Genetic Fortune: Winning or Losing Education, Income, and Health

Hyeokmoon Kweon, Casper A.P. Burik, Richard Karlsson Linnér, Ronald De Vlaming, Aysu Okbay,
Daphne Martschenko, K. Paige Harden, Thomas A. DiPrete, Philipp D. Koellinger

Contact
Philipp Koellinger
koellinger@wisc.edu

Introduction

Parents influence the starting-points of their children by providing them with social and genetic endowments that are due to luck in the sense that they are exogenously given rather than the result of one's own actions. Thus, inequalities of opportunity can partly arise from the outcomes of two family-specific "lotteries" that take place during conception — a "social lottery" that determines who our parents are, and a "genetic lottery" that determines which part of their genomes our parents pass on to us. The relative importance of social and genetic luck has policy relevance because the extent to which people are willing to tolerate inequality partially depends on whether they perceive that disparity originates from differences in effort and choice (e.g., working hard) or from differences in circumstances that are outside of one's control (e.g., luck in the social or genetic lotteries). With molecular genetic and family data, we show that the genetic lottery contributes to inequalities between siblings in education, income, and health. Partly, these effects work via educational attainment, i.e. a malleable environmental factor.

Data

Data source: UK Biobank (UKB)

- Sibling sample (N ≈35K) used to study the effects of the genetic lottery
- Non-sibling sample used to construct a polygenic score (PGS)
- Analyses restricted to only people with European ancestries

Health measures

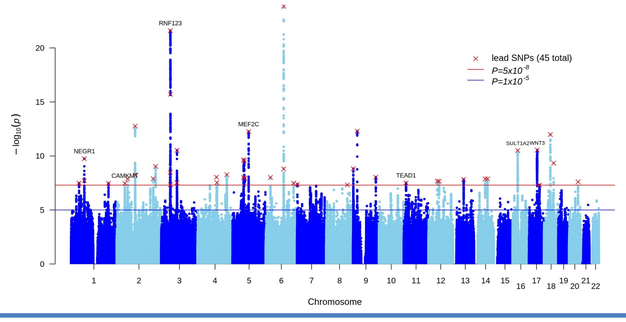
Waist-to-hip ratio, BMI, blood pressure, lung function, diagnosis records (hospitalization, death and cancer registries)

Socioeconomic measures

occupational wage (derived from standardized occupation codes), household income, regional income, neighbourhood score, education

Polygenic score (PGS) for income

Constructed from genome-wide association study (GWAS) of occupational wages with UKB's non-sibling sample (N≈253K).
PGS captures 1~3% of observed wages in independent samples.



Empirical strategy

Baseline model

For outcome y of individual *i* from family *j*:

$$y_{ij} = \beta s_{ij} + \mathbf{Z'_{ij}}\boldsymbol{\delta} + \alpha_j + \varepsilon_i$$

s_{ij} polygenic score, $\mathbf{Z}_i$ covariates, α_j family-specific effect for family *j*

Correcting for genetic nurture and population stratification

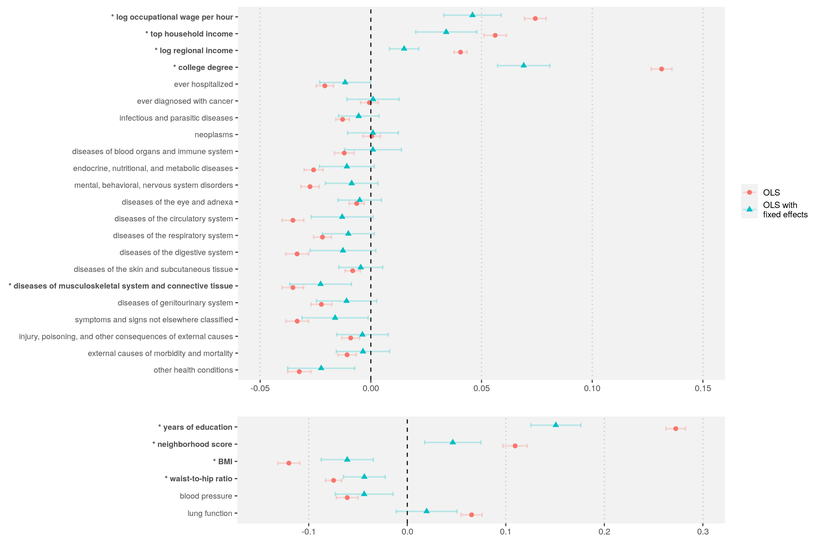
- By controlling for family fixed effects (FE) α_j, we exploit the random genetic differences between siblings to identify the causal consequences of the income PGS (s_{ij}) on socioeconomic and health outcomes (y_{ij}).

Correcting for measurement error in PGS

$\hat{\beta}$ estimated by OLS will contain attenuation bias due to measurement error in PGS arising from a finite GWAS sample size.
Solution: use genetic instrument variable (GIV) regression

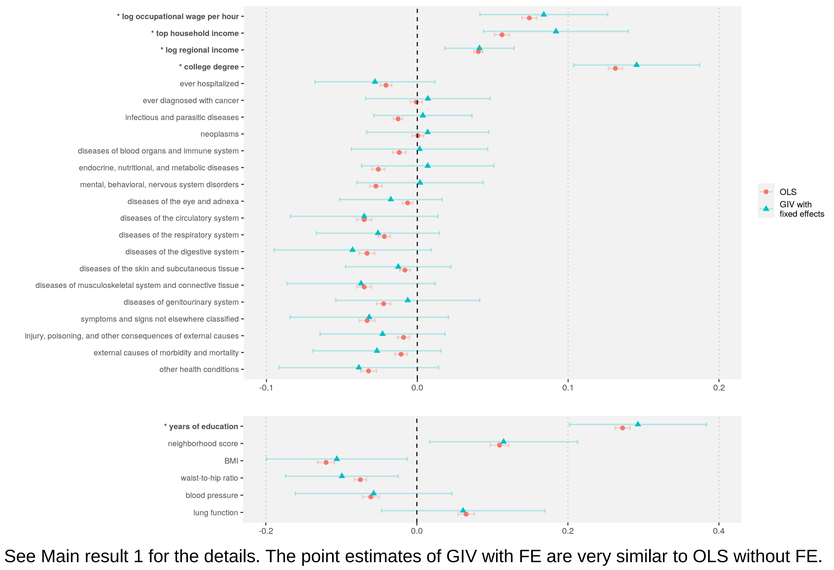
- Randomly split the GWAS sample into two independent subsamples that allow for constructing two indicators of PGS.
- Under the reasonable assumption that the error terms of both indicators are independent, one of them can be used as an instrument variable for the other.
- The obtained IV estimate corrects for measurement error in the PGS.

Main result 1: OLS with/without family fixed effects



The figure plots the regression coefficients for the standardized PGS from within-family analysis. Error bars are 95% confidence intervals. The upper panel shows the estimates measured on percentage scale. The lower panel plots the standardized estimates (*i.e.*, the outcomes and the PGS are both standardized). The asterisk indicates significance at the 5% family-wise error rate for the estimate with family fixed effects. Multiple testing is corrected for using Holm's method (Holm, 1979). Standard errors clustered by family.

Main result 2: OLS vs. GIV with FE

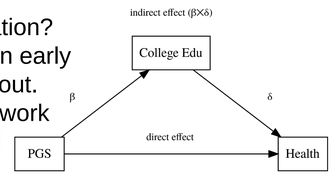


See Main result 1 for the details. The point estimates of GIV with FE are very similar to OLS without FE.

Decomposing the genetic effects

Is the genetic effect mediated by college education?

- College education because it is determined in early adulthood so reverse causality can be ruled out.
- If mediated, an example that genetic effects work via a channel at least partially environmental



	direct effect	indirect effect	total effect	indirect effect %
log occupational wage per hour	0.031***	0.014***	0.046***	31.7
(N=17,578)	(0.006)	(0.002)	(0.007)	
waist-to-hip ratio	-0.003**	0.0004**	-0.004***	11
(N=35,028)	(0.001)	(0.0001)	(0.001)	
BMI	-0.256***	-0.025*	-0.281***	8.8
(N=34,968)	(0.064)	(0.008)	(0.064)	
blood pressure	-0.546*	-0.077*	-0.622*	12.3
(N=31,372)	(0.210)	(0.027)	(0.209)	
lung function	0.013	0.005**	0.018	29.4
(N=29,844)	(0.014)	(0.002)	(0.013)	

FWE-corrected by Bonferroni, ***(0.1%), **(1%), *(5%). All include family FE. SE for indirect effects by delta method. SE clustered by family. PGS standardized

Conclusion

- Genetic fortune for higher income, in the form of random genetic differences between siblings, contributes to inequality throughout the life course. This raises questions about how much credit people can take for their success and health in life.
- Genes contribute to inequality, but this does not imply biological determinism or an irrelevance of policy. The causal pathways from genes to outcomes involve environmental and behavioural pathways that can be intervened upon (e.g. educational attainment).