers are coded.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

deviation, so that polygenic indexes are standardized and have zero mean given that each x_{il} is demeaned ($\bar{\hat{g}}_{ti} = 0$ and $sd(\hat{g}_{ti}) = 1$).

Polygenic weights $\hat{\gamma}_{tl}$ are transformations of the effects of genetic variants on trait t estimated in a GWAS, such that $\hat{\gamma}_{tl}$ minimizes $\mathbb{E}\left[(t_j - \sum_{l=1}^{L} x_{jl}\hat{\gamma}_{tl})^2\right]$. Therefore, a polygenic index of trait t is the best linear predictor of t given genetic variants $\{x_1, ...x_L\}$.[14]

In our analysis, we use polygenic indexes for Add Health respondents provided to Add Health by the Polygenic Index Repository (Becker et al., 2021).[15] Specifically, the Polygenic Index Repository uses a consistent methodology to construct polygenic indexes for 47 phenotypes in 11 datasets, including Add Health. The scores in the Repository are constructed using recent GWAS, as well as the UK Biobank and 23andMe GWAS. The Repository constructs weights for educational attainment using Lee et al. (2018) and 23andMe GWAS, and for cognitive performance using Trampush et al. (2017) and the UK Biobank GWAS. We use weights from the Repository to construct parental educational and cognitive performance polygenic indexes, given that parental polygenic indexes are not provided by Add Health.[16] We provide further details on the construction of polygenic indexes in Appendix C.

4.2.3 Polygenic Indexes as Measures of Educational Genetic Endowments

We now discuss how an educational polygenic index should be interpreted in our context. Our objective is to test how parents react to sibling differences in educational genetic endowments, and by educational genetic endowments we mean genetically

[14]Some genetic variants may have a positive effect on trait t and others may have a negative effect, so the weights may take positive or negative values.

[15]See the Add Health documentation (`https://cdr.lib.unc.edu/downloads/00000772p`) for details about the construction of polygenic indexes in this dataset.

[16]The full GWAS summary statistics for the 23andMe discovery data set has been made available through 23andMe under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please visit `https://research.23andme.com/collaborate/#dataset-access/` for more information and to apply to access the data.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

influenced initial ability for education. Hence, our analysis relies on the following assumption:

Assumption 1 *Educational genetic endowments are well proxied by the educational polygenic index.*

Genes affect certain cognitive and psychological characteristics, which in turn comprise initial ability. However, the skills that matter for school achievement are determined by both initial ability and subsequent investments that may in turn interact with initial ability. Therefore, the genetic component of educational attainment may capture the effects of initial genetic ability and the indirect effects of genes on education driven by environmental responses to the genetic ability (*e.g.*, parental or teachers' inputs).

In Appendix E we consider a model in which educational attainment is a function of educational genetic ability and of different inputs (including parental inputs). In turn, inputs are endogenously chosen and may depend on genetic characteristics that can differ from the educational genetic ability that directly affects educational attainment. First, we show that: (i) if inputs are only affected by educational genetic ability, and (ii) if there is not full compensation of genetic ability, estimators of β_1 and β_2 in equation (5) will provide unbiased estimates of the effects of educational genetic ability on parental investments. On the other hand, if genetic characteristics that affect inputs and educational genetic ability differ, the results provided in Appendix E indicate that compensatory inputs (parental and non-parental) may lead to underestimating the effect of genetic ability on parental investments, while reinforcing inputs may lead to overestimating this effect. However, even in the worst-case scenarios that we consider, the estimated effect will account for $65\% - 120\%$ of the true effect. The sign of the effect will be estimated correctly if there is not full compensation of genetic ability.

Full compensation would imply that the educational polygenic index has no positive association with genetic variants that affect cognitive function and brain-development processes. However, this is unlikely as it would be at odds with the results from the genetics literature according to which the educational polygenic index is positively

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

associated with brain volume, white-matter tract integrity, and neuronal development or function (Elliott et al., 2019; Demange et al., 2021).

Another issue is that educational polygenic indexes only capture part of the genetic component of educational attainment. First, polygenic indexes are *linear* functions of genetic variants, so they ignore potential non-linear effects of genes. While non-additive genetic factors may be potentially important, abundant theory and evidence from quantitative genetics indicates that most of the genetic variance in the analysis of population data can be captured by the additive genetic factors (Falconer and Mackay, 1996; Hill et al., 2008; Okbay et al., 2022; Palmer et al., 2021; Zhu et al., 2015).

Second, polygenic indexes are not based on all genetic variants but only on measured variants. In Appendix F we show that if measured genetic variants capture only half of the genetic component of educational attainment, as it is estimated by Cheesman et al. (2017), then our estimates will be deflated by a factor of $1/\sqrt{2} \approx 0.71$, assuming that unmeasured genetic variants have the same average effect sizes as measured genetic variants. The third challenge is related to the fact that polygenic weights $\hat{\gamma}_l$ are estimated in cross-sectional GWAS that do not control for parental genes, while the relevant polygenic index should be constructed from a GWAS that controls for parental genotypes. We discuss in Appendix E that this may deflate the estimated genetic effects by a factor of 0.7-0.8. In sum, the limitations listed above imply that the effect sizes that we estimate are likely to be conservative and should be interpreted as lower bounds of the true effects of individuals' genetic propensity for education.

An additional challenge we address is related to the fact that polygenic weights $\hat{\gamma}_{tl}$ from equation (6) are not estimated in the population but in a GWAS sample. This implies that polygenic indexes are subject to measurement error. The variance of the measurement error is generally larger in polygenic scores based on small GWAS, so it is recommended to rely on large GWAS. Hence, we favor the educational polygenic index, which relies on a GWAS of approximately 1 million individuals, over the cognitive performance polygenic index, which relies on a four times smaller GWAS.

Becker et al. (2021) show that when there is measurement error in a polygenic

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

index, its estimated effects will be attenuated towards zero. They also show that the variance of the measurement error can be inferred from the difference between the proportion of the variance of an outcome of interest (*e.g.*, educational attainment) explained by the set of genetic variants (referred to as SNP heritability and denoted by h^2_{SNP}) and the proportion of the variance of the outcome explained by the polygenic index (R^2 from the regression of the outcome on its polygenic index). Using estimates of R^2 and h^2_{SNP}, Becker et al. (2021) provide a measurement-error correction method for regressions that include a single polygenic index, assuming that the measurement error is uncorrelated with genetic variants and non-genetic variables.

In Appendix F, we follow Becker et al. (2021) and provide a measurement-error correction method to adjust the estimated genetic effects in models that include several polygenic indexes (*e.g.*, individuals', siblings', and parental polygenic indexes). This correction method requires estimates of h^2_{SNP}, R^2, and of parents-offspring and siblings genetic correlations, which we obtain in Appendix F.7. We use this method to correct for measurement error in our analysis.[17]

In Appendix F.5.1, we show that the standard errors are likely to be inflated due to measurement error, as it adds two additional sources of variation to the residual. The first source is due to the fact that the coefficients are biased, and the second is due to the additional variance introduced by the presence of measurement error in the genetic regressors. Therefore, when we discuss the statistical significance of our results, we note that our measurement-error-corrected standard errors are conservative, which may prevent us from concluding that some effects achieve standard levels of statistical significance if they are small.

4.3 Parental Investments

We use several alternative measures of parental investments based on questions about teenagers' relationship with their parents included in the in-home questionnaire administered in Wave I of Add Health. Adolescents were asked similar questions about their relationship with the mother and the father. In particular, we consider the fol-

[17] In Table H.10 of the Appendix, we provide sensitivity checks based on different estimates of h^2_{SNP}.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

lowing binary outcomes: *i) In the past 4 weeks went to a movie, play, museum, concert, or sports event with the mother (father); ii) In the past 4 weeks talked about school work or grades with the mother (father); iii) In the past 4 weeks worked on a project for school with the mother (father); iv) In the past 4 weeks talked about other things were doing in school with the mother (father); and v) On how many of the past 7 days was at least one of your parents in the room with you while you ate your evening meal?*.[18]

Using these variables, we construct three indicators: a parental investment index based on questions involving both parents, a maternal investment index based on questions involving the mother, and a paternal investment index based on questions involving the father. If only one parent is present in the household, the parental investment index only includes information regarding the teenagers' relationship with him/her. To construct summary indexes, we follow Kling et al. (2007) and compute each summary index variable Y^* as the unweighted average of standardized components:

$$Y^* = \frac{\sum_k Y_k^*}{K} \text{ , where } Y_k^* = \frac{Y_k - \bar{Y}_k}{sd(Y_k)}$$

, where Y_k is the k^{th} component of the index, $\bar{Y}_k$ denotes its mean and $sd(Y_k)$ is its standard deviation.

4.4 Estimation Sample Description

For our analysis we use the Add Health Sibling Pairs sample with available genetic data. We start with 3,139 sibling pairs. Next, we restrict this sample to 1092 pairs with available genetic information. Then, we eliminate non-European ancestry individuals, which leaves us with 619 sibling pairs. We then eliminate 15 pairs because the firstborn sibling or both twins had no information on parental investments. This restriction leaves us with an estimation sample of 604 sibling pairs (412 non-twin pairs and 192 twin pairs) from 557 families.

[18]Questions (i) and (v) are comparable to the questions *How often child goes to musical shows* and *Eats with mom/dad* used to measure parental investments in Cunha et al. (2010). Importantly, these measures have a high signal to noise ratio as a parental investment (see Table IIB Cunha et al. 2010).

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

In our baseline specification we control for children's age, age-squared, sibling differences in age (only in the non-twin subsample), a female dummy, and a female sibling dummy. We also assess the sensitivity of our results to the inclusion of additional controls. Specifically, in the extended list of controls we include a parental socio-economic status (SES) index (see Appendix G), a rural area indicator, and an indicator that both parents live in the household. Table H.1 in Appendix H displays summary statistics for the control variables. Figure H.1 of Appendix H plots kernel-smoothed densities of the educational polygenic indexes of individuals, siblings, and parents in our estimation sample. The distributions of the polygenic indexes are approximately normal.

Table H.2 in Appendix H describes the main outcomes measuring parental investments as well as their components. Our identification strategy requires that there is within-family variation in parental investments. We report across- and within-family standard deviations of parental investment indexes in Table H.3 of the Appendix.[19] Within-family standard deviations constitute 58%-59% of the across-family standard deviations. Interestingly, this share is lower for twins than for non-twins (49%-53% vs. 61%-62%), which is consistent with parents of twins being more likely to equally invest in their children, possibly because the public good component of parental investments is particularly important for children of the same age (see Section 2).

In Table H.4 of the Appendix we show that our parental investment indexes are positively associated with children's educational attainment measured in Wave IV. Specifically, a one standard deviation increase in the parental investment index is associated with an increase by 0.175 standard deviations of years of education ($p - value < 0.001$). This association remains positive (0.112) and statistically distinguishable from zero ($p - value < 0.001$) after we control for individual and family characteristics. Similarly, one standard deviation increase in the maternal and paternal investment indexes is associated with an increase by 0.106 ($p - value = 0.003$) and 0.133 ($p - value = 0.001$) standard deviations of years of education, respectively, after controlling for individual and family characteristics.

[19]We compute within-family standard deviations by removing family fixed effects from the outcome variables.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Our analysis relies on twins and firstborn children who have siblings with valid genetic data. Moreover, because of the relatively low predictive power of polygenic indexes for non-European ancestry groups, our analysis focuses on European ancestry individuals. In Table H.5 of the Appendix we compare our estimation sample to: (1) the Add Health representative sample (respondents with valid Wave I grand sample weights); and (2) a sample of only white respondents (with valid Wave I grand sample weights). Our estimation sample only includes firstborns or twins who are about one year older than individuals in the full sample of white respondents. Apart from age, the characteristics of our estimation sample and the full sample of white respondents are similar.

As for the comparison with the representative sample that includes other ancestry groups, Table H.5 of the Appendix shows that individuals in our estimation sample are more likely to live in rural areas and to live in two-parent households. However, these differences become negligible when the difference in the share of white individuals is accounted for. The average parental socioeconomic index and the parental investment indexes are not statistically distinguishable between our estimation sample and the full sample that includes non-European ancestry individuals.

Next, we show that the educational polygenic index is indeed associated with educational attainment and different measures of educational achievement.[20] To do so, we regress each measure of achievement on the educational polygenic index with and without controls while also applying the measurement-error correction method described in Appendix F. Panel A of Table 2 reports the effects (uncorrected and measurement-error-corrected) of the polygenic index on educational attainment (standardized to have mean 0 and standard deviation 1) in our pooled sample of firstborns and twins. Column (1) shows that the educational polygenic index explains 14.9% of the observed variance in educational attainment, and one standard deviation increase in the polygenic index increases educational attainment by 0.386 ($SE = 0.041$) stan-

[20]Educational attainment is measured using the question *"What is the highest level of education that you have achieved to date?"* addressed to Wave IV (age: 25-33) respondents.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

dard deviations.[21] Column (2) reports the same effect after controlling for the parental educational polygenic index, age, age squared, and sex. When these controls are included, the effect of the educational polygenic index amounts to 0.191 ($SE = 0.074$) standard deviations, and a standard deviation increase in the parental educational polygenic index is associated with 0.234 ($SE = 0.073$) standard deviations increase in educational attainment. Finally, in column (3) we include the same controls as in column (2) and we correct for measurement error, finding that a standard deviation increase in the educational polygenic index increases educational attainment by 0.369 ($SE = 0.076$) standard deviations.

Panels B and C of Table 2 show similar patterns when analysing other indicators related to educational achievement, such as Peabody Picture Vocabulary Test (PPVT) scores and grade point averages (GPA). While the value of the polygenic index is known to analysts, it is unlikely to be known by parents. Instead, parents at least partially observe some of the traits related to their children's polygenic indexes, and therefore their investment decisions can respond to these traits. Hence, it is important to ensure that the polygenic index is strongly associated not only with completed educational attainment, but also with earlier manifestations of ability, such as PPVT scores and GPA.

[Table 2 Here]

Since variation in children's polygenic indexes conditional on parental polygenic indexes resembles a lottery, one would expect sibling differences in polygenic indexes to be uncorrelated with individual and family characteristics when the parental polygenic index is controlled for. To check that this is indeed the case, we regress each of our control variables on sibling differences in the educational polygenic index, as well as on parental and children's own polygenic indexes. The results are reported in Table H.6 of the Appendix. In line with the idea that genetic variation across siblings is as good as random, sibling differences in the educational polygenic index are not significantly associated with any control variable.

[21]Note that the Repository educational polygenic index outperforms the explanatory power of educational polygenic indexes based on earlier GWAS (Okbay et al., 2016) that relied on smaller samples.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

5 Main Results

Our main results for the parental investments index are displayed in Table 3. The table reports the measurement-error-corrected coefficients from equation (5).[22]

Columns (1-3) show estimates of our main coefficients of interest, $\hat{\beta}_1$ and $\hat{\beta}_2$, obtained from estimating equation (5) in the pooled sample of twins and non-twins without controls (column 1), with baseline controls (column 2), and with additional controls (column 3). Columns (4-6) and (7-9) are based on the same specifications and show the estimated effects for non-twins and twins separately. Since adding controls barely alters coefficient estimates, we mainly focus our discussion on the results from columns (2), (5), and (8) that include baseline controls.

The baseline estimates in the pooled sample (column 2 of Table 3) indicate that parents, on average, display inequality aversion because the effect of sibling differences in the educational polygenic index is always negative and statistically distinguishable from zero at standard levels of testing. That is, after conditioning on parents' and children's own educational polygenic indexes, we find that parents invest less in children who are better endowed than their siblings. In particular, if sibling differences in the educational polygenic index increase by one standard deviation, the parental investment index decreases by 0.233 ($SE = 0.081$) standard deviations.

Column 5 of Table 3 shows that the estimated effect of sibling differences in the educational polygenic index for non-twins is also negative and statistically distinguishable from zero ($p - value = 0.007$), which is again consistent with parental inequality aversion. Specifically, in the sample of non-twins, a standard deviation increase in sibling differences in the educational polygenic index decreases the parental investment index by 0.270 ($SE = 0.100$) standard deviations.

[Table 3 Here]

This finding is very much in contrast with our evidence for twins (column 8 of Table 3), as we find that the effect of differences in the educational polygenic index between twins is small and imprecise (-0.043 with $SE = 0.123$). One potential explanation for our contrasting results for parents of twins and non-twin siblings may

[22]We report the uncorrected estimates in Table H.7 of the Appendix.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

be that it can be difficult for the former to invest differently across their children because they are exactly the same age. Intuitively, if a parent helps out with homework or plays with one twin, it is difficult to prevent the other twin from participating to some extent. This implies that, even if parents of twins were inequality averse, they may be unable to invest differently across their children.[23] In the extension of our theoretical model that incorporates externalities across siblings or a public good component of parental investments, we show that, when this component is strong, the effect of parental inequality aversion is smaller than when parental investments are perfectly separable (see Appendix A.1). An alternative explanation is that the mistaken belief potentially held by parents that non-identical twins have identical genetic endowments may make them less prone to perceiving, and therefore responding to the manifestation of differences in their twins' genetic propensity for education.

As for the price effect, the estimated effects of children's own educational polygenic index reported in Table 3 are positive and sizeable (0.201 with $SE = 0.077$). A positive price effect implies that, if it is large enough, even inequality averse parents may choose to follow a reinforcing or a neutral strategy. In fact, the sum of β_1 and β_2 is not statistically distinguishable from zero, which suggests that the price effect and the parental inequality aversion mechanisms offset each other in our sample.

Finally, in Table H.8 in the Appendix we investigate whether there is heterogeneity in the effect of sibling differences in the educational polygenic index on parental investments depending on the sex of the parent whose investment decisions are being analysed, as well as on the children's sex.

In order to analyse whether mothers and fathers respond differently to sibling differences in their genetic propensity for education, we estimate equation (5) using as outcomes the investment indexes that are solely based on maternal and paternal investments. The results of this analysis are shown in columns (1) and (2) of Table H.8 in the Appendix. We cannot reject that sibling differences in the polygenic index have the same effect on maternal and paternal investment decisions.

[23]This is in line with Bharadwaj et al. (2018), who empirically estimate that the public good component of parental investments is more important for children who are very close in age.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Several studies document that parental investments vary depending on children's sex (Lundberg, 2005; Baker and Milligan, 2016). However, the evidence is mixed and does not necessarily indicate that parents favor boys over girls. In columns (3) and (4), we separately estimate the effects of sibling differences in the educational polygenic index on parental investment decisions for boys and girls. The estimated effects are similar for the two groups, and we cannot reject the null hypothesis that parental responses to sibling differences in the educational polygenic index do not vary by children's sex. We also find that the effect of children's own educational polygenic index on parental investments is only statistically distinguishable from zero for girls, but we cannot rule out that the effect of children's own genetic propensity for education is the same for boys and girls.

6 Robustness Checks

6.1 Sensitivity Tests

We now conduct four sets of sensitivity tests. First, we check whether our results are robust to using an alternative measure of initial ability by re-estimating equation (5) using the cognitive performance polygenic index as a measure of genetic endowments. The measurement-error-corrected estimated effects are reported in Panel A of Table H.9 of the Appendix. In line with our main findings, the effect of sibling differences in the cognitive performance polygenic index is negative and similar in magnitude to the effect of sibling differences in the educational polygenic index. Moreover, the effect is statistically distinguishable from zero only for non-twin siblings, while the effect is small in magnitude and not distinguishable from zero for twins. Hence, our main results persist when using the cognitive performance rather than the educational polygenic index as a measure of genetic endowments.

A second potential concern is that there is no within-family variation in the educational polygenic index in families with monozygotic twins. Our identification of the effect of sibling differences in the educational polygenic index is based on the comparison of individuals who have the same value of the polygenic index but who differ in terms of their siblings' polygenic index (since we control for individuals' own

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

polygenic index). Therefore, in the case of monozygotic twins, we identify the effect of interest by comparing parental investments in a monozygotic twin with parental investments in some other child who has the same value of the polygenic index as the monozygotic twin but whose sibling has a different value of the polygenic index. If we restricted the sample only to monozygotic twins, identification would not be possible because monozygotic twins have identical DNA and therefore there would be no variation in sibling differences in the polygenic index. However, our analysis for twins also includes dizygotic twins. It is true, however, that the variance of sibling differences in the polygenic index is smaller in the sample of twins than in the sample of non-twins. Therefore, the effect for twins is likely to be estimated with lower precision than the effect for non-twins. Hence, in Panel B of Table H.9 of the Appendix we exclude monozygotic twins from the analysis. Panel B of Table H.9 shows that, in the subsample of dizygotic twins only, the effect of sibling differences in the educational polygenic index is close to zero and not statistically distinguishable from zero, which is consistent with our baseline results for twins reported in column 8 of Table 3.

Third, since genetic variants may be associated with the outcome because of population stratification —the presence of systematic differences in genetic variation between subpopulations—, the effect of the educational polygenic index may reflect these differences. This should not be an issue when one estimates the effect of children's educational polygenic index while controlling for their parents' educational polygenic index because the parental polygenic index captures the effect of population stratification. At any rate, in Panel C of Table H.9 of the Appendix we include the 20 principal components of the full genetic relatedness matrix as controls, which is a standard practice to control for population stratification (Price et al., 2006; Benjamin et al., 2012). The estimated effects of sibling differences in the educational polygenic index remain very similar to the baseline estimates reported in Table 3.

Fourth, our main estimation sample is limited to older siblings and twins in order to circumvent the issue that fertility decisions are endogenous (see Section 3.3). In Panel D of Table H.9 we report the estimated effects in an unrestricted sample that includes both older and younger siblings, as well as twins. The estimated effect of sibling differences in the educational polygenic index is negative but less precise in

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

the pooled sample ($p-value=0.051$) and in the sample of non-twins ($p-value=0.055$), and small and not statistically distinguishable from zero in the sample of twins ($p-value=0.981$).

Finally, in Table H.10 of the Appendix we analyse the sensitivity of the estimated effects of sibling differences in the educational polygenic index and the cognitive performance polygenic index to different assumptions about the proportion of the variance of educational attainment and cognitive performance explained by genetic variants (h^2_{SNP}) used in the measurement-error correction (see Appendix F). Columns (1)-(4) of Table H.10 in the Appendix show that, depending on the assumption about h^2_{SNP}, the estimated effect of a standard deviation increase of sibling differences in the educational polygenic index varies from -0.273 to -0.194. Columns (5)-(7) of Table H.10 show that the estimated effect of a standard deviation increase in sibling differences in the cognitive performance polygenic index varies from -0.280 to -0.186. All the effects are economically meaningful and consistent with the presence of parental inequality aversion.

6.2 Falsification Tests

In order to check that our results are not driven by chance, we run placebo tests. To obtain placebo versions of children's educational polygenic index, sibling differences in the polygenic index and parents' polygenic index, we replace their actual values with those from randomly drawn families from our sample. We then estimate equation (5) using these placebo values and including baseline covariates. As in our main analysis, we account for measurement error in the educational polygenic index in the placebo regressions. We repeat this procedure 1,000 times in order to obtain distributions of the estimated coefficients of placebo variables. We find that, in line with our results being genuine, the placebo effects are statistically significant at the 5% level in approximately 5% of our placebo regressions, which is what we expect to obtain by chance. Figure H.2 in the Appendix summarizes the results of these estimations by displaying the placebo $t-value$ distributions of the tests $\beta_1=0$, $\beta_2=0$, and $\beta_3=0$.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

6.2.1 "Too Early" Parental Responses

Parents cannot possibly observe differences between children at a very early age. Therefore, another placebo test of our main result consists in checking whether sibling differences in the educational polygenic index impact "too soon after birth" parental investment indicators, such as breastfeeding. We use a retrospective question from the Add Health parental questionnaire that asked mothers how long each of their children participating in the *in-home* interview was breastfed. We define an indicator variable which takes the value zero if the mother's answer is *"(He/ she) was not breastfed"* and one if she reports that the child was breastfed to some extent. Then we estimate equation (5) using this variable as an outcome and including the baseline controls listed in Table H.1 in the Appendix. As expected, the effect of sibling differences in the polygenic index on the probability of having been breastfed are very close to zero ($\hat{\beta}_1 = 0.001$ and $SE(\hat{\beta}_1) = 0.085$).

7 Conclusions

We take advantage of recent advances in behavioral genetics to revisit a longstanding question in economics, namely: how do parental investment decisions respond to adolescents' endowments and to endowment differences between siblings? In our estimation we rely on random variation of children's genotypes conditional on parental genotypes. We provide new evidence that American parents of European-ancestry non-twin adolescents display inequality aversion. In particular, we show that parents invest more (less) in children if their genetic predisposition for educational attainment is lower (higher) than that of their siblings controlling for the absolute level of parents' and children's genetic propensity for education. Hence, parental investments during adolescence may be able to reduce the effect of inequalities in genetic endowments. We also find evidence that the price effect is positive as parents invest more in adolescents with higher genetic propensity for education holding genetic differences among siblings constant, as well as parental genetic endowments and other characteristics. Interestingly, and in contrast to our findings for parents of non-twin adolescents, we find no evidence that American parents of European-ancestry twins

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

respond to differences in their twins' genetic endowments.

The limitation of using genetic data —and more specifically polygenic indexes— in our analysis is that they provide an imperfect measure of educational genetic propensity for education. First, polygenic indexes are linear combinations of genetic variants, so they ignore potential non-linear effects. Second, polygenic indexes are not based on all genetic variants but only on measured genetic variants. Third, polygenic indexes are computed using *estimated* weights, which implies that they are subject to measurement error. We take measurement error issues into consideration in our analysis, and we show that our estimated effects are likely to be deflated and should be interpreted as lower bounds of the true effects. Another limitation is that our analysis is conducted for individuals from European ancestries because polygenic indexes are poorer measures of educational genetic endowments for other ancestries. Further research is needed to obtain polygenic indexes for non-European ancestries. Hence, we cannot claim our results are generalizable to non-European ancestry individuals.

Our findings are important for evaluating the role of the family in shaping inequality as well as the effectiveness of compensatory policies. American parents of European-ancestry teenagers in our sample on average display inequality aversion in their investment decisions. Hence, well-targeted interventions that help lower-endowed adolescents and increase their endowments may in turn induce their inequality averse parents to invest less in their "compensated" children and act less as equalizing agents, attenuating the impact of compensatory programs.

Suppose instead that parents followed a reinforcing strategy. If the price effect is positive because it is less costly for parents to invest in highly endowed children than in their less well-endowed siblings, a reinforcing strategy may stem from inequality averse parents whose degree of inequality aversion could not offset the price effect, from neutral parents, or from parents who favor efficiency over equality. Well-targeted compensatory policies would be most effective in the latter case, as interventions that help lower-endowed adolescents will in turn induce parents who favor efficiency to invest more in them. The opposite would happen if parents were inequality averse, and no parental behavioral response would take place if parents displayed neutral preferences.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Taken as a whole, our evidence suggests that further research is needed to look into the black box of intra-household dynamics and the nature of parental investments for different types of families. As for the distinctive features of post-natal parental investments in twins versus singleton siblings, when incorporating externalities across siblings or a public good component of parental investments into our theoretical model, we show that, if this component is strong, the effect of parental inequality aversion is smaller than when parental investments are perfectly separable. Our findings that parents of singleton siblings display inequality aversion and no evidence for this effect for parents of twins are in line with this notion.

Finally, our results highlight the idea that early life conditions not only affect later-life outcomes directly, but also indirectly through intra-household allocation effects. This idea is not new (*e.g.*, Cunha et al. 2010 and Yi et al. 2015), but our paper is the first to provide direct evidence that sibling differences in educational genetic endowments shape parental investment decisions. In contrast, previous studies that rely on genetic data mainly focus on identifying the overall effect of children's own genetic endowments on parental investments, while ignoring the role of siblings' genetic endowments. For example, Breinholt and Conley (2020) and Houmark et al. (2020) find that parents invest more in young children with higher initial genetic endowments, which contrasts with our results that the overall effect —that is, the effect that combines the price and the inequality aversion effects— of children's genetic propensity for education on parental investments by the time children are adolescents is not statistically distinguishable from zero (controlling for parental and siblings' genetic propensity for education). One potential explanation for this difference is that parents might not respond the same way at different stages of childhood and adolescent development. Our results also show that siblings' genetic predisposition for education has a positive effect on parental investments, which implies that ignoring the role played by siblings' genetic endowments likely leads to a positive bias in the estimated effects of children's own and parental genetic endowments.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

References

Abufhele, Alejandra, Jere Behrman, and David Bravo, "Parental preferences and allocations of investments in children's learning and health within families," *Social Science & Medicine*, 2017, *194*, 76–86.

Aizer, Anna and Flávio Cunha, "The Production of Human Capital: Endowments, Investments and Fertility," Working Paper 18429, National Bureau of Economic Research September 2012.

Almond, Douglas and Bhashkar Mazumder, "Fetal origins and parental responses," *Annual Review of Economics*, 2013, *5* (1), 37–56.

_ **and Janet Currie**, "Killing me softly: The fetal origins hypothesis," *Journal of Economic Perspectives*, 2011, *25* (3), 153–72.

Baker, Michael and Kevin Milligan, "Boy-girl differences in parental time investments: Evidence from three countries," *Journal of Human Capital*, 2016, *10* (4), 399–441.

Barth, Daniel, Nicholas W. Papageorge, and Kevin Thom, "Genetic endowments and wealth inequality," *Journal of Political Economy*, 2020, *128* (4), 1474–1522.

_ **,** _ **,** _ **, and Mateo Velásquez-Giraldo**, "Genetic Endowments, Income Dynamics, and Wealth Accumulation Over the Lifecycle," Working Paper 30350, National Bureau of Economic Research 2022.

Beauchamp, Jonathan P., David Cesarini, Magnus Johannesson, Matthijs J.H.M. van der Loos, Philipp D. Koellinger, Patrick J.F. Groenen, James H. Fowler, J. Niels Rosenquist, A. Roy Thurik, and Nicholas A. Christakis, "Molecular Genetics and Economics," *Journal of Economic Perspectives*, 2011, *25* (4), 57–82.

Becker, Gary S. and Nigel Tomes, "Child endowments and the quantity and quality of children," *Journal of Political Economy*, 1976, *84* (4, Part 2), S143–S162.

Becker, Joel, Casper A.P. Burik, Grant Goldman, Nancy Wang, Hariharan Jayashankar, Michael Bennett, Daniel W. Belsky, Richard Karlsson Linnér,

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Rafael Ahlskog, Aaron Kleinman et al., "Resource profile and user guide of the Polygenic Index Repository," *Nature Human Behaviour*, 2021, pp. 1–15.

Behrman, Jere R., "Intrahousehold allocation of nutrients in rural India: Are boys favored? Do parents exhibit inequality aversion?," *Oxford Economic Papers*, 1988, *40* (1), 32–54.

_ , "Nutrition health birth order and seasonality: intrahousehold allocation among children in rural India.," *Journal of Development Economics*, 1988, *28* (1), 43–62.

_ **, Mark R. Rosenzweig, and Paul Taubman**, "Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment," *Journal of Political Economy*, 1994, *102* (6), 1131–1174.

_ **, Robert A. Pollak, and Paul Taubman**, "Parental preferences and provision for progeny," *Journal of Political Economy*, 1982, *90* (1), 52–73.

Benjamin, Daniel J., Christopher F. Chabris, Edward L. Glaeser, Vilmundur Gudnason, Tamara B. Harris, David I. Laibson, Lenore J. Launer, and Shaun Purcell, "Genoeconomics," in "Biosocial Surveys," National Academies Press (US), 2008.

_ **, David Cesarini, Christopher F. Chabris, Edward L. Glaeser, David I. Laibson, Gene/Environment Susceptibility-Reykjavik Study: Age, Vilmundur Guðnason, Tamara B. Harris, Lenore J. Launer, Shaun Purcell et al.**, "The promises and pitfalls of genoeconomics," *Annual Review of Economics*, 2012, *4* (1), 627–662.

Bhalotra, Sonia and Damian Clarke, "The Twin Instrument: Fertility and Human Capital Investment," *Journal of the European Economic Association*, 12 2019. jvz058.

Bharadwaj, Prashant, Juan Pedro Eberhard, and Christopher A. Neilson, "Health at birth, parental investments, and academic outcomes," *Journal of Labor Economics*, 2018, *36* (2), 349–394.

Bono, Emilia Del, John Ermisch, and Marco Francesconi, "Intrafamily resource allocations: a dynamic structural model of birth weight," *Journal of Labor Economics*, 2012, *30* (3), 657–706.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Breinholt, Asta and Dalton Conley, "Child-Driven Parenting: Differential Early Childhood Investment by Offspring Genotype," Working Paper 28217, National Bureau of Economic Research 2020.

Cameron, Stephen V. and James J. Heckman, "The dynamics of educational attainment for black, hispanic, and white males," *Journal of Political Economy*, 2001, *109* (3), 455–499.

Cheesman, Rosa, Saskia Selzam, Angelica Ronald, Philip S. Dale, Tom A. McAdams, Thalia C. Eley, and Robert Plomin, "Childhood behaviour problems show the greatest gap between DNA-based and twin heritability," *Translational Psychiatry*, 2017, *7* (12), 1–9.

Conley, Dalton and Jason Fletcher, *The Genome Factor: What the social genomics revolution reveals about ourselves, our history, and the future*, Princeton University Press, 2017.

Cunha, Flavio and James J. Heckman, "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation," *Journal of Human Resources*, 2008, *43* (4), 738–782.

_ , _ , and Susanne M. Schennach, "Estimating the technology of cognitive and noncognitive skill formation," *Econometrica*, 2010, *78* (3), 883–931.

_ , _ , Lance Lochner, and Dimitriy V. Masterov, "Interpreting the evidence on life cycle skill formation," *Handbook of the Economics of Education*, 2006, *1*, 697–812.

Currie, Janet, "Inequality at birth: Some causes and consequences," *American Economic Review*, 2011, *101* (3), 1–22.

Datar, Ashlesha, M. Rebecca Kilburn, and David S. Loughran, "Endowments and parental investments in infancy and early childhood," *Demography*, 2010, *47* (1), 145–162.

Demange, Perline A, Margherita Malanchini, Travis T. Mallard, Pietro Biroli, Simon R. Cox, Andrew D. Grotzinger, Elliot M. Tucker-Drob, Abdel Abdellaoui,

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Louise Arseneault, Elsje Van Bergen et al., "Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction," *Nature Genetics*, 2021, *53* (1), 35–44.

Ding, Weili, Steven F. Lehrer, J. Niels Rosenquist, and Janet Audrain-McGovern, "The Impact of Poor Health on Education: New Evidence Using Genetic Markers," Working Paper 12304, National Bureau of Economic Research June 2006.

Domingue, Benjamin W., Daniel W. Belsky, Dalton Conley, Kathleen Mullan Harris, and Jason D. Boardman, "Polygenic influence on educational attainment: New evidence from the National Longitudinal Study of Adolescent to Adult Health," *AERA open*, 2015, *1* (3).

Ejrnæs, Mette and Claus C. Pörtner, "Birth order and the intrahousehold allocation of time and education," *Review of Economics and Statistics*, 2004, *86* (4), 1008–1019.

Elliott, Maxwell L., Daniel W. Belsky, Kevin Anderson, David L. Corcoran, Tian Ge, Annchen Knodt, Joseph A. Prinz, Karen Sugden, Benjamin Williams, David Ireland et al., "A polygenic score for higher educational attainment is associated with larger brains," *Cerebral Cortex*, 2019, *29* (8), 3496–3504.

Falconer, D.S. and T.F.C. Mackay, "Introduction to quantitative genetics. Essex," *UK: Longman Group*, 1996.

Fletcher, Jason M. and Steven F. Lehrer, "Genetic lotteries within families," *Journal of Health Economics*, 2011, *30* (4), 647–659.

Frijters, Paul, David W. Johnston, Manisha Shah, and Michael A. Shields, "Intrahousehold resource allocation: do parents reduce or reinforce child ability gaps?," *Demography*, 2013, *50* (6), 2187–2208.

Halla, Martin and Martina Zweimueller, "Parental Response to Early Human Capital Shocks," Working Paper 7968, IZA Working Paper 2014.

Heckman, James J. and Stefano Mosso, "The economics of human development and social mobility," *Annual Review of Economics*, 2014, *6* (1), 689–733.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Hill, William G., Michael E. Goddard, and Peter M. Visscher, "Data and theory point to mainly additive genetic variance for complex traits," *PLoS Genetics*, 2008, *4* (2), e1000008.

Houmark, Mikkel, Victor Ronda, and Michael Rosholm, "The nurture of nature and the nature of nurture: How genes and investments interact in the formation of skills," Working Paper 13780, IZA Working Paper 2020.

Hsin, Amy, "Is biology destiny? Birth weight and differential parental treatment," *Demography*, 2012, *49* (4), 1385–1405.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental analysis of neighborhood effects," *Econometrica*, 2007, *75* (1), 83–119.

Kong, Augustine, Gudmar Thorleifsson, Michael L. Frigge, Bjarni J. Vilhjalmsson, Alexander I. Young, Thorgeir E. Thorgeirsson, Stefania Benonisdottir, Asmundur Oddsson, Bjarni V. Halldorsson, Gisli Masson et al., "The nature of nurture: Effects of parental genotypes," *Science*, 2018, *359* (6374), 424–428.

Lee, James J., Robbee Wedow, Aysu Okbay, Edward Kong, Omeed Maghzian, Meghan Zacher, Tuan Anh Nguyen-Viet, Peter Bowers, Julia Sidorenko, Richard Karlsson Linnér et al., "Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals," *Nature Genetics*, 2018, *50* (8), 1112–1121.

Lehrer, Steven F. and Weili Ding, "What is the role for molecular genetic data in public policy?," *, IZA World of Labor*, 2017, *2017: 395.*

Lundberg, Shelly, "Sons, daughters, and parental behaviour," *Oxford Review of Economic Policy*, 2005, *21* (3), 340–356.

Okbay, Aysu, Jonathan P Beauchamp, Mark Alan Fontana, James J. Lee, Tune H. Pers, Cornelius A. Rietveld, Patrick Turley, Guo-Bo Chen, Valur Emilsson, S. Fleur W. Meddens et al., "Genome-wide association study identifies 74 loci associated with educational attainment," *Nature*, 2016, *533* (7604), 539.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

_ , Yeda Wu, Nancy Wang, Hariharan Jayashankar, Michael Bennett, Seyed Moeen Nehzati, Julia Sidorenko, Hyeokmoon Kweon, Grant Goldman, Tamara Gjorgjieva et al., "Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals," *Nature Genetics*, 2022, *54* (4), 437–449.

Palmer, Duncan S., Wei Zhou, Liam Abbott, Nikolas Baya, Claire Churchhouse, Cotton Seed, Tim Poterba, Daniel King, Masahiro Kanai, Alex Bloemendal et al., "Analysis of genetic dominance in the UK Biobank," *bioRxiv*, 2021.

Papageorge, Nicholas W. and Kevin Thom, "Genes, education, and labor market outcomes: evidence from the health and retirement study," *Journal of the European Economic Association*, 2020, *18* (3), 1351–1399.

Pitt, Mark, Mark Rosenzweig, and Md Nazmul Hassan, "Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries," *American Economic Review*, 1990, *80* (5), 1139–56.

Plomin, Robert, John C. DeFries, Valerie S. Knopik, and Jenae M Neiderhiser, "Top 10 replicated findings from behavioral genetics," *Perspectives on Psychological Science*, 2016, *11* (1), 3–23.

Price, Alkes L., Nick J. Patterson, Robert M. Plenge, Michael E. Weinblatt, Nancy A. Shadick, and David Reich, "Principal components analysis corrects for stratification in genome-wide association studies," *Nature Genetics*, 2006, *38* (8), 904.

Restrepo, Brandon J., "Parental investment responses to a low birth weight outcome: who compensates and who reinforces?," *Journal of Population Economics*, 2016, *29* (4), 969–989.

Ronda, Victor, Esben Agerbo, Dorthe Bleses, Preben Bo Mortensen, Anders Børglum, David M. Hougaard, Ole Mors, Merete Nordentoft, Thomas Werge, and Michael Rosholm, "Family disadvantage, gender, and the returns to genetic human capital," *The Scandinavian Journal of Economics*, 2020.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Rosales-Rueda, Maria Fernanda, "Family investment responses to childhood health conditions: Intrafamily allocation of resources," *Journal of Health Economics*, 2014, *37* (C), 41–57.

Savelyev, Peter A., Benjamin C. Ward, Robert F. Krueger, and Matt McGue, "Health endowments, schooling allocation in the family, and longevity: Evidence from US twins," *Journal of Health Economics*, 2022, *81*, 102554.

Taubman, Paul, "The determinants of earnings: Genetics, family, and other environments: A study of white male twins," *The American Economic Review*, 1976, *66* (5), 858–870.

Terskaya, Anastasia, "Parental Human Capital Investment Responses to Children's Disabilities," *Journal of Human Capital*, 2023, *17* (1), 000–000.

Trampush, Joey W., Max Lam Zhan Yang, Jin Yu, Emma Knowles, Gail Davies, David C. Liewald, John M. Starr, Srdjan Djurovic, Ingrid Melle, K. Sundet et al., "GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium," *Molecular Psychiatry*, 2017, *22* (3), 336–345.

Wertz, Jasmin, Terrie E Moffitt, Jessica Agnew-Blais, Louise Arseneault, Daniel W. Belsky, David L. Corcoran, Renate Houts, Timothy Matthews, Joseph A. Prinz, Leah S. Richmond-Rakerd et al., "Using DNA from mothers and children to study parental investment in children's educational attainment," *Child Development*, 2019.

Yi, Junjian, James J. Heckman, Junsen Zhang, and Gabriella Conti, "Early Health Shocks, Intra-household Resource Allocation and Child Outcomes," *The Economic Journal*, 2015, *125* (588), 347–371.

Young, Alexander I., Seyed Moeen Nehzati, Chanwook Lee, Stefania Benonisdottir, David Cesarini, Daniel J. Benjamin, Patrick Turley, and Augustine Kong, "Mendelian imputation of parental genotypes for genome-wide estimation of direct and indirect genetic effects," *BioRxiv*, 2020, pp. 2020–07.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

_ , _ , Stefania Benonisdottir, Aysu Okbay, Hariharan Jayashankar, Chanwook Lee, David Cesarini, Daniel J. Benjamin, Patrick Turley, and Augustine Kong, "Mendelian imputation of parental genotypes improves estimates of direct genetic effects," *Nature Genetics*, 2022, *54*, 897–805.

Zhu, Zhihong, Andrew Bakshi, Anna A.E. Vinkhuyzen, Gibran Hemani, Sang Hong Lee, Ilja M. Nolte, Jana V. van Vliet-Ostaptchouk, Harold Snieder, Tonu Esko, Lili Milani et al., "Dominance genetic variation contributes little to the missing heritability for human complex traits," *The American Journal of Human Genetics*, 2015, *96* (3), 377–385.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Table 1: Parental Preferences for Children's Ability and Fertility Decisions

Endowment of 1st child	High				Low			
Parental preferences for high ability children	Strong		Indifferent		Strong		Indifferent	
Decision to have a 2nd child	no	no	maybe	maybe	yes	yes	maybe	maybe
Endowment of 2nd child (relative to the 1st)			higher	lower	higher	lower	higher	lower

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Table 2: Educational Polygenic Indexes, Years of Schooling and Other Education-Related Indicators

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Education attainment (standardized)			Panel B: PPVT (standardized)			Panel C: Overall GPA (standardized)		
	Uncorrected		Corrected	Uncorrected		Corrected	Uncorrected		Corrected
EA PGI	0.386	0.191	0.369	0.340	0.227	0.373	0.359	0.203	0.369
SE	(0.041)	(0.074)	(0.076)	(0.039)	(0.077)	(0.079)	(0.044)	(0.081)	(0.084)
$p-value$	[<0.001]	[0.010]	[<0.001]	[<0.001]	[0.003]	[<0.001]	[<0.001]	[0.013]	[<0.001]
Parental EA PGI		0.234	0.407		0.132	0.282		0.202	0.363
SE		(0.073)	(0.075)		(0.076)	(0.078)		(0.082)	(0.085)
$p-value$		[0.001]	[<0.001]		[0.083]	[<0.001]		[0.014]	[<0.001]
Baseline controls	*No*	*Yes*	*Yes*	*No*	*Yes*	*Yes*	*No*	*Yes*	Yes
N	604	604	604	572	572	572	452	452	452
R^2	0.149	0.188		0.115	0.126		0.126	0.185	

Note: GPA is the grade point average. PPVT is the Peabody Picture Vocabulary Test score. EA PGI is the educational attainment polygenic index. The table displays OLS coefficients obtained after regressing educational outcomes (standardized to have mean 0 and standard deviation 1) on EA PGI. The regressions in columns 2, 3, 5, 6, 8, 9 include age, age-squared, a female dummy, and parental EA PGI. EA PGI is always standardized to have mean 0 and standard deviation 1. Coefficient estimates and standard errors in columns 3, 6, 9 are measurement-error-corrected as described in Appendix F. Standard errors clustered at the family level are in parentheses.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01326

Table 3: The Effect of Educational Polygenic Index and Sibling Differences in Educational Polygenic Indexes on Parental Investments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Pooled Sample			Non-Twins			Twins		
EA PGI-Sibling's EA PGI	-0.232	-0.233	-0.226	-0.265	-0.270	-0.256	-0.050	-0.043	-0.052
SE	(0.084)	(0.081)	(0.082)	(0.101)	(0.100)	(0.102)	(0.134)	(0.129)	(0.138)
$p-value$	[0.006]	[0.004]	[0.006]	[0.009]	[0.007]	[0.012]	[0.708]	[0.737]	[0.710]
EA PGI	0.205	0.212	0.201	0.168	0.174	0.148	0.245	0.266	0.273
SE	(0.080)	(0.077)	(0.086)	(0.093)	(0.095)	(0.117)	(0.135)	(0.123)	(0.130)
$p-value$	[0.011]	[0.006]	[0.019]	[0.074]	[0.067]	[0.206]	[0.071]	[0.032]	[0.037]
Parental EA PGI	-0.001	-0.001	-0.002	0.011	0.023	0.008	-0.011	-0.014	-0.023
SE	(0.073)	(0.074)	(0.089)	(0.089)	(0.093)	(0.123)	(0.110)	(0.107)	(0.116)
$p-value$	[0.984]	[0.994]	[0.978]	[0.903]	[0.800]	[0.945]	[0.921]	[0.893]	[0.846]
N	604	604	604	412	412	412	192	192	192
Baseline controls	*No*	*Yes*	*Yes*	*No*	*Yes*	*Yes*	*No*	*Yes*	*Yes*
Additional controls	*No*	*No*	*Yes*	*No*	*No*	*Yes*	*No*	*No*	*Yes*

Note: The table reports the estimated effects of sibling differences in EA PGI, own EA PGI, and parental EA PGI on parental investments index (standardized to have mean 0 and standard deviation 1) in the pooled sample, in the sample of non-twins, and in the sample of twins. EA PGI is standardized to have mean 0 and standard deviation 1. Control variable are described in Table H.1. Coefficient estimates and standard errors are measurement-error-corrected as described in Appendix F. Standard errors clustered at the family level are in parentheses.