Trading social status for genetics in marriage markets: evidence from UK Biobank

Abdel Abdellaoui[*], Oana Borcan[†], Pierre Chiappori[‡] & David Hugh-Jones[§]

2022-03-30

Abstract

If social status and genetic variants are both assets in marriage markets, then the two will become associated in spouse pairs, and will be passed on to subsequent generations together. This process provides a new explanation for the surprising persistence of inequality across generations, and for observed genetic differences across the distribution of socio-economic status. We model Social-Genetic Assortative Mating (SGAM) and test for its existence in a large genetically-informed survey. We compare spouses of individuals with different birth order, which is known to affect socio-economic status and which is exogenous to own genetic endowments among siblings. Spouses of earlier-born siblings have more genetic variants that predict educational attainment. We provide evidence that this effect is mediated by individuals' own educational attainment and income. Thus, environmental shocks to socio-economic status are reflected in the DNA of subsequent generations. SGAM reveals a new aspect of the inheritance of inequality in contemporary and historical societies.

Introduction

Over the long run, inequality is surprisingly persistent across generations (Clark and Cummins 2015; Solon 2018). Intergenerational mobility is correlated with cross-sectional inequality (Becker et al. 2018; Krueger 2012), which has risen dramatically in high-income countries, at the same time as intergenerational absolute mobility has declined (Western, Bloome, and Percheski 2008; Chetty et al. 2017).[1] Assortative mating in marriage markets can increase the variance of human capital and income across families (Breen and Salazar 2011; Greenwood et al. 2014), and

[*]Department of Psychiatry, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands. Email: a.abdellaoui@amsterdamumc.nl

[†]School of Economics, University of East Anglia, Norwich, UK. Email: O.Borcan@uea.ac.uk

[‡]Department of Economics, University of Columbia, New York. Email: pc2167@columbia.edu

[§]Corresponding author. School of Economics, University of East Anglia, Norwich, UK. Email: D.Hugh-Jones@uea.ac.uk

[1]Though relative mobility has been stable (Chetty et al. 2014). In the United Kingdom the Gini coefficient has increased from 26% to 34.6% between 1977 and 2020. The United States has seen a 10 percentage point rise to 43.3% during 1962-2013.

increasing returns to human capital may explain rising inequality (e.g. Kaplan and Rauh 2013; Becker et al. 2018). It follows that how families are formed, and how they transmit traits and assets to their offspring, are critical for understanding inequality. These processes have been studied from both socio-economic and genetic angles. While educational homogamy is well established, genetic assortative mating has been demonstrated only recently (Hugh-Jones et al. 2016; Robinson et al. 2017). Similarly, wealthy families pass on advantages to their children through both genetic inheritance and environmental influence (Rimfeld et al. 2018; Björklund, Lindahl, and Plug 2006).[2]

This paper examines a plausible, but under-analysed, aspect of the spouse matching process: that both social status and genetics contribute to a person's attractiveness in marriage markets, and as a result, genetics and inherited social status may become associated in subsequent generations.[3] For example, suppose that wealth and intelligence are both positive assets in a potential spouse. Then wealthy people are more likely to marry intelligent people, and their children will inherit both wealth, and genetic variants associated with intelligence. We call this mechanism Social-Genetic Assortative Mating (SGAM). SGAM may be an important channel for the transmission of inequality. It leads to a hidden dimension of advantage for privileged families – hidden because most social science datasets do not include genetic information. This dimension may help to explain the surprising long-run persistence of inequality (Clark and Cummins 2015; Solon 2018). At the same time, this advantage is not an exogenous fact of biology, but endogenous to the social structure. Indeed, under SGAM, environmental shocks to an individual's social status may be reflected in the genetics of his or her children.

Below, we first outline a theoretical framework where attractiveness in the marriage market is a function of both socio-economic status (SES) and genetic variants. We show that social-genetic assortative mating in one generation increases the correlation between SES and genetic variants in the offspring generation. This result provides a new explanation of the association between genetics and socio-economic status (Belsky et al. 2018; Rimfeld et al. 2018; Björklund, Lindahl, and Plug 2006). While existing explanations have focused on meritocratic social mobility (genes cause SES), under SGAM causality goes both ways, from genes to SES and vice versa.

Next, using novel data on matched spouses born between 1935 and 1970 from the UK Biobank, we empirically test the hypothesis that an individual's higher social status attracts spouses with higher genetic potential for educational attainment. Our genetic measure, the Polygenic Score for Educational Attainment (PSEA), derives from large-scale genome-wide association studies (Lee et al. 2018) and is causally related to educational attainment itself, as well as to intelligence and labour market outcomes. It is already known that humans mate assortatively on PSEA, which

[2] See Sacerdote (2011) for a review of the behavioural genetics and economics literatures on the nature vs nurture debate; for a broader review of the studies on intergenerational transmission of income see Black and Devereux (2010).

[3] *Social status* refers to characteristics that an individual possesses in virtue of their social position. For example, my wealth is a fact about me that holds in virtue of my relationship to certain social institutions (bank deposits, title deeds et cetera). Other examples include caste, class, income, and educational qualifications. *Socio-economic status* (SES) is a specific type of social status which exists in economically stratified societies, and which refers to a combination of educational attainment, occupational class, income and wealth (e.g. White 1982).

makes it a likely candidate for detecting SGAM (Hugh-Jones et al. 2016; Robinson et al. 2017). We depart from the assumption that both socio-economic status and genetic traits can enhance the attractiveness of potential spouses, and are substitutable. This assumption has received support in recent economic studies of marriage markets, which suggest that people trade off physical characteristics for higher earnings (e.g. Chiappori, Oreffice, and Quintana-Domeque 2012) or matching social status (Banerjee et al. 2013).

The endogeneity of social status is the main challenge in identifying the causal effect of social status on the spouse's genetic endowment. For instance, individuals with high education qualifications tend to also have high educational attainment genes, and as mentioned above, they may take partners based on genomic similarity. In order to isolate the causal link from own socio-economic status to partner genes, we use the "accident of birth" as a source of exogenous variation in socio-economics status. Specifically, we use the birth order of individuals in the sample as a "treatment" which affects their partner choice through a range of mechanisms, of which the most salient one is own socio-economic status. It is well documented that earlier-born children enjoy higher parental investment and have better life outcomes, including measures of socio-economic status such as educational attainment and occupational status (Black, Devereux, and Salvanes 2011; Booth and Kee 2009; Lindahl 2008). At the same time, birth order is independent of siblings' genetic endowments, a fact guaranteed by the biological mechanism involved (the "lottery of meiosis").

While birth order is plausibly independent of siblings' genes, it cannot be used as a valid instrument for socio-economic status, because it may affect partner choice through alternative mechanisms. Hence, we steer away from a two-stage procedure and instead rely on a mediation analysis similar to Heckman, Pinto, and Savelyev (2013), who decompose the average treatment effect into effects of measured and unmeasured consequences of treatment. Specifically, we estimate a reduced-form model with spouse genes associated with educational attainment as the dependent variable, and own birth order as the main independent variable. We then estimate a model which also includes measures of own socio-economic status. In the latter model, socio-economic status can be interpreted, under certain assumptions, as a mediator of the effect of birth order on spouse genetics. We also include controls to balance covariates across individuals with different birth orders in different cohorts.

We find that later-born children have spouses with significantly lower polygenic scores for educational attainment in the reduced-form regressions. When we include university attendance as a measure of socio-economic status, birth order is no longer significant, while university attendance increases the spouse's genetic endowment at 0.1% significance. A similar pattern holds when we proxy socio-economic status with a measure of income, although the sample is reduced. Thus, SES appears to mediate the effect of birth order on spouse genetics. The results are robust to the inclusion of several own phenotype traits and a rich set of own genetic traits.

Our paper contributes to several literatures. Firstly, we highlight a novel mechanism of assortative mating. The economics literature on matching in marriage markets has typically focused on educational similarities (e.g. Pencavel 1998) or social class or caste (e.g. Abramitzky, Delavande, and Vasconcelos 2011; Banerjee et al. 2013), but also sorting based on age, physical traits and ethnicity (Hitsch, Hortaçsu, and Ariely 2010). Matching decisions on the marriage market have also been shown to follow multiple criteria, with some degree of substitutability between them.[4] For instance, Chiappori, Oreffice, and Quintana-Domeque (2012) showed that individuals trade off BMI for partners' income or education and that the marginal rate of substitution between these characteristics is different for males and females. The genetics literature has focused on genetic assortative mating (GAM), the phenomenon that people with similar genes marry each other. Recent research has confirmed the long-standing conjecture that GAM takes place in contemporary human populations (Howe et al. 2019; Hugh-Jones et al. 2016; Robinson et al. 2017). Geneticists have also developed the concept of cross-trait assortative mating (Beauchamp et al. 2010; Sundet et al. 2005), which refers to people with (genes for) e.g. height marrying people with (genes for) e.g. intelligence. As a result, the two types of variation become associated. In this paper we bring the two literatures together, extending the idea of cross-trait assortative mating to encompass both socially inherited status, and biologically inherited genetic variants. Our results confirm that individuals with higher social status are more likely to attract a spouse with higher innate cognitive ability.

Secondly, our findings have implications for understanding the sources of economic inequality and intergenerational mobility. Clark and Cummins (2015) show using a database of surnames that long-run intergenerational persistence of wealth is higher than simple parent-child correlations would predict. In particular, grandparents' wealth predicts grandchildren's wealth even after controlling for parents' wealth. Clark (2021) argues that the data can be explained by an underlying process where unobserved genetic variation determines wealth. We show below that SGAM could also generate these patterns. The mechanism again is unobserved genetic variation, but the interpretation is slightly different, since we view genetic endowments not an exogenous source of variation, but as an asset effectively "traded" in marriage markets in exchange for wealth and social status.

SGAM also affects cross-sectional inequality, like other forms of assortative mating (Fernández and Rogerson 2001). The standard mechanism is that when couples assort with respect to some characteristic, the resulting households will have more variance in that characteristic than if couples match randomly. This may then carry over into higher inter-individual inequality in the next generation (Fernandez, Guner, and Knowles 2005; Eika, Mogstad, and Zafar 2019). In particular, a likely driver of the rise in inequality is the increase in market returns to human capital (e.g. Kaplan and Rauh 2013; Eika, Mogstad, and Zafar 2019). In this context, the distribution of human capital is a key contributor to inequality, in addition to inherited wealth.[5] Of course, one part of human capital is acquired and the other is genetic. From twin studies, the heritability of occupational class and educational attainment, i.e. the proportion of variance

[4] Oreffice and Quintana-Domeque (2010) show that height and BMI are associated with spouse earnings. Dupuy and Galichon (2014) find spouse matching on multiple independent dimensions, including education, height, BMI and personality.

[5] See Adermon, Lindahl, and Waldenström (2018), Mulder et al. (2009).

explained by genetic differences between individuals, is around 50% (Tambs et al. 1989). Genome-wide Complex Trait Analysis shows that the family socio-economic status of 2-year-old children can be predicted from their genes (Trzaskowski et al. 2014). Studies comparing parent-child income and education associations between adoptees and non-adoptees show that both post-birth environment and pre-birth conditions (genetics and to a lesser extent prenatal environment) contribute to the transmission of wealth and human capital (e.g. Björklund, Lindahl, and Plug 2006). Thus, an important component of inequality is the association between SES and genetic variation – the “genes-SES gradient”.

The standard explanation for the genes-SES gradient is social meritocracy. Parents with higher ability reap higher market returns, and they may then pass both higher socio-economic status and their genes to their children, leading to an association between the two (Belsky et al. 2018). This mechanism depends on the level of meritocracy in social institutions (Branigan, McCallum, and Freese 2013; Heath et al. 1985); in a society where social status was ascribed rather than earned, it could not take effect. Indeed, after the fall of communism in Estonia, the heritability of SES increased, presumably because post-communist society allowed higher returns to talent (Rimfeld et al. 2018).

Social-genetic assortative mating provides a complementary explanation for the association between genes and SES, one which does not require social meritocracy. Even when social status is entirely ascribed, it may still become associated with certain genetic variants, so long as their associated phenotypes (and not only status) are prized assets in marriage markets. Since meritocracy is historically rare, while assortative mating is universal, this suggests that genes-SES gradients are likely to be historically widespread.

SGAM may increase social inequality overall, if there are complementarities between genetic and environmental components of human capital. For example, higher-ability parents may make more productive investments in children’s human capital (Cunha and Heckman 2007; Cunha, Heckman, and Schennach 2010; Heckman and Mosso 2014; Kong et al. 2018). Becker et al. (2018) demonstrated how inequality can arise in a model of intergenerational transmission of human capital where high income, high human capital parents are able to invest more in their children’s human capital than low income parents. Thus, by bringing “good genes” and enriched environments together, SGAM may increase inequality in the next generation.

Lastly, we contribute to a literature in economics that examines the relationship between genetic and economic variables. Benjamin et al. (2011) is an early review. Several more recent papers use polygenic scores, in particular polygenic scores for educational attainment (Barth, Papageorge, and Thom 2020; Papageorge and Thom 2020; Ronda et al. 2020). These papers – like the vast majority of the behavior genetics literature (see e.g. Plomin, DeFries, and McClearn 2008) – take genetic endowments as exogenous and examine how they affect individual outcomes, perhaps in interaction with the environment. We take a different approach by putting genetics on the “left hand side”. Thus, our paper challenges the assumption, in economics and beyond, that genetic endowment is exogenous to economic characteristics. While this may be tenable in within-generation studies, it ceases to hold in intergenerational models. Social-genetic assortative mating is a causal mechanism going from socio-economic status to genetic traits. Furthermore, our model shows that the strength of this mechanism varies with the structure of the society's marriage market. When both genes and status are both relevant in marriage markets, then they become associated in the next generation. When social status is irrelevant and only good genes matter — or vice versa — then genes and status do not become associated. Social structure matters for the redistribution of genes, just as for other forms of capital.

The observations behind SGAM are not new. That status and physical attractiveness assort in marriage markets is a commonplace, and a perennial theme of literature. In the Iliad, powerful leaders fight over the beautiful slave-girl Bryseis. In Jane Austen's novels, wealth, attractiveness and "virtue" all make a good match. Marx (1844) wrote "the effect of ugliness, its repelling power, is destroyed by money." And Donald Trump claimed: "part of the beauty of me is that I am very rich." The literature on mate preference from evolutionary psychology (Buss and Barnes 1986; Buss 1989; Buss and Schmitt 2019) confirms that attractive mate characteristics include aspects of social status ("high earning capacity," "professional status") as well as traits that are partly under genetic influence ("intelligent," "tall," "kind," "physically attractive"). Despite this, we have found almost no previous work in genetics or economics that analyses SGAM or its consequences.[6]

Model

People in the marriage market have two characteristics: $x = (x_1, x_2) \hookrightarrow \mathcal{N}\left(\begin{matrix} 0 \\ 0 \end{matrix}, \begin{matrix} s^2 & \sigma \\ \sigma & S^2 \end{matrix}\right)$.

We interpret x_1 as a genetic measure, and x_2 as a measure of SES or of social status more generally. The correlation between x_1 and x_2 is then

$$Corr = \frac{\sigma}{sS}.$$

People's attractiveness is given by

$$i(x) = ax_1 + (1-a)x_2$$

where a is a parameter reflecting the relative importance of genetics to SES in the marriage market. Note that since

[6]Halsey (1958) showed in a two-class model that social mobility combined with assortative mating might increase the association between genetics and social class. Belsky et al. (2018) offer three reasons for the association between education-linked genetics and SES, but do not consider SGAM.

the variance of the shocks to x_1 and x_2 (see below) has been normalized to 1, a also reflects this variance. That is, a large variance of SES shocks (compared to genetic shocks) translates into a being large.

Attractiveness i is distributed $\mathcal{N}\left(0, \sigma_I^2\right)$, where

$$\sigma_I^2 = a^2 s^2 + (1-a)^2 S^2 + 2a(1-a)\sigma.$$

People form matches with transferable utility, where the surplus for a match between x and y is $S\left(i\left(x\right), i\left(y\right)\right)$ such that $\partial^2 S/\partial i \partial j > 0$, i.e. S is supermodular. As a result there is positive assortative mating on attractiveness: x matches with y only if $i\left(x_1, x_2\right) = j\left(y_1, y_2\right)$. We describe this as social-genetic assortative mating (SGAM).

Benchmark: couples under random matching

We also consider random matching (RM) as a benchmark to compare against SGAM. Under RM, the distribution of couples' characteristics is normal with mean 0 and covariance matrix

$$\mathbb{C}\begin{pmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} s^2 & \sigma & 0 & 0 \\ \sigma & S^2 & 0 & 0 \\ 0 & 0 & s^2 & \sigma \\ 0 & 0 & \sigma & S^2 \end{pmatrix}$$

Couples under SGAM

Proposition 1. *Under SGAM, the distribution of couples' characteristics is normal, mean 0 and covariance matrix*

$$\mathbb{C}\begin{pmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} s^2 & \sigma & A^2 & AC \\ \sigma & S^2 & AC & C^2 \\ A^2 & AC & s^2 & \sigma \\ AC & C^2 & \sigma & S^2 \end{pmatrix}$$

where:

$$\begin{aligned} A &= \frac{as^2 + (1-a)\sigma}{\sqrt{a^2 s^2 + (1-a)^2 S^2 + 2a(1-a)\sigma}} \quad \textit{and} \\ C &= \frac{a\sigma + (1-a) S^2}{\sqrt{a^2 s^2 + (1-a)^2 S^2 + 2a(1-a)\sigma}}, \end{aligned}$$

Proof. *See Appendix* ■

In particular, we consider the distribution of couples' SES. Under RM this has mean 0 and variance $2S^2$. Under SGAM, the variance is:

$$V(x_2 + y_2) = 2S^2 + 2C^2 > 2S^2$$

Note, however, that

$$V(x_2 + y_2) \leq 4S^2$$

which would be reached if $a = 0$, i.e. if people only matched on SES. Thus, SGAM increases inequality but less so than pure matching on SES.

Children

Couples each have 2 children. Assume that a child's characteristics are given by:

$$\begin{aligned} x'_1 &= \frac{\tau}{2}(x_1 + y_1) + \varepsilon \\ x'_2 &= \frac{\theta}{2}(x_2 + y_2) + \eta \end{aligned} \tag{1}$$

where x and y are the child's parents, and ε and η are normal random shocks with mean 0 and variance 1.

Parameter τ reflects genetic inheritance. Under standard biological assumptions $\tau = 1$ and characteristics show no regression to the mean. In our model this leads the variance of x_1 to grow without limit over generations. In reality, we expect $\tau < 1$ because very extreme characteristics are selected against, a process known as stabilizing selection.

Parameter θ reflects inheritance of SES. Unlike τ it may vary between societies. θ is high when there is high intergenerational transmission of SES. If we interpret x_2 narrowly as wealth, $(1-\theta)$ can be thought of as the rate of inheritance taxation.

For the time being, we assume that a person's genetic endowment has no impact on their SES. Technically, thus, x'_2 does not directly depend on x'_1. In a meritocratic society we would expect this to be violated. We show that even absent meritocracy, correlations between x'_1 and x'_2 can arise.

We can now calculate the covariance matrix for $x' = (x'_1, x'_2)$ as:

$$
\begin{aligned}
\mathbb{C} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \frac{\tau}{2} & 0 & \frac{\tau}{2} & 0 \\ 0 & \frac{\theta}{2} & 0 & \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} s^2 & \sigma & A^2 & AC \\ \sigma & S^2 & AC & C^2 \\ A^2 & AC & s^2 & \sigma \\ AC & C^2 & \sigma & S^2 \end{pmatrix} \begin{pmatrix} \frac{1}{2}\tau & 0 \\ 0 & \frac{1}{2}\theta \\ \frac{1}{2}\tau & 0 \\ 0 & \frac{1}{2}\theta \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2}A^2\tau^2 + \frac{1}{2}s^2\tau^2 + 1 & \frac{1}{2}\theta\sigma\tau + \frac{1}{2}AC\theta\tau \\ \frac{1}{2}\theta\sigma\tau + \frac{1}{2}AC\theta\tau & \frac{1}{2}C^2\theta^2 + \frac{1}{2}S^2\theta^2 + 1 \end{pmatrix}
\end{aligned}
$$

We now explore two issues.

- One is that under SGAM, genetic characteristics are no longer exogenous; because of assortative matching, they are (partly) socially determined. In practice, we shall show that even if genetics and SES are uncorrelated among parents, the expected genetic endowment of the child is positively related to parental SES.
- As a result, in the long run a correlation appears between traits; that is, high SES people have ‘better’ genes.

Regarding point 1, we compute the expected genetic characteristic of the child, conditional on parental SES:

$$
\mathbb{E}\left[\frac{\tau}{2}(x_1 + y_1) + \varepsilon \mid x_2 = v, y_2 = w\right]
$$

Given the symmetry of the model, this conditional expectation only depends on the parents’ total wealth, i.e. $v + w$.

Claim 1. *Under random matching, the expected genetic endowment of the children is proportional to the parents’ SES and to the covariance between SES and genetics for the parents. In particular, if $\sigma = 0$ (i.e. genetics and SES uncorrelated for the parents), then the expected genetic endowment of the children does not depend on parental SES.*

XXX put in appendix. *Under RM, the joint distribution of $\left(\frac{\tau}{2}(x_1 + y_1) + \varepsilon, x_2, y_2\right)$ is normal with mean $(0, 0, 0)$ and covariance*

$$
C = \begin{pmatrix} \frac{\tau^2}{2}\left(s^2 + \sigma\right) + 1 & \frac{\tau}{2}\sigma & \frac{\tau}{2}\sigma \\ \frac{\tau}{2}\sigma & S^2 & 0 \\ \frac{\tau}{2}\sigma & 0 & S^2 \end{pmatrix}
$$

Therefore

$$\begin{aligned}\mathbb{E}\left[\frac{\tau}{2}(x_1+y_1)+\varepsilon \mid x_2=v, y_2=w\right] &= \begin{pmatrix} \frac{\tau}{2}\sigma & \frac{\tau}{2}\sigma \end{pmatrix} \begin{pmatrix} S^2 & 0 \\ 0 & S^2 \end{pmatrix}^{-1} \begin{pmatrix} v \\ w \end{pmatrix} \\ &= \frac{\sigma\tau}{2S^2}(v+w)\end{aligned}$$

In particular, if $\sigma = 0$, this expectation is equal to 0. ■

Claim 2. *Under SGAM, if $\sigma = 0$ (i.e. genetics and SES are uncorrelated for the parents), then the expected genetic endowment of the children is increasing in parental SES. Moreover, the relationship increases with the ratio of genetic variance to wealth variance.*

Proof. *Here, the joint distribution of $\left(\frac{\tau}{2}(x_1+y_1)+\varepsilon, x_2, y_2\right)$ is normal with mean $(0,0,0)$ and covariance*

$$\Sigma = \begin{pmatrix} \frac{1}{2}\tau^2(A^2+s^2)+1 & \frac{\tau}{2}(\sigma+AC) & \frac{\tau}{2}(\sigma+AC) \\ \frac{\tau}{2}(\sigma+AC) & S^2 & C^2 \\ \frac{\tau}{2}(\sigma+AC) & C^2 & S^2 \end{pmatrix}$$

Therefore

$$\begin{aligned}\mathbb{E}\left[\frac{\tau}{2}(x_1+y_1)+\varepsilon \mid x_2=v, y_2=w\right] &= \begin{pmatrix} \frac{\tau}{2}(\sigma+AC) & \frac{\tau}{2}(\sigma+AC) \end{pmatrix} \begin{pmatrix} S^2 & C^2 \\ C^2 & S^2 \end{pmatrix}^{-1} \begin{pmatrix} v \\ w \end{pmatrix} \\ &= \frac{1}{2}\tau\frac{\sigma+AC}{C^2+S^2}(v+w) \qquad (2)\end{aligned}$$

In particular, if $\sigma = 0$, we have

$$\begin{aligned} A &= \frac{as^2}{\sqrt{a^2s^2+(1-a)^2S^2}} \quad \textit{and} \\ C &= \frac{(1-a)S^2}{\sqrt{a^2s^2+(1-a)^2S^2}}, \end{aligned}$$

and (2) becomes

$$\begin{aligned}\mathbb{E}\left[\frac{\tau}{2}(x_1+y_1)+\varepsilon \mid x_2=v, y_2=w\right] &= \frac{1}{2}\tau\frac{a(1-a)s^2}{a^2s^2+2(1-a)^2S^2}(v+w) \\ &= \frac{1}{2}\tau\frac{a(1-a)\lambda}{a^2\lambda+2(1-a)^2}(v+w)\end{aligned}$$

where $\lambda = s^2/S^2$. ■

The coefficient $\frac{a(1-a)\lambda}{a^2\lambda+2(1-a)^2}$ is plotted below (as a function of a and λ).

`![](trading-genetics_files/figure-latex/plot-model-1.pdf) <!---->`

Note that

- the coefficient cannot exceed $0.119\sqrt{2\lambda}$
- the maximum is reached for $a = \frac{2-\sqrt{2\lambda}}{2-\lambda}$,

(Plots omitted)

Point 2: 'Why are wealthy people smart ?'

Question: What is the relationship between an individual's wealth and genetic endowment?

To answer, we study the correlation between children's traits as a function of σ, the covariance of parents' traits. We first consider the general case, then concentrate on $\sigma = 0$ (i.e., traits initially uncorrelated).

Correlation between characteristics: the Random Matching benchmark

Under RM, the correlation matrix for children's characteristics is:

$$
\begin{aligned}
\mathbb{C} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \frac{\tau}{2} & 0 & \frac{\tau}{2} & 0 \\ 0 & \frac{\theta}{2} & 0 & \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} s^2 & \sigma & 0 & 0 \\ \sigma & S^2 & 0 & 0 \\ 0 & 0 & s^2 & \sigma \\ 0 & 0 & \sigma & S^2 \end{pmatrix} \begin{pmatrix} \frac{1}{2}\tau & 0 \\ 0 & \frac{1}{2}\theta \\ \frac{1}{2}\tau & 0 \\ 0 & \frac{1}{2}\theta \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2}s^2\tau^2 + 1 & \frac{1}{2}\theta\sigma\tau \\ \frac{1}{2}\theta\sigma\tau & \frac{1}{2}S^2\theta^2 + 1 \end{pmatrix}
\end{aligned}
$$

so that the correlation between charactersitics for children is:

$$
Corr\left(x_1', x_2'\right) = \frac{\frac{1}{2}\theta\sigma\tau}{\sqrt{\frac{1}{2}\tau^2 s^2 + 1}\sqrt{\frac{1}{2}\theta^2 S^2 + 1}}
$$

Two things can be noted. First, $\sigma = 0$ (uncorrelated characteristics for the parents) gives a zero correlation for children as well. Second, the correlation is less than that of the parents.

Claim 3. *Under random matching, the correlation between characteristics is smaller for children than for parents. In particular, if genetics and wealth are uncorrelated for the parents, then they are uncorrelated for the children.*

Correlation between characteristics under PAM

Now the correlation between children's traits:

$$Corr\left(x_1', x_2'\right) = \frac{\frac{1}{2}\theta\tau\left(\sigma + AC\right)}{\sqrt{\frac{1}{2}\tau^2\left(A^2+s^2\right)+1}\sqrt{\frac{1}{2}\theta^2\left(C^2+S^2\right)+1}}$$

In particular, if $\sigma = 0$ then

$$A = \frac{as^2}{\sqrt{a^2s^2+\left(1-a\right)^2S^2}}, C = \frac{\left(1-a\right)S^2}{\sqrt{a^2s^2+\left(1-a\right)^2S^2}},$$

and the correlation is

$$Corr\left(x_1', x_2'\right) = \frac{1}{2}\frac{\theta\tau}{a^2s^2+\left(1-a\right)^2S^2}\frac{a\left(1-a\right)s^2S^2}{\sqrt{\frac{1}{2}\tau^2\left(\frac{as^2}{a^2s^2+\left(1-a\right)^2S^2}+s^2\right)+1}\sqrt{\frac{1}{2}\theta^2\left(\frac{\left(1-a\right)S^2}{a^2s^2+\left(1-a\right)^2S^2}+S^2\right)+1}} > 0$$

Claim 4. *Under PAM, if genetics and wealth uncorrelated for the parents, then they are positively correlated for the children.*

Whether characteristics are more or less correlated for children than for parents depends on whether the initial correlation between parents' characteristics is larger or smaller than the asymptotic one, derived later on.

Empirical application

As an empirical application, we assume that $\sigma = 0$ and $s = S = 1$, and we set $\tau = .95$. Then

$$A = \frac{a}{\sqrt{a^2+\left(1-a\right)^2}} \text{ and } C = \frac{\left(1-a\right)}{\sqrt{a^2+\left(1-a\right)^2}},$$

The correlation between children's characteristics, as a function of θ and a, is given in the following graph

(omitted)

Extension: asymptotics

The RM benchmark

Main result

Under random matching, the distribution converges to an asymptotic distribution that is normal with mean zero. Its covariance is given by the following result:

Proposition 2. *Under RM, the dynamics converges to a stationary distribution that is normal with mean zero and covariance matrix*

$$C = \begin{pmatrix} \frac{2}{2-\tau^2} & 0 \\ o & \frac{2}{2-\theta^2} \end{pmatrix}$$

In particular, the traits are asymptotically uncorrelated.

Proof. *The fixed point condition on the covariance matrix is*

$$\begin{pmatrix} s^2 & \sigma \\ \sigma & S^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}s^2\tau^2 + 1 & \frac{1}{2}\theta\sigma\tau \\ \frac{1}{2}\theta\sigma\tau & \frac{1}{2}S^2\theta^2 + 1 \end{pmatrix}$$

which gives

$$s^2 = \frac{2}{2-\tau^2}, S^2 = \frac{2}{2-\theta^2}, \sigma = 0$$

■

Conditional expectation

One can then compute the asymptotic conditional expectation of children's genetics given parental wealth. Here:

$$\mathbb{E}\left[\frac{\tau}{2}(x_1 + y_1) + \varepsilon \mid x_2 = v, y_2 = w\right] = 0$$

reflecting the asymptotic absence of correlation between traits

The PAM case

We now study the asymptotic properties of the population dynamics created by the equations above.

Main result

Proposition 3. *For $\theta < 1$ and $\tau < 1$, the dynamics converges to a stationary distribution that is normal with mean zero and covariance matrix*

$$C = \left(\left(\begin{array}{cc} \bar{s}^2 & \bar{\sigma} \\ \bar{\sigma} & \bar{S}^2 \end{array} \right) \right)$$

For $\theta = 1$, the dynamics diverges and S^2 goes to $+\infty$; for $\tau = 1$, the dynamics diverges and s^2 goes to $+\infty$

Proof. *Start by characterizing the invariant distribution. This must satisfy:*

$$\left(\begin{array}{cc} \bar{s}^2 & \bar{\sigma} \\ \bar{\sigma} & \bar{S}^2 \end{array} \right) = \left(\begin{array}{cc} \frac{1}{2}\tau^2 \left(\bar{A}^2 + \bar{s}^2\right) + 1 & \frac{1}{2}\theta\tau \left(\bar{\sigma} + \bar{A}\bar{C}\right) \\ \frac{1}{2}\theta\tau \left(\bar{\sigma} + \bar{A}\bar{C}\right) & \frac{1}{2}\theta^2 \left(\bar{C}^2 + \bar{S}^2\right) + 1 \end{array} \right)$$

where

$$\begin{aligned} \bar{A} &= \frac{a\bar{s}^2 + (1-a)\,\bar{\sigma}}{\sqrt{a^2\bar{s}^2 + (1-a)^2\,\bar{S}^2 + 2a\,(1-a)\,\bar{\sigma}}} \text{ and} \\ \bar{C} &= \frac{a\bar{\sigma} + (1-a)\,\bar{S}^2}{\sqrt{a^2\bar{s}^2 + (1-a)^2\,\bar{S}^2 + 2a\,(1-a)\,\bar{\sigma}}}, \end{aligned}$$

Therefore

$$\begin{aligned} \bar{s}^2 &= \frac{1}{2}\tau^2 \left(\frac{\left(a\bar{s}^2 + (1-a)\,\bar{\sigma}\right)^2}{a^2\bar{s}^2 + (1-a)^2\,\bar{S}^2 + 2a\,(1-a)\,\bar{\sigma}} + \bar{s}^2 \right) + 1 \\ \bar{S}^2 &= \frac{1}{2}\theta^2 \left(\frac{\left(a\bar{\sigma} + (1-a)\,\bar{S}^2\right)^2}{a^2\bar{s}^2 + (1-a)^2\,\bar{S}^2 + 2a\,(1-a)\,\bar{\sigma}} + \bar{S}^2 \right) + 1 \\ \left(1 - \frac{1}{2}\theta\tau\right)\bar{\sigma} &= \frac{1}{2}\theta\tau \frac{\left(a\bar{\sigma} + (1-a)\,\bar{S}^2\right)\left(a\bar{s}^2 + (1-a)\,\bar{\sigma}\right)}{a^2\bar{s}^2 + (1-a)^2\,\bar{S}^2 + 2a\,(1-a)\,\bar{\sigma}} \end{aligned} \tag{3}$$

Define

$$\lambda = \bar{s}^2/\bar{S}^2 \text{ and } \mu = \bar{\sigma}/\bar{S}^2$$

The last equation gives

$$\left(1 - \frac{1}{2}\theta\tau\right)\mu = \frac{1}{2}\theta\tau \frac{(a\mu + (1-a))\,(a\lambda + (1-a)\,\mu)}{a^2\lambda + (1-a)^2 + 2a\,(1-a)\,\mu}$$

which is linear in λ; therefore

$$\lambda = \frac{\mu\left(\frac{1}{2}\theta\tau - 1\right)\left((a-1)^2 - 2a\mu\,(a-1)\right) - \frac{1}{2}\theta\tau\mu\,(a-1)\,(-a + a\mu + 1)}{a^2\mu\left(1 - \frac{1}{2}\theta\tau\right) - \frac{1}{2}a\theta\tau\,(1 - a + a\mu)} \tag{4}$$

The first two give:

$$1-\frac{1}{2}\theta^2\left(\frac{(a\mu+(1-a))^2}{a^2\bar{s}^2+(1-a)^2\bar{S}^2+2a(1-a)\bar{\sigma}}+1\right)=\lambda-\frac{1}{2}\tau^2\left(\frac{(a\lambda+(1-a)\mu)^2}{a^2\bar{s}^2+(1-a)^2\bar{S}^2+2a(1-a)\bar{\sigma}}+\lambda\right)$$

where λ is given by (4). This is equivalent to $F(\mu)=0$, where

$$F(\mu)=\lambda-\frac{1}{2}\tau^2\left(\frac{(a\lambda+(1-a)\mu)^2}{a^2\lambda+(1-a)^2+2a(1-a)\mu}+\lambda\right)-\left(1-\frac{1}{2}\theta^2\left(\frac{(a\mu+(1-a))^2}{a^2\lambda+(1-a)^2+2a(1-a)\mu}+1\right)\right)=0$$

This equation is quadratic in μ; thus it has two closed form solutions, of which one, denoted $\phi(a,\theta,\tau)$, is positive. Then

$$\lambda=\psi(a,\theta,\tau)=\frac{\phi(a,\theta,\tau)\left(\frac{1}{2}\theta\tau-1\right)\left((a-1)^2-2a\phi(a,\theta,\tau)(a-1)\right)-\frac{1}{2}\theta\tau\phi(a,\theta,\tau)(a-1)(-a+a\phi(a,\theta,\tau)+1)}{a^2\phi(a,\theta,\tau)\left(1-\frac{1}{2}\theta\tau\right)-\frac{1}{2}a\theta\tau(1-a+a\phi(a,\theta,\tau))}$$

Finally, the second equation in (3) gives

$$1=\frac{1}{2}\theta^2\left(\frac{(a\phi(a,\theta,\tau)+(1-a))^2}{a^2\psi(a,\theta,\tau)+(1-a)^2+2a(1-a)\phi(a,\theta,\tau)}+1\right)+\frac{1}{\bar{S}^2}$$

which gives $\bar{S}^2$, then $\bar{s}^2$ and $\bar{\sigma}$ follow. ■

Conditional expectation

Here:

$$\mathbb{E}\left[\frac{\tau}{2}(x_1+y_1)+\varepsilon \mid x_2=v, y_2=w\right] = \frac{1}{2}\tau\frac{\bar{\sigma}+\bar{A}\bar{C}}{\bar{C}^2+\bar{S}^2}(v+w) \quad (5)$$

$$= \mu\frac{\theta\bar{S}^2}{2\bar{S}^2-2} \quad (6)$$

Numerical application

In what follows, we impose

$$\tau=.95$$

The three components of the asymptotic covariance matrix (s^2, S^2, σ) are given in the following graph

(omitted)

Lastly, the asymptotic correlation between characteristics is:

$$Corr = \frac{\sigma}{sS} = \frac{\mu}{\sqrt{\lambda}} = \frac{\phi\left(a, \theta, \tau\right)}{\sqrt{\psi\left(a, \theta, \tau\right)}}$$

(graph omitted)

and the conditional expectation of genetics given wealth is proportional to wealth, coefficient:

Note that both $\bar{S}^2$ and $\bar{\sigma}$, as well as the correlation between characteristics and the conditional expectation of genetics given wealth, are increasing in θ, i.e. decreasing in the tax rate. Higher taxation reduces the asymptotic variance of wealth (not surprisingly), but also the correlation between genetic and wealth. I.e., the correlation should smaller in Sweden than in the US ...

$$
\begin{aligned}
S^2 &= G(a,\theta,\tau) = \frac{1}{1-\frac{1}{2}\theta^2\left(\frac{(a\phi(a,\theta,\tau)+(1-a))^2}{a^2\psi(a,\theta,\tau)+(1-a)^2+2a(1-a)\phi(a,\theta,\tau)}+1\right)} \\
s^2 &= g(a,\theta,\tau) = G(a,\theta,\tau)\,\psi(a,\theta,\tau) \\
\sigma &= h(a,\theta,\tau) = G(a,\theta,\tau)\,\phi(a,\theta,\tau)
\end{aligned}
$$

$$
\begin{aligned}
\bar{A} &= S\frac{a\lambda+(1-a)\mu}{\sqrt{a^2\lambda+(1-a)^2+2a(1-a)\mu}} \text{ and} \\
\bar{C} &= S\frac{a\mu+(1-a)}{\sqrt{a^2\lambda+(1-a)^2+2a(1-a)\mu}}, \\
CE &= \frac{1}{2}\tau\frac{\mu+\frac{(a\lambda+(1-a)\mu)(a\mu+(1-a))}{a^2\lambda+(1-a)^2+2a(1-a)\mu}}{\frac{(a\mu+(1-a))^2}{a^2\lambda+(1-a)^2+2a(1-a)\mu}+1} = \frac{S^2}{2}\frac{\mu\tau\theta^2+\theta(2-\theta\tau)\mu}{2S^2-2} \\
&= .475\frac{\mu\left(a^2\lambda+(1-a)^2+2a(1-a)\mu\right)+(a\lambda+(1-a)\mu)(a\mu+(1-a))}{(a\mu+(1-a))^2+\left(a^2\lambda+(1-a)^2+2a(1-a)\mu\right)}
\end{aligned}
$$

$$
\lambda = \frac{1}{2}\tau^2\left(\frac{(a\lambda+(1-a)\mu)^2}{a^2\lambda+(1-a)^2+2a(1-a)\mu}+\lambda\right)+\frac{1}{S^2}
$$

$$
\frac{1}{\theta^2}\left(2-\frac{2}{S^2}\right) = \frac{(a\mu+(1-a))^2}{a^2\lambda+(1-a)^2+2a(1-a)\mu}+1 \tag{7}
$$

$$
\frac{(2-\theta\tau)}{\theta\tau}\mu = \frac{(a\mu+(1-a))(a\lambda+(1-a)\mu)}{a^2\lambda+(1-a)^2+2a(1-a)\mu} \tag{8}
$$

$$
\frac{\phi(a,\theta,\tau)}{2}\frac{\theta(G(a,\theta,\tau))^2}{(G(a,\theta,\tau))^2-1}
$$