Trading social status for genetics in marriage markets: Evidence from Great Britain and Norway






Abdel Abdellaoui *Amsterdam UMC*

Oana Borcan *University of East Anglia*

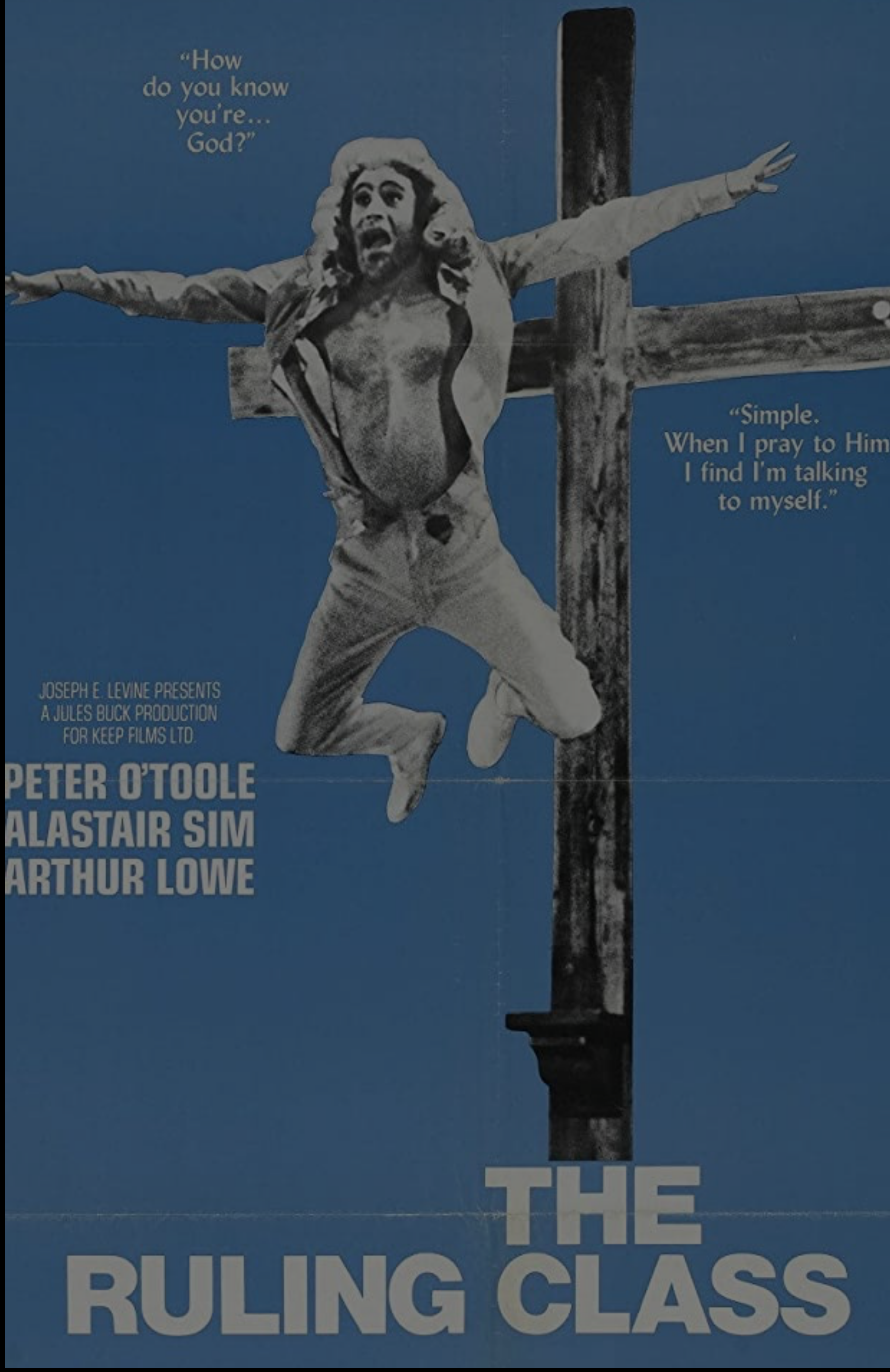
Pierre Chiappori *Columbia*

David Hugh-Jones *Idle Bum*

Fartein Ask Torvik *Norwegian Institute of Public Health*

Eivind Ystrøm *Norwegian Institute of Public Health*

moba



Background

Genetics explains c. 50% of variation in many human characteristics.

What's a social scientist to do?

One answer: **make genetics the dependent variable.**

Assortative mating

- Hugh-Jones, Abdellaoui et al. (2016). Assortative Mating on Education Leads to Genetic Spousal Resemblance for Causal Alleles. *Intelligence* 59.
- This paper…

Migration

- Abdellaoui, Hugh-Jones, …, Visscher (2019). Genetic Correlates of Social Stratification in Great Britain. *Nature Human Behaviour* 3.

Natural selection

- Hugh-Jones and Abdellaoui (2022). Human Capital Mediates Natural Selection in Contemporary Humans. *Behavior Genetics* 52:4.

Goals of this paper

In increasing order of ambition:

- Explain a puzzle about the **intergenerational persistence of inequality**.
- Provide a new explanation of the **genes-SES (socio-economic status) gradient**.
- Rethink the **nature of inequality** in historical human societies.
- Change how we think about **genetic variation**.

Many genetic measures, including polygenic scores for education and health outcomes, differ between people of low and high socio-economic status (SES).

The leading explanation for this **genes-SES gradient** is meritocracy: genetic variants that cause success in *labour markets* lead to upward mobility.

An alternative explanation: both some genetic variants, and high SES, are desirable qualities in **marriage markets**.

If you are rich or privileged, you may marry someone intelligent or good-looking.

Both SES and genetics are then inherited by the next generation.

Under **Social-Genetic Assortative Mating**:

- Shocks to SES are reflected in the DNA of subsequent generations.
- The size of the genes-SES gradient depends on social structure, e.g. on persistence of inherited wealth.
- The genes-SES gradient is likely historically widespread, beyond modern meritocracies.

Literature

Large, mostly separate, literatures on assortative mating in economics and genetics.

Genetics: genetic assortative mating (GAM) (Rao, Morton, and Yee 1976; Heath and Eaves 1985; Otto, Christiansen, and Feldman 1995), including cross-trait assortative mating (Beauchamp et al. 2010; Sundet et al. 2005; Border et al. 2022).

Economics: models and empirics on assortative mating and inequality, including cross-trait assortative mating (Fernández and Rogerson 2001; Fernandez, Guner, and Knowles 2005; Eika, Mogstad, and Zafar 2019; Chiappori, Dias, and Meghir 2018).

Genetic theory predicted in the 1970s that genetically and culturally transmitted traits could become associated in the population (Rao, Morton, and Yee 1976).

Despite this, previous work has not drawn conclusions for the genes-SES gradient (e.g. Rimfeld et al. 2018) or made the link with social structure.

- Ours is the first post-genomic revolution empirics to directly demonstrate SGAM.
- Our model provides a microeconomic foundation for SGAM, which might help bridge the two literatures.

Model

Traits x_1 (genetic) and x_2 (SES) are normally distributed in the population.

Attractiveness is $ax_1 + (1-a)x_2$ where $0 \leq a \leq 1$ is the relative importance of genes.

People mate assortatively.

Children inherit

$$x_1' = \tau\,(x_1 + y_1)/2 + \varepsilon$$

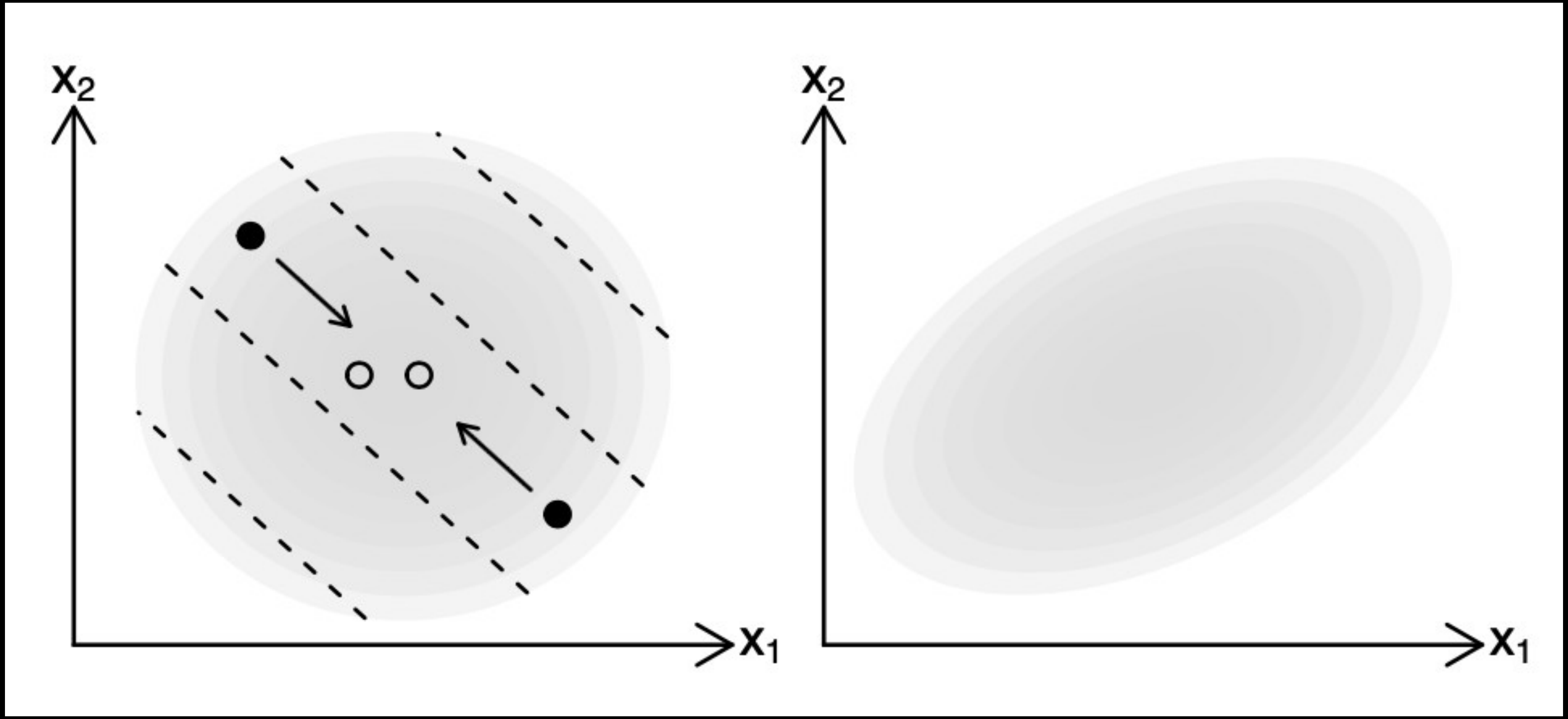
$$x_2' = \theta\,(x_2 + y_2)/2 + \eta$$

Where:

τ is close to 1 (genetic inheritance).

θ reflects persistence of SES (e.g. inheritance tax rate $1 - \theta$).

Intuition



Parents (●) mate along iso-attractiveness curves.
Their children (○) are between them.
As a result, the children's distribution is squashed along the attractiveness gradient.

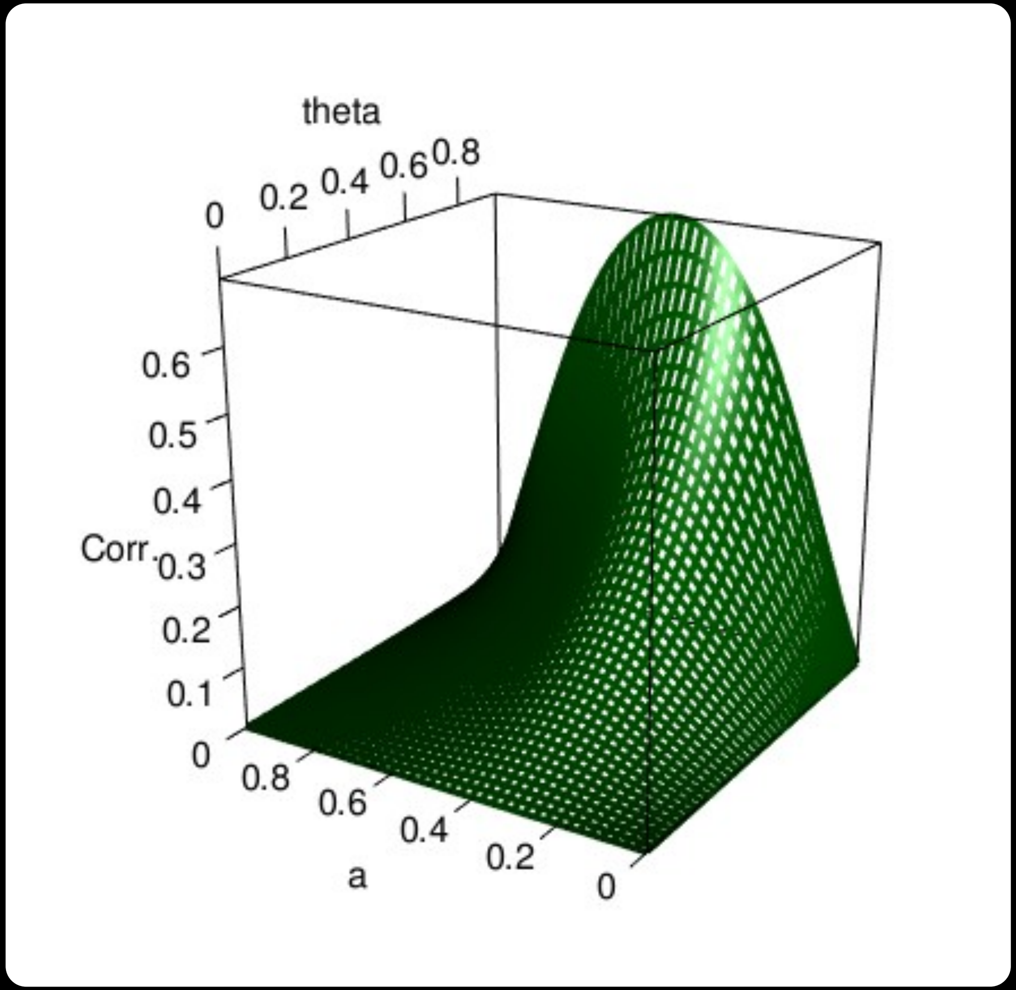
Result

If parents' x_1 and x_2 are independent, children's x_1' and x_2' are positively correlated for $0 < a < 1$.

The correlation increases in persistence of SES (θ).

The same holds for the long-run distribution.

Extensions



Long-run correlation between x_1 and x_2, by a and θ

Data

- UK Biobank, a study of about 500,000 individuals born 1935-1970.

  UKBB has no explicit information on spouse pairs. We match pairs manually, and check using genetic children of the pairs. This gives 35,682 spouse pairs.
- MOBA, a study of about 90,000 mothers and babies born in Norway 1998-2008. About 70,000 fathers are included.

Polygenic scores – a primer

Individual DNA contains about 3 billion base pairs, on 23 chromosomes. Each base pair can be one of four letters (A, C, T or G). About 0.1% of these pairs will be different between any two individuals.

A Single Nucleotide Polymorphisms (SNPs) is a base pair where individuals' genomes vary by a single "letter". Individuals have 0, 1 or 2 copies of the rare *allele*.

DNA *array data* records many SNPs of an individual's genome (in UKBB, about 500,000 SNPs).

Polygenic scores – a primer

Many heritable traits are highly *polygenic*: the sum of small effects from many loci on the genome.

A polygenic score is estimated using many SNPs. Bivariate correlations are estimated:

$$Trait_i = \beta_j SNP_{i,j} + \varepsilon_i$$

The score is then

$$PGS_i = \sum_j \beta_j SNP_{i,j}$$

Intuition: a PGS is like a regression $\hat{y}$, but with “naively” estimated βs.

β_j may be corrected for correlation by various methods.

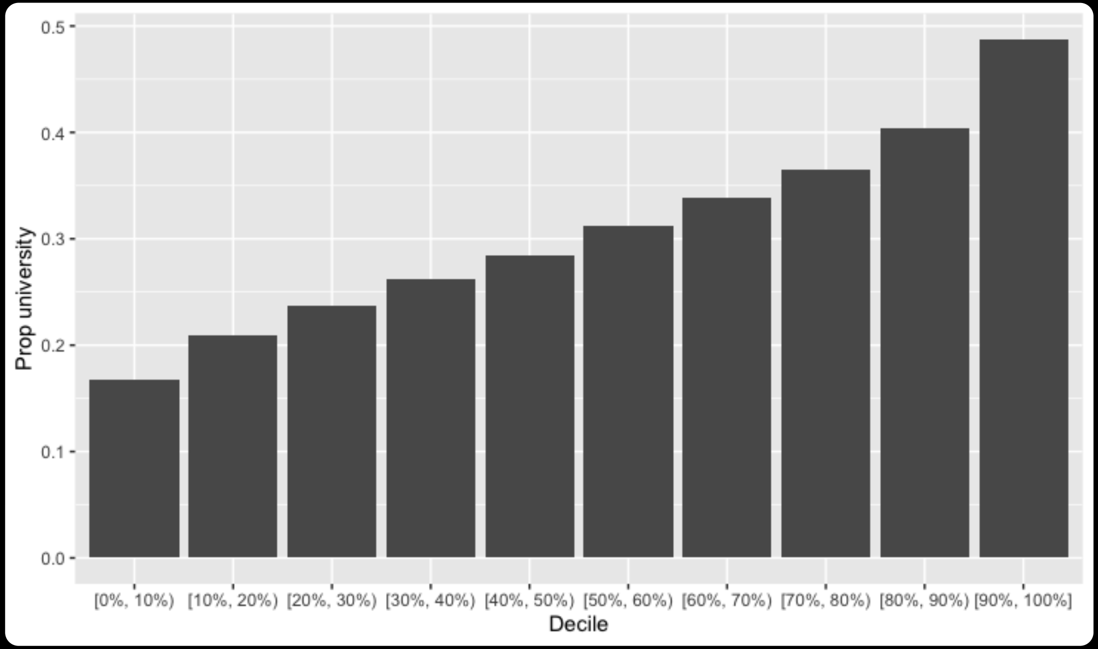
Polygenic Score for Educational Attainment

Our dependent variable is spouse's **Polygenic Score for Educational Attainment** (PSEA).

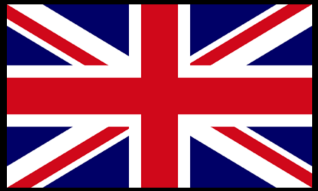
Linear regression of probability of attending university on PSEA:
1 sd PSEA = 9.2 percentage points.

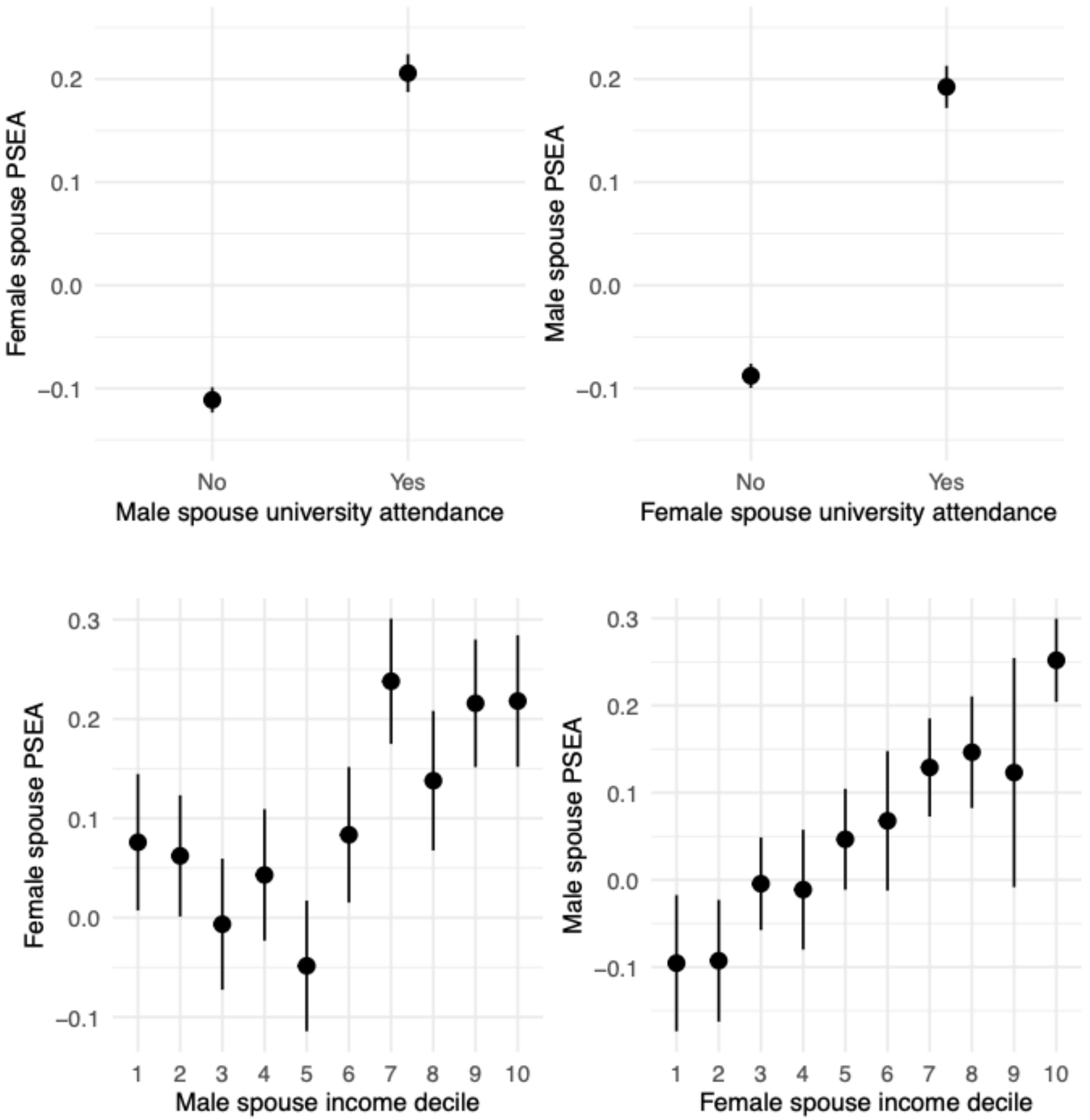
Within-siblings regression (causal):
4.5 percentage points.

Substantial confounding with family environment, but large causal effects remain.



University attendance by PSEA decile




Female spouse PSEA
0.2
0.1
0.0
−0.1
No
Yes
Male spouse university attendance
Male spouse PSEA
0.2
0.1
0.0
−0.1
No
Yes
Female spouse university attendance
Female spouse PSEA
0.3
0.2
0.1
0.0
−0.1
1 2 3 4 5 6 7 8 9 10
Male spouse income decile
Male spouse PSEA
0.3
0.2
0.1
0.0
−0.1
1 2 3 4 5 6 7 8 9 10
Female spouse income decile

These results could be confounded by the focal individual's own genetics.

- We already know that there is assortative mating on PSEA (Hugh-Jones et al. 2016).

We need an independent variable which

- affects SES;
- is independent of genetics;
- varies across a large enough N.
  - Polygenic scores and causes of variation in SES are noisy
  - The spouse matching process is unpredictable (Shakespeare 1595)

We use **birth order**.

- Siblings have the same expected polygenic scores, by the **"lottery of meiosis"**.
- Early-born siblings receive more parental care and have better life outcomes, including **socio-economic status (SES)**.

Estimation strategy

Hard to justify instrumental variables:

- Birth order affects other things than SES.
- We only have imperfect measures of SES (estimated income, educational attainment).

Instead we run a **mediation analysis**:

- Does birth order affect spouse's PSEA?
- Is the effect mediated by measures of SES?

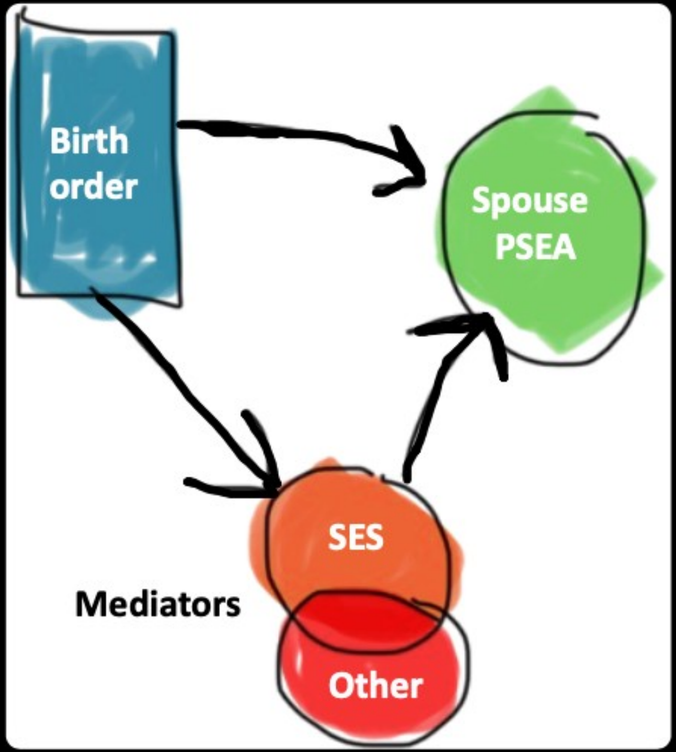
Estimation strategy

Hard to justify instrumental variables:

- Birth order affects other things than SES.
- We only have imperfect measures of SES (estimated income, educational attainment).

Instead we run a **mediation analysis**:

- Does birth order affect spouse's PSEA?
- Is the effect mediated by measures of SES?



Controls and mediators

SES mediators

University attendance

Earnings (of first job, guesstimated from SOC code, in UK; decile at age 30 in Norway)

Non-SES mediators

Height

BMI

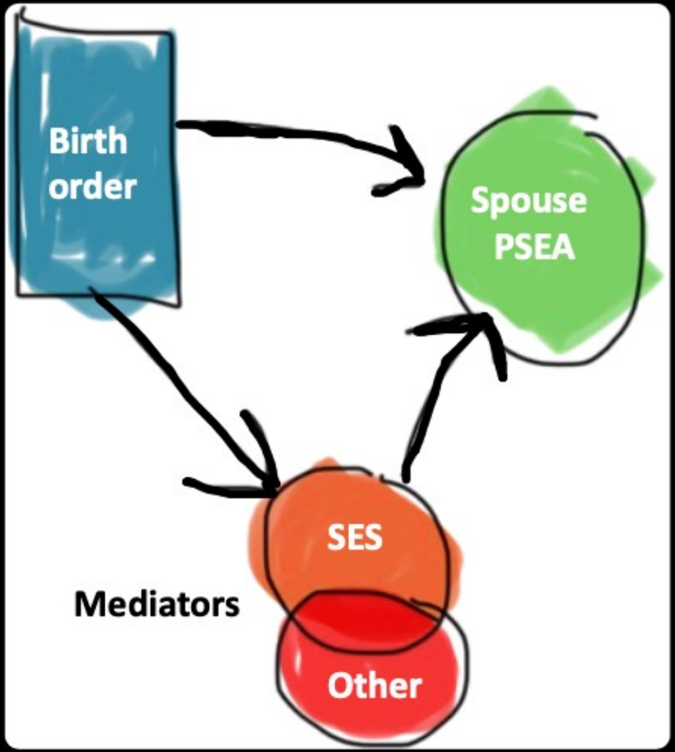
Fluid IQ & self-reported health (only in UKBB)

Controls

Family size

Month of birth, year of birth

Parent’s age at birth (only available for some respondents)



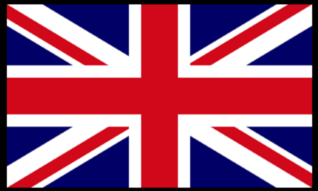
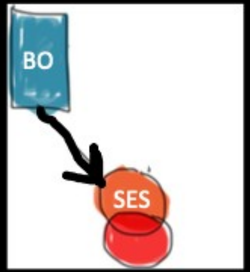


Table 1: Regressions of mediators on birth order

	University	Income	Fluid IQ	Height	BMI	Health
Birth order	−0.0790 ***	−1.0899 *	−0.2733 ***	−0.7012 ***	0.1907 **	−0.0430 ***
	(0.0067)	(0.4264)	(0.0304)	(0.1355)	(0.0662)	(0.0103)
PSEA	0.0889 ***	1.5144 ***	0.3180 ***	0.1970 *	−0.4281 ***	0.0533 ***
	(0.0046)	(0.3307)	(0.0200)	(0.0921)	(0.0456)	(0.0068)
Parents' age at birth	0.0163 ***	0.2623 ***	0.0588 ***	0.1514 ***	−0.0989 ***	0.0110 ***
	(0.0012)	(0.0722)	(0.0053)	(0.0241)	(0.0117)	(0.0018)
Family size dummies	Yes	Yes	Yes	Yes	Yes	Yes
Birth month dummies	Yes	Yes	Yes	Yes	Yes	Yes
Birth year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	10220	3412	10220	10220	10220	10220
R2	0.074	0.026	0.058	0.017	0.023	0.018



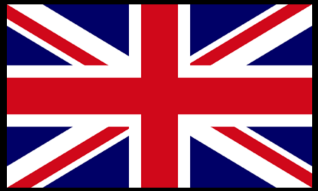
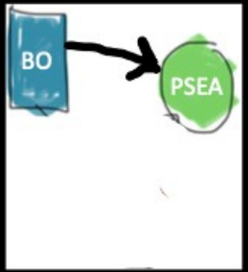
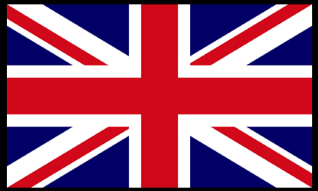


Table 2: Regressions of spouse PSEA on birth order

	(1)	(2)	(3)
Birth order	−0.0091	−0.0075	−0.0314 *
	(0.0074)	(0.0074)	(0.0146)
Own PSEA		0.0650 ***	0.0573 ***
		(0.0065)	(0.0100)
Parents' age at birth			0.0116 ***
			(0.0026)
Family size dummies	Yes	Yes	Yes
Birth month dummies	No	Yes	Yes
Birth year dummies	No	Yes	Yes
N	23840	23797	10206
R2	0.003	0.010	0.013



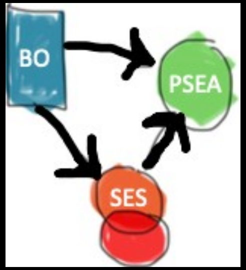


SES mediators

Non-SES mediators

Table 3: Regressions of spouse PSEA on birth order and mediators (Great Britain)

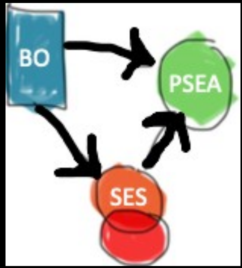
	(1)	(2)	(3)	(4)
Birth order	−0.0314 *	−0.0045	−0.0106	−0.0042
	(0.0145)	(0.0145)	(0.0270)	(0.0270)
University		0.2179 ***		0.1538 ***
		(0.0225)		(0.0375)
Income			0.0037 ***	0.0031 **
			(0.0011)	(0.0011)
Fluid IQ		0.0172 **	0.0201 *	0.0112
		(0.0052)	(0.0092)	(0.0096)
Height		0.0029 **	0.0046 *	0.0043 *
		(0.0011)	(0.0020)	(0.0020)
BMI		−0.0109 ***	−0.0114 **	−0.0109 **
		(0.0022)	(0.0040)	(0.0040)
Self-reported health		0.0181	0.0145	0.0077
		(0.0152)	(0.0275)	(0.0273)
Own PSEA	0.0573 ***	0.0263 *	0.0218	0.0118
	(0.0120)	(0.0121)	(0.0202)	(0.0203)
Parents’ age at birth	0.0116 ***	0.0053 *	0.0091 +	0.0078 +
	(0.0026)	(0.0026)	(0.0047)	(0.0047)
Family size dummies	Yes	Yes	Yes	Yes
Birth month dummies	Yes	Yes	Yes	Yes
Birth year dummies	Yes	Yes	Yes	Yes
N	10206	10206	3407	3407
R^2	0.013	0.032	0.030	0.034



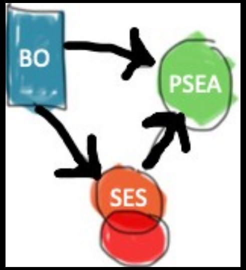
Zoom

Table 4: Percent of birth order effects accounted for by mediators (Great Britain)

	Model 2 (%)	Model 3 (%)	Model 4 (%)
University	54.9		38.7
Income		13.0	10.6
Fluid IQ	15.0	17.6	9.7
Height	6.6	10.4	9.5
BMI	6.6	7.0	6.6
Self-reported health	2.5	2.0	1.1



Robustness



	(1)	(2)	(3)	(4)
Birth order	−0.0684 ***	−0.0426 ***	−0.0652 ***	−0.0428 ***
	(0.0051)	(0.0054)	(0.0054)	(0.0054)
University		0.3049 ***		0.3091 ***
		(0.0075)		(0.0076)
Income			0.0053 ***	−0.0043 **
			(0.0013)	(0.0014)
Height		0.0033 ***	0.0061 ***	0.0032 ***
		(0.0004)	(0.0004)	(0.0004)
BMI		−0.0206 ***	−0.0240 ***	−0.0208 ***
		(0.0009)	(0.0009)	(0.0009)
Own PSEA	0.1193 ***	0.0699 ***	0.1068 ***	0.0704 ***
	(0.0034)	(0.0036)	(0.0035)	(0.0036)
Parents' age at birth	0.0132 ***	0.0085 ***	0.0125 ***	0.0085 ***
	(0.0008)	(0.0008)	(0.0008)	(0.0008)
Family size dummies	Yes	Yes	Yes	Yes
Birth month dummies	Yes	Yes	Yes	Yes
Birth year dummies	Yes	Yes	Yes	Yes
N	89308	80465	80503	80465
R^2	0.032	0.060	0.041	0.060

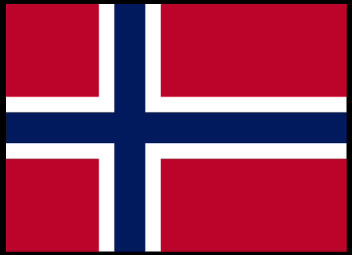
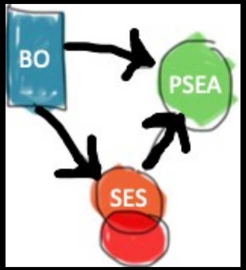


Table 7: Percent of birth order effects accounted for by mediators (Norway)

	Model 2 (%)	Model 3 (%)	Model 4 (%)
University	35.5		36.0
Income		0.9	−0.7
Height	2.3	4.3	2.3
BMI	0.7	0.8	0.7



Socio-Genetic Assortative Mating

Explain a puzzle about the **intergenerational persistence of inequality**.

- Inequality can persist because of unmeasured genetic variation (Clark 2021). Genetics can be a mediator, not just a confound, for transmission of SES over generations.

Provide a new explanation of the **genes-SES gradient**.

- In modern meritocracies, genes affect SES.
- Under SGAM, in all societies, SES can affect genes.
- Shocks to SES are reflected in the DNA of subsequent generations.
- The size of the genes-SES gradient is affected by social institutions.
  - Some evidence that income matters more in Great Britain than Norway.

Rethink the **nature of inequality** in historical human societies.

- Prediction: a genes-status gradient should be visible in ancient DNA.

Change how we think about **genetic variation**.

- Yes, genes are "biological"…
- But across generations, **genetic variation is a social outcome**.

Results from human genetics have been controversial, and will likely continue to be so.

Rather than banning research we think is harmful, perhaps it would help to take the perspective above:

"Genes are not special."

Thank you!

History is the true natural history of man

Marx

Extensions

The basic result extends to non-normal/discontinuous distributions, and non-linear attractiveness.

We can allow “meritocracy”, where adult SES depends directly on own genes.

- Genes-SES correlation may either increase or decrease in θ.

The a parameter can differ for men and women.

- Genes-SES correlation is maximized when parameters are most different. (Intuition: then e.g. high-status males almost always match high-genetics females.)

Spouse pairs

Some respondents in the Biobank sample have a genetic child who is also in the sample.

Among our spouse pairs, 511 have a genetic child of at least one partner in the sample.

For 86% (441) of these, the child is the genetic child of both partners.

Comparison: 11% of families with dependent children included a stepchild in England and Wales in 2011 (National Statistics 2014).

Robustness

Extra mediators: BMI, self-reported health.

Birth order is independent of 33 different polygenic scores.

Results are qualitatively robust…

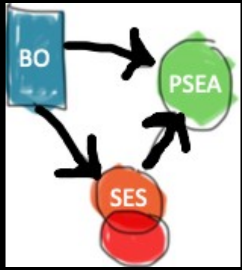
- … if we use birth order dummies: strongest effect for first child versus subsequent children.
- … using age left full-time education as the key mediator
- … for males and females only (initial birth order coefficient is not significant)
- … for couples with children

SES mediators

Non-SES mediators

Table 3: Regressions of spouse PSEA on birth order and mediators

	(1)	(2)	(3)	(4)
Birth order	−0.0314 *	−0.0045	−0.0106	−0.0042
	(0.0146)	(0.0146)	(0.0270)	(0.0270)
University		0.2179 ***		0.1538 ***
		(0.0225)		(0.0377)
Income			0.0037 ***	0.0031 **
			(0.0011)	(0.0011)
Fluid IQ		0.0172 **	0.0201 *	0.0112
		(0.0053)	(0.0094)	(0.0097)
Height		0.0029 **	0.0046 *	0.0043 *
		(0.0011)	(0.0020)	(0.0019)
BMI		−0.0109 ***	−0.0114 **	−0.0109 **
		(0.0022)	(0.0040)	(0.0040)
Self-reported health		0.0181	0.0145	0.0077
		(0.0151)	(0.0272)	(0.0271)
Own PSEA	0.0573 ***	0.0263 **	0.0218	0.0118
	(0.0100)	(0.0101)	(0.0184)	(0.0185)
Parents' age at birth	0.0116 ***	0.0053 *	0.0091 +	0.0078 +
	(0.0026)	(0.0026)	(0.0047)	(0.0047)



Bibliography

Abdellaoui, Abdel, David Hugh-Jones, Löıc Yengo, Kathryn E Kemper, Michel G Nivard, Laura Veul, Yan Holtz, et al. 2019. “Genetic Correlates of Social Stratification in Great Britain.” *Nature Human Behaviour* 3 (12): 1332–42.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2011. “Older and Wiser? Birth Order and IQ of Young Men.” *CESifo Economic Studies* 57 (1): 103–20.

Booth, Alison L, and Hiau Joo Kee. 2009. “Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment.” *Journal of Population Economics* 22 (2): 367–97.

Clark, Gregory. 2021. “For Whom the Bell Curve Tolls: A Lineage of 400,000 English Individuals 1750-2020 Shows Genetics Determines Most Social Outcomes.” Working Paper. http://faculty.econ.ucdavis.edu/faculty/ gclark/ClarkGlasgow2021.pdf.

Clark, Gregory, and Neil Cummins. 2015. “Intergenerational Wealth Mobility in England, 1858–2012: Surnames and Social Mobility.” *The Economic Journal* 125 (582): 61–85.

Eika, Lasse, Magne Mogstad, and Basit Zafar. 2019. “Educational Assortative Mating and Household Income In- equality.” *Journal of Political Economy* 127 (6): 2795–835.

Fernandez, Raquel, Nezih Guner, and John Knowles. 2005. “Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality.” *Quarterly Journal of Economics* 120 (1): 273–344.

Fernández, Raquel, and Richard Rogerson. 2001. “Sorting and Long-Run Inequality.” *Quarterly Journal of Economics* 116 (4): 1305–41.

Furnham, Adrian. 1993. “Just World Beliefs in Twelve Societies.” *Journal of Social Psychology* 133 (3): 317–29.

Gramsci, Antonio. 1971. Selections from the Prison Notebooks. Lawrence; Wishart London.

Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos. 2014. “Marry Your Like: Assortative Mating and Income Inequality.” *American Economic Review* 104 (5): 348–53.

Halsey, AH. 1958. “Genetics, Social Structure and Intelligence.” *British Journal of Sociology* 9 (1): 15–28.

Hugh-Jones, David, Karin JH Verweij, Beate St Pourcain, and Abdel Abdellaoui. 2016. “Assortative Mating on Educational Attainment Leads to Genetic Spousal Resemblance for Polygenic Scores.” *Intelligence* 59: 103–8.

Hugh-Jones, David and Abdel Abdellaoui. 2022. “Human capital mediates natural selection in contemporary humans “. Working paper.

Lindahl, Lena. 2008. “Do Birth Order and Family Size Matter for Intergenerational Income Mobility? Evidence from Sweden.” *Applied Economics* 40 (17): 2239–57.

National Statistics. 2014. “Stepfamilies in 2011.” https://webarchive.nationalarchives.gov.uk/20160105222243/http:// www.ons.gov.uk/ons/rel/family-demography/stepfamilies/2011/stepfamilies-rpt.html.

Rimfeld, Kaili, Eva Krapohl, Maciej Trzaskowski, Jonathan R. I. Coleman, Saskia Selzam, Philip S. Dale, Tonu Esko, Andres Metspalu, and Robert Plomin. 2018. “Genetic Influence on Social Outcomes During and After the Soviet Era in Estonia.” *Nature Human Behaviour* 2 (4): 269–75.

Schwartz, Christine R, and Robert D Mare. 2005. “Trends in Educational Assortative Marriage from 1940 to 2003.” *Demography* 42 (4): 621–46.

Shakespeare, William. 1595. *A Midsummer Night’s Dream*.

Solon, Gary. 2018. “What Do We Know so Far about Multigenerational Mobility?” *The Economic Journal* 128 (612): F340–52.

Tambs, Kristian, Jon Martin Sundet, Per Magnus, and K re Berg. 1989. “Genetic and Environmental Contributions to the Covariance Between Occupational Status, Educational Attainment, and IQ: A Study of Twins.” *Behavior Genetics* 19 (2): 209–22.

Trzaskowski, Maciej, Nicole Harlaar, Rosalind Arden, Eva Krapohl, Kaili Rimfeld, Andrew McMillan, Philip S. Dale, and Robert Plomin. 2014. “Genetic Influence on Family Socioeconomic Status and Childrens Intelligence.” *Intelligence* 42 (January): 83–88