23. Genes and heritability in intergenerational transmission

Tilbe Atav, Cornelius A. Rietveld and Hans van Kippersluis

INTRODUCTION

Heritability studies have shown that virtually all human traits are to some extent heritable (Turkheimer, 2000; Polderman et al., 2015). This 'First Law' of behavioural genetics has important implications for how differences and inequalities between individuals perpetuate through generations, a question that motivates studies of intergenerational mobility. However, despite the obvious relationship, the literatures on heritability and intergenerational mobility have developed largely independently. In this chapter, we connect these two strands of literature.

We start by introducing a theoretical framework in section 1, which we use to link the concepts of heritability and intergenerational mobility. We explain that conventional measures of heritability and intergenerational mobility both quantify the contribution of specific factors to differences between families.

In section 2, we provide an overview of methods and empirical findings concerning the contribution of genetic factors, making a distinction between family-based and molecular genetics studies. For similar reviews of standard measures of intergenerational mobility we refer to the chapters in Part I of this *Handbook*.[1] Together, these overviews showcase how empirical estimations of heritability and intergenerational mobility often do not account for interplays between different factors or simply assume independence.

In section 3, we discuss the difficulties related to isolating the role played by specific factors, addressing gene–environment interplay in particular. Furthermore, we highlight studies that evaluate such interplays by combining insights from the fields of behavioural genetics and intergenerational mobility.

Throughout this chapter, we focus on the heritability and intergenerational transmission of *socioeconomic status* (SES), defined as the position of an individual in society measured by their educational attainment, income level and occupation. Although we cover all three outcomes in our discussion, the emphasis differs per section depending on the knowledge base.

1. THE INTERGENERATIONAL TRANSMISSION OF SOCIOECONOMIC STATUS

In the economic literature on intergenerational mobility, the Becker-Tomes (Becker & Tomes, 1979, 1986) model often serves as the theoretical backbone for empirical analyses. This model sketches a variety of channels through which intergenerational transmission of socioeconomic characteristics occurs. In the model, parents maximize their utility by choosing how much of their income to consume and how much of it to invest in their child. In addition to these

investments, the capital of the child depends on the endowed capital and luck. In this chapter, we focus on the role of the endowed capital. For a more exhaustive review of the model and its extensions we refer to Chapter 3 by Lindahl in this *Handbook*.

The endowed capital consists of a social endowment (such as the reputation and connection of families), a genetic endowment and other family commodities. Parents cannot invest in the endowment of their children. Instead, the endowment of a child depends on the endowment of the parent, and the heritability thereof (i.e., the fraction that is transmitted to children). A key prediction of the theory is that when parental endowments are higher, the utility of the child is higher both due to a higher level of inherited endowments, as well as due to higher transfers of financial capital. When there are complementarities between endowments and parental investments, additional parental investments could be a final channel through which parental endowments influence a child's outcomes.

Most empirical studies on intergenerational mobility quantify the transmission from parents to children described by the Becker-Tomes model by evaluating the relationship between parental income, education or occupation (Y_p) and the respective child outcome (Y_c).[2] The most widely used measure is the intergenerational elasticity (IGE), which equals the percent change in, for example, the child's income that is associated with a one percent change in parental income.[3] The IGE can subsequently be used to compute the degree of intergenerational mobility ($1 - \text{IGE}$) and the intergenerational correlation ($\text{IGC} = \text{IGE}\frac{\sigma_{Y_p}}{\sigma_{Y_c}}$, with σ denoting the standard deviation).

In an overview paper, Björklund and Jäntti (2012) conclude that these conventional intergenerational associations, often below 0.3 for earnings and around 0.4 for educational attainment, narrowly capture the effect of a single characteristic of a family, typically based on a single parent. Such associations fail to measure the full set of family attributes that are relevant to the persistence of inequality, and, as a result, they underestimate the broader role of family background. A further investigation of the black box of family background or child endowments, therefore, appears crucial for our understanding of how outcomes are related across generations. These assertions are still relevant and are echoed in more recent literature (Björklund & Jäntti, 2020; Cholli & Durlauf, 2022; Mogstad & Torsvik, 2021).

Studies that do intend to capture family background more broadly frequently use sibling correlations to demonstrate that the IGC underestimates the role of family background. To illustrate, Solon (1999) posits a basic model where an outcome Y_{ij} of individual j in family i is decomposed into a shared family component a_i and an individual-specific component b_{ij}:

$$Y_{ij} = a_i + b_{ij} \tag{23.1}$$

Since a_i and b_{ij} are independent by assumption, the variance in the outcome (σ_Y^2) can be decomposed into a shared family component, and a sibling specific variance component driven by idiosyncratic factors: $\sigma_Y^2 = \sigma_a^2 + \sigma_b^2$. It immediately follows that the sibling correlation between sibling j and j' measures the proportion of the variance that is due to shared factors:

$$Corr\left(Y_{ij}, Y_{ij'}\right) = \frac{Cov\left(Y_{ij}, Y_{ij'}\right)}{\sigma_Y^2} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2} \tag{23.2}$$

Intuitively, if shared family characteristics play a larger role in shaping socioeconomic outcomes, siblings should be more similar in their obtained outcomes. Therefore, stronger sibling

correlations imply a perpetuation of inequalities between families as they signify the share of the variance within a generation that is due to overall differences between families.[4]

For the US, estimates of such correlations in earnings and wages were around 0.5 for brothers and 0.35 for sisters. The correlation in educational attainment is around 0.6 for both genders (Mazumder, 2008). Lower sibling correlations were found in the Nordic countries and in Australia (Björklund et al., 2002; Björklund & Jäntti, 2012; Deutscher & Mazumder, 2021). See Solon (1999) and Jäntti and Jenkins (2015) for more comprehensive reviews.

Without further assumptions and data, sibling correlations as in equation (23.2) reflect a mixed bag of genetic and environmental characteristics that siblings share. After all, siblings on average share 50 percent of their genetic make-up, and share many family-related characteristics other than, for example, parental income. Hence, the sibling correlation by itself is uninformative about the contribution of specific factors.

To make inferences about such factors, intergenerational mobility studies further decompose the family component a_i. The intergenerational mobility literature typically decomposes a_i into parental income (or another specific parental characteristic) versus all other family background characteristics uncorrelated with parental income. More specifically, Solon (1999) shows that the sibling correlation can be expressed as the sum of the IGE squared and a component resulting from factors unrelated to parental income. Expectedly, IGE estimates are moderately positively correlated with sibling correlations, but obviously capture only a fraction of the total sibling correlation (Deutscher & Mazumder, 2021). Mazumder (2008) finds that the share of the sibling correlation that is accounted for by the IGE of income equals roughly 36 percent for men in the US. Björklund and Jäntti (2012) find a share of only 8 percent in Sweden, implying that differences between countries in the relative contribution of different factors can be substantial.

While intergenerational mobility studies have extensively assessed descriptive associations and have used advanced statistical techniques to isolate the causal contribution of specific social transmission mechanisms of intergenerational persistence, the role of genetic transmission is often left implicit. In the next section we move on to a discussion of studies that attempt to quantify the relative importance of the genetic component of the variance in an outcome, namely heritability studies.

2. THE HERITABILITY OF SOCIOECONOMIC STATUS

Heritability studies investigate to what extent observed differences in a particular trait can be explained by genetic differences between individuals. Over time, different methodologies have been developed and used to estimate heritability. Early studies are closely related to the sibling correlation framework presented in Section I, as they assess genetic influences indirectly by comparing similarities in outcomes between different types of family members. More recent studies draw on molecular genetic data to estimate the relationship between the genotype of an individual and a specific outcome.

2.1 Family-Based Studies

While outcomes such as educational attainment, income and occupation are observed, not all genetic and environmental factors that affect them are. Family-based studies, therefore, treat

genetic and environmental influences as latent factors. Heritability is estimated by exploiting different levels of environmental and genetic relatedness between family members, often using the so-called ACE model.

The ACE model

Fisher's seminal paper 'The correlation between relatives on the supposition of Mendelian inheritance' can be seen as the foundation of family-based heritability studies (Visscher & Goddard, 2019). Fisher (1918) finds that when a large number of genes contribute to variation in outcomes between individuals, this variance can be decomposed into a genetic component and an environmental component using the trait correlations between relatives. In other words, we can use observed correlations in outcomes between family members to evaluate what share of the differences in outcomes can be explained by differences in genes and what share of the differences can be explained by differences in environment, without directly measuring genetic or environmental factors.

To do so, and using the common notation in the behavioural genetics literature, equation (23.3) further decomposes the shared family component (a_i) and individual-specific component (b_{ij}) of equation (23.1) into additive genetic influences (A), shared environmental influences (C) and non-shared environmental influences (E). The shared environment represents all family, community, social or neighbourhood factors that are shared by a particular type of relative, and that are independent from the genetic factors. The non-shared environmental component captures any variance left unexplained by the genetic and shared environmental components. Therefore:

$$Y = aA + cC + eE \tag{23.3}$$

Under the assumption that the genetic and environmental influences are independent, the variance in outcome Y equals the sum of the variances of the elements on the right-hand side of the equation:

$$\sigma_Y^2 = a^2\sigma_A^2 + c^2\sigma_C^2 + e^2\sigma_E^2 \tag{23.4}$$

Heritability equals the fraction of the variance in outcome Y that can be attributed to the variance in genes: $\frac{a^2}{a^2+c^2+e^2}$ (or simply a^2 when Y is standardized).[5] As A, C, and E are unobserved, the parameters of the model are estimated by comparing covariances between relatives with varying degrees of genetic relatedness. The covariance in Y between relative 1 and 2 equals:

$$\sigma_{Y_1Y_2} = a^2\sigma_{A_1A_2} + c^2\sigma_{C_1C_2} \tag{23.5}$$

The non-shared environmental component, i.e., the residual, disappears from this equation as the covariance between pairs of individuals for this component is zero by construction.

Further assumptions about the genetic ($\sigma_{A_1A_2}$) and environmental ($\sigma_{C_1C_2}$) covariances between related individuals are necessary to estimate a^2 and c^2. Different assumptions can be made for twins, regular siblings, adoptive siblings, and parent–child combinations, etc. The most popular application of the model is the classical twin study, where monozygotic (MZ) twins, who are genetically identical – i.e., $\sigma_{A_1A_2} = 1$ – are compared to dizygotic (DZ) twins, who are as genetically equal as regular siblings. These studies rely on the *random mating* assumption,

which entails that parents have not selected each other based on genetic similarities (i.e., no assortative mating). The implication of this assumption is that $\sigma_{A_1 A_2} = 0.5$ for DZ twins. In addition, classical twin studies rely on the *equal environments* assumption, which states that both twin types have an identical shared environment, i.e., $\sigma_{C_1 C_2} = 1$ for both MZ and DZ twins.

These assumptions ensure that, on average, the only difference in covariance between MZ and DZ twins is the assumed genetic difference. Substituting for the assumptions into equation (23.5) and comparing differences between MZ and DZ twins allows for the estimation of the variance components a^2 and c^2.[6] Studies have shown, however, that such estimates are highly sensitive to changes in the assumptions of the ACE model (Goldberger, 1979). Therefore, these assumptions have been greatly questioned and scrutinized. We refer to Goldberger (1978), Sacerdote (2011), and Evans and Martin (2000) for further discussion of the assumptions and their validity.

Empirical findings

One of the first studies in which the ACE model was applied to socioeconomic outcomes was performed by Taubman (1976). Using a sample of American white male twins, he finds that genetic variation explains 6–50 percent of the variation in earnings and 23–46 percent of the variation in education. The shared environment, on the other hand, accounts for 4–18 percent and 27–30 percent of the variation in earnings and schooling, respectively.

Studies using much larger datasets with information about different types of Swedish siblings find a lower heritability of earnings. Using an age adjusted multi-year average for earnings, Björklund et al. (2005) find that the genetic component accounts for approximately 20 percent of the earnings variation in men and a mere 10 percent in women. The majority of the variation in earnings, around 64 percent, is explained by unshared environmental factors. Using a similar analysis, Cesarini (2010) finds heritability estimates of 0.27, 0.55, 0.72 and 0.42 for income, educational attainment, cognitive and non-cognitive skills, respectively. The shared environmental components were quite substantial at approximately 0.15 for educational attainment and cognitive and non-cognitive skills. This factor is found to be modest for income (0.05) in this sample. Furthermore, a lifecycle analysis shows that the difference between MZ and DZ correlations is fairly stable between the ages 26 to 59. At the ages 20 to 25 the differences are somewhat larger, resulting from a stronger impact of the shared environment.

Heritability estimates can vary substantially across outcomes, and across different measures of the same SES outcome (Sacerdote, 2007). For example, heritability estimates for educational attainment in studies that use test scores diverge widely. Test scores for a variety of subjects in primary school were shown to be between 44–73 percent heritable. The shared environment component estimates were between 10–23 percent (De Zeeuw et al., 2015). Pooling several samples of twins, Silventoinen et al. (2020) find a genetic variance component of 43 percent and a shared environmental component of 31 percent for educational attainment in terms of years of schooling. Heritability was higher for men and for older cohorts (born before 1950), but relatively similar across regions. An earlier meta-analysis did find some differences between countries (Branigan et al., 2013). These findings suggest that various factors affect the heritability of a trait.

2.2 Molecular Genetics Studies

The family-based studies discussed thus far were designed to work around the very limited availability of direct measurements of genes. The completion of the Human Genome project in the early 2000s marked a fundamental change in the accessibility of such data. The increasing availability of genotyped samples allows for the implementation of molecular genetic studies in fields within the social sciences (Benjamin et al., 2012; Conley, 2016).

Methodology

Like the family-based studies, molecular genetic studies exploit genetic differences across individuals. While the human genome, comprising approximately three billion nucleotide pairs, is for a large part identical between individuals, some population variation exists. S*ingle nucleotide polymorphisms* (SNPs) are the most commonly analysed type of genetic variation. A SNP entails a difference in nucleotides at a specific location in the human genome. That is, a SNP is a location in the DNA where one individual has, for example, inherited the adenine nucleotide, while another individual has inherited the cytosine nucleotide. We refer to Beauchamp et al. (2011) and Biroli et al. (2022) for more information about SNPs.

The toolkit for genomic data analysis has grown rapidly as many new methods have been developed in the past two decades. Importantly, Yang et al. (2010) developed Genome-based Restricted Maximum Likelihood (GREML) estimation to assess the proportion of variance in a trait that is explained by genotyped SNPs. Rather than drawing on genetic relationships between family members, Yang et al. (2010) exploit pairwise genetic relationships between *unrelated* individuals as measured by SNP-based similarity. Analysing unrelated individuals neutralizes the impact of shared environmental influences. Under the assumption that there are no remaining confounding environmental influences, GREML estimates represent a lower bound of the trait heritability as they only capture additive genetic effects from commonly genotyped SNPs. Rare variants and non-additive genetic influences are not accounted for. GREML studies have convincingly shown that heritability estimates based on common SNPs equal approximately 50 percent of twin-based heritability estimates (Yang et al., 2017).

Genome-Wide Association Studies (GWASs), on the other hand, attempt to find associations between individual SNPs across the entire genome and a particular outcome using a simple regression as in equation (23.6). Here, y_i represents the outcome, e.g., education, occupation or income. SNP x_{ij} can take values 0, 1, and 2, reflecting the number of minor alleles individual *i* carries of SNP *j.* The minor allele is the nucleotide that is least common in a population. Thus, β_j is the association between SNP *j* and outcome *y*. Z_i is a vector of controls and ε_i is the effect of residual factors which are assumed to be exogenous.

$$y_i = \beta_0 + \beta_j x_{ij} + Z_i\gamma_i + \varepsilon_i \tag{23.6}$$

Because of the large number of SNPs in the human genome, GWASs run this regression for each SNP separately (Visscher et al., 2017). While it is possible that a single SNP causes a particular trait, most traits are polygenic, i.e., influenced by many different SNPs each with a tiny effect size (Chabris et al., 2015). After controlling for the correlation between SNPs (SNPs adjacent to each other on a chromosome tend to be inherited together), these effect sizes are typically summarized into a polygenic index (PGI) by taking the weighted sum of SNP with

weights proportional to the association between the SNP and the outcome of interest as shown in equation (23.7):

$$PGI_i = \sum_{j=1}^{J} \beta_j x_{ij} \tag{23.7}$$

The PGI represents the best linear genetic predictor for a particular outcome (Becker et al., 2021). However, due to measurement error stemming from finite sample bias in the underlying GWAS, the explanatory power of a PGI is usually considerably lower than the SNP-based heritability estimate GREML provides.

In the paragraphs below, we discuss findings of the GREML studies and GWASs. Although we touch on some limitations of molecular genetic studies too, we refer to Benjamin et al. (2012) and Conley (2016) for more comprehensive overviews.

Empirical findings

Rietveld et al. (2013) conducted the first successful GWAS on a SES outcome (education). The authors found three genome wide significant SNPs, and their associations could all be replicated in an independent sample (Rietveld et al., 2014). Each SNP explains about 0.02 percent (approximately one month) of the variance in years of education. The PGI accounts for approximately 2–3 percent of the variation in educational attainment. The effect sizes of individual SNPs are much smaller than what is generally found for physical traits, which suggests that genetic influences on behavioural traits such as educational attainment are driven by a large number of SNPs with relatively smaller effects. This has been coined the 'Fourth Law' of behavioural genetics (Chabris et al., 2015).

Such small effect sizes make it difficult to detect associations in small samples. Subsequent GWASs on educational attainment have shown that, indeed, more SNPs can be identified when using larger samples (Okbay et al., 2016; Lee et al., 2018; Okbay et al., 2022). For example, using a sample of approximately three million individuals, Okbay et al. (2022) find 3,952 significant SNPs which together explain 12–16 percent of the variation in educational attainment. While these estimates are still smaller than GREML estimates of the heritability that centre around 25 percent for educational attainment (e.g., Davies et al., 2016; Okbay et al., 2016; Tropf et al., 2017), the expectation is that by simply expanding the GWAS discovery sample, the predictive power of the PGI will converge to the GREML estimate.

It bears repetition that the SNP-based heritability estimate (25 percent for educational attainment) is often substantially below the heritability estimate obtained using family-based designs (typically around 50 percent for educational attainment). This difference is referred to as the 'missing heritability' (e.g., Mills et al., 2020; Tropf et al., 2017). It is an active and ongoing area of research to understand the causes of this divergence in heritability estimates, and some of the difference is thought to be explained by rare genetic variants not studied in GWAS, and violations of the equal environment assumption in twin studies (Harden, 2021).

Similar gaps exist in studies on measures of occupation and income. While the GREML estimate for self-employment is 25 percent, the PGI for this trait did not capture a significant share of the variation in self-employment (Van der Loos et al., 2013). The GREML estimate for household income is 11 percent (Hill et al., 2016), which is about half of the heritability of income typically found using family-based designs. PGIs from recent GWASs on household

income and individual income, on the other hand, capture 2–3 percent of the variation in these measures (Hill et al., 2019; Kweon et al., 2020).

The PGIs of education have been found to be associated with several different educational, occupational and other behavioural outcomes (Belsky et al., 2016; Mills & Tropf, 2020; Papageorge & Thom, 2020). These associations can be driven by phenotypic as well as genetic relationships. For example, genetic covariance was found between educational attainment and cognitive and non-cognitive skills, as well as a set of anthropometric and behavioural traits (Demange et al., 2021; Marioni et al., 2014; Okbay et al., 2016). Furthermore, studies have found genetic associations between income measures and traits such as educational attainment, health, personality and cognitive ability (Hill et al., 2019; Kweon et al., 2020). Such genetic associations with SES indicators are believed to be mediated through the partial heritability of other traits (Hill et al., 2016), also referred to as vertical pleiotropy.[7]

3. GENES AND ENVIRONMENT IN INTERGENERATIONAL TRANSMISSIONS

In estimating heritability, both family-based and molecular genetic studies explicitly or implicitly assume an additive separation of genetic and environmental factors. Isolating the contribution of genetic and other family and social factors to the intergenerational transmission of an outcome, however, is difficult, as such factors may correlate and interact. Below, we first discuss the implications of such interplay between genetic and environmental factors for heritability estimates. Subsequently, we discuss studies that attempt to isolate the genetic component in intergenerational associations.

3.1 Gene–Environment Interplay

Several studies suggest that genetic factors that are associated with education, occupation and income are correlated to (gene–environment correlation, or rGE) or interacting with (gene–environment interaction, or G×E) environmental factors that affect these outcomes. The implications of this so-called gene-environment (G–E) interplay are complex, and essentially mean that the environment can *confound*, *mediate* and *moderate* genetic effects.

Gene–environment correlations

First, environmental factors may *confound* the genetic effect. The Mendelian laws of inheritance explain that genetic variants of children are mechanically correlated with genetic variants of their parents. What recent studies have shown is that these genetic variants of the parents shape the rearing environment of children, a phenomenon known as 'genetic nurture' (Kong et al., 2018), which can be categorized as a passive rGE process. When not controlling for parental genotype, the rearing environment is a confounder of the genetic association, which can lead to biased estimates of genetic influences. Indeed, as shown by Lee et al. (2018), within-family association estimates were approximately 40 percent lower than between-family GWAS estimates. More generally, confounding of genetic effects can stem from population stratification if differences in the prevalence of genetic factors across (sub)populations and families are correlated with the presence of environmental circumstances. Correspondingly, Domingue et al. (2015) show that controlling for maternal education and neighbourhood atten-

uates the positive relationship between the PGI from Rietveld et al. (2013) and educational attainment.

Second, genetic effects are likely to be *mediated* by environments. Boardman et al. (2012), for example, find genetic homophily within friendship networks. This suggests that genes may drive individuals to self-select into certain environments, also called active rGE. Such selection, however, does not need to be a decision by the individual, and may instead work through institutional factors or from responses of others in the environment to heritable traits, also called evocative rGE.

More generally, any genetic association is always contingent on the environmental and social context where the association is found (Biroli et al., 2022). This implies that genetic effects should not be thought of as purely biological and immutable effects, but instead as a reflection of the prevailing environmental situation that can be mutable.

Gene–environment interactions

Thirdly, environmental factors may *moderate* the effect of genes. Genetic effects can vary greatly between individuals brought up in different types of environments. A meta-analysis on the interaction between genetic effects and childhood SES finds support for the so-called Scarr-Rowe hypothesis, which predicts that the genetic component will be larger for individuals raised in favourable environments (Tucker-Drob & Bates, 2016). Cross-national differences in the estimates imply that institutional factors may drive such interaction effects and suggest that genetic effects are mutable by policy changes. Correspondingly, Heath et al. (1985) find a higher heritability of educational attainment for males subjected to changes in educational policies in Norway, Knigge et al. (2022) show in a Dutch sample that the timing of tracking in secondary education can affect the heritability of academic performance, and Barcellos et al. (2021) find a negative interaction effect between the educational attainment PGI and an increase in the school-leaving age in the UK on wages.

Direction of research

The analysis of such G-E interactions is currently an active area of research (see Pereira et al. (2022) for a review and Mills and Tropf (2020) for an overview of methodological challenges), but the studies reviewed here illustrate how difficult it is to define factors as strictly genetic or strictly environmental: the presence of rGE or G×E severely complicates the interpretation of genetic effects estimated by both family-based and molecular genetic studies. For this reason, a promising direction is the use of family-based molecular genetic studies. These studies conduct, for example, within-family GWASs (Howe et al., 2022) or within-family G×E analyses (e.g., Muslimova et al., 2020), and combine the best of family-based and molecular genetic designs to more cleanly identify the relative contributions of genetic and environmental characteristics.

3.2 The Genetic Component of Intergenerational Associations

Typically, studies in intergenerational transmission explicitly and transparently acknowledge that they are studying intergenerational *correlations*, leaving the need to distinguish between genetic and environmental channels superfluous. Alternatively, scholars have used innovative research designs (see Holmlund et al. (2011), for an overview and comparison of methods) to circumvent the need to quantify the genetic contribution to intergenerational transmission.

Studies that explicitly quantify the genetic component in intergenerational transmission of SES are scarce.

Bowles and Gintis (2002) provide an early exception. The authors explain that the IGC in outcome y can be decomposed into additive contributions of different factors (x):

$$IGC = \sum r_{y_p x} \beta_{y_c x} \tag{23.8}$$

Using a twin-based estimate of heritability and an assumed degree of assortative mating for earnings, they compute the normalized genetic influence on earnings, $\beta_{y_c g_c}$, and the correlation between parental earnings and child genotype, $r_{y_p g_c}$. Their calculations suggest that the genetic component may account for almost a third of the IGC in earnings.

More recent studies use molecular genetic data and PGIs for both parents and children (Conley et al., 2015; Isungset et al., 2022; Liu, 2018). Including the child's educational attainment PGI to estimations of the IGC of educational attainment, these studies find that genetic factors explain 14 to 20 percent of the parent–child correlation, depending on country and measure of educational attainment. Ultimately, these studies confirm earlier findings that both genetic and social parental factors are responsible for intergenerational transmissions.

The use of explicit genetic data in intergenerational transmission studies also opens the door to explore complex interactions between parental and children's genotype and characteristics. Liu (2018) and Isungset et al. (2022) both find that even after controlling for the child's PGI, there is a significant positive effect of the parent's PGI on offspring educational attainment (i.e., so-called genetic nurture). However, while the effect of parental education remained significant and similar in size, conditional on the child's PGI, the effect of the parental PGI disappeared in the Norwegian setting (Isungset et al., 2022) and was attenuated severely in the US setting (Liu, 2018). Furthermore, Conley et al. (2015) find positive interaction effects between the PGIs of the mother and the child in the US, indicating that the advantage of a higher PGI score is larger when the mother's PGI score is high as well. In the Norwegian setting, however, Isungset et al. (2022) do not find any strong interaction effects between child's PGI and parental education or genotype.

In an alternative application, Wertz et al. (2020) find a positive correlation between non-financial parental investments and the education attainment PGI of the mother (child), even after controlling for the PGI of the child (mother), providing evidence for active and evocative rGE. The relationship between educational attainment and such investments was reduced by approximately 8 percent when the child's PGI was controlled for. Interestingly, the results suggest that the family environment is shaped by both parental and child genes.

CONCLUSION

Intergenerational correlations and sibling correlations represent a mixed bag of genetic and environmental effects, making these measures hard to interpret. For this reason studies in both intergenerational mobility as well as behavioural genetics have moved beyond these simple correlations. Intergenerational mobility studies seek to estimate causal effects of parental characteristics in settings using exogeneous shocks or settings in which genetic transmission is absent (e.g., adoption designs). Behavioural genetics studies have moved beyond compari-

sons of simple sibling correlations to richer family designs of twins and extended families to separate genetic and environmental influences.

The first attempts to integrate these two literatures by explicitly adding genetic data to intergenerational mobility studies emerged only very recently. Current evidence suggests that 14 to 20 percent of intergenerational correlations in educational attainment run through the genetic channel. Building upon existing studies, molecular genetic data can be used to address further questions about which environmental and genetic factors drive intergenerational correlations and to what extent. Furthermore, a better understanding of how genes and environment interact is needed. Thus, the role of gene–environment interplay in intergenerational persistence seems particularly pertinent to investigate in future studies.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge funding from NORFACE through the Dynamics of Inequality Across the Life Course (DIAL) programme (GEIGHEI 462–16–100) and from the European Research Council (GEPSI 946647).

NOTES

1. Chapters 3 by Lindahl and 5 by Nybom on, respectively, transmission of human capital and intergenerational income mobility are relevant in particular.
2. See Chapters 2 by Boliver and Capsada-Munsech, 3 by Lindahl, 4 by Heath and Li and 5 by Nybom of this Handbook for a review of the literature regarding the intergenerational transmission of education, human capital, social class, and income, respectively.
3. The intergenerational transmission of an outcome can be defined, measured and evaluated in various ways. For an overview and comparison of alternative measures we refer to Deutscher and Mazumder (2021).
4. Sibling correlations still represent a lower bound estimate of the importance of family background as they do not account for factors that are not shared by siblings. Examples are unshared genetic factors, differential parental treatment and changes across time that might affect siblings differently (Björklund & Jäntti, 2012).
5. In the literature, a^2 is also commonly referred to as h^2 or g^2.
6. We refer to Knopik et al. (2018) for an introduction to the underlying statistical methodology.
7. Pleiotropy entails that one gene can hold direct (horizontal pleiotropy), indirect (vertical pleiotropy) or spurious relationships with multiple traits (Mills & Tropf, 2020).

REFERENCES

Barcellos, S., Carvalho, L., & Turley, P. (2021). The effect of education on the relationship between genetics, early-life disadvantages, and later-life SES. *NBER Working Paper Series,* 28750. https://doi.org/10.3386/w28750

Beauchamp, J. P., Cesarini, D., Johannesson, M., Van der Loos, M. J. H. M., Koellinger, P. D., Groenen, P. J. F. & Christakis, N. A. (2011). Molecular genetics and economics. *Journal of Economic Perspectives*, *25*(4), 57–82. https://doi.org/10.1257/jep.25.4.57

Becker, J., Burik, C. A., Goldman, G., Wang, N., Jayashankar, H., Bennett, M., ... Okbay, A. (2021). Resource profile and user guide of the Polygenic Index Repository. *Nature Human Behaviour*, *5*(1), 1744–1758. https://doi.org/10.1038/s41562-021-01119-3

Becker, G. S., & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy, 87*(6), 1153–1189. https://doi.org/10.1086/260831

Becker, G. S., & Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of Labor Economics*, *4*(3, Part 2), S1–S39. https://doi.org/10.1086/298118

Belsky, D. W., Moffitt, T. E., Corcoran, D. L., Domingue, B., Harrington, H., Hogan, S., … Caspi, A. (2016). The genetics of success: How single-nucleotide polymorphisms associated with educational attainment relate to life-course development. *Proceedings of the National Academy of Sciences*, *27*(7), 957–972. https://doi.org/10.1177/0956797616643070

Benjamin, D. J., Cesarini, D., Chabris, C. F., Glaeser, E. L., Laibson, D. I., Guðnason, V. & Lichtenstein, P. (2012). The promises and pitfalls of genoeconomics. *Annual Review of Economics*, *4*(1), 627–662. https://doi.org/10.1146/annurev-economics-080511-110939

Biroli, P., Galama, T. J., Von Hinke, S., Van Kippersluis, H., Rietveld, C. A., & Thom, K. (2022). The economics and econometrics of gene-environment interplay. *arXiv*, 2203.00729. https://doi.org/10.2139/ssrn.4048082

Björklund, A., Eriksson, T., Jäntti, M., Raaum, O., & Österbacka, E. (2002). Brother correlations in earnings in Denmark, Finland, Norway and Sweden compared to the United States. *Journal of Population Economics*, *15*(4), 757–772. https://doi.org/10.1007/s001480100095

Björklund, A., & Jäntti, M. (2012). How important is family background for labor-economic outcomes? *Labour Economics*, *19*(4), 465–474. https://doi.org/10.1016/j.labeco.2012.05.016

Björklund, A., & Jäntti, M. (2020). Intergenerational mobility, intergenerational effects, sibling correlations, and equality of opportunity: A comparison of four approaches. *Research in Social Stratification and Mobility*, *70*, 100455. https://doi.org/10.1016/j.rssm.2019.100455

Björklund, A., Jäntti, M., & Solon, G. (2005). Influences of nature and nurture on earnings variation: A report on a study of various sibling types in Sweden. In S. Bowles, H. Gintis & M. O. Groves (Eds), *Unequal chances: Family background and economic success* (pp. 145–164). Princeton University Press. https://doi.org/10.1515/9781400835492.145

Boardman, J. D., Domingue, B. W., & Fletcher, J. M. (2012). How social and genetic factors predict friendship networks. *Proceedings of the National Academy of Sciences*, *109*(43), 17377–17381. https://doi.org/10.1073/pnas.1208975109

Bowles, S., & Gintis, H. (2002). The inheritance of inequality. *Journal of Economic Perspectives,* 16(3), 3–30. https://doi.org/10.1257/089533002760278686

Branigan, A. R., McCallum, K. J., & Freese, J. (2013). Variation in the heritability of educational attainment: An international meta-analysis. *Social Forces, 92*(1), 109–140. https://doi.org/10.1093/sf/sot076

Cesarini, D. (2010). *Essays on genetic variation and economic behavior*. MIT Press.

Chabris, C. F., Lee, J. J., Cesarini, D., Benjamin, D. J., & Laibson, D. I. (2015). The fourth law of behavior genetics. *Current Directions in Psychological Science*, *24*(4), 304–312. https://doi.org/10.1177/0963721415580430

Cholli, N. A., & Durlauf., S. N. (2022). Intergenerational mobility. *NBER Working Paper Series*, 29760. https://doi.org/10.3386/w29760

Conley, D. (2016). Socio-genomic research using genome-wide molecular data. *Annual Review of Sociology, 42*(1), 275–299. https://doi.org/10.1146/annurev-soc-081715-074316

Conley, D., Domingue, B. W., Cesarini, D., Dawes, C., Rietveld, C. A., & Boardman, J. D. (2015). Is the effect of parental education on offspring biased or moderated by genotype? *Sociological Science*, *2*(1), 82–105. https://doi.org/10.15195/v2.a6

Davies, G., Marioni, R. E., Liewald, D. C., Hill, W. D., Hagenaars, S. P., Harris, S. E., … Deary, I. J. (2016). Genome-wide association study of cognitive functions and educational attainment in UK Biobank (*N*=112 151). *Molecular Psychiatry, 21*(6), 758–767. https://doi.org/10.1038/mp.2016.45

De Zeeuw, L. E. J., De Geus, E. J. C., & Boomsma, D. I. (2015). Meta-analysis of twin studies highlights the importance of genetic variation in primary school educational achievement. *Trends in Neuroscience and Education*, *4*(3), 69–79. https://doi.org/10.1016/j.tine.2015.06.001

Demange, P. A., Malanchini, M., Mallard, T. T., Biroli, P., Cox, S. R., Grotzinger, A. D., … Nivard, M. G. (2021). Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nature Genetics*, *53*(1), 35–44. https://doi.org/10.1038/s41588-020-00754-2

Deutscher, N. & Mazumder, B. (2021). Measuring intergenerational income mobility: A synthesis of approaches. *Federal Reserve Bank of Chicago Working Paper*, 2021–09. https://doi.org/10.21033/wp-2021-09

Domingue, B. W., Belsky, D., Conley, D., Harris, K. M., & Boardman, J. D. (2015). Polygenic influence on educational attainment: New evidence from the National Longitudinal Study of Adolescent to Adult Health. *AERA Open*, *1*(3), 1–13. https://doi.org/10.1177/2332858415599972

Evans, D. M. & Martin, N. G. (2000). The validity of twin studies. *Genescreen*, *1*(2), 77–79. https://doi.org/10.1046/j.1466-9218.2000.00027.x

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, *52*(2), 399–433. https://doi.org/10.1017/S0080456800012163

Goldberger, A. S. (1978). The genetic determination of income: Comment. *American Economic Review, 68*(5), 960–969.

Goldberger, A. S. (1979). Heritability. *Economica, 46*(184), 327–347. https://doi.org/10.2307/2553675

Harden, K. P. (2021). *The genetic lottery: Why DNA matters for social equality*. Princeton University Press. https://doi.org/10.1515/9780691226705

Heath, A., Berg, K., Eaves, L., Solaas, M. H., Corey, L. A., Sundet, J., … Nance, W. E. (1985). Education policy and the heritability of educational attainment. *Nature*, *314*(6013), 734–736. https://doi.org/10.1038/314734a0

Hill, W. D., Davies, N. M., Ritchie S. J., Skene, N. G., Bryois, J., Bell, S., … Deary, I. J. (2019). Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nature Communications*, *10*, 5741. https://doi.org/10.1038/s41467-019-13585-5

Hill, W. D., Hagenaars, S. P., Marioni, R. E., Harris, S. E., Liewald, D., Davies, G., … Deary, I. J. (2016). Molecular genetic contributions to social deprivation and household income in UK Biobank. *Current Biology*, *26*(22), 3083–3089. https://doi.org/10.1016/j.cub.2016.09.035

Holmlund, H., Lindahl, M., & Plug, E. (2011). The causal effect of parents' schooling on children's schooling – A comparison of estimation methods. *Journal of Economic Literature*, *49*(3), 615–651. https://doi.org/10.1257/jel.49.3.615

Howe, L. J., Nivard, M. G., Morris, T. T., Hansen, A. F., Rasheed, H., Cho, Y., ... & Davies, N. M. (2022). Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nature Genetics, 54*(5), 581–592. https://doi.org/10.1038/s41588-022-01062-7

Isungset, M. A., Conley, D., Zachrisson, H. D., Ystrøm, E., Havdahl, A., Njølstad, P. R., & Lyngstad, T. H. (2022). Social and genetic associations with educational performance in a Scandinavian welfare state. *Proceedings of the National Academy of Sciences, 119*(25), e2201869119. https://doi.org/10.1073/pnas.2201869119

Jäntti, M. & Jenkins, S. P. (2015). Income mobility. In A. B. Atkinson & F. Bourguignon (Eds), *Handbook of income distribution.* vol. 2 (pp. 807–935). North-Holland. https://doi.org/10.1016/B978-0–444-59428-0.00011-4

Knigge, A., Maas, I., Stienstra, K., De Zeeuw, E. L. & Boomsma, D. I. (2022). Delayed tracking and inequality of opportunity: Gene-environment interactions in educational attainment. *NPJ Science of Learning*, *7*, 6. https://doi.org/10.1038/s41539-022-00122-1

Knopik, V. S., Neiderhiser, J. M., DeFries, J. C., & Plomin, R. (2018). *Behavioral genetics*. Worth Publishers.

Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., … Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science, 359*(6374), 424–428. https://doi.org/10.1126/science.aan6877

Kweon, H., Burik, C. A. P., Linnér, R. K., De Vlaming, R., Okbay, A., … Koellinger, P. D. (2020). Genetic fortune: Winning or losing education, income and health. *Tinbergen Institute Discussion Paper*, TI 2020–053/V. https://doi.org/10.2139/ssrn.3682041

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., … Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, *50*(28), 1112–1121. https://doi.org/10.1038/s41588-018-0147-3

Liu H. (2018). Social and genetic pathways in multigenerational transmission of educational attainment. *American Sociological Review, 83*(2), 278–304. https://doi.org/10.1177/0003122418759651

Marioni, R. E., Davies, G., Hayward, C., Liewald, D., Kerr, S. M., Campbell, A., … Deary, I. J. (2014). Molecular genetic contributions to socioeconomic status and intelligence. *Intelligence*, *44*(100), 26–32. https://doi.org/10.1016/j.intell.2014.02.006

Mazumder, B. (2008). Sibling similarities and economic inequality in the US. *Journal of Population Economics*, *21*(3), 685–701. https://doi.org/10.1007/s00148-006-0127-2

Mills, M. C., Barban, N., & Tropf, F. C. (2020). *An introduction to statistical genetic data analysis*. MIT Press.

Mills, M. C. & Tropf, F. C. (2020). Sociology, genetics and the coming of age of sociogenomics. *Annual Review of Sociology*, *46*(1), 553–581. https://doi.org/10.1146/annurev-soc-121919-054756

Mogstad, M., & Torsvik, G. (2021). Family background, neighborhoods and intergenerational mobility. *NBER Working Paper Series*, 28874. https://doi.org/10.3386/w28874

Muslimova, D., Van Kippersluis, H., Rietveld, C. A., Von Hinke, S., & Meddens, S. F. W. (2020). Dynamic complementarity in skill production: Evidence from genetic endowments and birth order. *arXiv*, 2012.05021. https://doi.org/10.2139/ssrn.3748468

Okbay, A., Beauchamp, J., Fontana, M., Lee, J. J., Pers, T. H., Rietveld, C. A., … Benjamin, D. J. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*, *533*(7604), 539–542.

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Moeen, S., … Young, A. I. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics, 54*(4), 437–449. https://doi.org/10.1038/s41588-022-01016-z

Papageorge, N. W. & Thom, K. (2020). Genes, education, and labor market outcomes: Evidence from the health and retirement study. *Journal of the European Economic Association*, *18*(3), 1351–1399. https://doi.org/10.1093/jeea/jvz072

Pereira, R. D., Biroli, P., Galama, T., Von Hinke, S., Van Kippersluis, H., Rietveld, C., & Thom, K. (2022). Gene-environment interplay in the social sciences. In A. Dixit, S. Edwards, and K. Judd (Eds), *Oxford Research Encyclopedia of Economics and Finance*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190625979.013.804

Polderman, T., Benyamin, B., de Leeuw, C., Sullivan, P. F., Van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*(7), 702–709. https://doi.org/10.1038/ng.3285

Rietveld, C. A., Conley, D., Eriksson, N., Esko, T., Medland, S. E., Vinkhuyzen, A. A. E., … Social Science Genetics Association Consortium. (2014). Replicability and robustness of genome-wide-association studies for behavioral traits. *Psychological Science*, *25*(11), 1975–1986. https://doi.org/10.1177/0956797614545132

Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., … Koellinger, P. D. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science*, *340*(6139), 1467–1471.

Sacerdote, B. (2007). How large are the effects from changes in family environment? A study of Korean American adoptees. *Quarterly Journal of Economics, 122*(1), 119–157. https://doi.org/10.1162/qjec.122.1.119

Sacerdote, B. (2011). Nature and nurture effects on children's outcomes: What have we learned from studies of twins and adoptees? In J. Benhabib, A. Bisin & M. O. Jackson (Eds), *Handbook of Social Economics* vol. 1, (pp. 1–30). North-Holland. https://doi.org/10.1016/B978-0–444-53187-2.00001-2

Silventoinen, K., Jelenkovic, A., Sund, R., Latvala, A., Honda, C., Inui, F., … Kaprio, J. (2020). Genetic and environmental variation in educational attainment: An individual-based analysis of 28 twin cohorts. *Scientific Reports, 10*(1), 12681. https://doi.org/10.1038/s41598-020-69526-6

Solon, G. (1999). Intergenerational mobility in the labor market. In O.C. Ashenfelter & D. Card (Eds), *Handbook of labor economics* vol. 3A (pp. 1761–1800). North-Holland. https://doi.org/10.1016/S1573-4463(99)03010-2

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

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

Tucker-Drob, E. M. & Bates, T. C. (2016). Large cross-national differences in gene×socioeconomic status interaction on intelligence. *Psychological Science*, *27*(2), 138–149. https://doi.org/10.1177/0956797615612727

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science*, *9*(5), 160–164. https://doi.org/10.1111/1467-8721.00084

Van der Loos, M. J. H. M., Rietveld, C. A., Eklund, N., Koellinger, P. D., Rivadeneira, F., Abecasis, G. R., … Thurik, A. R. (2013) The molecular genetic architecture of self-employment. *PLOS ONE*, *8*(4), e60542. https://doi.org/10.1371/journal.pone.0060542

Visscher, P. M. & Goddard, M. E. (2019). From R.A. Fisher's 1918 paper to GWAS a century later. *Genetics*, *211*(4), 1125–1130. https://doi.org/10.1534/genetics.118.301594

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 years of GWAS discovery: biology, function, and translation. *American Journal of Human Genetics, 101*(1), 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005

Wertz, J., Moffitt, T. E., Agnew-Blais, J., Arseneault, L., Belsky, D. W., Corcoran, D. L. …, Caspi, A. (2020). Using DNA from mothers and children to study parental investment in children's educational attainment. *Child Development, 91*(5), 1745–1761. https://doi.org/10.1111/cdev.13329

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., ... Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, *42*(7), 565–569. https://doi.org/10.1038/ng.608

Yang, J., Zeng, J., Goddard, M. E., Wray, N. R., & Visscher, P. M. (2017). Concepts, estimation and interpretation of SNP-based heritability. *Nature Genetics*, *49*(9), 1304–1310. https://doi.org/10.1038/ng.3941