# Accepted Manuscript

Genetic architecture of socioeconomic outcomes: Educational attainment, occupational status, and wealth

Hexuan Liu

PII: S0049-089X(19)30126-7

DOI: https://doi.org/10.1016/j.ssresearch.2019.04.008

Reference: YSSRE 2295

To appear in: *Social Science Research*

Received Date: 8 February 2019

Revised Date: 19 March 2019

Accepted Date: 18 April 2019

Please cite this article as: Liu, H., Genetic architecture of socioeconomic outcomes: Educational attainment, occupational status, and wealth, *Social Science Research* (2019), doi: https://doi.org/10.1016/j.ssresearch.2019.04.008.

# Genetic Architecture of Socioeconomic Outcomes: Educational Attainment, Occupational Status, and Wealth

Hexuan Liu[1,2]

[1] *School of Criminal Justice, the University of Cincinnati*

[1] *Institute for Interdisciplinary Data Science, the University of Cincinnati*

Corresponding author:

Hexuan Liu
CECH, School of Criminal Justice
2610 McMicken Circle
660 Dyer Hall
PO Box 210389
Cincinnati, OH 45221-0389
Phone: (513) 556-5827
Email: hexuan.liu@uc.edu

Key words: socioeconomic status; socio-genomics

**ACKNOWLEDGMENTS**

Many thanks go to Guang Guo, Kathleen M. Harris, Francois Nielsen, Michael J. Shanahan, Yang C. Yang, J.C. Barnes, Stephanie Moller, the *SSR* editor, and two reviewers for their helpful comments on the manuscript. An earlier version of this manuscript was presented at the 2016 Integrating Genetics and Social Sciences annual meeting. This research uses data from the HRS, which is sponsored by National Institute on Aging Grants NIA U01AG009740, RC2AG036495, and RC4AG039029, and conducted by the University of Michigan. This research benefits from GWAS results and polygenic scores made publicly available by the Social Science Genetic Association Consortium (SSGAC).

# Genetic Architecture of Socioeconomic Outcomes: Educational Attainment, Occupational Status, and Wealth


## ABSTRACT

This study takes a socio-genomic approach to examine the complex relationships among three important socioeconomic outcomes: educational attainment, occupational status, and wealth. Using more than 8,000 genetic samples from the Health and Retirement study, it first estimates the collective influence of genetic variants across the whole human genome to each of the three socioeconomic outcomes. It then tests genetic correlations among three socioeconomic outcomes, and examines the extent to which genetic influences on occupational status and wealth are mediated by educational attainment. Analyses using the genomic-relatedness-matrix restricted maximum likelihood method show significant genetic correlations among the three outcomes, and provide evidence for both mediated and independent genetic influences. A polygenic score analysis demonstrates the utility of findings in socio-genomic studies to address genetic confounding in causal relationships among the three socioeconomic outcomes.


**Key words**: socioeconomic status; socio-genomics

## INTRODUCTION

As one central concept in social science research, socioeconomic status (SES) is a complex construct comprising multiple traits or outcomes that portray one's position in society, including educational attainment, occupational status, income, and wealth. Conventionally, positive correlations between socioeconomic outcomes have been used to infer causal mechanisms of social reproduction and mobility (Blau and Duncan 1967; Bourdieu and Passeron 1977; Featherman and Hauser 1976; Hout 1988; Mare 1993; Sewell *et al.* 1969; Teachman 1987; Torche 2005). The validity of such causal inferences, however, is threatened by unmeasured between-individual heterogeneity. One important source of between-individual heterogeneity is genetic variation (Stigler 2005). If there is a genetically-related variable that is an independent cause of both outcomes, empirical estimates of causal relationships are likely to be biased (i.e., genetic confounding). Consequently, this may result in ineffective or even detrimental policy recommendations and tremendous wastes of resources (DiPrete *et al.* 2018).

Genetic confounding between two traits assumes three preconditions. First, both traits are heritable. Second, their heritability is, to a large extent, attributable to a common set of genetic factors. Third, shared genetic influences on both traits operate through independent pathways. If all these preconditions are satisfied, ignoring genetic measures would result in omitted variable biases in the analysis. Assessing genetic confounding has proven challenging as few available datasets provide the measures to do so (Barnes *et al*. 2014).

Recent advances in the genomic sciences and technology have enabled researchers to collect genetic markers from large, representative samples. This offers sociologists the opportunity to assess genetic confounding in causal analyses. Drawing on approximately 8,000 genetic samples from the Health and Retirement Study, this study examines the three

preconditions of genetic confounding in relationships among three important socioeconomic outcomes: educational attainment, occupational status, and wealth. Using the genomic-relatedness-matrix restricted maximum likelihood method (GREML), analyses are conducted to (1) estimate the collective contribution of genotypes across the whole genome to each of the three socioeconomic outcomes, (2) assess genetic correlations among the three socioeconomic outcomes, and (3) examine the extent to which genetic influences on occupational status and wealth are mediated by educational attainment. Moreover, polygenic score analyses are conducted to illustrate the utility of polygenic scores in correcting for genetic confounding in causal relationships among the three socioeconomic outcomes.

**BACKGROUND AND AIMS OF THE STUDY**

The causal relationship between education and other socioeconomic outcomes is one of social science's central themes. A large and influential body of research has demonstrated that education produces knowledge and skills that can be translated into higher productivity in the market place; it also helps develop social capital (i.e., relationships and networks) that increases the likelihood of socioeconomic success in later life (Schultz 1961; Becker 1964; Sewell et al. 1969; Jencks 1972, 1979; Sewell *et al.* 1975; Collins 1979; Coleman 1988; Farkas *et al.* 1997; Lin 1999; Grodsky and Pager 2001; Heckman *et al.* 2006). In addition to these explanations, it is possible that links between education and other socioeconomic outcomes partially reflect unmeasured between-individual heterogeneity due to genetic factors (Conley 2001).

How do genes affect one's SES? First, genetic differences may contribute to variation in individual characteristics that are associated with socioeconomic achievement, such as cognitive ability, self-control, interpersonal skills, and financial decision-making abilities (Belsky *et al.* 2016; Barth *et al.* 2018). Second, one's own or proximate others' genes may affect his or her

selection of the environment (i.e., gene-environment correlation, or rGE) (Plomin *et al.* 1977; Scarr and McCartney 1983; Jaffee and Price 2007; Fletcher and Conley 2013; Belsky *et al.* 2018; Liu 2018; Wagner *et al.* 2013). For example, parents may create an educational environment for their children that reflects their own heritable characteristics (Conley and Fletcher 2017). There is also evidence that individuals tend to make friends with, or marry, those who are genetically similar to them (Fowler *et al.* 2011; Boardman *et al.* 2012a; Guo *et al.* 2014; Domingue *et al.* 2014; Conley *et al.* 2016; Domingue *et al.* 2018).

Genetic factors may confound the relationship between two traits under three preconditions: (1) both traits are heritable; (2) the heritability of both traits is largely attributable to common genetic factors; and (3) genetic influences on two traits operate through independent pathways. Accordingly, to investigate genetic confounding in the relationship between educational attainment and other socioeconomic outcomes, it is important to assess the heritability of these outcomes, the extent to which that their heritability is attributed to shared genetic influences, and the extent to which genetic influences on other outcomes are mediated by education.

Genetic influences has been conventionally assessed using twins, adoptees, and other family data (e.g.,(Boardman *et al.* 2010; Boardman *et al.* 2012b; Guo and Stearns 2002; Nielsen 2006; Nielsen and Roos 2015; Turkheimer *et al.* 2003). This approach, however, relies on critical assumptions such as equal environments for identical and fraternal twins and an absence of assortative mating (Goldberger 1979). Recent advances in genomic science and technology have produced a tremendous amount of molecular genetic data. Such data are increasingly available in large-scale social science datasets (e.g., the Health and Retirement Study, the Fragile Families Study, the National Longitudinal Study of Adolescent to Adult Health, and the Wisconsin

Longitudinal Study), and provide social scientists unprecedented opportunities to study heritability without having to rely on family data. Recently, researchers have started using these data in studies of genetic contributions to various traits related to SES, including intelligence and educational outcomes (Domingue *et al.* 2016; Krapohl *et al.* 2015; Lee *et al.* 2018; Marioni *et al.* 2014; Rietveld *et al.* 2013a; Okbay *et al.* 2016; Trzaskowski *et al.* 2014).

*Aim 1 of this study is to assess genetic contributions to educational attainment, occupational status, and wealth using genome-wide genotype data.*

Quantitative geneticists have developed the concept of genetic correlation (rG) to assess the extent to which two traits share the same genetic causes (Neale and Maes 1996). The estimate of an rG between two traits ranges from -1 to 1. An rG of 0 indicates that the effects of genetic variants on one trait are independent of the other; an rG of 1 indicates that all genetic effects on the two traits are identical (i.e., two traits are influenced by overlapping or linked genetic variants); and an rG of -1 indicates that all genetic effects on two traits are completely in opposite directions (i.e., genotypes that increase one trait will decrease the other). rG is important as it provides information that cannot be captured by heritability estimates (Plomin *et al.* 1993). Although two traits may individually have a larger heritability, their genetic correlation can still be low if, for example, both traits are affected by non-overlapping and non-linked genetic variants.

*Aim 2 of this study is to assess genetic correlations among educational attainment, occupational status, and wealth.*

An important source of genetic correlation is pleiotropy, namely that one genetic variant influences multiple traits (Solovieff *et al.* 2013; Pickrell *et al.* 2013). Pleiotropy is complex, and it can occur for various reasons. It is likely that a single gene has multiple but unrelated

biological effects (i.e., biological pleiotropy). For example, a particular genetic variant can be a risk variant for both prostate cancer and colorectal cancer (Manak *et al.* 2009; Wasserman *et al.* 2010). It is also possible that a gene modifies one trait, which in turn, affects another trait (i.e., mediated pleiotropy). As an example, some genetic variants can be associated with both lung cancer (Ahrens *et al.* 2008) and nicotine dependence (Thorgeirsson *et al.* 2008). The genetic variants are likely to influence lung cancer by altering the level of nicotine dependence (Chanock and Hunter 2008).

Understanding pleiotropic effects is critical for causal analyses (Boardman *et al*. 2015; Liu and Guo 2016; Wedow *et al*. 2018). Figure 1 demonstrates four scenarios of pleiotropic effects on three socioeconomic outcomes. Panel A represents a scenario in which genetic causes for three outcomes are disjoint (i.e., no pleiotropy). In this scenario, causal analyses of relationships among the three socioeconomic outcomes will be unbiased if genetic factors are ignored. In the second scenario represented by Panel B, a causal analysis of the effect of education on occupation or wealth also would not suffer from the omitted variable bias if genetic factors are ignored as the genetic influences on occupation and wealth are completely accounted for by education. In the third scenario, represented by Panel C, genetic influences on three socioeconomic outcomes operate independently. Thus, a failure to control for the genetic factors in the analysis will lead to omitted variable biases. In the fourth scenario represented by Panel D, genetic influences on occupation and wealth are partially mediated by education, and partially independent of education. Ignoring genetic factors will also result in omitted variable biases.

Importantly, if the genetic influence on occupation or wealth operates through education, social policies aimed at equalizing educational opportunities should also reduce inequality in occupational status or wealth. If, however, genetic influences on occupation and wealth are

largely independent of education, such policies would be less effective. A good understanding of the pleiotropic effects is useful for predicting the effects of such policies.

*Aim 3 of this study is to examine the extent to which genetic influences on occupational status and wealth are mediated by education.*

[Figure 1 about here]

Recent developments in genome-sequencing technology have opened up a new field of scientific inquiry—socio-genomics (Robinson *et al.* 2005). Socio-genomic studies have provided social scientists with great insight. They have identified specific molecular-level genetic variants associated with individual traits that are related, directly or indirectly, to SES. In particular, three genome-wide association studies (GWAS) have identified more than 1,000 single nucleotide polymorphisms (SNPs) associated with educational outcomes (Rietveld *et al.* 2013a; Okbay *et al.* 2016; Lee *et al.* 2018). Based on the results of these GWA studies, polygenic scores (PGSs) have been developed as compound measures that aggregate estimates of multiple genetic effects on education (Belsky *et al.* 2018; Belsky *et al.* 2016; Conley *et al.* 2015; Conley *et al.* 2016; Conley and Domingue 2016; Domingue *et al.* 2018; Liu and Guo 2015; Trejo *et al.* 2018; Wedow *et al.* 2018). These PGSs offer social scientists opportunities to assess complex causal relationships among socioeconomic variables (Fletcher 2018; Freese 2018).

*Aim 4 of this study is to demonstrate the utility of polygenic scores in addressing genetic confounding in the relationships among educational attainment, occupational status, and wealth.*

Heritability and genetic correlations can be complicated by population stratification. Non-European ancestry samples in HRS are insufficient for GREML analyses that typically require large samples to produce robust results (Guo *et al.* 2015). Moreover, the polygenic score analysis in this study is based on the GWAS findings for individuals of European ancestry. It is

uncertain whether these findings are replicable in other ancestral populations. To minimize confounding effects of population stratification, the analytic sample is limited to individuals who self-reported as non-Hispanic Whites and whose genotypes are consistent with European-ancestry populations.

**DATA**

Data for this study come from the Health and Retirement Study (HRS). HRS is a longitudinal study of Americans over age 50 conducted every two years from 1992 to 2016. It collected information on economic, health, social, and other factors relevant to aging and retirement. HRS includes six birth cohorts: the Study of Assets and Health Dynamics Among the Oldest Old (AHEAD [born before 1924]); Children of Depression (CODA [born 1924-1930]); HRS [born 1931-1941]; War Babies (WB [born 1942-1947]); Early Baby Boomers (EBB [born 1948-1953]); and Mid Baby Boomers (MBB [born 1954-60]).

*Genetic Samples.* DNA samples used in this study were collected from HRS participants between 2006 and 2008 using the Illumina Human Omni-2.5 Quad beadchip, which covers approximately 2.5 million single nucleotide polymorphisms (SNPs). Of these samples, 12,507 passed standardized quality control processes of the University of Washington Genetics Coordinating Center (GCC). Among these samples, 8,653 were from participants of European ancestry.

*Socioeconomic Status Measures.* This study focuses on three SES outcomes: educational attainment, occupational status, and wealth. Educational attainment is measured using years of education ("What is the highest grade of school or year of college you completed?"). Occupational status is based on respondents' current job or the longest job for retirees. Occupational categories are transformed into occupational prestige scores (NORC scale) before

analysis. Wealth is based on household wealth (sum of all types of assets, pensions, etc.). Income and wealth measures are both available in HRS. Wealth was chosen over income because research shows that it is a more accurate measure of SES among older adults (Allin *et al.* 2009).

The SES measures are time-sensitive. For example, a high school degree might indicate high socioeconomic status for older cohorts, but medium/low socioeconomic status for younger cohorts. To address this issue, within-cohort standardization was implemented. Specifically, all three SES measures were recoded into relative indicators based on a baseline sample. The baseline sample includes onset measures for all respondents (whether they provided DNA or not).[1] The baseline sample was used for standardization instead of the analytic sample to minimize potential biases due to sample attrition. Respondents were divided into 10 categories on the basis of the 9 deciles of each of the three SES measures within each birth cohort in the baseline sample.[2]

**METHODS**

*Genomic-relatedness-matrix Restricted Maximum Likelihood Method*

The GREML method is used to estimate the proportion of variance in the outcome that can be explained by SNPs (i.e., SNP heritability) (Yang *et al.* 2011a). This method is based on genetic similarity between unrelated individuals estimated using SNP information. GREML has also been used to estimate the SNP heritability of health and social outcomes including height (Yang *et al.* 2010), body mass index (Yang *et al.* 2011b), schizophrenia (Lee *et al.* 2012a), intelligence (Davies *et al.* 2011), personality traits (Vinkhuyzen *et al.* 2012), subjective well-

[1] Sensitivity analyses were conducted based on wealth measured in different years. The major findings remain the same.

[2] Sensitivity analyses were conducted based on outcomes using different coding strategies (e.g., respondents were divided in to more than 10 categories or less). The major findings remain, suggesting the findings in this paper are robust to different coding strategies.

being (Rietveld *et al.* 2013b), and economic and political preferences (Benjamin *et al.* 2012). The GREML approach has been extended to estimate the genetic correlation (rG) between different traits (Lee *et al.* 2012b). Details of the GREML method can be found in Section A of the Appendix.

*Polygenic Score Method*

While the GREML method is powerful in estimating SNP heritability and genetic correlation, it relies on a latent quantity (genetic variance) instead of an observed measure of genotypes. Therefore, it cannot be used directly to correct for genetic confounding in causal analyses. In contrast, polygenic scores (PGS) are observed measures that aggregate estimates of multiple genetic effects on a phenotype. They can be utilized directly as control variables in regression analyses. PGS are typically based on existing GWAS results. PGSs in this study were constructed using results from the most powerful GWAS on education to date (Lee *et al.* 2018). PGSs are normally distributed in HRS. Larger PGSs are associated with higher levels of educational attainment. Table 1 shows the distribution of the education PGS in HRS. Details of the PGS method can be found in Section B of the Appendix.

[Table 1 about here]

**ANALYTIC STRATEGY**

GREML analyses were conducted to assess the three preconditions of genetic confounding in relationships among the three socioeconomic outcomes. First, univariate GREML analyses were performed to estimate SNP heritability of educational attainment, occupational status, and wealth. Second, bivariate GREML analyses were conducted to estimate genetic correlations between socioeconomic outcomes. Two likelihood ratio tests (LRT) were performed to compare the fitted model and a null model assuming no genetic correlation (i.e.,

genetic effects on the two traits are independent) and the fitted model and a null model assuming perfect genetic correlation (i.e., genetic effects on the two traits are identical). Third, GREML analyses were conducted to estimate genetic variances ($\sigma_g^2$) of occupational status and wealth controlling for educational attainment. A significant drop in the genetic variance would indicate that genetic influences on the other two socioeconomic outcomes are mediated by education. Little to no change in the estimate would indicate that genetic influences on occupation and wealth are independent of education.

To demonstrate the utility of polygenic scores in addressing genetic confounding in causal analyses, analyses were conducted based on the education PGS. The education PGS was included in regression models estimating the association between two socioeconomic outcomes. A significant change in the association would indicate the importance of using genetic measures to reduce the omitted variable bias.

To account for potential population stratification, all the analyses were adjusted for the first 10 principal components computed from the genome-wide SNP data (Price *et al.* 2006). Moreover, GREML analyses require genetically unrelated individuals. Because genetically related individuals often share living environments, including them in the analysis could lead to biased estimates of genomic and environmental contributions (Yang *et al.* 2011). Thus, individuals with a genetic relationship greater than .025 were removed from the analytic sample. That resulted in a sample size of 7,861 in GREML analyses. In the PGS analysis, robust standard error estimates were used to adjust for clustering in HRS (e.g., spousal pairs).

**RESULTS**

*Bivariate correlations*

Table 2 presents bivariate correlations among the key study variables. Consistent with previous research, the three socioeconomic outcomes are positively correlated with each other. Specifically, correlations between education and occupation, education and wealth, and occupation and wealth are respectively .48, .29, and .23 ($p < .01$). The education PGS is significantly associated with all three socioeconomic outcomes.

[Table 2 about here]

*GREML Analysis 1: SNP Heritability of Three Socioeconomic Outcomes*

Table 3 displays univariate GREML results for estimating SNP heritability of educational attainment, occupational status, and wealth. The estimated SNP heritability for education is 40% (*SE* = .06). As indicated by the significant p value ($p < .01$) at the bottom, dropping the genetic component causes a significant loss of information in the model. The estimated SNP heritability is 35% for occupational status (*SE* = .07, $p < .01$) and 31% for wealth (*SE* = .07, $p < .01$). The SNP heritability estimate, however, may vary by sample due to measurement error, gene-environment interaction, or other factors (Boardman *et al.* 2015; Conley *et al.* 2015; Domingue *et al.* 2016; de Vlaming *et al.* 2017; Krapohl and Plomin 2015; Okbay *et al*. 2016; Lee *et al*. 2018; Marioni *et al.* 2014; Rietveld *et al.* 2013a; Tropf *et al.* 2017; Trzaskowski *et al.* 2014).

[Table 3 about here]

*GREML Analysis 2: Genetic Correlations among Three Socioeconomic Outcomes*

Table 4 demonstrates results of bivariate GREML models estimating genetic correlations between socioeconomic outcomes. As the results show, 40% (*SE* = .06, $p < .01$), 35% (SE = .07, $p < .01$), and 32% (*SE* = .07, $p < .01$) of the variances in educational attainment, occupational status, and wealth, respectively, are explained by genome-wide SNPs. These estimates are consistent with the SNP heritability estimates in Table 3 produced by univariate GREML

models. The genetic correlation between education and occupation is estimated to be .73 (*SE* = .10, $p < .01$ in the test for rG = 0 and $p < .01$ in the test for rG = 1), between education and wealth .82 (*SE* = .13, $p < .01$ in the test for rG = 0 and $p = .10$ in the test for rG = 1), and between occupation and wealth .60 (*SE* = .14, $p < .01$ in the test for rG = 0 and $p < .01$ in the test for rG = 1). All genetic correlations are significantly greater than 0, suggesting that the three socioeconomic outcomes are influenced by a common set of genetic variants. Moreover, most genetic correlations are significantly smaller than 1 (with the exception of the genetic correlation between education and wealth), suggesting that genetic effects on the three socioeconomic outcomes are not identical. Some genetic variants that have relatively larger effects on some outcomes may have smaller or no effects on others.

[Table 4 about here]

*GREML Analysis 3: Pleiotropic Effects*

Table 5 shows results of testing for pleiotropic effects. After controlling for education, the estimated genetic variance of occupational status is reduced by 49%, but it remains statistically significant (see Column 1). This result indicates both mediated and independent genetic influences. Specifically, about half of the genetic influence on occupational status operates through education, and the rest is independent of education. The SNP heritability drops to 24% (*SE* = .07, $p < .01$), which means that given education, about a quarter of the remaining variation in occupational status can be explained by genome-wide SNPs.

Similarly, results in Column 2 show that the estimated genetic variance of wealth drops by 43% after controlling for education but remains statistically significant. Again, this result indicates both mediated and independent genetic influences on wealth. The SNP heritability drops to 20% (*SE* = .07, $p < .01$), suggesting that given education, one-fifth of the remaining

variation in wealth can be explained by genome-wide SNPs. The results do not change much when occupational status is included as an additional control variable in the model (see Column 3).

Based on results from the GREML analyses, all three preconditions of genetic confounding are satisfied. It implies that ignoring genetic influences when estimating causal relationships among the three socioeconomic measures will result in omitted variable biases. The next section of results will demonstrate how polygenic scores can be used to correct such biases.

[Table 5 about here]

*Polygenic Score Analysis*

Tables 6 displays results of PGS analyses for occupational status. Consistent with prior studies and the bivariate results, Models 6.1 and 6.2 show that both educational attainment and the education PGS are significantly associated with occupational status ($p < .01$). The Sobel test was conducted to test the significance of the mediating effect of education on the association between the education PGS and occupation (Sobel 1982). After education is controlled, the effect size of the education PGS drops by 70% (Sobel test $p < .01$), but its coefficient remains significant (see Model 6.3). When the education PGS is controlled, the estimate of the association between education and occupation drops by 5%.

[Table 6 about here]

Tables 7 demonstrates results of PGS analyses for wealth. Models 7.1-7.3 show that educational attainment, occupational status, and the education PGS are significantly associated with wealth ($p < .01$). Education mediates about 41% of the association between the education PGS and wealth, and occupation medicates 20% of the association (Sobel test $p < .01$) (see Models 7.4 and 7.5). When both education and occupation are controlled, the coefficient of the

education PGS is reduced by 43% (Sobel test $p < .01$), but remains significant (see Model 7.7). Accordingly, the influence of education-related genotypes on wealth is partially independent of education and occupation. In this case, the association between education or occupation and wealth drops by about 15% after controlling for the education PGS.

[Table 7 about here]

*Supplementary Analysis*

When genotypes were collected in HRS, AHEAD respondents were at least 83 years old and CODA were at least 76 years old. This mean that a member of older cohorts had to live to a much more advanced age to be included in the analytic sample. The retention of healthier, more affluent participants in older cohorts may lead to results that cannot be generalized to a broader population.

To examine the influence of sample attrition, two sensitivity analyses were conducted. First, all models were re-estimated based on a subset of the analytic sample excluding AHEAD and CODA that have a higher proportion of robust survivors. Then, inverse probability weights were calculated for participants in the analytic sample based on variables including birth year, sex, age at entry, educational attainment, smoking, and overall health. Models were re-estimated adjusting for the inverse probability weights (Domingue *et al.* 2017). As a result, most of the adjusted results were similar to the original results (see Tables A1-A5 in the Appendix).

Moreover, GREML analyses were conducted using combined samples of all ancestries. The main results are similar to those based on only samples of European ancestry (see Tables A1-A3 in supplementary materials). This is likely due to the overwhelming majority of European descents in the sample. The relatively small samples of other ancestral groups do not afford sufficient statistical power for stratified GREML analysis. In addition, analyses were also

conducted for males and females separately, but no evidence of systematic gender differences was found.

Finally, research has shown that the association between PGS and outcomes may be attributed to parenting effects (Kong *et al.* 2018; Liu 2018). In this study, associations between the education PGS and three socioeconomic outcomes may be confounded by parental influences related to genetics such as parental education. To examine this possibility, PGS analyses were re-conducted controlling for parental education. Indeed, the effect size of the education PGS was reduced after controlling for parental education. Yet, the results of mediation and genetic confounding analyses were similar to original results (see Tables A4-A5).

**DISCUSSION AND CONCLUSIONS**

This study makes significant substantive and methodological contributions to social stratification and mobility research. Sociologists have long been interested in structural influences on stability and mobility of socioeconomic status over the life course and across generations (Blau and Duncan 1967; Bourdieu and Passeron 1977; Featherman and Hauser 1976; Hout 1988; Mare 1993; Sewell *et al.* 1969; Teachman 1987; Torche 2005). Understanding the genetic architecture of socioeconomic outcomes is crucial for such research. Importantly, the genetic architecture of socioeconomic outcomes is socially structured. In an open society with more opportunities for mobility, the heritability of socioeconomic outcomes is expected to be higher. In other words, the heritability of socioeconomic outcomes is a measure of social mobility (Adkins and Guo 2008; Guo and Stearns 2002; Nielsen 2006; Nielsen and Roos 2015). Moreover, causal inference of relationships among socioeconomic outcomes is subject to genetic confounding. If two different socioeconomic outcomes are genetically correlatd, then the estimate of causal relationship between them is likely to be biased.

This study first examines the overall contribution of genetic variants across the human genome to the three outcomes. GREML analyses show that genome-wide SNPs explain respectively 40%, 35%, and 33% of the variations in educational attainment, occupational status, and wealth in HRS. Since education is typically completed in early adulthood and occupation and wealth play a more important role in determining SES later in life, the declining SNP heritability may imply that as an individual ages, his or her SES is more likely a consequence of contextual opportunities and constraints, rather than intrinsic characteristics determined by genes. This hypothesis, however, cannot be directly tested due to a lack of consistent life-course SES measures in HRS.[3] Future research may test this hypothesis using data sources that include such measures. Another possible source of the variations in the SNP heritability estimates might be measurement errors in the outcomes. Larger measurement errors may result in downward bias in the estimation of SNP heritability.

Second, this study investigates genetic correlations among different socioeconomic outcomes. The results demonstrate that educational attainment, occupational status, and wealth are, to a large extent, associated with overlapping or linked genetic variants. In particular, the genetic correlation between education and occupation (.73) and that between education and wealth (.82) are greater that the genetic correlation between occupation and wealth (.60). Despite the potential influence of measurement errors, this result highlights the central role of education in an individual's socioeconomic trajectory across the life course. Education involves the navigation of a complex institutional setting during childhood, adolescence, and young adulthood. Importantly, this is likely the time in life when one's genes may evoke particular

[3] HRS includes longitudinal measures of income and wealth in late adulthood, but no such measures are available for earlier stages in the life course (i.e., before age 50).

responses and subsequent environment. This is consistent with recent research showing that genetic correlation is situated in and shaped by social context (Wedow *et al.* 2018).

Third, this study provides evidence that, after controlling for education, estimates of genetic variances of occupation and wealth drop by 40-50%, yet remain significant. These results are indicative of a mixture of mediated and independent genetic influences in relationships among the three outcomes (see Panel D in Figure 1). Put together, these findings suggest that all three preconditions of genetic confounding are satisfied in the relationships among educational attainment, occupational status, and wealth. It implies that omitting genetic control variables may result in biases in the estimation of the causal relationships among the three socioeconomic outcomes.

These findings also have important implications for social mobility research. While previous studies have consistently found strong correlations between educational attainment and socioeconomic outcomes in later life, there is a large component in the residual of later-life socioeconomic outcomes not predicted by education (e.g., 75% of the variance in occupational status and 90% of the variance in wealth are not explained by educational attainment in HRS). The analysis in the present study shows that about one-fifth of the variances in the residual of occupational status and wealth (i.e., variances not explained by educational attainment) can be ascribed to genetic factors, and the rest are ascribed to non-genetic factors or measurement errors.

The fourth contribution is methodological. As genetic factors were unobservable, previous analytical strategies developed to correct for genetic confounding typically treated genetic factors as latent variables based on critical assumptions (Conley and Rauscher 2013;

Fujiwara and Kawachi 2009; Kohler *et al.* 2011).[4] Recently available genome-wide genotype data provide sociologists with the opportunity to advance causal methods. This study demonstrates the utility of PGS based on these recent socio-genomic studies. The PGS constructed based on the study of Lee *et al.* (2018) explains up to 15% of the associations among educational attainment, occupational status, and wealth. This indicates that a remarkable proportion of the statistical associations among three socioeconomic outcomes can be attributed to the confounding effects of the genetic variants included in the calculation of the education PGS.

Yet, the PGS in this study is based on genetic variants associated with educational attainment. It is likely that other genetic variants not covered by this PGS may play a role in relationships among the socioeconomic outcomes. More advanced methods may be needed to address genetic confounding (see(Vansteelandt *et al.* 2009; Conley *et al.* 2015; DiPrete *et al.* 2018). In addition, it is important to incorporate non-genetic environmental confounders in the analysis. There are also non-genetic strategies such as instrumental variables to identify a causal effect.

This study can be expanded and improved in several ways. Research has shown that in addition to pleiotropy, genetic correlations may have other causes such as genetic nurturing effects (Kong *et al.* 2018), and cross-trait assortative mating (Keller *et al.* 2013). Genetic nurturing effects refer to effects of parental genotypes that operate through the nurturing environment. Genetic correlation between different traits may rise if these traits are affected independently by parenting factors. Cross-trait assortative mating happens when individuals

[4] The models are typically in a fixed-effects model framework assuming that the dependent variable and the independent variable are influenced by common genotypes and the genetic influence is time invariant.

mate based on different traits. If these traits are heritable, cross-trait assortative mating may cause a statistical correlation between the increasing alleles across the traits, inducing a positive genetic correlation. To examine these possibilities, future research needs to conduct intergenerational analyses using genetic measures from multiple generations.

Moreover, measured SNPs may be differently correlated with the causal variants due to variations in linkage disequilibrium (LD) across ancestral populations (e.g., one measured SNP may be highly correlated with the true but unmeasured causal variant in one population but not in another) (Reich *et al.* 2001). Thus, estimates of SNP heritability and genetic correlation based on the same SNPs might vary across different ancestral populations. There is also evidence that genetic influences and genetic correlations may vary by age, gender, country, birth cohort, and socioeconomic status (Branigan *et al*. 2013; Heath *et al.* 1985; de Vlaming *et al.* 2017; Ge *et al.* 2017; Guo and Stearns 2002; Lee *et al.* 2018; Perry 2016; Tropf *et al*. 2017; Turkheimer *et al.* 2003; Wedow *et al.* 2018). The analytic sample in this study does not afford sufficient statistical power for robust moderation analyses. Future studies may extend the current analysis to investigate social moderation of genetic influence and genetic correlation when more data become available.

Finally, specific mechanisms of genetic effects on SES remain unclear. Although genotypes are fixed at birth, socio-environmental factors may lead to epigenetic modifications that regulate gene expression. Changes in epigenetic mechanisms may lead to changes in individual behavior, which in turn, contribute to changes in SES. Currently large-scale human studies of epigenetic mechanisms are still rare, and the relationship between socioeconomic outcomes and epigenetic modifications is not well understood. In the future, developments in epigenetic technology in conjunction with increasingly available large-scale data will facilitate

our understanding of specific mediating mechanisms through which changes in social context shape genetic influences on traits related to SES.

**REFERENCES**


Adkins, Daniel E., and Guang Guo. 2008. “Societal Development and the Shifting Influence of the Genome on Status Attainment.” *Research in Social Stratification and Mobility* 26(3):235–55.

Ahrens, Wolfgang, Frank Skorpen, Pagona Lagiou, Chu Chen, Triantafillos Liloglou, *et al.* 2008. "A Susceptibility Locus for Lung Cancer Maps to Nicotinic Acetylcholine Receptor Subunit Genes on 15q25." *Nature* 452(7187):633-637.

Allin, Sara, Cristina Masseria, and Elias Mossialos. 2009. "Measuring Socioeconomic Differences in Use of Health Care Services by Wealth Versus by Income." *American Journal of Public Health* 99(10):1849-1855.

Barnes, J. C., Brian B. Boutwell, Kevin M. Beaver, Chris L. Gibson, and John P. Wright. 2014. "On the Consequences of Ignoring Genetic Influences in Criminological Research." *Journal of Criminal Justice* 42(6):471-482.

Barth, Daniel, Nicholas W. Papageorge, and Kevin Thom. 2018. "Genetic Endowments and Wealth Inequality." NBER Working Paper No. w24642. Available at SSRN: https://ssrn.com/abstract=3185896

Becker, G. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* New York: Columbia University Press.

Belsky, Daniel W., Benjamin W. Domingue, Robbee Wedow, Louise Arseneault, Jason D. Boardman, *et al.* 2018. "Genetic Analysis of Social-Class Mobility in Five Longitudinal Studies." *Proceedings of the National Academy of Sciences of the United States of America* 115(31):E7275-7284.

Belsky, Daniel W., Terrie E. Moffitt, David L. Corcoran, Benjamin Domingue, HonaLee Harrington, *et al.* 2016. "The Genetics of Success: How Single-Nucleotide Polymorphisms Associated with Educational Attainment Relate to Life-Course Development." *Psychological Science* 27(7):957-972.

Benjamin, Daniel. J , David Cesarini, Matthijs J. H. M. van der Loos, Christopher T. Dawes, Philipp D. Koellinger, *et al*. 2012. “The Genetic Architecture of Economic and Political Preferences.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (21):8026–8031.

Blau, P. M. and O. D. Duncan. 1967. *The American Occupational Structure*. New York: Wiley.

Boardman, Jason D., Casey L. Blalock, and Fred C. Pampel. 2010. "Trends in the Genetic Influences on Smoking." *Journal of Health and Social Behavior* 51(1):108-123.

Boardman, Jason D., Benjamin W. Domingue, and Jonathan Daw. 2015. "What Can Genes Tell Us About the Relationship between Education and Health?" *Social Science and Medicine* 127:171-80.

Boardman, Jason D. , Benjamin Domingue, W., and Jason Fletcher, M. . 2012a. "How Social and Genetic Factors Predict Friendship Networks." *Proceedings of the National Academy of Sciences* 109(43):17377-17381.

Boardman, Jason D., Michael E. Roettger, Benjamin W. Domingue, Matthew B. McQueen, Brett C. Haberstick, *et al.* 2012b. "Gene–Environment Interactions Related to Body Mass: School Policies and Social Context as Environmental Moderators." *Journal of Theoretical Politics* 24(3):370-388.

Bourdieu, Pierre and Jean-Claude Passeron. 1977. *Reproduction in Education, Society, and Culture.* Beverly Hills, CA: Sage.

Branigan, Amelia R., Kenneth J. McCallum, and Jeremy Freese. 2013. "Variation in the Heritability of Educational Attainment: An International Meta-Analysis." *Social Forces* 92(1):109-140.

Chanock, Stephen J. and David J. Hunter. 2008. "When the Smoke Clears." *Nature* 452(7187):537.

Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94(1988):S95-S120.

Collins, Randall. 1979. *The Credential Society : An Historical Sociology of Education and Stratification*. New York: The Academic Press.

Conley, D. and J. M. Fletcher. 2017. *The Genome Factor:What the Social Genomics Revolution Reveals About Ourselves, Our History, and the Future*. NJ: Princeton University Press.

Conley, D., T. Laidley, D. W. Belsky, J. M. Fletcher, J. D. Boardman, *et al.* 2016. "Assortative Mating and Differential Fertility by Phenotype and Genotype across the 20th Century." *Proceedings of the National Academy of Sciences of the United States of America* 113(24):6647-6652.

Conley, Dalton. 2001. "Decomposing the Black□White Wealth Gap: The Role of Parental Resources, Inheritance, and Investment Dynamics." *Sociological Inquiry* 71(1):39-66.

Conley, Dalton and Benjamin Domingue. 2016. "The Bell Curve Revisited: Testing Controversial Hypotheses with Molecular Genetic Data." *Sociological Science* 3:520-539.

Conley, Dalton, Benjamin W. Domingue, David Cesarini, Christopher Dawes, Cornelius A. Rietveld, *et al.* 2015. "Is the Effect of Parental Education on Offspring Biased or Moderated by Genotype?" *Sociological Science* 2: 82-105.

Conley, Dalton and Emily Rauscher. 2013. "Genetic Interactions with Prenatal Social Environment: Effects on Academic and Behavioral Outcomes." *Journal of Health and Social Behavior* 54(1):109-127.

Davies, G., A. Tenesa, A. Payton, J. Yang, S. E. Harris, D. Liewald, *et al*. 2011. "Genome-wide Association Studies Establish That Human Intelligence Is Highly Heritable and Polygenic." *Molecular Psychiatry* 16(10):996–1005.

de Vlaming, R., A. Okbay, C. A. Rietveld, M. Johannesson, P. K. Magnusson, *et al*. 2017. "Meta-GWAS Accuracy and Power (MetaGAP) Calculator Shows that Hiding Heritability Is Partially Due to Imperfect Genetic Correlations across Studies." *PLoS Genetics* 13(1): e1006495.

DiPrete, Thomas. A., Casper A. P. Burki, and Philipp D. Koellinger. 2018. "Genetic Instrumental Variable Regression: Explaining Socioeconomic And Health Outcomes In Non Experimental Data." *Proceedings of the National Academy of Sciences of the United States of America.*115(22):E4970-E4979.

Domingue, Benjamin W., Daniel W. Belsky, Jason Fletcher, Dalton Conley, Jason D. Boardman, *et al.* 2018. "The Social Genome of Friends and Schoolmates in the National Longitudinal Study of Adolescent to Adult Health." *Proceedings of the National Academy of Sciences of the United States of America* 115(4):702-707.

Domingue, Benjamin W., Jason Fletcher, Dalton Conley, and Jason D. Boardman. 2014. "Genetic and Educational Assortative Mating among Us Adults." *Proceedings of the National Academy of Sciences* 111(22):7996-8000.

Domingue, Benjamin W., Daniel W. Belsky, Amal Harrati, Dalton Conley, David Weir, *et al.* 2017. "Mortality Selection in a Genetic Sample and Implications for Association Studies." *International Journal of Epidemiology* 46(4):1285-1294.

Domingue, Benjamin. W., Robbee Wedow, Dalton Conley, Matt McQueen, Thomas J. Hoffmann, and Jason D. Boardman. 2016. "Genome-wide Estimates of Heritability for Social Demographic Outcomes." *Biodemography and Social Biology* 62(1): 1-18.

Farkas, George, Paula England, Keven Vicknair, and Barbara Stanek Kilbourne. 1997. "Cognitive Skill, Skill Demands of Jobs, and Earnings among Young European American, African American, and Mexican American Workers." *Social Forces* 75(3):913-938.

Featherman, David L. and Robert M. Hauser. 1976. "Changes in the Socioeconomic Stratification of the Races, 1962-73." *American Journal of Sociology* 82(3):621-651.

Fletcher, Jason M. 2018. "Economics and Genomics." O*xford Research Encyclopedia of Economics and Finance* DOI: 10.1093/acrefore/9780190625979.013.14.

Fletcher, Jason M. and Dalton Conley. 2013. "The Challenge of Causal Inference in Gene-Environment Interaction Research: Leveraging Research Designs from the Social Sciences." *American Journal of Public Health* 103(1):S42-S45.

Fowler, James H., Jaime E. Settle, and Nicholas A. Christakis. 2011. "Correlated Genotypes in Friendship Networks." *Proceedings of the National Academy of Sciences of the United States of America* 108(5):1993-1997.

Freese, Jeremy. 2018. "The Arrival of Social Science Genomics." *Contemporary Sociology: A Journal of Reviews* 47(5): 524-536.

Fujiwara, Takeo and Ichiro Kawachi. 2009. "Is Education Causally Related to Better Health? A Twin Fixed-Effect Study in the USA." *International Journal of Epidemiology* 38(5):1310-1322.

Ge, T., C. Y. Chen, B. M. Neale, M. R. Sabuncu, and J. W. Smoller. 2017. "Phenome-Wide Heritability Analysis of the Uk Biobank." *PLoS Genetics* 13(4):1-21.

Goldberger, Arthur S. 1979. "Heritability." *Economica* 46(184):327-347.

Grodsky, Eric and Devah Pager. 2001. "The Structure of Disadvantage: Individual and Occupational Determinants of the Black-White Wage Gap." *American Sociological Review* 66(4):542-567.

Guo, Guang., Ling Wang, Hexuan Liu, and Thomas. Randall. 2014. "Genomic Assortative Mating in Marriages in the United States." *PloS One* 9(11):e112322.

Guo, Guang and Elizabeth Stearns. 2002. "The Social Influences on the Realization of Genetic Potential for Intellectual Development." *Social Forces* 80(3):881-910.

Guo, Guang, Hexuan Liu, Ling Wang, Haipeng Shen, and Wen Hu. 2015. "The Genome-wide Influence on Human BMI by Physical Activity, Life Course, and Historical Period." *Demography* 52(5):1651-1670.

Heath, A. C., K. Berg, L. J. Eaves, M. H. Solaas, L. A. Corey, *et al.* 1985. "Education Policy and the Heritability of Educational Attainment." *Nature* 314(6013):734-736.

Heckman, James J, Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24(3):411-482.

Hout, Michael. 1988. "More Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s." *American Journal of Sociology* 93(May 88):1358-1400.

Jaffee, S. R. and T. S. Price. 2007. "Gene-Environment Correlations: A Review of the Evidence and Implications for Prevention of Mental Illness." *Molecular Psychiatry* 12(5):432-442.

Jencks, C. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books, Inc.

Jencks, Christopher. 1979. *Who Gets Ahead? The Determinants of Economic Success in America*. New York: Basic Books, Inc.

Keller, Matthew C., Christine E. Garver-Apgar, Margaret J. Wright, Nicholas G. Martin, and Robin P. Corley *et al.* 2013. "The Genetic Correlation Between Height and IQ: Shared Genes or Assortative Mating." *PLoS Genetics* 9 (4) e1003451.

Kohler, Hans-Peter, Jere R. Behrman, and Jason Schnittker. 2011. "Social Science Methods for Twins Data: Integrating Causality, Endowments, and Heritability." *Biodemography and Social Biology* 57(1):88-141.

Kong, A., G. Thorleifsson, M. L. Frigge, B. J. Vilhjalmsson, and A. I. Young, et al. 2018. "The Nature of Nurture: Effects of Parental Genotypes." Science 359(6374): 424-428.

Krapohl, E. and R. Plomin. 2015. "Genetic Link between Family Socioeconomic Status and Children's Educational Achievement Estimated from Genome-Wide SNPs." *Molecular Psychiatry* 21:437-443.

Lee, James J., Robbee Wedow, Aysu Okbay, Edward Kong, Omeed Maghzian, *et al.* 2018. "Gene Discovery and Polygenic Prediction from a Genome-Wide Association Study of Educational Attainment in 1.1 Million Individuals." *Nature Genetics* 50(8):1112-1121.

Lee, S. H., T. R. Decandia, S. Pipke, J. Yang; P. F. Sullivan, *et al*. 2012a. "Estimating the Proportion of Variation in Susceptibility to Schizophrenia Captured by Common SNPs." *Nature Genetics* 44 (3):247–50.

Lee, S. H., J. Yang, M. E. Goddard, P. M. Visscher, and N. R. Wray. 2012b. "Estimation of Pleiotropy between Complex Diseases Using Single-Nucleotide Polymorphism-Derived Genomic Relationships and Restricted Maximum Likelihood." *Bioinformatics* 28(19):2540-2542.

Lin, Nan. 1999. "Social Networks and Status Attainment." *Annual Review of Sociology* 25(1):467-487.

Liu, Hexuan. 2018. "Social and Genetic Pathways in Multigenerational Transmission of Educational Attainment." *American Sociological Review* 83(2):278-304.

Liu, Hexuan and Guang Guo. 2015. "Lifetime Socioeconomic Status, Historical Context, and Genetic Inheritance in Shaping Body Mass in Middle and Late Adulthood." *American Sociological Review* 80(4):705-737.

—. 2016. "Opportunities and Challenges of Big Data for the Social Sciences: The Case of Genomic Data." *Social Science Research* 59(2016):13-22.

Manak, J. Robert, Ramesh Shivdasani, Matt Davis, Brian E. Henderson, Gerhard A. Coetzee, *et al.* 2009. "The 8q24 Cancer Risk Variant Rs6983267 Shows Long-Range Interaction with Myc in Colorectal Cancer." *Nature Genetics* 41(8):882-884.

Mare, Robert D. 1980. "Social Background and School Continuation Decisions." *Journal of the American Statistical Association* 75(370):295-305.

—. 1993. "Educational Stratification on Observed and Unobserved Components of Family Background." in *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*, edited by Y.Shavit and H. P. Blossfeld. Boulder, CO: Westview.

Marioni, R. E., G. Davies, C. Hayward, D. Liewald, S. M. Kerr, *et al.* 2014. "Molecular Genetic Contributions to Socioeconomic Status and Intelligence." *Intelligence* 44(1):26-32.

McGue, Matt and Kaare Christensen. 2002. "The Heritability of Level and Rate-of-Change in Cognitive Functioning in Danish Twins Aged 70 Years and Older." *Experimental Aging Research* 28(4):435-451.

Neale, Michael and Hermine Maes. 1996. *Methodology for Genetic Studies of Twins and Families (6th Edition)*. Dordrecht, The Netherlands: Kluwer Academic Publishers B.V.

Nielsen, François. 2006. "Achievement and Ascription in Educational Attainment: Genetic and Environmental Influences on Adolescent Schooling." *Social Forces* 85(1):193-216.

Nielsen, François and J. Micah Roos. 2015. "Genetics of Educational Attainment and the Persistence of Privilege at the Turn of the 21st Century." *Social Forces* 94(2):535-561.

Okbay, Aysu, Jonathan P. Beauchamp, Mark Alan Fontana, James J. Lee, Tune H. Pers, *et al.* 2016. "Genome-Wide Association Study Identifies 74 Loci Associated with Educational Attainment." *Nature* 533(7604):539-542.

Perry, Brea L. 2016. "Gendering Genetics: Social and Biological Contingencies in the Protective Effects of Social Integration for Men and Women." *American Journal of Sociology* 121(6):1655-1696.

Pickrell, Joseph K., Tomaz Berisa, Jimmy Z. Liu, Laure Segurel, Joyce Y. Tung, and David A. Hinds. 2016. "Detection and Interpretation of Shared Genetic Influences on 42 Human Traits." *Nature Genetics* 48(7):709–717.

Plomin, R., J. C. DeFries, and J.C. Loehlin. 1977. "Genotype-environment Interaction and Correlation in the Analysis of Human Behavior". *Psychological Bulletin* 84 (2): 309–322.

Plomin, Robert, Jerome Kagan, Robert N. Emde, J. Steven Reznick, Julia M. Braungart, *et al.* 1993. "Genetic Change and Continuity from Fourteen to Twenty Months: The Macarthur Longitudinal Twin Study." *Child Development* 64(5):1354-1376.

Price, Alkes. L., Nick. J. Patterson, Robert. M. Plenge, Michael. E. Weinblatt, Nancy. A. Shadick, *et al.* 2006. "Principal Components Analysis Corrects for Stratification in Genome-Wide Association Studies." *Nature Genetics* 38(8):904-909.

Reich, David E., Michele Cargill, Stacey Bolk, James Ireland, Pardis C. Sabeti, Daniel J. Richter, Thomas Lavery et al. 2001. "Linkage disequilibrium in the human genome." *Nature* 411 (6834): 199-204.

Rietveld, Cornelius A., Arpana Agrawal, Johan G. Eriksson, Eva Albrecht, Behrooz Z. Alizadeh, *et al.* 2013a. "Gwas of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment." *Science* 340(6139):1467-1471.

Rietveld, Cornelius A., David Cesarinib, Daniel J. Benjamin, Philipp D. Koellinger, Jan-Emmanuel De Neve, *et al*. 2013b. "Molecular Genetics and Subjective Well-Being." *Proceedings of the National Academy of Sciences of the United States of America* 110(24):9692-9697.

Robinson, Gene E., Christina M. Grozinger, and Charles W. Whitfield. 2005. "Sociogenomics: Social Life in Molecular Terms." *Nature Reviews Genetics* 6(4):257-270.

Scarr, Sandra and Kathleen McCartney. 1983. "How People Make Their Own Environments: A Theory of Genotype → Environment Effects." *Child Development* 54(2):424-435.

Scarr, Sandra and Richard A. Weinberg. 1978. "The Influence of "Family Background" on Intellectual Attainment." *American Sociological Review* 43(5):674-692.

Schultz, Theodore W. 1961. "Investment in Human Capital " *American Economic Review* 51(1):1-17.

Sewell, William H., Archibald O. Haller, and Alejandro Portes. 1969. "The Educational and Early Occupational Attainment Process." *American Sociological Review* 34(1):83-92.

Sewell, William H., Robert M. Hauser, and Madison Dept of Rural Sociology Wisconsin Univ. 1975. *Education, Occupation, and Earnings. Achievement in the Early Career*. New York: The Academic Press.

Sobel, Michael E. 1982. "Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models." *Sociological Methodology* 13(290-312.

Solovieff, Nadia, Chris Cotsapas, Phil H. Lee, Shaun M. Purcell, and Jordan W. Smoller. 2013. "Pleiotropy in Complex Traits: Challenges and Strategies." *Nature Reviews Genetics* 14(7):483-495.

Stigler, S. M. 2005. "Correlation and Causation: A Comment." *Perspectives in Biology and Medicine* 48(1 Supplement):88–94.

Teachman, Jay D. 1987. "Family Background, Educational Resources, and Educational Attainment." *American Sociological Review* 52(4):548-557.

Thorgeirsson, T. E., F. Geller, P. Sulem, T. Rafnar, A. Wiste, *et al.* 2008. "A Variant Associated with Nicotine Dependence, Lung Cancer and Peripheral Arterial Disease." *Nature* 452(7187):638-642.

Torche, Florencia. 2005"Unequal But Fluid: Social Mobility in Chile in Comparative Perspective." *American Sociological Review* 70 (3): 422-450.

Trejo, Sam, Daniel W. Belsky, Jason D. Boardman, Jeremy Freese, Kathleen M. Harris, Pam Herd, Kamil Sicinski, and Benjamin W. Domingue. 2018. "Schools as Moderators of Genetic Associations with Life Course Attainments: Evidence form the WLS and Add Health" *Sociological Science* 5:513-540.

Tropf, F. C., S. H. Lee, R. M. Verweij, G. Stulp, P. J. van der Most, R. de Vlaming, *et al.* 72017. "Hidden Heritability Due to Heterogeneity across Seven Populations." *Nature Human Behaviour* 1(10): 757-765.

Trzaskowski, M., N. Harlaar, R. Arden, E. Krapohl, K. Rimfeld, *et al*. 2014. "Genetic Influence on Family Socioeconomic Status and Children's Intelligence. *Intelligence* 42:83–88.

Turkheimer, Eric, Andreana Haley, Mary Waldron, Brian D'Onofrio, and Irving I. Gottesman. 2003. "Socioeconomic Status Modifies Heritability of Iq in Young Children." *Psychological Science* 14(6):623-628.

Vansteelandt, Stijn, Sylvie Goetgeluk, Sharon Lutz, Irwin Waldman, Helen Lyon, *et al.* 2009. "On the Adjustment for Covariates in Genetic Association Analysis: A Novel, Simple Principle to Infer Direct Causal Effects." *Genetic Epidemiology* 33(5):394-405.

Vinkhuyzen, A. A. E, N. L. Pedersen, J. Yang, S. H. Lee, P. K. E. Magnusson, *et al*. 2012. "Common SNPs Explain Some of the Variation in the Personality Dimensions of Neuroticism and Extraversion." *Translational Psychiatry* 2 (e102): 0.1038/tp.2012.49.

Wagner, Brandon, Jiang Li, Hexuan Liu, and Guang Guo. 2013. "Gene-Environment Correlation: Difficulties and a Natural Experiment-Based Strategy." *American Journal of Public Health* 103(1):S167-73.

Wasserman, Nora F., Ivy Aneas, and Marcelo A. Nobrega. 2010. "An 8q24 Gene Desert Variant Associated with Prostate Cancer Risk Confers Differential in Vivo Activity to a Myc Enhancer." *Genome Research* 20(9):1191-1197.

Wedow, Robbee, Meghan Zacher, Brooke M. Huibregtse, Kathleen Mullan Harris, Benjamin W. Domingue, *et al.* 2018. "Education, Smoking, and Cohort Change: Forwarding a Multidimensional Theory of the Environmental Moderation of Genetic Effects." *American Sociological Review* 83(4):802-832.

Yang, J., A., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, *et al*. 2010. “Common SNPs Explain a Large Proportion of the Heritability for Human Height.” *Nature Genetics* 42 (7):565–9.

Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher. 2011a. "GCTA: A Tool for Genome-Wide Complex Trait Analysis." *The American Journal of Human Genetics* 88(1):76-82.

Yang, J., T. A. Manolio, L. R. Pasquale, E. Boerwinkle, N. Caporaso, *et al.* 2011b. “Genome Partitioning of Genetic Variation for Complex Traits Using Common SNPs.” *Nature Genetics* 43 (6):519–25.

**Table 1. Summary Statistics of Key Variables in the HRS Genetic Sample**

| Variable | Mean | SD |
|---|---|---|
| Education (years of school) | 13.16 | 2.57 |
| Occupation (Occupational prestige score) | 45.12 | 12.75 |
| Wealth (thousand dollars) | 527.30 | 1782.49 |
| Education Polygenic Score | -.31 | .16 |

*Note*: Summary statistics of the four variables are based on raw values before standardization.

**Table 2. Bivariate Correlations between Key Variables**

| | Education | Occupation | Wealth | Education PGS |
|---|---|---|---|---|
| Education | 1 | | | |
| Occupation | .48** | 1 | | |
| Wealth | .29** | .23** | 1 | |
| Education PGS | .33** | .22** | .21** | 1 |

*Note*: Bivariate correlations are calculated using standardized variables.
*$p < .05$; **$p < .01$ (two-tailed tests).

**Table 3. Genomic and Environmental Contributions to Three Socioeconomic Outcomes**

| | Education | Occupation | Wealth |
|---|---|---|---|
| Genetic Variance | 3.58(.65) | 2.22(.48) | 2.37(.56) |
| Residual Variance | 5.42(.45) | 4.20(.33) | 5.22(.39) |
| Phenotypic Variance | 9.00(.25) | 6.42(.18) | 7.59(.21) |
| SNP Heritability | .400(.06) | .35(.07) | .31(.07) |
| logL | -12007.98 | -10775.52 | -11460.42 |
| logL0 | -12024.63 | -10766.71 | -11470.17 |
| LRT | 33.30 | 22.38 | 19.51 |
| p-value | .00 | .00 | .00 |
| Sample Size | 7,861 | 7,861 | 7,861 |

*Note*: All models control for the largest 10 principal components for adjusting population stratification. Numbers in parentheses are standard errors. LRT represents the likelihood ratio test statistic for testing if adding genetic measures to the model improves the model fit.

**Table 4. Genetic Correlations among Three Socioeconomic Outcomes**

| | T1: Education T2: Occupation (1) | T1: Education T2: Wealth (2) | T1: Occupation T2: Wealth (3) |
|---|---|---|---|
| Genetic variance | | | |
| T1 | 3.61(.65) | 3.60(.65) | 2.24(.48) |
| T2 | 2.21(.48) | 2.40(.55) | 2.42(.56) |
| Cov(T1, T2) | 2.05(.44) | 2.41(.44) | 1.40(.37) |
| Residual variance | | | |
| T1 | 5.40(.45) | 5.41(.44) | 4.18(.33) |
| T2 | 4.21(.33) | 5.20(.39) | 5.19(.39) |
| Cov(T1, T2) | 1.76(.30) | .40(.31) | .45(.26) |
| Phenotypic variance | | | |
| T1 | 9.01(.25) | 9.01(.25) | 6.43(.18) |
| T2 | 6.42(.18) | 7.60(.21) | 7.61(.21) |
| SNP Heritability | | | |
| T1 | .40(.06) | .40(.06) | .35(.07) |
| T2 | .35(.07) | .32(.07) | .32(.07) |
| rG Test | | | |
| rG | .73(.10) | .82(.13) | .60(.147) |
| p-value (rG=0) | .00 | .00 | .00 |
| p-value (rG=1) | .00 | .10 | .01 |
| Sample Size | 7,861 | 7,861 | 7,861 |

*Note*: All models control for the largest 10 principal components for adjusting population stratification. Numbers in parentheses are standard errors. Two likelihood ratio tests (LRT) are performed to compare the fitted model and a null model assuming no genetic correlation (i.e., rG=0), and the fitted model and a null model assuming perfect genetic correlation (i.e., rG=1).

**Table 5. Pleiotropic Effects on Three Socioeconomic Outcomes**

| | Occupation (controlling for education) | Wealth (controlling for education) | Wealth (controlling for education and occupation) |
|---|---|---|---|
| Genetic Variance | 1.13(.36) | 1.35(.50) | 1.32(.50) |
| Residual Variance | 3.68(.26) | 5.38(.36) | 5.33(.35) |
| Phenotypic variance | 4.81(.14) | 6.73(.19) | 6.65(.19) |
| SNP Heritability | .24 (.07) | .20(.07) | .20(.07) |
| logL | -9778.46 | -11143.74 | -11100.88 |
| logL0 | -9783.45 | -11147.62 | -11104.71 |
| LRT | 9.98 | 7.75 | 7.65 |
| p-value | .00 | .00 | .00 |
| Sample Size | 7,861 | 7,861 | 7,861 |

*Note*: All models control for the largest 10 principal components for adjusting population stratification. Numbers in parentheses are standard errors. LRT represents the likelihood ratio test statistic for testing if adding genetic measures to the model improves the model fit.

**Table 6. Coefficients (Standard Error) of Regression Models Predicting Occupation Status Using Educational Attainment and Polygenic Scores**

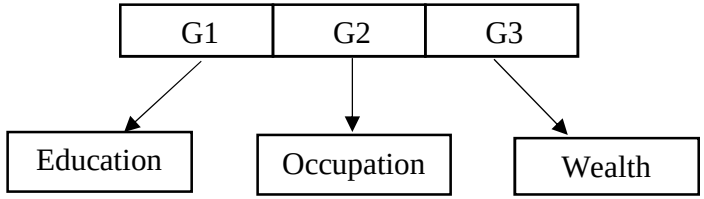
| | Model 6.1 | Model 6.2 | Model 6.3 |
|---|---|---|---|
| Education | .41(.01)** | | .39(.01)** |
| Education PGS | | .52(.03)** | .16(.03)** |
| Adjusted $R^2$ | .23 | .05 | .23 |
| Sample Size | 7,867 | 7,867 | 7,867 |

*Note*: The education PGS is residualized on the first ten principal components to adjust for population stratification. Standard error estimates were corrected for clustering in HRS.
*$p < .05$; **$p < .01$ (two-tailed tests).

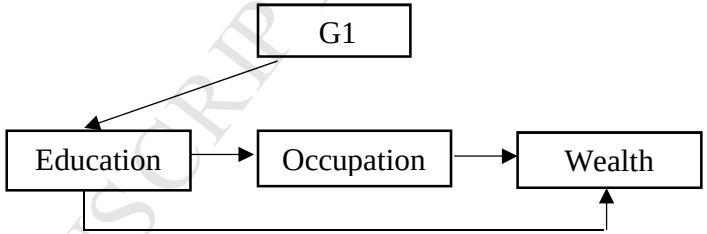
**Table 7. Coefficients (Standard Error) of Regression Models Predicting Wealth Using Educational Attainment, Occupational Status, and Polygenic Scores**

| | Model 7.1 | Model 7.2 | Model 7.3 | Model 7.4 | Model 7.5 | Model 7.6 | Model 7.7 |
|---|---|---|---|---|---|---|---|
| Education | .27(.01)** | | | .23(.01)** | | .22(.01)** | .19(.01)** |
| Occupation | | .25(.01)** | | | .21(.01)** | .13(.01)** | .12(.01)** |
| Education PGS | | | .54(.03)** | .32(.03)** | .43(.03)** | | .31(.03)** |
| Adjusted $R^2$ | .08 | .05 | .04 | .10 | .08 | .10 | .11 |
| Sample Size | 7,867 | 7,867 | 7,867 | 7,867 | 7,867 | 7,867 | 7,867 |

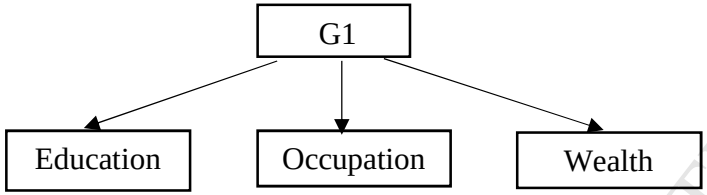
*Note*: The education PGS is residualized on the first ten principal components to adjust for population stratification. Standard error estimates were corrected for clustering in HRS.
*$p < .05$; **$p < .01$(two-tailed tests).



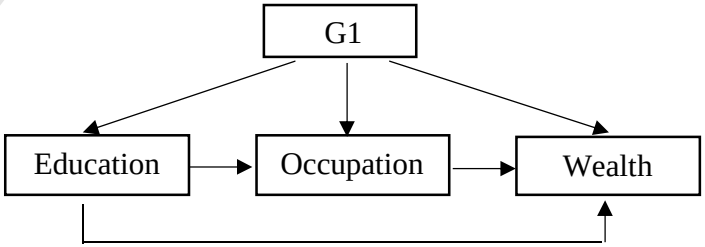

Panel A: No Pleiotropy




Panel B: Mediated Genetic Influences




Panel C: Independent Genetic Influences




Panel D: Mediated and Independent Genetic Influences

**Figure 1. Pleiotropic Effects on Education, Occupation, and Wealth**