SUPPLEMENTAL MATERIAL TO GENETIC ANALYSIS OF SOCIAL-CLASS MOBILITY

DW Belsky & BW Domingue et al.

SECTION 1. SUPPLEMENTAL METHODS

1.1 The Environmental Risk (E-Risk) Longitudinal Twin Study tracks the development of a birth cohort of 2,232 British participants. The sample was drawn from a larger birth register of twins born in England and Wales in 1994-1995 (1). Full details about the sample are reported elsewhere (2). Briefly, the E-Risk sample was constructed in 1999-2000, when 1,116 families (93% of those eligible) with same-sex 5-year-old twins participated in home-visit assessments. This sample comprised 56% monozygotic (MZ) and 44% dizygotic (DZ) twin pairs; sex was evenly distributed within zygosity (49% male). Families were recruited to represent the UK population of families with newborns in the 1990s, on the basis of residential location throughout England and Wales and mother's age. Teenaged mothers with twins were over-selected to replace high-risk families who were selectively lost to the register through non-response. Older mothers having twins via assisted reproduction were under-selected to avoid an excess of well-educated older mothers. The study sample represents the full range of socioeconomic conditions in the UK, as reflected in the families' distribution on a neighborhood-level socioeconomic index (ACORN [A Classification of Residential Neighborhoods], developed by CACI Inc. for commercial use) (3): 25.6% of E-Risk families lived in "wealthy achiever" neighborhoods compared to 25.3% nationwide; 5.3% vs. 11.6% lived in "urban prosperity" neighborhoods; 29.6% vs. 26.9% lived in "comfortably off" neighborhoods; 13.4% vs. 13.9% lived in "moderate means" neighborhoods, and 26.1% vs. 20.7% lived in "hard-pressed" neighborhoods. E-Risk underrepresents "urban prosperity" neighborhoods because such households are likely to be childless.

Home-visits assessments took place when participants were aged 5, 7, 10, 12 and, most recently, 18 years, when 93% of the participants took part. At ages 5, 7, 10, and 12 years, assessments were carried out with participants as well as their mothers (or primary caretakers); the home visit at age 18 included interviews only with participants. Each twin was assessed by a different interviewer. These data are supplemented by searches of official records and by questionnaires that are mailed, as developmentally appropriate, to teachers, and co-informants nominated by participants themselves. The Joint South London and Maudsley and the Institute of Psychiatry Research Ethics Committee approved each phase of the study. Parents gave informed consent and twins gave assent between 5-12 years and then informed consent at age 18.

Genotyping and Imputation. We used Illumina HumanOmni Express 12 BeadChip arrays (Version 1.1; Illumina, Hayward, CA) to assay common single-nucleotide polymorphism (SNP) variation in the genomes of cohort members. The resulting database was restricted to SNPs called successfully in >98% of the cohort and in Hardy-Weinberg equilibrium ($p>0.001$). We imputed additional SNPs using the IMPUTE2 software (Version 2.3.1; https://mathgen.stats.ox.ac.uk/impute/impute_v2.html; (4)) and the 1000 Genomes Phase 3 reference panel (5). Imputation was conducted on autosomal SNPs appearing in dbSNP (Version 140; http://www.ncbi.nlm.nih.gov/SNP/; (6)) that were "called" in more than 98% of the samples. Invariant SNPs were excluded. The E-Risk cohort contains monozygotic twins, who are genetically identical; we therefore empirically measured genotypes of one randomly-selected twin per monozygotic pair and assigned these data to their monozygotic co-twin. We directly measured genotypes of both members of dizygotic twin pairs. Prephasing and imputation were

conducted using a 50-million-base-pair sliding window. The resulting genotype databases included genotyped SNPs and SNPs imputed with 90% probability of a specific genotype among the European-descent members of the E-Risk cohort (N=1,999 participants in 1,011 families). We used the same procedure to construct a genetic database for the twins' mothers (N=859).

Polygenic Scoring**.** We calculated polygenic scores for European-descent E-Risk participants using the PRsice software [v1.22, http://prsice.info/ (7)] based on the Social Science Genetic Association Consortium's most recent published genome wide association study (GWAS) results for educational attainment (https://www.thessgac.org/data) (8). Scores were calculated following the method described by Dudbridge (9) according to the procedure used in previous studies (10, 11). Briefly, SNPs in the E-Risk genotype database were matched to published GWAS results. For each SNP, a loading was calculated as the number of education-associated alleles multiplied by the effect-size estimated in the original GWAS. Loadings were then averaged across the SNP set to calculate the polygenic score. Polygenic scores were calculated with data from all SNPs; no statistical significance threshold was applied to restrict SNPs included in polygenic score analysis.

Social Origins and Social Attainments**.** We analyzed educational mobility in the E-Risk cohort. E-Risk Study members were aged 18 at the time of the most recent follow-up, precluding analysis of occupational outcomes. Parents' educational attainment was measured from structured interviews with E-Risk children's mothers when the children were aged 5 years. Parent educational attainment was measured according to General Certificate of Education (GCSE) level (1,2, or 3) or university degree. We included in the level-3 GCSE category, known as the A-level, parents who held a Higher National Certificate of Higher National Diploma following recommendations of England's Quality Assurance Agency for Higher Education (12). We coded educational origins as the highest level of education achieved by either parent. E-Risk Study members reported their own level of education at the age-18 interview. Education was measured according to General Certificate of Education (GCSE) level (1,2, or 3). Parent and child education values were Z-transformed to have M=0 SD=1 for analysis. Educational mobility analysis regressed Study members' educational attainments on their polygenic scores and their parents' educational attainments.

E-Risk analysis included Study members with available genetic and attainment data (N=1,860) and their mothers (N=804).

1.2 The National Longitudinal Study of Adolescent to Adult Health (Add Health) is an ongoing, nationally-representative longitudinal study of the social, behavioral, and biological linkages in health and developmental trajectories from early adolescence into adulthood (13, 14). The cohort was drawn from a probability sample of 152 middle and high schools and is representative of American adolescents in grades 7-12 in 1994-1995. Since the start of the project, participants have been interviewed in home at four data collection waves (numbered I-IV), most recently in 2007-2008, when 15,701 Study members took part (15).

Genotyping**.** At the Wave IV interview in 2007-2008, saliva and capillary whole blood were collected from respondents. 15,159 of 15,701 individuals interviewed consented to genotyping, and 12,254 agreed to genetic data archiving. DNA extraction and genotyping was conducted on this archive sample using two platforms (Illumina Omni1 and Omni2.5). After quality controls, genotype data were available for 9,974 individuals. We restricted analysis to participants with relatively homogenous European American (N=5,690) and African American (N=1,938) genetic ancestry. Imputation was conducted on SNPs "called" in more than 98% of the samples with minor allele frequency >1% using the Michigan Imputation Server (http://imputationserver.readthedocs.io/en/latest/pipeline/). For the European-descent subsample imputation analysis used the Haplotype Reference Consortium (HRC) reference panel (16). For the African American subsample, imputation analysis used the 1000 Genomes version-3 reference panel (5).

Polygenic Scoring**.** Polygenic scores were calculated for European-descent Add Health participants by the SSGAC using the LDpred software (17) and educational attainment GWAS summary statistics from a meta-analysis that excluded the Add Health data (8). Polygenic scores were calculated for African American Add Health participants using the LDpred software (17) and the public-release version of the educational attainment GWAS summary statistics (https://www.thessgac.org/data). (No African American Add Health participants were included in the educational attainment GWAS.) Polygenic scores were calculated with data from all SNPs; no statistical significance threshold was applied to restrict SNPs included in polygenic score analysis.

Social Origins & Attainments**.** We measured social origins of Add Health Study members from information about their families collected from Add Health participants' parents at the Wave I interview, when participants were still in middle or high school. We examined reports of parental education, parental occupation, household income, and household receipt of public assistance. These four measures were correlated (r>0.19). We conducted principal components analysis of the measures to produce a factor score (18). The first principal component explained 53% of the variance. We used loadings on this component to compute the Social Origins Factor Score. Values were Z-transformed to have M=0 SD =1 for analysis.

We measured social attainments of Add Health Study members from their education and occupations reported on the Wave IV survey, when participants were in their late 20s and early 30s. *Education* was coded in years of completed schooling according to the procedure used by the Social Science Genetic Association Consortium (19). *Occupational attainment* was coded using the Hauser and Warren Occupational Income and Occupational Education scales

(20). The scale scores were averaged to compute a total socioeconomic index. At the Wave IV interview in 2007-8, Add Health respondents reported their current or most recent occupation, which was coded by Add Health using the SOC2000 coding scheme. We converted SOC2000 codes to Occupational Income and Occupational Education scale values using the cross-walk provided by the Wisconsin Center for Demography and Ecology (21). Values were Z-transformed to have M=0 SD =1 for analysis.

Add Health analysis included all Study members with available genetic and attainment data (European-descent N=5,526; African American-descent N=1,814).

1.3 The Dunedin Study is a longitudinal investigation of health and behavior in a complete birth cohort. Study members (N=1,037; 91% of eligible births; 52% male) were all individuals born between April 1972 and March 1973 in Dunedin, New Zealand (NZ), who were eligible based on residence in the province and who participated in the first assessment at age 3. The cohort represents the full range of socioeconomic status on NZ's South Island and matches the NZ National Health and Nutrition Survey on key health indicators (e.g., BMI, smoking, GP visits) (22). The cohort is primarily white; fewer than 7% self-identify as having non-Caucasian ancestry, matching the South Island (22). Assessments were carried out at birth and ages 3, 5, 7, 9, 11, 13, 15, 18, 21, 26, 32, and, most recently, 38 years, when 95% of the 1,007 study members still alive took part. At each assessment, each study member is brought to the research unit for a full day of interviews and examinations.

Genotyping and Imputation**.** The Dunedin Study used Illumina HumanOmni Express 12v1.1 BeadChip arrays (Illumina CA, USA) to assay common Single Nucleotide Polymorphism (SNP) variation in the genomes of our cohort members. The resulting database was restricted to SNPs called successfully in >98% of the cohort and in Hardy-Weinberg equilibrium (p>0.001). Additional SNPs were imputed using the impute2 software (version 2.3.1, https://mathgen.stats.ox.ac.uk/impute/impute_v2.html; (4)) and 1000 Genomes version-3 reference panel (5). Imputation was conducted on autosomal SNPs appearing in dbSNP (v140) that were called in >98% of the Dunedin Study samples. Invariant SNPs were excluded. Pre-phasing and imputation were conducted using a 50M base-pair sliding window. The resulting genotype database included genotyped SNPs and SNPs imputed with 90% probability of a specific genotype among the non-Maori members of the Dunedin cohort (n=918).

Polygenic Scoring**.** We calculated polygenic scores for Dunedin Study participants using the PRsice software [v1.22, http://prsice.info/ (7)] based on the Social Science Genetic Association Consortium's most recent published genome wide association study (GWAS) results for educational attainment (https://www.thessgac.org/data) (8). Scores were calculated following the method described by Dudbridge (9) according to the procedure used in previous studies (10, 11). Briefly, SNPs in the E-Risk genotype database were matched to published GWAS results. For each SNP, a loading was calculated as the number of education-associated alleles multiplied by the effect-size estimated in the original GWAS. Loadings were then

averaged across the SNP set to calculate the polygenic score. Polygenic scores were calculated with data from all SNPs; no statistical significance threshold was applied to restrict SNPs included in polygenic score analysis.

Social Origins and Social Attainments**.** We conducted social mobility analysis in the Dunedin Study using the New Zealand Socioeconomic Index (NZSEI), which was computed for Study member's families when they were children and for the Study members themselves when they were aged 38 years. The NZSEI is a parallel measure to the Hauser and Warren TSEI analyzed in Add Health. NZSEI scores occupations based on income and educational levels of persons with that job in the New Zealand Census. The scale has a range of 1-6 (1 = unskilled laborer, 6 = professional). We used the 1976 NZSEI (23) to code parental occupations assessed when Study members were born and at eight subsequent assessments through their 15th birthdays. The highest occupational status of either parent was averaged across the childhood assessments to compute Study members' social origins NZSEI (24). We used the 2006 NZSEI (25) to code Study members own occupations in adulthood (10). Social origins and attainment score values were Z-transformed to have M=0 SD =1 for analysis.

Dunedin analysis included Study members with available genetic and attainment data (N=831).

1.4 The Wisconsin Longitudinal Study (WLS) is a survey based on a 1/3 sample of all 1957 Wisconsin high school graduates and one randomly-selected sibling (26). The graduate respondents were originally empaneled with an in-person questionnaire at age 18 in 1957, which was followed with data collection at ages 25, 36, 54, 65, and finally 72 in 2011, over which time 18,129 graduates and siblings contributed data. A subset of WLS siblings were first empaneled in 1977 and the remainder in 1994. We included data from siblings born within 15 years of their index graduate sibling's birth year. The WLS includes a wide range of administrative and prospectively collected data from early life through adulthood.

Genotyping & Imputation**.** In 2006-2007 WLS first collected saliva samples from respondents using Oragene kits and a mail-back protocol patterned closely on a previous study (27). An additional sample collection was completed in 2011 for those who did not submit samples in 2006-7. Response rates for DNA collection were similar to those observed for collection of other data in the WLS (28). After quality control, genetic data were available for N=9,012 respondents. Of these respondents, 64% are graduate members of the sample and 36% are siblings of graduates. Full details are available at www.ssc.wisc.edu/wlsresearch/gwas. Imputation was conducted following the same procedure described for Add Health.

Polygenic Scoring**.** Polygenic scores were calculated for European-descent WLS participants by the SSGAC using the LDpred software (17) and educational attainment GWAS summary statistics from a meta-analysis that excluded the WLS data. Polygenic scores were calculated with data from all SNPs; no statistical significance threshold was applied to restrict SNPs included in polygenic score analysis.

Social Origins & Attainments**.** We measured social origins of WLS members using a composite index including parents' education and father's occupation in 1957, and household income during 1957-1960 (29). Parental education and occupation data were reported by participants. Income data was obtained from tax records. Full documentation is available from the WLS (https://www.ssc.wisc.edu/wlsresearch/documentation/appendices/L/cor689.asc). Values were Z-transformed to have M=0 SD=1 for analysis.

We measured socioeconomic attainment for WLS members using measures of occupational attainment collected through 2005 and measures of wealth collected in 2005 and 2011. *Occupational attainment* was coded using the Hauser and Warren Occupational Income and Occupational Education scales (20). Scale scores were averaged to compute a total socioeconomic index. We measured lifetime occupational attainment as the maximum value across all jobs reported. Values were Z-transformed to have M=0 SD =1 for analysis. *Wealth* was measured from structured interviews about assets conducted in 2005 and 2011. Dollar amounts were inflated to 2012 dollars according to the Consumer Price Index (https://www.bls.gov/cpi/) and inverse-hyperbolic-sine transformed to reduce skew (30). Values were averaged across measurements and Z-transformed to have M=0 SD =1 to calculate the final wealth value used for analysis.

WLS analysis included all Study members with available genetic and attainment data (European-descent N=7,111).

1.5 The Health and Retirement Study (HRS) is a longitudinal survey of a representative sample of Americans aged >50 and their spouses initiated in 1992 (31). HRS is administered biennially and includes over 26,000 persons in 17,000 households. Respondents are interviewed about economic, social, and health issues (http://hrsonline.isr.umich.edu/index.php). We analyzed respondents from the 6 primary HRS birth cohorts born 1890-1959.

Genotyping & Imputation**.** The educational attainment polygenic score analyzed in the HRS was computed by SSGAC based on the version-1 HRS genotype dataset posted to dbGap (Study Accession phs000428.v1.p1). Imputation was conducted following the same procedure described for Add Health.

Polygenic Scoring**.** Polygenic scores were calculated for European-descent HRS participants by the SSGAC using the LDpred software (17) and educational attainment GWAS summary statistics from a meta-analysis that excluded the HRS data. Polygenic scores were calculated for African American HRS participants using the LDpred software (17) and the public-release version of the educational attainment GWAS summary statistics (https://www.thessgac.org/data). (No African American HRS participants were included in the educational attainment GWAS.) Polygenic scores were calculated with data from all SNPs; no statistical significance threshold was applied to restrict SNPs included in polygenic score analysis.

Social Origins and Attainments**.** We measured social origins of HRS members from retrospective reports made by the Study members about their childhoods. We examined reports of parental education (coded based on years of reported education as <12, 12-15, or 16+), father's occupation (coded as manual, professional, or management), perceptions of the families socioeconomic circumstances (well-off, average, or poor or variable), and a count of number of hardships the family experienced (father absent from the home, father unemployed, family moved because of financial problems, family received financial assistance, coded as 0, 1, or more than one). These four measures were correlated in our analysis sample (r>0.11). We conducted factor analysis of these four indices to compute a social origins score (18). We conducted principal components analysis of the measures to produce a factor score. The first principal component explained 42% of the variance. We used loadings on this component to compute the Social Origins Factor Score. Values were Z-transformed to have M=0 SD =1 for analysis.

We measured social attainment in the HRS using data on education and household wealth collected from Study members during structured interviews at each of the eleven biannual HRS data collections between 1992 and 2012 and prepared by RAND Corporation (32). Households provided a median of 8 observations of wealth (IQR=5-10). Dollar amounts were inflated to 2012 dollars according to the Consumer Price Index (https://www.bls.gov/cpi/), inverse-hyperbolic-sine transformed to reduce skew (30), and standardized by age. Wealth values were then averaged across measurements and Z-transformed to have M=0 SD =1 to calculate the final value used for analysis.

HRS analysis included Study members with available genetic and attainment data (European-descent N=8,533; African American N=1,609).

1.6 Population Stratification

Population stratification, the non-random patterning of alleles across populations of different ancestry, is a potential confound in genetic association studies (33) and, by extension, polygenic score analysis (34, 35). We addressed population stratification in two ways. First, we conducted analysis separately within groups of participants with homogenous European-American and African-American genetic ancestry. Second, we sought to adjust ancestry-stratified analysis for any residual population stratification using principal components computed from the genome-wide SNP data (36). The first several principal components estimated from genome-wide SNP data are thought to reflect genome-wide patterning of allele frequencies by shared ancestry (37). We adjusted analysis for the first 10 principal components by regressing computed polygenic scores on the principal components within our two population strata and predicting residual values (38). These residuals were then standardized within each stratum to have M=0, SD=1 for analysis. We conducted all analyses separately within each population stratum.

1.7 Analysis

We conducted analysis using regression models. Regression models were fitted separately to data from each study. To test genetic associations with socioeconomic attainments we regressed measures of attainment on the polygenic score and study-specific covariates (Eq1). To test gene-environment correlation with social origins, we used regression in which the dependent variable was the Study member's polygenic score and the predictor of interest was their social origins score (Eq2). To test genetic associations with social mobility, we followed the method used in our previous work (10); we repeated analysis of attainment, adding the social origins measure as a covariate, (Eq3).

1) Attainment: *Socioeconomic Attainment* = b_1*PGS* + X + e
2) Gene-Environment Correlation: *PGS* = b_1*Social Origins* + X + e
3) Social Mobility: *Socioeconomic Attainment* = b_1*PGS* + b_2*Social Origins* + X + e

The notation "X" refers to a set of study-specific covariates and "e" refers to the study-specific error term. Covariates were as follows: Analysis for all studies included covariate adjustment for sex. Add Health, WLS, and HRS analysis included a series of dummy variables encoding birth year as covariates (birth year did not vary for the Dunedin and E-Risk studies). Add Health and WLS included school-based sampling in their designs. Analysis of attainment and mobility in Add Health and WLS included a series of dummy variables encoding the secondary school the participant attended to account for any unmeasured environmental differences at the school level. WLS comprised 1957 high school graduates and their siblings. To account for different sample selection processes for graduates and their siblings, WLS analysis included a dummy variable encoding whether the WLS participant was sampled as a graduate or sibling.

Add Health, WLS and E-Risk analyses included siblings. Analysis of these cohorts used clustered standard errors at the family level to account for non-independence of sibling data (39). HRS analyses included spouses. HRS analysis used clustered standard errors at the family level to account for non-independence of spousal data.

Sibling-difference Analysis. We conducted sibling difference analyses using family-fixed-effects regression (40) (Eq4, in which the subscript 'i' denotes the individual, the subscript 'f' denotes the family, and the bar above the variable name indicates the average value).

4) Sibling Difference: $Socioeconomic\ Attainment_{if} - \overline{Socioeconomic\ Attainment}_f = \beta\left(PGS_{if} - \overline{PGS}_f\right) + (\varepsilon_{if} - \bar{\varepsilon}_f)$

For analyses described in equations 1-3, key predictors and outcomes were standardized to have M=0, SD=1 so that coefficient estimates represent effect-sizes equivalent to Pearson's r. For analysis described in equation 4, we used these same standardizations so that effect-sizes can be compared directly between the full-sample and sibling-difference models.

Percentile-rank Mobility Analysis. We conducted additional mobility analysis based on percentile rankings of participants' social origins and attainments (Eq5). We first computed percentile ranks for measures of social origins and measures of attainments. Next, we computed percentile-rank mobility as the difference between attainment percentile-rank and

social-origins percentile-rank. Finally, we fitted regressions of percentile-rank mobility on the polygenic score, social-origins rank, and study-specific covariates above.

5) Percentile-Rank Mobility: *(Attainment Percentile Rank – Social Origins Percentile Rank)* = b_1*PGS* + b_2*Social Origins Percentile Rank* + X + e

To visualize this analysis descriptively, we constructed matrices of transition probabilities following the approach previously used by Isaacs and colleagues to illustrate effects of educational attainment on income mobility (41). This analysis assigned each participants' social origin score and attainment score a quantile rank within the distribution of the sample in which they were observed. Proportions of participants with each attainment quantile rank were then calculated for each quantile of social origins. Quantiles were defined by GCSE-levels of participants and their parents for E-Risk analysis, by quintiles of participants' social-origins and occupational attainment scores for Add Health and WLS analysis, by NZSEI-levels of participants and their parents for Dunedin analysis, and by participants' social-origins and wealth quintiles for HRS analysis. Transition matrices are reported in **Supplemental Figure 1**.

Mother-child Social Genetic Analysis. We conducted mother-child social genetic analysis to test potential social transmission of genetic effects. The model takes the form

6) Mother-child Social Genetic Analysis: $Attainment_{child} = b_1 PGS_{child} + b_2 PGS_{mother} + X + e$

where the social-genetic effect of mother's genetics on her child's attainment is estimated by b_2, the residual association between a mother's polygenic score and her child's attainment after covariate adjustment for her child's polygenic score.

The mother-child social genetic analysis design exploits the fact that the polygenic scores of mothers and their children are based on measurements of the same genotypes. Typical covariate adjustment is not sufficient to block confounding when the covariate is measured with error. However, the mother-child social genetic design represents a special case of covariate adjustment in which the variable of interest (the mother's polygenic score) and the covariate (the child's polygenic score) are identical measurements; i.e. they are sums of the same genotypes transformed using the same weights. Under the assumption that the weights are measured with the same error for the purposes of scoring the genomes of mothers and their children, any error in measurement will be identical between a mother's polygenic score and her child's polygenic score. In parallel, any true signal measured by the mother's polygenic score will also be measured by her child's polygenic score. Under the assumption that mother and offspring polygenic scores are measured with identical error, the residual association of the mother's polygenic score with her child's attainment is independent of genetics transmitted directly from mother to child.

Mother and offspring polygenic scores might not be measured with the same error because SNPs analyzed in GWAS and used to construct polygenic scores may not be causal variants, i.e. DNA sequence differences that alter the organism's biology in ways that produce differences in educational outcomes. Instead, measured SNPs may be proxy measures or "tags" for such causal variants that are not observed in the SNP data. When a SNP measured in a

GWAS is not a causal variant, the signal detected in the GWAS derives from correlation between the measured SNP and the unmeasured casual variant. To the extent this correlation is <1, it is possible that a SNP measured in a mother may be a better or worse proxy for a given causal variant than the same SNP measured in her child. Such differences may bias estimates of the residual association between a mother's polygenic score and her child's attainment. Under the condition $r(\text{tag-SNP, Casual Variant})_{mother} > r(\text{tag-SNP, Casual Variant})_{child}$, the bias is away from the null and accumulates additively across all loci for which the condition holds. Under the condition $r(\text{tag-SNP, Casual Variant})_{mother} < r(\text{tag-SNP, Casual Variant})_{child}$, an opposite bias should arise.

SECTION 2. Supplemental Analysis

2.1 Supplemental Analysis of Variation in Magnitudes of Genetic Associations with Social Attainment Depending on Social Origins

We tested if the magnitude of associations between Study members' education polygenic scores and their socioeconomic attainments varied depending on their social origins. We re-estimated social mobility models for occupational attainment, this time adding a product term to test the interaction between education polygenic sore and social origins. We included product terms for interactions of the polygenic score and social origins with other model covariates as controls (42). In the Add Health Study, the interaction was positive, indicating the magnitude of association between the polygenic score and attainment was larger for children with better-off social origins as compared to peers born into less-well-off families; each standard-deviation increase in social origins score was associated with a 43% increase in the magnitude of the genetic effect ($p<0.001$). In the Dunedin, WLS, and HRS cohorts, product terms were not statistically different from zero (Dunedin $p=0.459$; WLS $p=0.133$, HRS $p=0.210$; **Supplemental Table S6**).

2.2 Supplemental Analysis in Wisconsin Longitudinal Study subsamples of graduates and participants with non-farm origins

The WLS has a somewhat restricted distribution of educational attainment; participants are high school graduates and their siblings (who may or may not be high school graduates). This compressed educational distribution may influence findings for occupational attainment. To evaluate how such effects might operate, we excluded siblings and repeated analysis only in the subsample of WLS members who were part of the high school graduating class of 1957 (n=5,239 with genetic and attainment data). Consistent with a restriction in the variation of attainment outcomes, effect-sizes estimated for this subsample were somewhat smaller (**Supplemental Table S8**). Thus, the WLS sample, being restricted to high school graduates and their siblings, may provide somewhat attenuated estimates of genetic associations with social mobility.

The WLS also has a larger proportion of participants growing up on farms as compared to our other cohorts. To evaluate how this difference might affect results, we repeated analysis in the non-farm subsample (n=5,630 with genetic and attainment data). Results were similar (**Supplemental Table S8**).

2.3 Supplemental Analysis of Polygenic Score Associations with Attainment and Mobility in African American Participants in the Add Health Study and HRS

We restricted our primary analysis to European-descent participants in the E-Risk, Add Health, Dunedin, HRS, and WLS samples because the educational attainment GWAS was conducted in European-descent individuals. Matching population ancestry between a GWAS used to develop a polygenic score and the sample in which the polygenic score is analyzed is important because ancestry-related genetic variation may confound polygenic prediction (35). Specifically, genetic variants that are not causally related to a phenotype may nevertheless associate with that phenotype in a GWAS because they "tag", or correlate with, a nearby segment of the genome that is causally related to the phenotype. When population ancestry is similar between a GWAS and a polygenic score test sample, GWAS coefficients for such genetic variants still are valid for polygenic score construction. This is because a variant that tags a causal genomic region in the GWAS is likely to do the same in the polygenic score test sample. However, when population ancestry is different between a GWAS and a polygenic score test sample, GWAS coefficients can no longer be assumed to be valid because measured GWAS variants may tag different genomic regions in the test sample as compare to the GWAS sample (43). At minimum, this introduces measurement error into polygenic score analysis.

We conducted exploratory analysis to test if polygenic scores for educational attainment based on GWAS of educational attainment in European-descent individuals (19) would predict variation in social attainment and mobility in African Americans using data from African American participants in the Add Health (N=1,814) and HRS studies (N=1,608). Polygenic score analysis cannot be used to compare genetic distributions between different ancestry groups (34). However, within-population analysis may be possible, with the caveat that measurement error will attenuate estimates. Our analysis aimed to test if this attenuation would be modest enough that future analyses of attainment processes in non-European populations could make use of the education polygenic score. In a previous analysis of siblings in the Add Health Study, we estimated that effect-sizes for the 2013 version of the education polygenic score (44) were attenuated by about half when studied in African Americans (45). Using data from African American Add Health and HRS participants, we repeated our analysis, including tests of polygenic associations with educational and occupational attainment and mobility and gene-environment correlations with social origins.

Across analyses, effect-sizes for polygenic-score associations with attainment and mobility among African Americans were roughly one third to one half as large as effect-sizes estimated in the European-descent Participants. Results are reported in **Supplemental Table S9**.

REFERENCES

1. Trouton A, Spinath FM, Plomin R (2002) Twins Early Development Study (TEDS): A multivariate, longitudinal genetic investigation of language, cognition and behavior problems in childhood. *Twin Res* 5:444–448.

2. Moffitt TE, E-risk Team (2002) Teen-aged mothers in contemporary Britain. *J Child Psychol Psychiatry* 43:727–742.

3. Odgers CL, Caspi A, Bates CJ, Sampson RJ, Moffitt TE (2012) Systematic social observation of children's neighborhoods using Google Street View: a reliable and cost-effective method. *J Child Psychol Psychiatry* 53(10):1009–17.

4. Howie BN, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *Plos Genet* 5:e1000529.

5. 1000 Genomes Project Consortium, et al. (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* 491(7422):56–65.

6. Sherry ST, et al. (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 29(1):308–311.

7. Euesden J, Lewis CM, O'Reilly PF (2015) PRSice: Polygenic Risk Score software. *Bioinforma Oxf Engl* 31(9):1466–1468.

8. Lee J, et al. (2018) Gene discovery and polygenic prediction from a 1.1-million-person GWAS of educational attainment. *Nat Genet* In Press.

9. Dudbridge F (2013) Power and Predictive Accuracy of Polygenic Risk Scores. *PLOS Genet* 9(3):e1003348.

10. Belsky DW, et al. (2016) The Genetics of Success: How Single-Nucleotide Polymorphisms Associated With Educational Attainment Relate to Life-Course Development. *Psychol Sci* 27(7):957–972.

11. Wertz J, et al. (2018) Genetics and Crime: Integrating New Genomic Discoveries Into Psychological Research About Antisocial Behavior. *Psychol Sci*:956797617744542.

12. Quality Assurance Agency for Higher Education, Council for Curriculum, Examinations and Assessment, and Qualifications, Curriculum and Assessment Authority (2000) Finding your way around: A leaflet about the National Qualifications Framework.

13. Harris KM, et al. (2013) Social, Behavioral, and Genetic Linkages from Adolescence Into Adulthood. *Am J Public Health* 103(Suppl 1):S25–S32.

14. Harris KM, Halpern CT, Haberstick BC, Smolen A (2013) The National Longitudinal Study of Adolescent Health (Add Health) Sibling Pairs Data. *Twin Res Hum Genet Off J Int Soc Twin Stud* 16(1):391–398.

15. Harris KM (2013) The Add Health Study: Design and Accomplishments. *Carol Popul Cent Univ N C Chap Hill*. Available at: https://pdfs.semanticscholar.org/6185/b332bff5b1aca7a030150beacd2a8d76fc57.pdf.

16. Consortium the HR, et al. (2016) A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 48(10):1279–1283.

17. Vilhjálmsson BJ, et al. (2015) Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *Am J Hum Genet* 97(4):576–592.

18. Stacklies W, Redestig H, Scholz M, Walther D, Selbig J (2007) pcaMethods--a bioconductor package providing PCA methods for incomplete data. *Bioinforma Oxf Engl* 23(9):1164–1167.

19. Okbay A, et al. (2016) Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 533(7604):539–542.

20. Hauser RM, Warren JR (1997) Socioeconomic indexes for occupations: A review, update, and critique. *Sociol Methodol* 27(1):177–298.

21. Frederick C (2010) *A Crosswalk for using Pre-2000 Occupational Status and Prestige Codes with Post-2000 Occupation Codes* (Cetner for Demography and Ecology, University of Wisconsin-Madison) Available at: https://www.ssc.wisc.edu/cde/cdewp/2010-03.pdf [Accessed September 27, 2016].

22. Poulton R, Moffitt TE, Silva PA (2015) The Dunedin Multidisciplinary Health and Development Study: overview of the first 40 years, with an eye to the future. *Soc Psychiatry Psychiatr Epidemiol* 50(5):679–693.

23. Elley WB, Irving JC (1976) Revised socioeconomic index for New-Zealand. *N Z J Educ Stud* 11(1):25–36.

24. Poulton R, et al. (2002) Association between children's experience of socioeconomic disadvantage and adult health: a life-course study. *Lancet* 360(9346):1640–1645.

25. Milne BJ, Byun U, Lee A (2013) *New Zealand socio-economic index 2006.* (Statistics New Zealand, Wellington, NZ).

26. Herd P, Carr D, Roan C (2014) Cohort profile: Wisconsin longitudinal study (WLS). *Int J Epidemiol* 43(1):34–41.

27. Rylander-Rudqvist T, Håkansson N, Tybring G, Wolk A (2006) Quality and Quantity of Saliva DNA Obtained from the Self-administrated Oragene Method—A Pilot Study on the Cohort of Swedish Men. *Cancer Epidemiol Prev Biomark* 15(9):1742–1745.

28. Hauser RM (2005) Survey Response in the Long Run: The Wisconsin Longitudinal Study. *Field Methods* 17(1):3–29.

29. Hauser R, Sheridan J, Warren J (1999) Socioeconomic Achievements of Siblings in the Life Course: New Findings from the Wisconsin Longitudinal Study. *Res Aging* 21(2):338–378.

30. Friedline T, Masa RD, Chowa GAN (2015) Transforming wealth: Using the inverse hyperbolic sine (IHS) and splines to predict youth's math achievement. *Soc Sci Res* 49(Supplement C):264–287.

31. Sonnega A, et al. (2014) Cohort Profile: the Health and Retirement Study (HRS). *Int J Epidemiol* 43(2):576–585.

32. Chien S, et al. (2014) RAND HRS data documentation, version n. *Labor Popul Program RAND Cent Study Aging*.

33. Hamer D, Sirota L (2000) Beware the chopsticks gene. *Mol Psychiatry* 5(1):11–13.

34. Belsky DW, Israel S (2014) Integrating genetics and social science: genetic risk scores. *Biodemography Soc Biol* 60(2):137–155.

35. Martin AR, et al. (2017) Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am J Hum Genet* 100(4):635–649.

36. Price AL, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904–909.

37. Price AL, Zaitlen NA, Reich D, Patterson N (2010) New approaches to population stratification in genome-wide association studies. *Nat Rev Genet* 11(7):459–463.

38. Conley D, et al. (2016) Assortative mating and differential fertility by phenotype and genotype across the 20th century. *Proc Natl Acad Sci* 113(24):6647–6652.

39. Williams RL (2000) A note on robust variance estimation for cluster-correlated data. *Biometrics* 56:645–646.

40. Fletcher JM (2010) Adolescent depression and educational attainment: results using sibling fixed effects. *Health Econ* 19(7):855–871.

41. Isaacs JB, Haskins R, Sawhill IV (2008) *Getting Ahead or Losing Ground: Economic Mobility in America* (Brookings) Available at: https://www.brookings.edu/research/getting-ahead-or-losing-ground-economic-mobility-in-america/ [Accessed January 11, 2018].

42. Keller MC (2014) Gene × Environment Interaction Studies Have Not Properly Controlled for Potential Confounders: The Problem and the (Simple) Solution. *Biol Psychiatry* 75(1):18–24.

43. Gabriel SB, et al. (2002) The structure of haplotype blocks in the human genome. *Science* 296:2225–9.

44. Rietveld CA, et al. (2013) GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* 340(6139):1467–71.

45. Domingue BW, Belsky DW, Conley D, Harris KM, Boardman JD (2015) Polygenic Influence on Educational Attainment: New Evidence From the National Longitudinal Study of Adolescent to Adult Health. *AERA Open* 1(3):2332858415599972.

SECTION 3. SUPPLEMENTAL RESULTS

Supplemental Table S1. Summary Statistics of Analysis Variables

		European-Descent Sample			African American Sample		
Study	**Measure**	**M**	**SD**	**N**	**M**	**SD**	**N**
E-Risk	Educational Attainment (GSCE Level)			1,860			
	None	4%					
	Level-1	19%					
	Level-2	30%					
	Level-3	47%					
	Parent's Education (GCSE Level/ University Degree)			950			
	None	12%					
	Level-1	12%					
	Level-2	35%					
	Level-3	22%					
	University Degree	19%					
Add Health	Attainment (socioeconomic index score 2008)	96	37	5,526	88	36	1,814
	Education (years)	15	2	5,526	14	2	1,814
	Social Origins (factor score)	0.18	0.89	5,321	-0.31	1.10	1,661
	Parents' Education (years)	14	2	5,503	14	2	1,791
Dunedin	Educational Attainment (Highest Degree)			831			
	None	13%					
	School Leaving Certificate	15%					
	Sixth-From Certificate/ Bursary	42%					
	University Degree	30%					
	Attainment (New Zealand Socioeconomic Index Score 2012)	3.82	1.12	831			
	Social Origins (Parents' New Zealand Socioeconomic Index Score 1972-1987)	3.79	1.43	831			
WLS	Attainment (socioeconomic index score 1972-2005)	107.71	39.73	7,111			
	Social Origins (factor score)	16.42	11.11	7,111			
	Household Wealth (2005-2011)	1,068,358	3,062,678	7,007			
HRS	Household Wealth (1992-2012)	$604,723	$1,045,601	8,546	$145,558	$286,200	1,608
	Education (years of schooling)	13.15	2.54	8,533	11.62	3.26	1,608
	Social Origins (factor score)	0.17	1.25	8,546	-0.33	1.14	1,608

Supplemental Table S2. Effect-size estimates for analysis of social attainment and mobility [1]

Table S2 on following page.

[1] Social-attainment analysis was conducted according to Eq 1 (Supplemental Information Section 1.7). Social-mobility analysis was conducted according to Eq 3 (Supplemental Information Section 1.7). Add Health models included school and birth-year fixed effects and covariate adjustment for sex; standard errors were clustered at the family level to account for non-independence of sibling data. E-Risk models included covariate adjustment for sex; standard errors were clustered at the family level to account for non-independence of twin data. WLS models included school and birth-year fixed effects and covariate adjustment for sex and whether the participant was sampled as a 1957 high school graduate or the sibling of a graduate; standard errors were clustered at the family level to account for non-independence of sibling data. Dunedin models included covariate adjustment for sex. HRS models included birth-year fixed effects and covariate adjustment for sex; standard errors were clustered at the household level to account for non-independence of observations of spouses.

					Estimates with Adjustment for Educational Attainment		
		Effect-Size	Test-statistic	N	Effect-Size	Test-statistic	N
Education							
E-Risk							
	Genetic Association with Educational Attainment	0.27	10.83	1,860			
	Genetic Association with Educational Mobility	0.16	6.72				
Add Health							
	Genetic Association with Educational Attainment	0.28	22.39	5,526			
	Genetic Association with Educational Mobility	0.22	18.06				
Dunedin							
	Genetic Association with Educational Attainment	0.30	9.17	771			
	Genetic Association with Educational Mobility	0.27	8.47				
HRS							
	Genetic Association with Educational Attainment	0.32	31.69	8,150			
	Genetic Association with Educational Mobility	0.26	26.55				
Occupational Attainment							
Add Health							
	Genetic Association with Occupational Attainment	0.20	15.09	5,526	0.06	5.07	5,526
	Genetic Association with Upward Social Mobility	0.15	10.87		0.05	3.81	
Dunedin							
	Genetic Association with Occupational Attainment	0.26	8.59	831	0.11	3.81	831
	Genetic Association with Upward Social Mobility	0.21	7.12		0.10	3.63	
WLS							
	Genetic Association with Occupational Attainment	0.16	13.29	7,111	0.03	2.46	7,047
	Genetic Association with Upward Social Mobility	0.13	11.54		0.02	2.36	
Wealth							
WLS	Genetic Association with Household Wealth	0.12	9.74	7,007	0.06	4.82	6,947
	Genetic Association with Upward Social Mobility	0.10	8.41		0.06	4.71	
HRS	Genetic Association with Household Wealth	0.22	19.97	8,546	0.11	10.14	8,532
	Genetic Association with Upward Social Mobility	0.19	17.28		0.10	9.40	

Supplemental Table S3. Effect-sizes from sibling-difference analysis of attainment [2]

		Effect-Size	Test-statistic	N Families
Educational Attainment				
E-Risk				
	Genetic Association with Educational Attainment	0.32	9.14	
	Sibling Difference Estimate	0.13	2.53	388
Add Health				
	Genetic Association with Educational Attainment	0.30	6.72	
	Sibling Difference Estimate	0.15	2.33	352
Occupational Attainment				
Add Health				
	Genetic Association with Occupational Attainment	0.21	4.91	
	Sibling Difference Estimate	0.07	1.04	352
WLS				
	Genetic Association with Occupational Attainment	0.17	9.69	
	Sibling Difference Estimate	0.15	5.15	1,779
Wealth				
WLS				
	Genetic Association with Wealth	0.12	6.54	
	Sibling Difference Estimate	0.14	4.61	1,778

[2] Sibling-difference analysis was conducted according to Eq 4 (Supplemental Information Section 1.7). Add health models included family and birth-year fixed effects and covariate adjustment for sex. E-Risk models included family fixed-effects. All E-Risk twins were the same biological sex. Only dizygotic twins were included in analysis. WLS models included family and birth-year-fixed effects and covariate adjustment for sex and whether the participant was sampled as a 1957 high school graduate or as the sibling of a graduate.

Supplemental Table S4. Effect-sizes for polygenic score associations with social origins and social origins associations with attainment [3]

Study	Measure	Estimate	Effect-size	Test-statistic	N
Add Health					
	Family Socioeconomic Status Factor-Score	Association with Study Member's Polygenic Score	0.29	21.09	5,321
		Association with Study Member's Attainment	0.28	19.05	
Dunedin					
	Parents' Occupational Attainment	Association with Study Member's Polygenic Score	0.15	4.64	831
		Association with Study Member's Attainment	0.33	10.72	
WLS					
	Family Socioeconomic Status Factor-Score	Association with Study Member's Polygenic Score	0.12	8.93	7,111
		Association with Study Member's Attainment	0.25	20.74	
HRS					
	Family Socioeconomic Status Factor-Score	Association with Study Member's Polygenic Score	0.17	15.53	8,546
		Association with Study Member's Household Wealth	0.23	20.46	

[3] Social-origins analysis was conducted according to Eq 2 (Supplemental Information Section 1.7). Add Health models included birth-year fixed effects and covariate adjustment for sex; standard errors were clustered at the family level to account for non-independence of sibling data. WLS models included birth-year fixed effects and covariate adjustment for sex and whether the participant was sampled as a 1957 high school graduate or the sibling of a graduate; standard errors were clustered at the family level to account for non-independence of sibling data. Dunedin models included covariate adjustment for sex. HRS models included birth-year fixed effects and covariate adjustment for sex; standard errors were clustered at the household level to account for non-independence of observations of spouses.

Supplemental Table S5. Percentile-rank social-mobility estimates.[4]

	Social Origins Ranking Metric	Attainment Ranking Metric	Mobility Estimate	Test-statistic	N
E-Risk	Education	Education	1.62	1.88	1,860
Add Health	Social Origins Score	Occupational Attainment	3.42	8.56	5,321
Dunedin	Occupational Attainment	Occupational Attainment	6.22	7.09	831
WLS	Social Origins Score	Occupational Attainment	4.15	11.51	7,111
HRS	Social Origins Score	Wealth	5.53	18.13	8,546

[4]Percentile-rank mobility analysis was conducted according to Eq 5 (Supplemental Information Section 1.7). The mobility estimate corresponds to the number of percentile ranks a person is expected to move up the social ladder relative to their parents per standard deviation increment in their education polygenic score. Add Health models included school and birth-year fixed effects and covariate adjustment for sex; standard errors were clustered at the family level to account for non-independence of sibling data. WLS models included school and birth-year fixed effects and covariate adjustment for sex and whether the participant was sampled as a 1957 high school graduate or the sibling of a graduate; standard errors were clustered at the family level to account for non-independence of sibling data. Dunedin models included covariate adjustment for sex. HRS models included birth-year fixed effects and covariate adjustment for sex; standard errors were clustered at the household level to account for non-independence of observations of spouses.

Supplemental Table S6. Tests of interaction between polygenic score and social origins in models predicting socioeconomic attainments. [5]

	Est	t-statistic	p-value
AH			
Polygenic Score (PGS)	0.13	7.46	9.97E-14
Social Origins Score (SO)	0.23	12.18	1.14E-33
PGSxSO Interaction	0.06	4.55	5.49E-06
Dunedin			
Polygenic Score (PGS)	0.16	3.73	2.02E-04
Social Origins Score (SO)	0.30	6.45	1.93E-10
PGSxSO Interaction	-0.02	-0.74	0.459
WLS			
Polygenic Score (PGS)	0.13	8.14	4.87E-16
Social Origins Score (SO)	0.24	14.87	2.98E-49
PGSxSO Interaction	-0.02	-1.50	0.133
HRS			
Polygenic Score (PGS)	0.18	11.46	4.45E-30
Social Origins Score (SO)	0.18	10.90	1.97E-27
PGSxSO Interaction	0.01	1.14	0.255

[5] Models included the following covariates: For the Add Health Study, WLS, and HRS, birth year, sex, and product-terms modeling interactions between these variables and both the social origins and polygenic score variables; for the Dunedin Study, sex and product terms modeling interactions between sex and both the social origins and polygenic score variables. Add Health and WLS models included school fixed effects and clustered standard errors at the family level to account for non-independence of sibling data. For the HRS, models clustered standard errors at the household level to account for non-independence of spouse data.

Supplemental Table S7. Estimates of polygenic-score effect-sizes disattenuated for measurement error. [6]

				Estimates Adjusted for Education	
SNP h2=	0.25	Disattenuated Pearson r	Disattenuated R-squared	Disattenuated Pearson r	Disattenuated R-squared
Educational Attainment					
E-Risk	Education	0.50	0.25		
	Mobility	0.29	0.09		
Add Health	Education	0.50	0.25		
	Mobility	0.40	0.16		
Dunedin	Education	0.50	0.25		
	Mobility	0.44	0.20		
HRS	Education	0.50	0.25		
	Mobility	0.40	0.16		
Occupational Attainment					
Add Health	Occupational Attainment	0.35	0.12	0.11	0.01
	Mobility	0.26	0.07	0.08	0.01
Dunedin	Occupational Attainment	0.44	0.19	0.18	0.03
	Mobility	0.36	0.13	0.17	0.03
WLS	Occupational Attainment	0.41	0.17	0.07	0.005
	Mobility	0.35	0.12	0.07	0.004
Wealth					
WLS	Occupational Attainment	0.32	0.10	0.16	0.03
	Mobility	0.27	0.07	0.15	0.02
HRS	Occupational Attainment	0.35	0.12	0.18	0.03
	Mobility	0.29	0.09	0.16	0.03

[6] We computed disattenuated effect-sizes based on an assumption of SNP heritability (h2) of 0.25 using the equation described by Tucker-Drob (https://www.biorxiv.org/content/early/2017/07/19/165472)

Supplemental Table 8. Effect-sizes for polygenic score associations with attainment and mobility in WLS 1957 high school graduates and WLS members with non-farm social origins [7]

				Adjusted for Educational Attainment	
	Effect-Size	**Test-statistic**	**N**	**Effect-Size**	**Test-statistic**
Graduates					
Genetic Association with Occupational Attainment	0.14	10.15	5,239	0.02	1.31
Genetic Association with Upward Social Mobility	0.11	8.65		0.01	1.18
Genetic Association with Social Origins	0.12	8.71			
Non-Farm					
Genetic Association with Occupational Attainment	0.17	12.86	5,630	0.04	3.24
Genetic Association with Upward Social Mobility	0.14	10.85		0.04	3.04
Genetic Association with Social Origins	0.15	10.08			

[7] Models included school and birth-year fixed effects and covariate adjustment for sex; models fitted for WLS participants who had non-farm social origins additionally included a dummy variable encoding whether the participant was sampled as a 1957 high school graduate or as the sibling of a graduate. In models fitted for WLS participants who had non-farm social origins, standard errors were clustered at the family level to account for non-independence of sibling data.

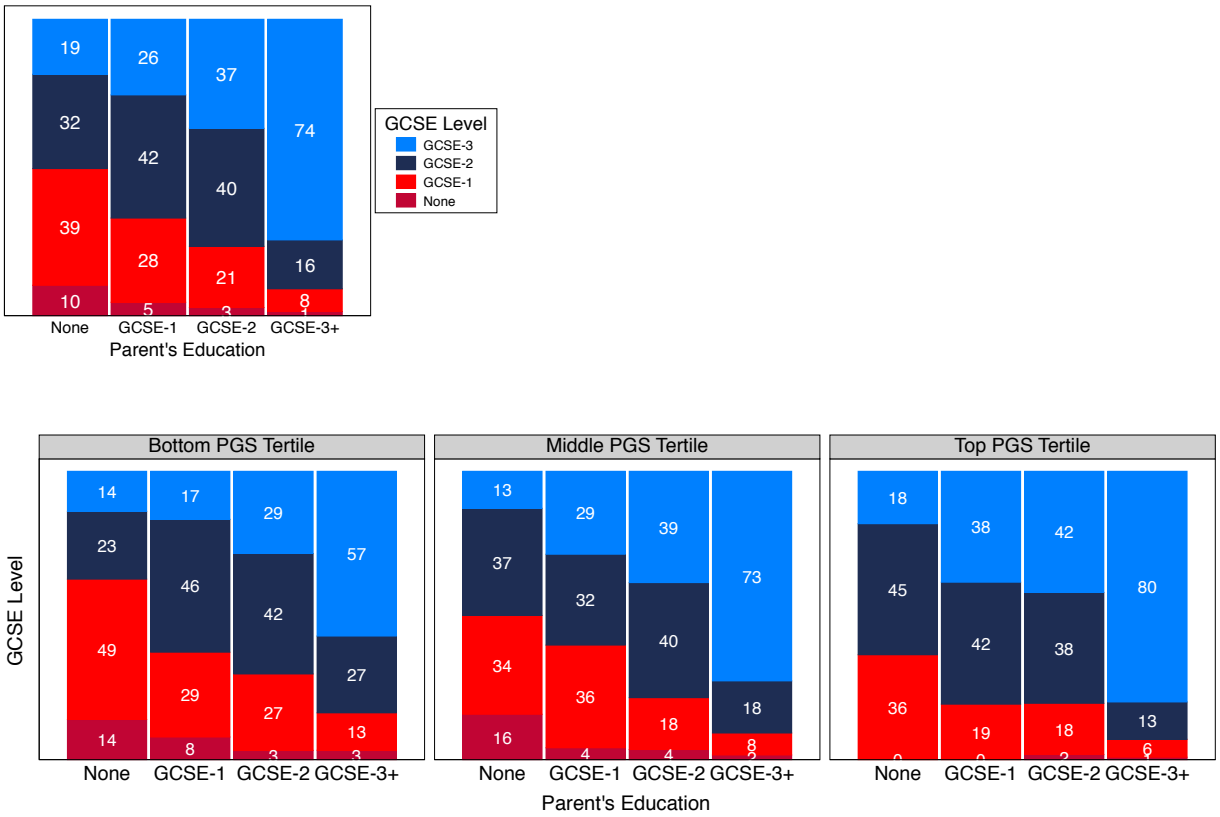
Supplemental Table S9. Effect-sizes for exploratory analysis of polygenic score associations with attainment and mobility among African Americans in the Add Health Study and the HRS [8]

					Estimates with Adjustment for Educational Attainment		
		Effect-Size	Test-statistic	N	Effect-Size	Test-statistic	N
Add Health							
Education	Genetic Association with Educational Attainment	0.08	3.47	1,814			
	Genetic Association with Educational Mobility	0.05	2.41				
Occupational Attainment	Genetic Association with Occupational Attainment	0.10	4.30	1,814	0.06	2.89	1,814
	Genetic Association with Upward Social Mobility	0.08	3.50		0.06	2.73	
Social Origins	Association with Study Member's Polygenic Score	0.10	4.13	1,660			
	Association with Study Member's Attainment	0.26	10.25				
HRS							
Education	Genetic Association with Educational Attainment	0.09	3.76	1,392			
	Genetic Association with Educational Mobility	0.09	4.16				
Wealth	Genetic Association with Household Wealth	0.08	3.54	1,609	0.05	2.27	1,608
	Genetic Association with Upward Social Mobility	0.09	3.60		0.05	2.28	
Social Origins	Association with Study Member's Polygenic Score	-0.02	-1.02	1,608			
	Association with Study Member's Household Wealth	0.07	3.12				

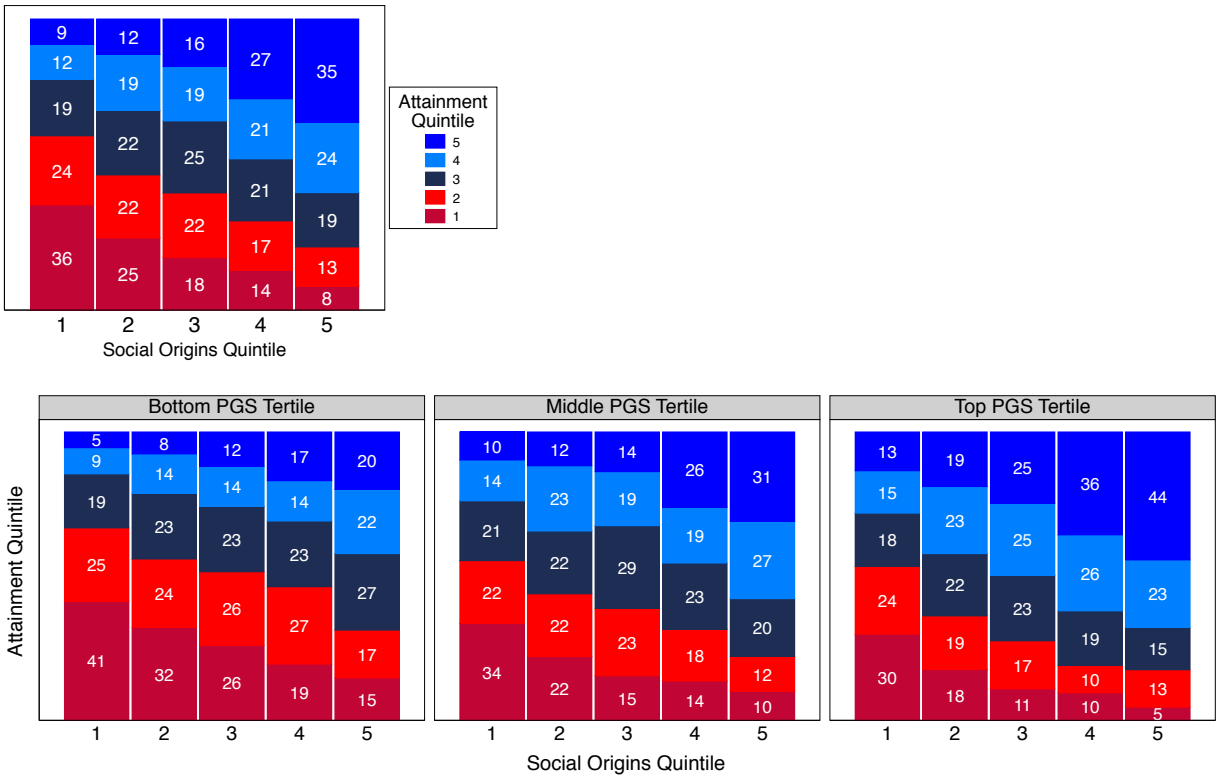
[8] Add Health models included school- and birth-year fixed effects and covariate adjustment for sex; standard errors were clustered at the family level to account for non-independence of sibling data. HRS models included birth-year fixed effects and covariate adjustment for sex; standard errors were clustered at the household data to account for non-independence of observations of spouses.

Supplemental Figure 1. Quantile-Rank Mobility Transition Matrices. Figure shows social mobility as quantile rank transition probabilities in the E-Risk, Add Health, Dunedin, WLS, and HRS samples. In the graphs, each bar corresponds to a social origin quantile. Each bar is subdivided into sections corresponding to each attainment quantile. Sections are sized to reflect the proportions of participants in a given social origin quantile who achieved each quantile of attainment. Each panel of the figure shows a sample-wide transition probability matrix and transition probability matrices for sub-samples defined by polygenic score tertile. E-Risk transitions (Panel A) are computed from categories of educational attainment for participants and their parents. Add Health transitions (Panel B) are computed from quintile ranks of participants' occupational attainments and their social-origins scores. Dunedin transitions (Panel C) are computed from categories of occupational attainment for participants' and their parents. WLS transitions (Panel D) are computed from quintile ranks of participants' occupational attainments and their social origins scores. HRS transitions (Panel E) are computed from quintile ranks of participants' wealth and their social origins scores.

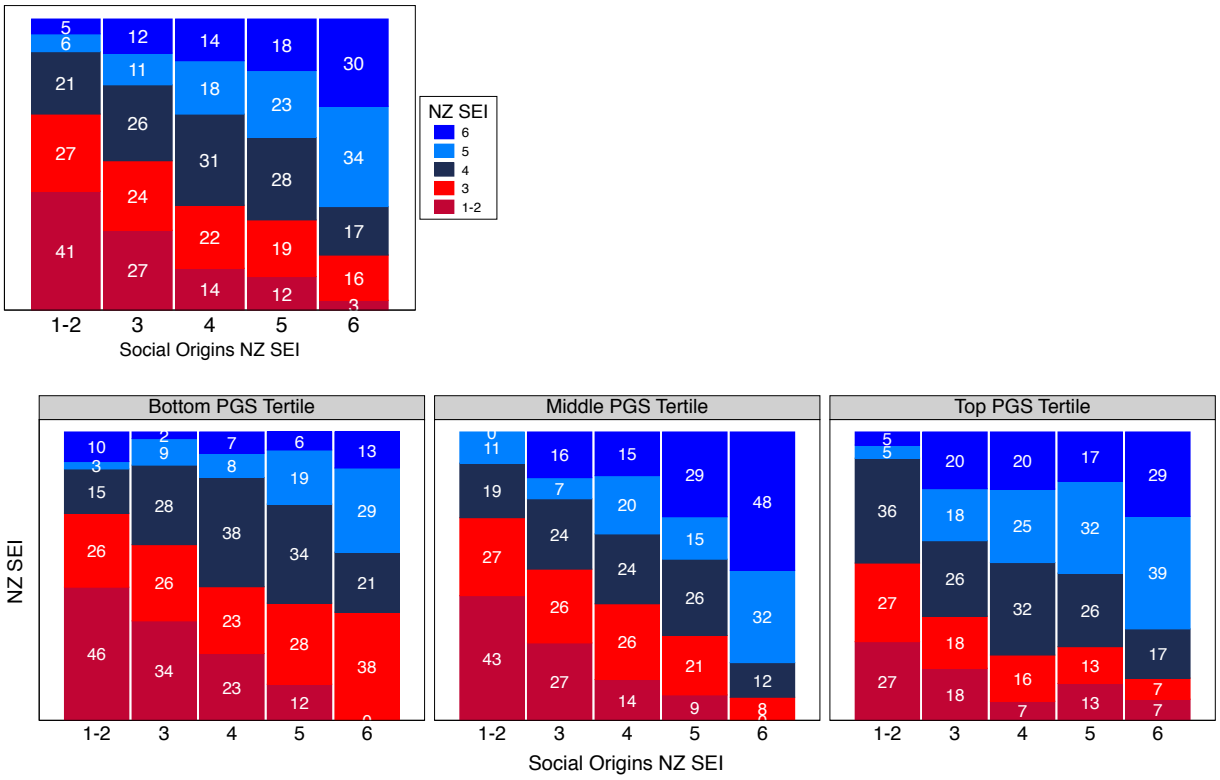
Panel A. E-Risk



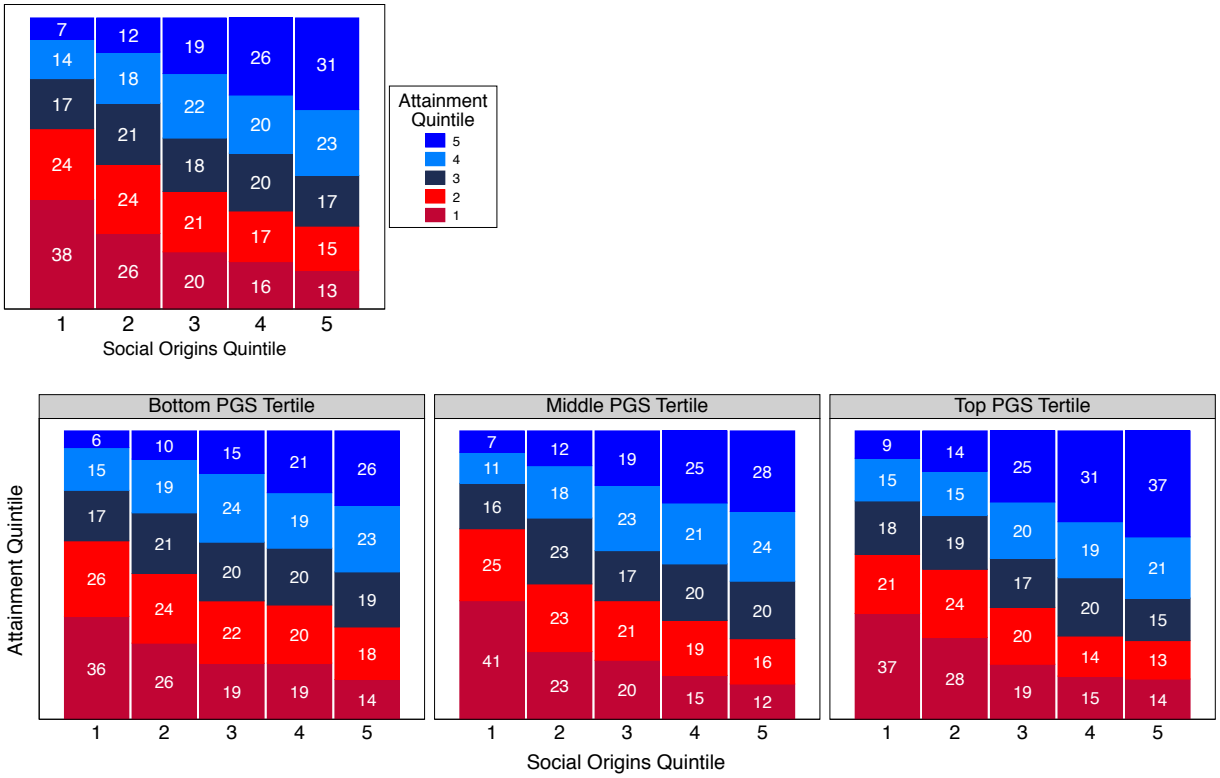
Panel B. Add Health



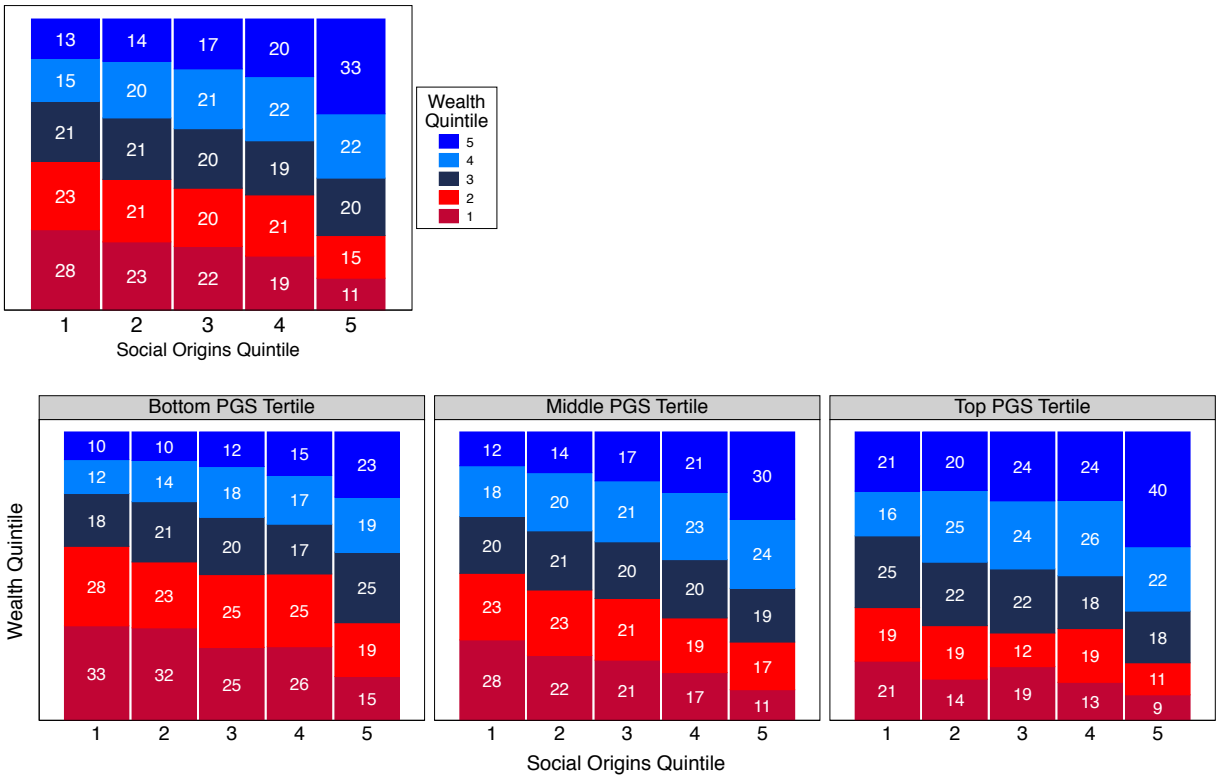
Panel C. Dunedin



Panel D. WLS


7
14
17
24
38
12
18
21
24
26
19
22
18
21
20
26
20
20
17
16
31
23
17
15
13
1
2
3
4
5
Social Origins Quintile
Attainment Quintile
5
4
3
2
1
Bottom PGS Tertile
Middle PGS Tertile
Top PGS Tertile
Attainment Quintile
6
15
17
26
36
10
19
21
24
26
15
24
20
22
19
21
19
20
20
19
26
23
19
18
14
7
11
16
25
41
12
18
23
23
23
19
23
17
21
25
21
20
19
15
28
24
20
16
12
9
15
18
21
37
14
15
19
24
28
25
20
17
20
19
31
19
20
14
15
37
21
15
13
14
1
2
3
4
5
Social Origins Quintile

Panel E. HRS


Wealth Quintile
5
4
3
2
1
Social Origins Quintile
Bottom PGS Tertile
Middle PGS Tertile
Top PGS Tertile
Wealth Quintile
Social Origins Quintile