For Whom the Bell Curve Tolls: A Lineage of 400,000 English Individuals 1750-2020 shows Genetics Determines most Social Outcomes[1]

Gregory Clark, University of California, Davis and LSE (March 1, 2021)

Economics, Sociology, and Anthropology are dominated by the belief that social outcomes depend mainly on parental investment and community socialization. Using a lineage of 402,000 English people 1750-2020 we test whether such mechanisms better predict outcomes than a simple additive genetics model. The genetics model predicts better in all cases except for the transmission of wealth. The high persistence of status over multiple generations, however, would require in a genetic mechanism strong genetic assortative in mating. This has been until recently believed impossible. There is however, also strong evidence consistent with just such sorting, all the way from 1837 to 2020. Thus the outcomes here are actually the product of an interesting genetics-culture combination.

1. Introduction

It is widely believed that while social status - measured as occupational status, income, health, or wealth – is correlated between parents and children, this correlation is driven by parental investments in children, or by cultural transmission.[2] This belief has profound influence on peoples' perception of the fairness of social rewards, of the need for government intervention in the lives of disadvantaged children, and of the social value of education.

In this paper I test whether culture/human capital or genetics offers a better explanation of the inheritance of social attributes, using a lineage of 402,000 English individuals 1750-2020. To do so we have to specify both a general model of cultural/human capital inheritance, and one of genetic inheritance. There is already a well established model of additive genetic inheritance, formulated by Fisher in 1918. This I test against the data below. Specifying a model of cultural/human capital transmission as an alternative is more difficult. The ways culture/human capital has been hypothesized to operate are many and varied.

[1] This paper is very much a working paper for discussion at a seminar. It is a work in progress, incompletely references, and awkwardly expressed at points, and combines results from a number of different working papers.

[2] Studies of adoptions and of twins suggest that this belief is not well founded. Such studies suggest that genetic transmission explains the majority of social outcomes, but leave room for substantial social influences. See, for example, Sacerdote, 2007.

A Socio-Genetic Model of Inheritance

Most complex human traits are influenced genetically not by single variants in the DNA sequence, but by the additive effect of many locations in the DNA where there are variants in the base pairs, where each location itself has a very small effect on the trait in question. Height, for example, is inherited in just this polygenic way. For each location on the genome, i, where variants have an influence on height we can assign a score of d_i = 0, 1 or 2, depending on whether a person has 0, 1 or 2 copies of the tall variant across *n* locations. We can then assign a predicted height to the person

$$H = \sum_{i=1}^{n} a_i d_i \qquad (1)$$

where a_i reflects the effect of the polymorphism i on predicted height. For height there are believed to be at least 300 locations that matter in height determination. Table 1 illustrates the process:

Table 1: Determination of Genotypic Value with Additive Inheritance

Locus	Allele Value	Weight	Effect
1	$v_1 = 0,1, or 2$	w_1	$w_1 \times v_1$
2	$v_2 = 0,1, or 2$	w_2	$w_2 \times v_2$
3	$v_3 = 0,1, or 2$	w_3	$w_3 \times v_3$
…..			
n	$v_4 = 0,1, or 2$	w_n	$w_n \times v_n$
All			$H = \sum_{i=1}^{n} w_i v_i$

This type of genetic inheritance is called additive because the effect at each location is just the additive sum of how many height positive variants a person has at that location. Also the overall genetic effect is just the addition of the effects at each location. There are no important supplementary effects from dominant or recessive genes, and no interactive effects from combinations across the different loci on the genome. Thus this is a very simple and clear model of the intergenerational transmission of social status. In a famous paper in 1918 Ronald A. Fisher showed what the implications of this type of transmission of status would be for the correlation between all members of a family tree under different types of assortment by parents (Fisher, 1918). The key assumption of this model is that the environment has little independent impact on outcomes.

When we come to social outcomes the idea here will be that people inherit a set of abilities that determine, whatever their parents' circumstances, their ultimate outcome in terms of occupational status, education, health or longevity. For wealth, where there is an actual transfer between generations, we would not expect the Fisher rules to hold.

To model this additive genetic transmission we denote the genotype value of a person on any trait as x, and the corresponding phenotype as y. For any individual the phenotype is generated from the genotype, with some random component. This random component can be very small, as with the 10 digit ridge count on your fingers. Or it can be very large, as with adult longevity.

$$y_t = x_t + u_t \tag{2}$$

If we denote the genotypic value of the fathers as x_f and of the mother as x_m, with then the genotype of a child will be[3]

$$x_c = \left(\frac{x_f}{2} + \frac{x_m}{2}\right) + e_t = \bar{x}_p + e_t \tag{3}$$

In this case the hereditability of the trait, the proportion of phenotype variance explained by the genotype variance, will be the regression coefficient of the child phenotype on the average of the parent phenotype which is

$$h^2 = \frac{var(\bar{x}_p)}{var(\bar{x}_p)+var(\bar{u}_p)} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2} = \frac{\sigma_x^2}{\sigma_y^2} \tag{4}$$

Note that with respect to the average of the parents the genotype does not regress to the mean for children. But for individual parents there will be regression to the mean, however, which will depend on the degree of genetic assortment in mating.[4]

---

[3] The random error component e represents the randomness of which alles are inherited from each parent, and from the fact that there is not in fact a 50:50 split in allele inheritance by parent.
[4] A person of high or low genotypic value for a trait will usually match with someone of lower genotypic value.

Let the correlation between the genotypes of the parents on the relevant characteristics be *m*, and let the matching of parents be based solely on the genotype. Then the correlation in genotype between child and a single parent will be

$$\left(\frac{1+m}{2}\right)$$

Consequently the intergenerational correlation between a single parent and a child will be

$$h^2\left(\frac{1+m}{2}\right)$$

Note that this will also be the correlation between siblings.

If, however, parents are matching on the phenotype for this trait, and that phenotype correlation is r, the single parent-child correlation is now

$$h^2\left(\frac{1+r}{2}\right)$$

Note also that the genotype correlation between the parents will now be $m = h^2r$, and so is less than the phenotype correlation. Now the correlation between siblings is

$$h^2\left(\frac{1+rh^2}{2}\right) = h^2\left(\frac{1+m}{2}\right) \tag{5}$$

Table 2 shows the expected correlation between various relatives on either of these two types of assortment by parents. These correlations can be extended all the way down the family tree. It is notable that these correlations depend only on h^2, and m or r, as well as the degree of genetic distance between members of the tree.

Figure 1 shows the pattern of correlations as a function of genetic distance in the family tree, for the case where the parents match on the genotype. The vertical axis shows the correlation between family members, on a logarithm scale. The horizontal axis shows how many steps apart the parties are on the family tree. Figure 2 shows the equivalent pattern for the case where parents match based on the phenotype.

Note that in each case the decline of correlation across generations is by a factor of $\left(\frac{1+m}{2}\right)$. With random marriage in terms of genetics that factor will be 0.5. In that case the influence of earlier generations on current outcomes will soon decline to effectively 0. But there is evidence that m is in fact high, in the range 0.6-0.8. We discuss the consequences of this below.

Table 2: Correlations between relatives with assortative mating

Relative to Child	Parents Match on Genotype	Parents Match on Phenotype
Average of parents	h^2	h^2
Single parent	$h^2\left(\frac{1+m}{2}\right)$	$h^2\left(\frac{1+r}{2}\right)$
Sibling	$h^2\left(\frac{1+m}{2}\right)$	$h^2\left(\frac{1+m}{2}\right)$
Half-sibling	$h^2\left(\frac{1+m}{2}\right)^2$	$h^2\left(\frac{1+m}{2}\right)\left(\frac{1+r}{2}\right)$
Grandparent	$h^2\left(\frac{1+m}{2}\right)^2$	$h^2\left(\frac{1+m}{2}\right)\left(\frac{1+r}{2}\right)$
Uncle	$h^2\left(\frac{1+m}{2}\right)^2$	$h^2\left(\frac{1+m}{2}\right)^2$
Great-grandparent	$h^2\left(\frac{1+m}{2}\right)^3$	$h^2\left(\frac{1+m}{2}\right)^2\left(\frac{1+r}{2}\right)$
Cousin	$h^2\left(\frac{1+m}{2}\right)^3$	$h^2\left(\frac{1+m}{2}\right)^3$

Note: m is the correlation of parents on the relevant genotype, r the correlation on the relevant phenotype. h^2 is the regression coefficient of the child phenotype on the average of the parents phenotypes. Where the parents match on the phenotype the genotype correlation will be $m = rh^2$.
Source: Fisher, 1918, Crow and Felsenstein, 1968, Nagylaki, 1978.

Figure 1: Outcome Correlations as a Function of Genetic Distance, Matching on Genotype

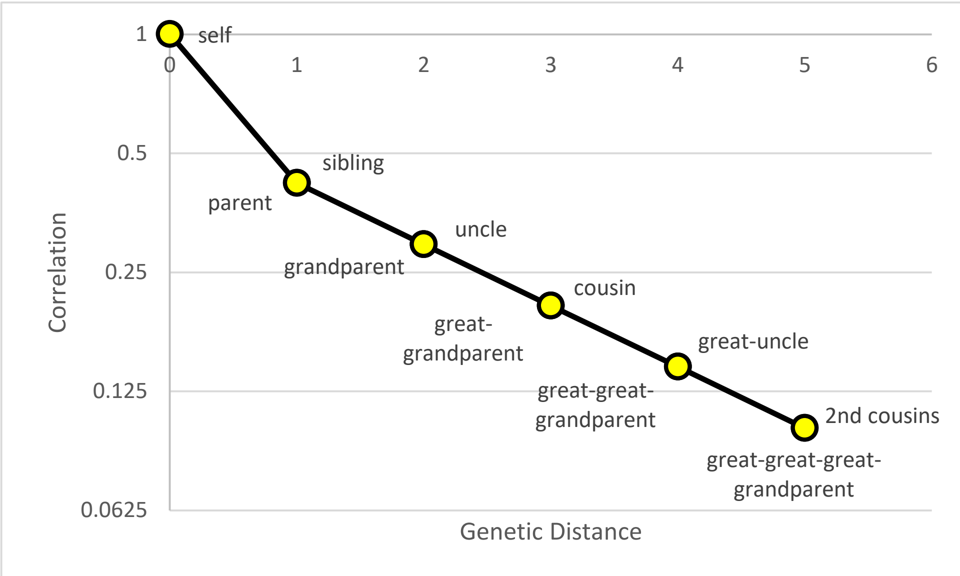
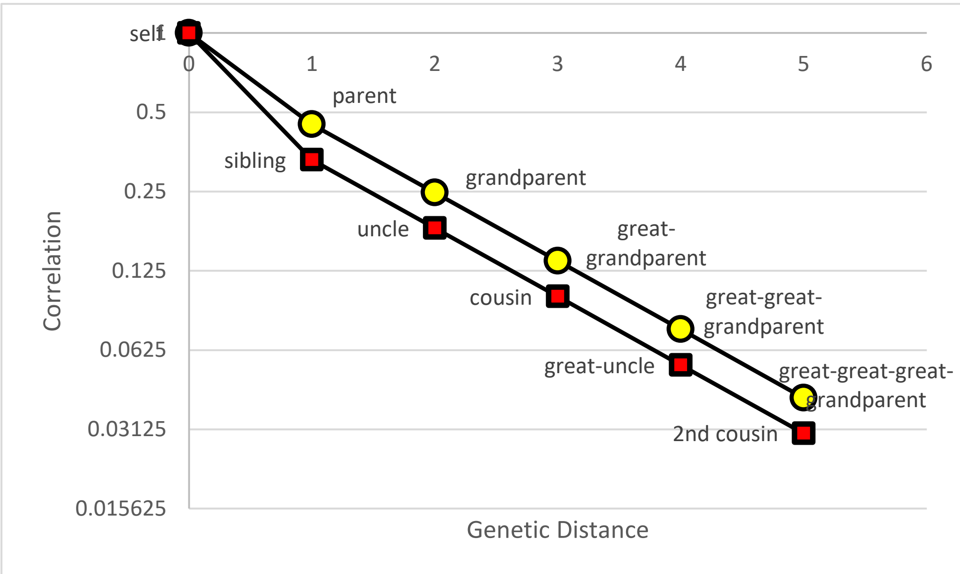


Figure 2: Outcome Correlations as a Function of Genetic Distance, Matching on Phenotype



The reason I have labelled this a **socio-genetic** model of social status is that the crucial parameter m, marital assortment, is completely socially determined. There is no intrinsic reason that people should match in marriage based on their social abilities. They could match purely on physical characteristics, or on personality traits unrelated to social and economic outcomes. They do match, in some societies, on whatever cousins are available of the appropriate age and gender. Interestingly, though, if matching is just to a random cousin then in equilibrium in such a society m will be quite low at around 0.23, whereas in England the evidence for m, as mentioned, is in the order of 0.6-0.8.

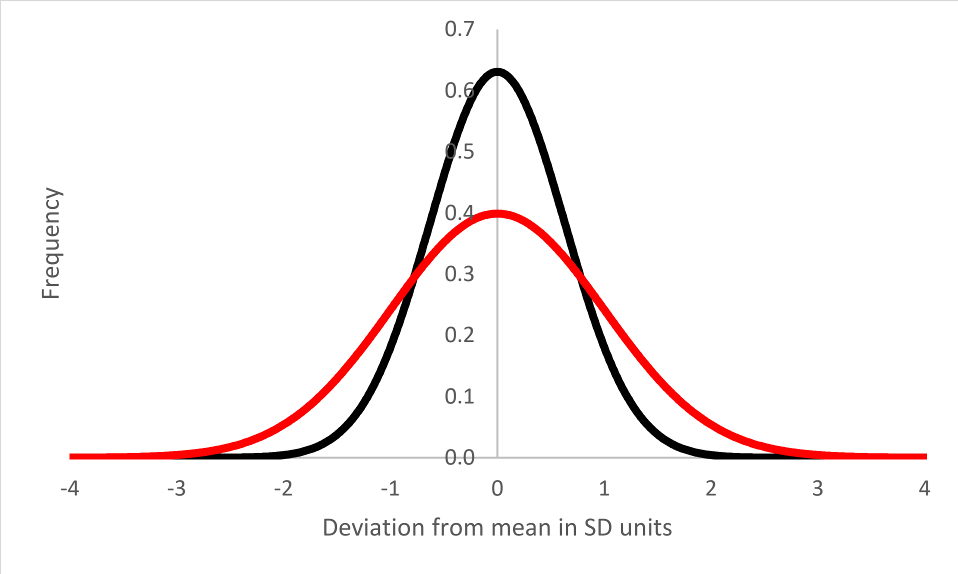
This social choice around marriage has profound implication in a world of genetic transmission for the overall individual distribution of social abilities in society, and for the rate of social mobility. In equilibrium the distribution of the genes for overall social abilities will be

$$\sigma_x^2 = \frac{\sigma_e^2}{1-\left(\frac{1+m}{2}\right)} \tag{6}$$

where σ_e^2 is from equation (3) the randomness in the shuffling of genes between generations. Figure 3 shows what will happen to the distribution as m increases. These effects are initially modest, but get very substantial as m rises. This for m = 0.6 compared to m = 0, the standard deviation of social abilities will be in equilibrium 1.58 times the no assortment value. If m got as high as 0.8 the standard deviation would be 2.24 times the value in absence of assortment.

Again as we increase m from 0 to 0.8 it would make the long run correlation in underlying status rise from 0.5 to 0.9.

Figure 3: Marital Assortment and the Social Distribution of Abilities



Testing the Socio-Genetic Model

Additive genetic transmission of status generates a number of interesting and testable implications, that will be examined below using the Families of England database.

1. Pattern of Correlations in a Family Tree

The first implication is that the correlation between relatives in a family tree declines at a constant rate $\frac{(1+m)}{2}$ where m is the genotype correlation of the parents. So long run social mobility rates depend purely on the degree of genetic assortment in marriage. Also the decline in correlations should follow a very regular pattern. For example, if we take people in a family tree who are roughly contemporaneous, siblings, cousins, 2nd cousins, 3rd cousins and so on, then the predicted correlations between them under either type of marital assortment would be

Siblings	$h^2\left(\frac{1+m}{2}\right)^1$
Cousins	$h^2\left(\frac{1+m}{2}\right)^3$
2nd Cousins	$h^2\left(\frac{1+m}{2}\right)^5$
3rd Cousins	$h^2\left(\frac{1+m}{2}\right)^7$

We can test whether we observe such a pattern of correlations for a variety of outcomes.

2. Siblings

The second interesting implication is that whatever the mechanism of marital assortment, the parent child correlation in status will equal or exceed that of siblings. This is a surprising and unexpected conclusion from the perspective of those who would give a large role to family environments in social outcomes. Children share the same parents, life in the same house, are located in the same community. They do not share parents with their parents, or the same house, or the same community. Further we observe regression to the mean between individual parents and their offspring in terms of social outcomes. If this is to be driven by family environments then there have to be systematic differences in family environment between parents and children. So on any environmental account of social accounts the correlation between children should be greater than that between parent and child. So this constitutes an interesting test of these competing accounts of social outcomes.

3. Type of Parental Assortment

The pattern of correlations among relatives will also reveal whether parental assortment is through the phenotype or the genotype. With assortment on the phenotype the parent child correlation will exceed that of siblings. If assortment is on the genotype then these correlations will be equal. Also with assortment on the phenotype the grandparent-child correlation will exceed the uncle/aunt-child correlation. With assortment on the genotype these correlations will again be equal.

4. Gender Symmetry

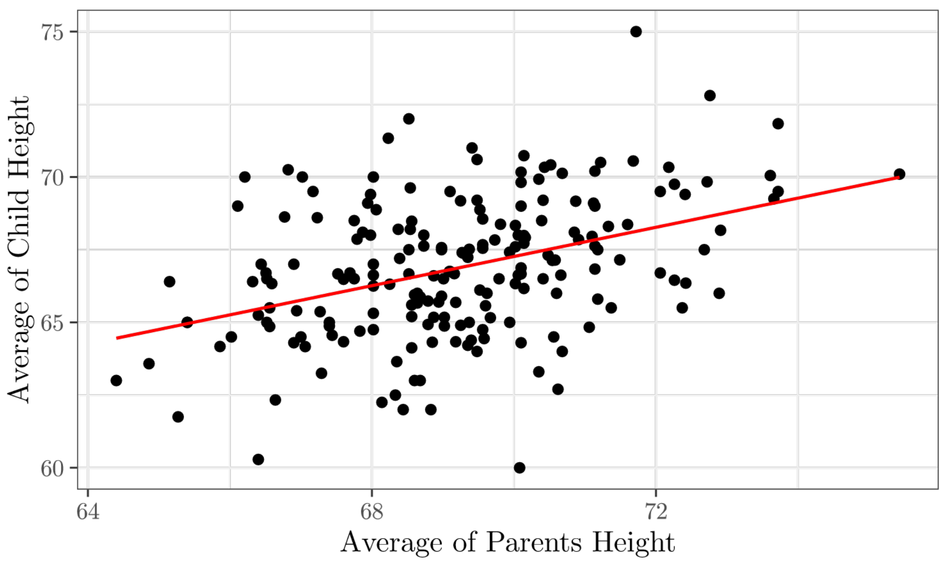
With genetic transmission of social outcomes there should be a strict symmetry of correlations between the paternal and maternal side of the family. The correlation, for example, in any outcome between the paternal grandfather and child should equal that between the maternal grandfather and child. In contrast with social transmission of social outcomes we can imagine significant asymmetry. Property may descend more through the male than the female line. Occupational opportunities can be linked more closely to the maternal line than the maternal, where there are family businesses.

5. Linearity of Regression to the Mean

Another implication of the additive genetic model is that there will be linearity in the regression to the mean observed between generations. The rate of movement to the mean will be the same all across the distribution of parental characteristics. It will be the same for the bottom 1% as for the top 1%. There will thus be no "wealth traps" or "poverty traps" where children of parents at the extremes of the social distribution show unusual persistence in their characteristics, as would potentially appear in social mechanisms of inheritance.

A good model for additive genetic inheritance in modern high-income societies is height, which is largely genetically inherited, and is the outcome of at least 300 genes each of which exerts very modest influence. Height inheritance certainly meets the second stipulation for the correlation pattern set out in table 1. Regression to the mean is indeed, at least to a first approximation, linear across the whole range of parent heights as figure 1 shows. The figure shows the heights of parents and children in Galton's pioneering study of the inheritance of heights. Height inheritance thus fits well this additive genetic model. Also, plausibly, mating actually sorts on the height phenotype rather than the underlying height genotype. So for height the long run intergenerational correlation should be close to 0.5.

Figure 4: Linearity of Regression to the Mean with Height



Source: Galton (1886)

6. Family Size

Exogenous shocks to family size should produce no effects on child outcomes. In England the evidence is that for marriages earlier than 1880 there was no attempt to control fertility. From the perspective of the parents family size was random.

7. Birth Order

Outcomes should be independent of birth order, except potentially for wealth where there are social rules of inheritance that can favor the first born, or all males.

8. Living Versus Dead Relatives

The correlations between individuals in a lineage should be unaffected by whether they ever interact. The status of living grandparents should predict child outcomes no better than dead relatives. In this model status is inherited as a first order Markov process, with only the parents genetics mattering. The relatives merely provide information on the underlying genotype of the parents. The amount of information they provide will be a function of their genetic distance from the parents.

9. Detecting the Underlying Genetic Correlations

A further implication of the pattern of correlations found in table 1 is that if we estimate the correlations between relatives in table 1 for one measure of social status, but instrument with instead the status of a relative, then when the OLS correlation is

$$h^2(\frac{1+m}{2})^n$$

the instrumented correlation will be

$$(\frac{1+m}{2})^n$$

An Economic Model of Social Status?

It would be great if there was a well accepted economic model of social status determination whose predictions we could contrast with the predictions of the socio-genetic model above. But economics has made little progress in developing a testable model of social status. The workhorse model of intergenerational social mobility in economics is that of Becker and Tomes (1979), now more than 40 years old. That model, as explicated by Solon (1999), assumes a parent (generation $t-1$) and one child (generation t), where the parent allocates their lifetime earnings y_{t-1} between their own consumption C_{t-1} and investment H_{t-1} in the child's earnings capacity. Parents cannot borrow on behalf of their children to invest in their human capital because of imperfect capital markets. Children also inherit abilities E from the parent. With this specification

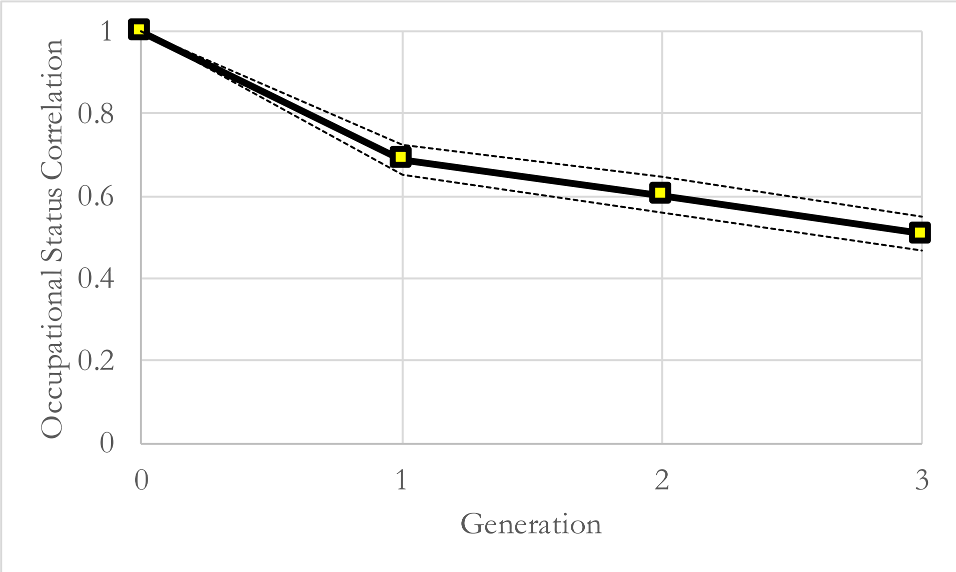
$$y_t = (1+r)H_{t-1} + E_t \tag{7}$$

where r is the return to human capital investment, and E_t is child ability. As Goldberger (1989) points out the predictions of this model are not distinguishable across from a model of simple regression to the mean. Since the model assumes one child it can say nothing about what to expect is the correlation in outcomes of siblings. With multiple children it could be that parents invest more human capital in the one with less talent to compensate, or it could be that they invest more in the talented offspring because the returns are greater. The only essential takeaway is that investment in human capital matters, and that social institutions which provide that investment socially should reduce the parent-child correlation of outcomes.

This means there are some predictions that we can test with the lineage data. The first is that positive shocks to wealth in the parent generation should transmit to greater human capital in the child generation (and thus higher occupational status). The second is that, though this steps outside the model which assumes one child per family, shocks that increase family sizes should under most assumptions, reduce child outcomes since there will be fewer resources per child. Another prediction would be that early deaths of either parent will reduce the resource flow to children and hence reduce their educational and occupational outcomes.

But as we shall see there is nothing in the Becker/Tomes framework that predicts or captures some unexpected patterns that appear clearly in the lineage data, such as that of figure 5. This figure shows exactly the pattern predicted in the socio-genetic model (figures 1 and 2) of correlations across multiple generations. To have an alternative to the socio-genetic model we need to have at least a model that would produce figure 5.

Figure 5: Occupational Status Correlations across Multiple Generations, England, births 1780-1930



Note: The dotted lines show the 1% confidence intervals around each estimated correlation.

A Cultural Model of Inheritance

One model that would capture the intergenerational pattern of figure 5 is as below. Suppose that social outcomes, *y*, are the product of the family culture/environment, *z*, and some random component, *u*, so that

$$y_{it} = z_{it} + u_{it} \,.^{5} \tag{8}$$

Suppose also that the family culture/environment is regressing to the mean at rate (1-b). Then

$$z_{it} = bz_{it-1} + e_{it} \,. \tag{9}$$

e_t is a random component that must exist to keep the dispersion of family culture *z* constant across families across generations.

With this structure the average correlation of social outcomes between parent and child will be

[5] In all cases in this paper variables are measured with mean 0, so that we can dispense with intercept terms in the equations.

$$\hat{\beta} = b\frac{\sigma_z^2}{\sigma_z^2+\sigma_u^2} = \theta b \tag{10}$$

The correlation between siblings, sharing a common environment, z_t , will on average be

$$\hat{\rho} = \frac{\sigma_z^2}{\sigma_z^2+\sigma_u^2} = \theta \tag{11}$$

Thus the sibling correlation will exceed the parent-child correlation on this cultural inheritance model. Note that this occurs even though we have left plenty of room for random influences on the outcomes among siblings exposed to common family culture. Between parent and child there is always the extra element of difference in that culture that does not exist for children. With b = .8, the sibling correlations will exceed the parent-child by 25%.

With the structure of inheritance embodies in equations (8) and (9) we can predict the correlations of any two relatives in a lineage. Thus

Parent	θb
Sibling	θ
Uncle/Aunt	θb
Grandparent	θb^2
Cousins	θb^2
Great Grandparent	θb^3

In particular the correlation between children and their parents should be less than that between children. The correlation of children and their parents should be the same as between children and their aunts and uncles. Also the correlation between children and their grandparents should equal that of children and their cousins. This is a different pattern of correlations than predicted in the socio-genetic model.

Status Determination in an English Lineage

We can test whether the cultural or additive genetic model of inheritance works predicts better using a lineage under construction for English families 1750-2020, that shows all familial links, plus a variety of social outcomes. So far we have the familial connections of all the people in the lineage (401,999), but only social outcomes for a subgroup of people. Figure 6 shows a sample lineage for one couple and some of their descendants from the database. The figure illustrates the richness of the set of family links that the database contains. In this case the lineage covers 7 generations. But what matters is the set of social outcomes we can associate with the members of the lineage. Table 3 summarizes the data currently available.[6] The social status indicators we have are wealth at death, occupation, educational attainment, schooling and training 11-20, and age at death.[7] The ones we employ here are wealth at death, occupational status, and higher educational attainment. Because of the time period covered by these measures, which is births 1750-1929, we include only men in the sample.

Wealth at Death: For England and Wales the Principal Probate Registry records whether someone was probated, and the value of their estate for all deaths in England 1858-2021. For 1799-1857 we also get from the Canterbury and York courts estate values for those higher in the wealth distribution (top 4% of men). For this measure we have women also, and so could potentially extend the study to study women also.

Occupation Status: Occupations are given in the censuses of 1841-1911 as well as the population register of 1939. There are also occupation statements in some marriage registers for both grooms and the fathers of the marriage parties, for fathers in birth registers, for the deceased in death registers, and also in some years for the deceased or for executors in probate records. We translated these various occupational statements into 348 occupational categories – carpenter, laborer, solicitor, dealer, stockbroker etc. We gave these occupations a social status score between 0 and 100. That score was created as an equally weighted average of three elements: average normalized ln wealth at death by occupation, average fraction of people in each occupation with a university degree or equivalent, and average fraction of males in each occupation who were in school or in training when observed ages 11-20 in the censuses of 1811-1911, and the population register of 1939.

[6] We expect to be able to add much more information on occupations and schooling.
[7] In recent years in England first names are a strong indicator of social status.

Figure 6: An Illustrative Portion of a Family Lineage, Lineage Database

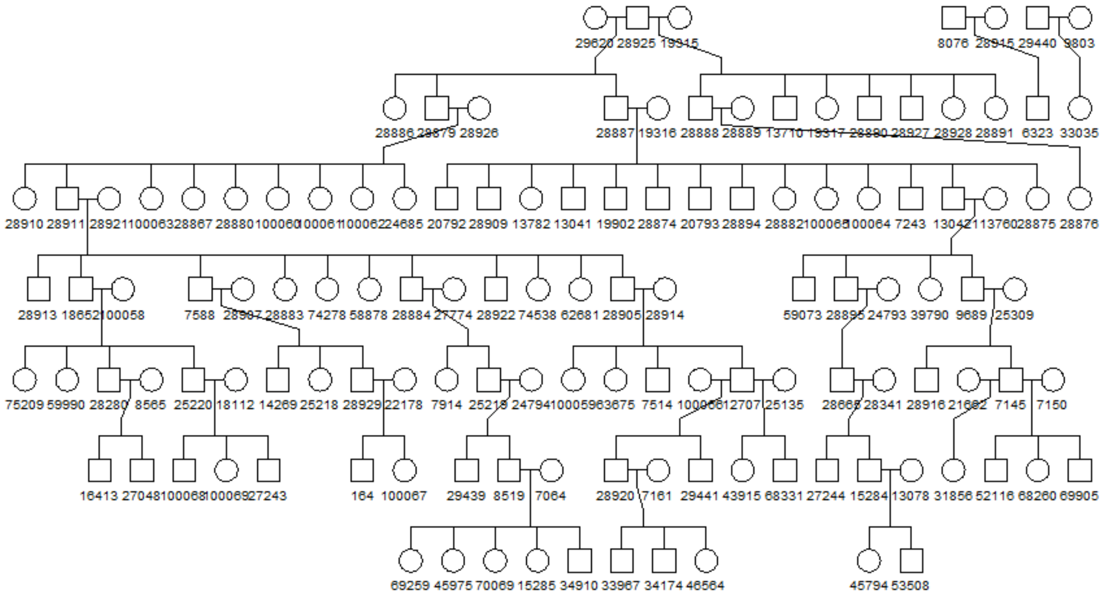


Table 3: Data Availability in the Families of England Lineage - Correlations

Relative	Genetic Distance*	Higher education	Occupational Rank	Wealth
Father	1	34,179	35,879	24,299
Brother	1	52,712	51,636	36,060
Uncle	2	89,698	80,188	64,760
Grandfather	2	26,603	24,780	17,849
Cousin	3	44,701	43,672	31,084
Great Grandfather	3	17,784	16,134	9,705
Great-Great Grandfather	4	9,965	8,673	3,486
2nd Cousin	5	45,644	45,746	31,016
Great-Great-Great Grandfather	5	5,494	4,758	698
3rd Cousin	7	37,899	39,935	26,730

*assuming marital assortment on the genotype

Higher Education: This is an indicator variable with a value 1 if the person achieved a higher educational status. Complete records are available for attendees Oxford and Cambridge (1750-2018), the Royal Military Academy Woolwich (1790-1839) and and the Royal Military College Sandhurst (1800-1946). Complete records are available for the UK Medical Registers, 1859-2017, UK, Civil Engineer Lists, 1818-1930, UK, Electrical Engineer Lists, 1871-1930, UK, Mechanical Engineer Records, 1847-1930, UK, Articles of Clerkship (attorneys), 1756-1874.

Table 2 shows the information available on each category of relative in our database currently. It also shows the share of genes that on average will be identical by descent between the various relatives. As can be seen, if there is not significant genetic assortment in marriage the share of genes relations like 3rd to 5th cousins will share will be extremely small, and thus any correlation in outcomes explained by genetics trivially small.

Table 3 shows the correlations of relatives of different degrees on three of the status characteristics, as well as the closeness of the genetic connection. If the degree of genetic distance is n, then the correlation in genetics will be

$$\left(\frac{1+m}{2}\right)^n$$

If m = 0 then that correlation for 5th cousins will be .0005. If, however, m = 0.6, then the correlation would be 0.09. As noted above the Fisher equation implies that the logarithm of the intergenerational correlation of status on any measure will be linear with respect to genetic distance. In particular where ρ_n is the correlation between relatives n steps apart genetically

$$\ln(\rho_n) = \ln(h^2) + nln\left(\frac{1+m}{2}\right)$$

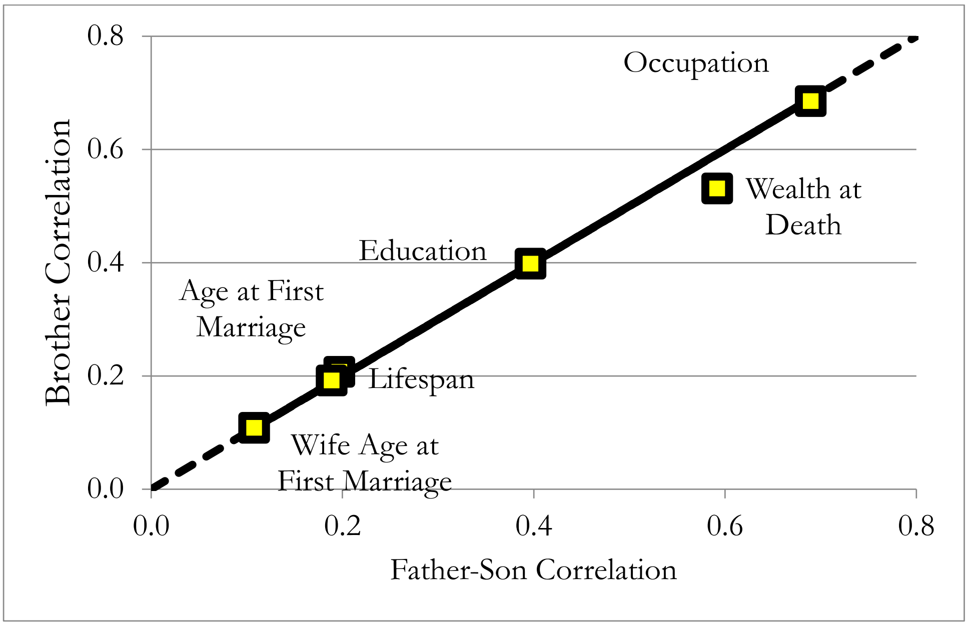
The most dramatic contrast between the cultural and the genetic model of transmission concerns the relative correlation of siblings versus parent-child in outcomes. These correlations are shown in the second and third lines of table 3. Figure 7 shows the relative father-son and brother correlations for the three attributes in table 7, as well as for lifespan, age at first marriage and wife's age at first marriage. As can be seen these correlations are near identical. This outcome is inconsistent both with cultural transmission, and also even with genetic transmission, but with assortment based on the phenotype (assuming strong assortment). Something is causing an unexpected degree in variation in the outcomes of siblings (from a cultural or parental investment perspective). Additive genetic transmission has a build in explanation of this pattern.

Table 4: Family Tree Correlations, Males

Relative	Genetic Distance	Occupational Rank	Higher Education	Wealth
Father	1	0.706	0.478	0.563
		(.004)	(.004)	(.004)
Brother	1	0.691	0.463	0.523
		(.004)	(.005)	(.005)
Uncle	2	0.607	0.316	0.439
		(.003)	(.004)	(.003)
Grandfather	2	0.621	0.394	0.481
		(.006)	(.006)	(.005)
Cousin	3	0.580	0.376	0.409
		(.004)	(.004)	(.005)
Great Grandfather	3	0.555	0.317	0.398
		(.008)	(.008)	(.015)
Great-Great	4	0.464	0.227	0.354
Grandfather		(.012)	(.011)	(.028)
2nd Cousin	5	0.377	0.295	0.278
		(.017)	(.004)	(.006)
Great-Great-Great	5	0.445	0.192	0.294
Grandfather		(.004)	(.015)	(.069)
3rd Cousin	7	0.338	0.253	0.192
		(.005)	(.005)	(.006)

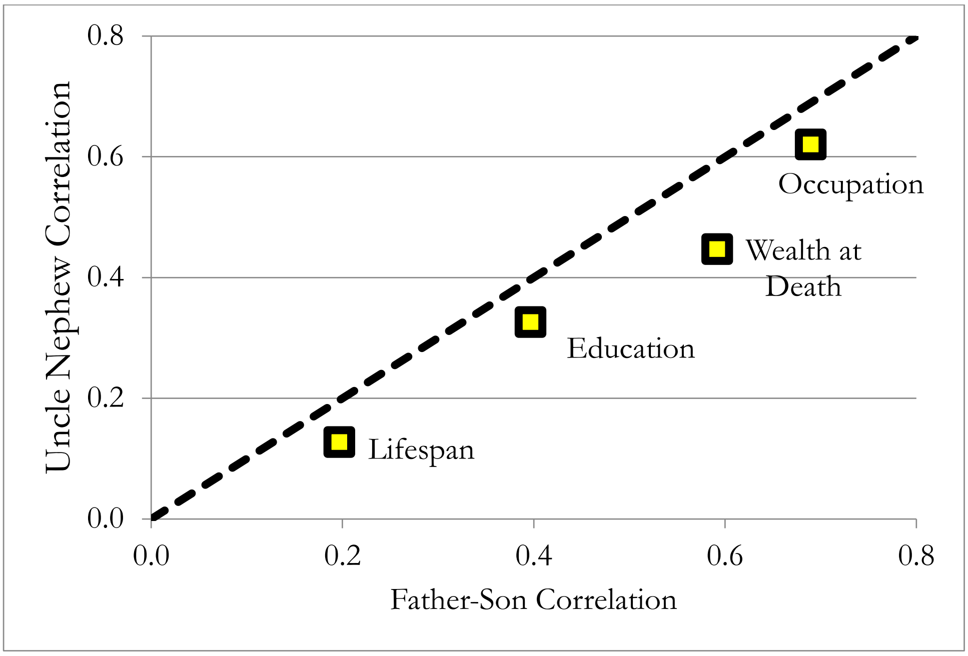
*assuming marital assortment on the genotype
Note: Pearson Correlations. Standard errors in parentheses.

Figure 7: Comparative Father-Son and Brother Correlations



Another implication of the cultural model outlined above is that the correlation of children in status with their fathers should equal that with their uncles. With genetic transmission the uncle correlation is lower by a factor of $\left(\frac{1+m}{2}\right)$. Figure 8 shows these comparative correlations compared with the prediction of the cultural model, which is that the correlations will fall along the 45° line. The pattern of correlations is again systematically at variance with the simple cultural model, and in line with the genetic model. The uncle-nephew correlations are all smaller than those of father-son.

Figure 8: Father versus Uncle Correlations



This suggests that we need a modified cultural transmission model, which increases the correlation between father and son, relative to brothers. Suppose that we have, as before,

$$y_{it} = z_{it} + u_{it} \tag{12}$$

But now,

$$z_{it} = b(z_{it-1} + u_{it}) + e_{it} \ . \tag{13}$$

The idea here is that any deviation in brother outcomes from that predicted by their family environment as children gets embedded in the environment of their children. Thus now the outcomes of children are more closely linked to their fathers than to their brothers. Now the correlation between father and son rises from θb to b, while that between brothers remains at θ. With the right parameters we can have $\theta = b$, so that the father and sibling correlations are the same.

However, while this model will capture this observed feature of social mobility, it fails completely to capture other significant features. One is that whatever the short run mobility rates is for any aspect of social status, the long run mobility is the same. Now the long run mobility rate has to equal that in the short run. So this crucial aspect of the social mobility process is missing.

Figure 9, for example, plots all the correlations in table 4 for occupational rank, against the expected genetic distance between relatives. Since genetic distance has a predicted multiplicative effect in reducing correlations the correlation is shown on a logarithmic scale on the vertical axis. The dotted line shows the fitted relationship, where the implied long run intergenerational correlation of occupational status is 0.88, the implied genetic correlation of status in marriage is 0.76, and the R^2 of the fit is 0.96. Note how well genetic distance predicts the correlation of occupational status. The correlation ion occupational status across 3rd cousins, who would rarely know or interact with each other, is still 0.34, just half the correlation of brothers.

Figure 10 shows again the pattern of correlations, this time for educational status (measured as an indicator variable for attaining higher educational qualifications). Again the implied long run intergenerational correlation of occupational status is 0.88, and the implied genetic correlation of status in marriage is thus 0.76. But in this case the R^2 of the fit is only 0.64.

Figure 11 plots all the correlations in table 4 for wealth at death. Now b is estimated at 0.85, m at 0.69, and the R^2 is 0.98. Wealth involves for richer families a physical transfer of property. But even here the recipients have choices about whether to spend this transferred wealth, or accumulate it. And their decisions will be influences by their current wage earnings. So it is quite possible for there to be important genetic components underlying wealth at death.

Thus the empirical evidence here from the correlations is all consistent with genetic transmission of social status. But it requires a high degree of genetic assortment in marriage, in the region of 0.7 to 0.75. We consider below if that is feasible.

Figure 9: Occupational Status Correlations

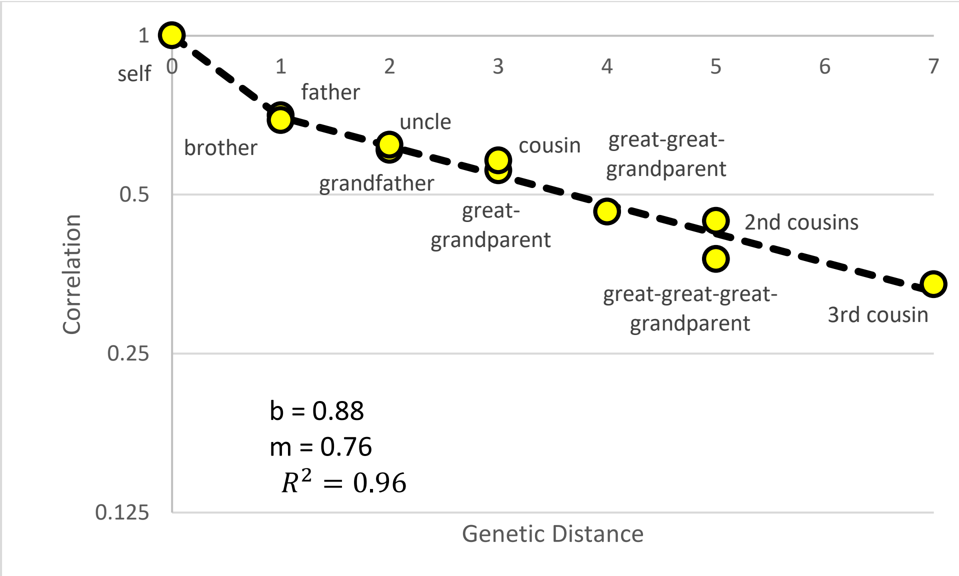


Figure 10: Educational Status Correlations

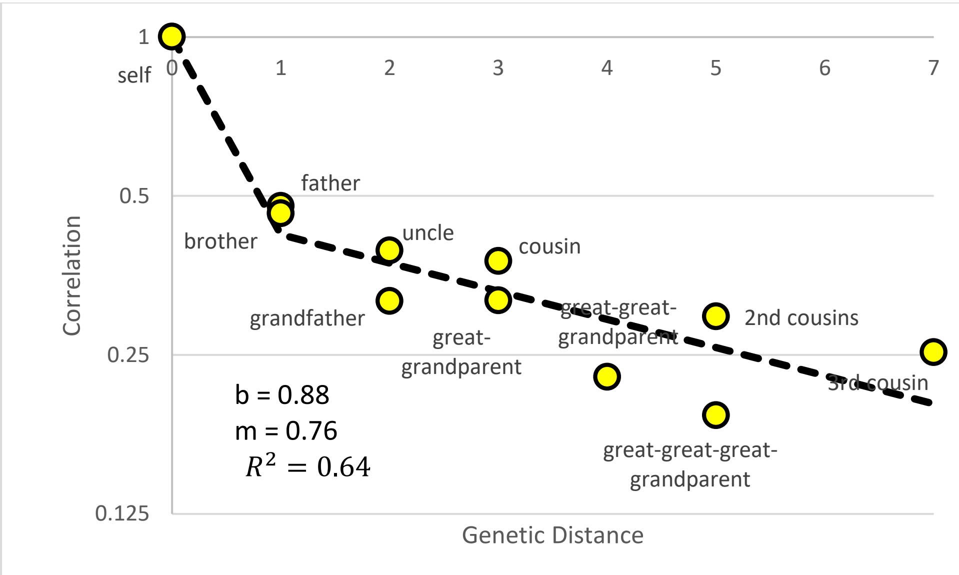
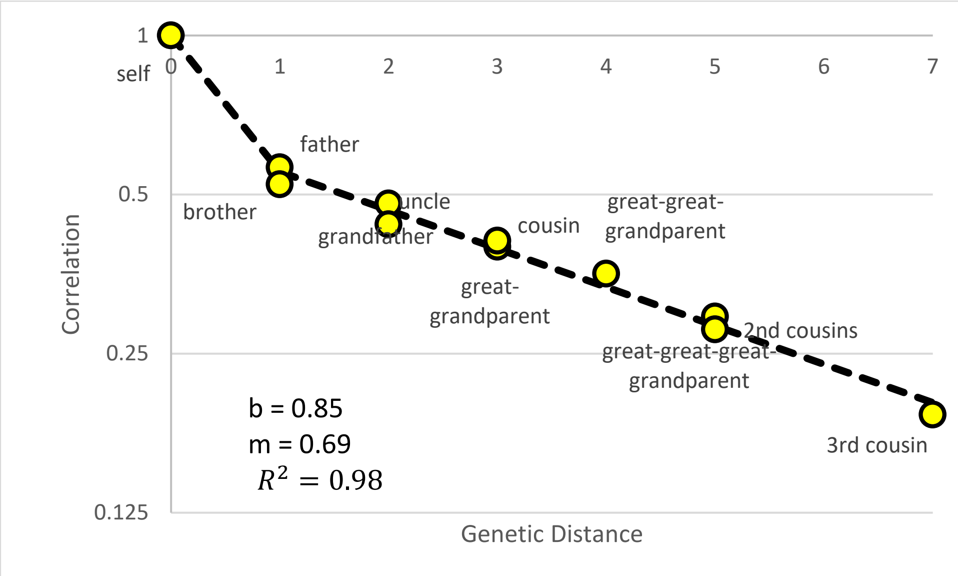


Figure 11: Wealth at Death Correlations and Genetic Distance



Other Tests of the Socio-Genetic Model

Gender Symmetry

In this model despite their social disabilities in much of the period of the FOE database, such as barriers to formal education and employment, the socio-genetic model would predict that women play a symmetrical role with that of men in determining their childrens' social outcomes. The way we can empirically test for this is by looking at the predictive influence of the maternal and paternal grandfathers on grandchild outcomes, since for much of the period women did not have formal educational qualifications or occupations. To test for symmetry we can estimate the coefficients in the regression

$$y_c = a + b_f y_{gf} + b_w y_{gw} + e$$

where y_c is the outcome for the grandson, y_{gf} the outcome for the paternal grandfather, and y_{gw} the outcome for the maternal grandfather. Table 5 shows the estimates.

Table 5: The Comparative Influence of Paternal and Maternal Grandfathers

	ln Wealth	Occupational Rank	Higher Education†
Grandfather Status	(1)	(2)	(3)
ln Wealth			
Paternal	.298***		
	(.029)		
Maternal	.105***		
	(.030)		
Occupational Rank			
Paternal		.346***	
		(.018)	
Maternal		.312***	
		(.019)	
Higher Education†			
Paternal			.271***
			(.020)
Maternal			.289***
			(.020)
Observations	577	2,509	2,299
R^2	.292	.498	.252
Note:			*p<0.05; **p<0.01; ***p<0.001 † Indicator Variable.

As can be seen for wealth there is a distinctive asymmetry, where paternal lineage wealth is nearly three times as influential in predicting grandchild wealth as is maternal lineage wealth. But when we turn to occupational status or to attainment of higher education there is no difference in the predictive power of the paternal versus the maternal grandfathers. For these attributes women were just as important as men in predicting outcomes. This is consistent with the model where it was parental genetics which were the transmitters of social status.

Genetics and Environment Interactions

The assumption above that underlies the Fisher formula for the correlation of relatives that "*Genes and environment are uncorrelated, or the environment has little independent impact on outcomes*" would seem to be obviously violated in the case of social traits. The familial environment of families does clearly vary, and that variation will be correlated with the genetics that help determine family social status. However, there is good information from the *Families of England* database that the second part of the condition largely holds, and that family environments do indeed have *little independent impact on outcomes.*

The lineage used in this paper also allows us to test for the effects of elements of family environment on social outcomes, because for part of the period covered by the lineage, marriages 1780-1880, family size was largely random. In this period there was great variation in completed family size, numbers of children reaching age 21, with the size range in the sample for men ranging from 1 to 18. There was no correlation between family size and any measure of social status for fathers. There was also very weak correlation between brothers, and between fathers and sons, in terms of either births or completed family size. That correlation was in the range 0.03-0.05. Since brothers and fathers and sons correlate very strongly on an underlying latent variable for social status, which would correlate with lifestyles and choices on family size, this implies that both the number of births, and also childhood mortality, were mainly random in this interval, and not the product of individual decisions.

We just summarize the effects of the family size and birth order on social outcomes for marriages here, since we have another paper devoted to this substantial topic (Clark and Cummins, 2017). The families in the lineage can be separated into those lines where average wealth at death circa 1850 was high, and those where wealth at death then was average or non-existent. In the high wealth families servants anyway provided much of child care, so the effects of size might be expected to be less. In poorer families, this was a period where there was mostly no compulsory education. A legal requirement of school attendance to age 10 was only introduced in 1881. Thus in poorer families parents had to make an important decision about whether to support the children in schooling ages 11-20 that would be affected potentially by family size.

Table 6 summarizes the effect of family size, measured as either births (N0) or children reaching age 21 (N21) per father, on various social outcomes for marriages 1780-1879. Mostly results for sons are shown, except for the case of wealth at death, since in this period only sons have occupations and educational attainments. In each case the elasticity of the measure with respect to family size is given. In most cases there is a negative effect of size on outcomes. But if we look at adult outcomes – wealth at death, occupational status at age 40, attainment of higher educational qualifications, and child mortality for children – then only in the case of wealth are there significant effects, and here only for families from wealthy lineages. For occupational status, for example, the elasticities range from -0.03 to -0.09, implying that a 10% increase in family size reduces occupational status by 0.3%-0.9%.

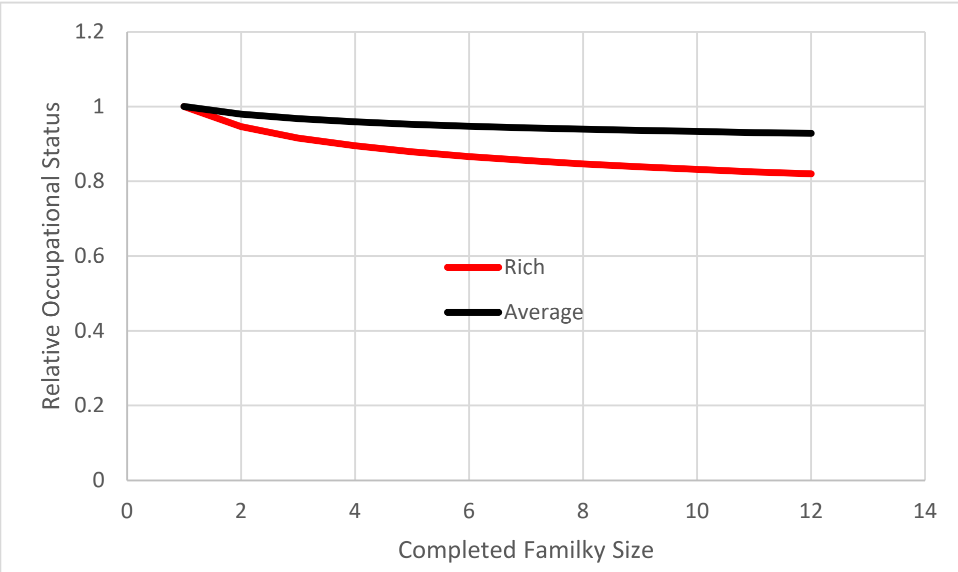
Table 6: Elasticities of Outcomes with respect to Family Size, Marriages 1780-1879

Variable	Rich	.	Poor/Avg.	.
	N0	N21	N0	N21
At School, 11-20, Boys	−.07	−.05	−.24***	−.13*
At School, 11-20, Girls	−.04	−.04	−.26***	−.14*
Higher Education	−.17***	−.12**	.32	.25
Occupational Status	−.09***	−.08***	−.05**	−.03
Child Survival, Children	.01	.01	−.01	.01
Wealth	−.56***	−.50***	−.25***	−.10

Note: *p<0.1; **p<0.05; ***p<0.01

To illustrate how modest are the size effects in Table 12 consider figure 12, which shows relative adult occupational status as a function of completed family size for average families, and for the families of the rich. For families in the lineages of average or poor social status expanding family size from 1 to 12 reduced the adult occupational status of children by 7%. For rich families the estimated effect was greater, an 18% reduction. But in both cases the overwhelmingly strong predictor of social outcomes was the social status of the father, not the numbers of children in the family.

Figure 12: Family Size and Occupational Status of Children, marriages before 1880



Marital Assortment

We see above that marital assortment at the genetic level is key to explaining the substantial persistence of status across multiple generations if genes are the transmission mechanism. To explain the observed persistence the genetic correlation has to be 0.6 or greater. But all measures of marital assortment in terms of measurable characteristics suggest much less assortment. Table 7, for example, shows some estimates of the correlation of partners in marriage on observables. Even for years of education the observed correlation is only 0.50.

Table 7: Phenotypic Correlations between Spouses

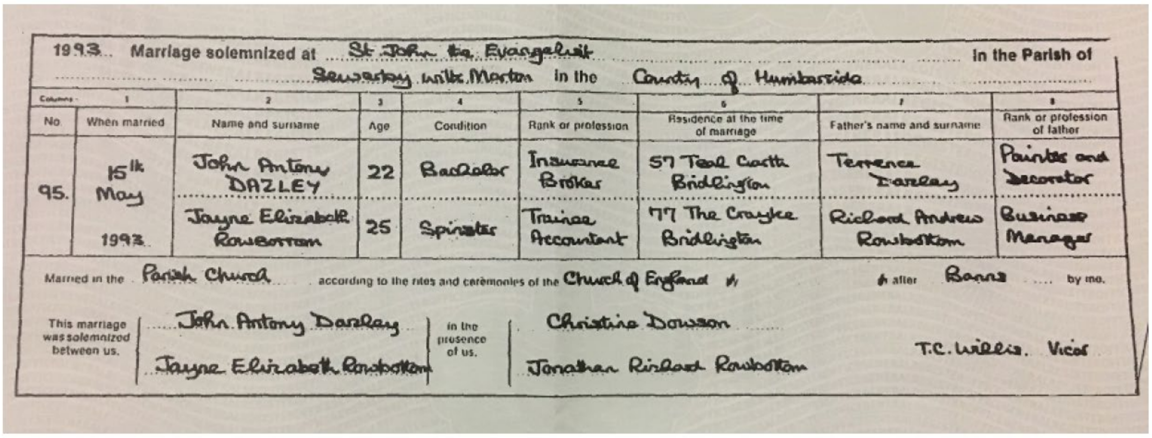
Characteristics	Correlation	Source
Height	0.29	McManus and Mascie-Taylor, 1984
Education	0.50	Watkins and Meredith, 1981
Income	0.34	Watkins and Meredith, 1981
Occupational Status	0.12	Watkins and Meredith, 1981
IQ	0.20-0.45	Mascie-Taylor, 1989
BMI	0.28	Abrevaya and Tang, 2011
Personality Traits	0.15	Mascie-Taylor, 1989

However a recent study from the UK Biobank, which has a collection of genotypes of individuals together with measures of their social characteristics, supports the idea that there is strong genetic assortment in mating. Robinson et al. (2017) look at the phenotype and genotype correlations for a variety of traits – height, BMI, blood pressure, years of education - using data from the biobank. For most traits they find as expected that the genotype correlation between the parties is less than the phenotype correlation. But there is one notable exception. For years of education, the phenotype correlation across spouses is 0.41 (0.011 SE). However, the correlation across the same couples for the genetic predictor of educational attainment is significantly higher at 0.654 (0.014 SE) (Robinson et al., 2017, 4). Thus couples in marriage in recent years in England were sorting on the genotype as opposed to the phenotype when it comes to educational status.

It is not mysterious how this happens. The phenotype measure here is just the number of years of education. But when couples interact they will have a much more refined sense of what the intellectual abilities of their partner are: what is their general knowledge, ability to reason about the world, and general intellectual ability. Somehow in the process of matching modern couples in England are combining based on the weighted sum of a set of variations at several hundred locations on the genome, to the point where their correlation on this measure is 0.65.

We can empirically test for the idea that couples are matching on some underlying social and intellectual ability using an interesting set of data that has been created by a somewhat arbitrary act of Parliament back in 1837. This is the marriage certificate in England and Wales, which has been largely unchanged from 1837 to the present. An oddity of the marriage certificate is that it records the occupation of the bride and groom, as well as their fathers. Figure 13 shows an example of such a marriage certificate from 1993. Note that in the years before 1960 the occupation of the bride was most commonly left blank.

Figure 13: Sample Marriage Certificate, 1993


1993. Marriage solemnized at St John the Evangelist in the Parish of Sewerby with Marton in the County of Humberside

No.	When married	Name and surname	Age	Condition	Rank or profession	Residence at the time of marriage	Father's name and surname	Rank or profession of father
95.	15th May 1993	John Antony DAZLEY	22	Bachelor	Insurance Broker	57 Teal Garth Bridlington	Terrence Dazley	Painter and Decorator
		Jayne Elizabeth Rowbottom	25	Spinster	Trainee Accountant	77 The Croyke Bridlington	Richard Andrew Rowbottom	Business Manager

Married in the Parish Church according to the rites and ceremonies of the Church of England after Banns by me.

This marriage was solemnized between us. John Antony Dazley, Jayne Elizabeth Rowbottom

in the presence of us. Christine Dowson, Jonathan Richard Rowbottom

T.C. Willis. Vicar

Suppose we assume that the parties in marriage are really combining based on underlying genetic characteristics which generate these occupational reports as a noisy signal of their underlying social abilities? What would the pattern of correlations between the parties look like. If we take m as the correlation in genetics between the spouses, and b as the correlation between bride and groom and their fathers (where in equilibrium $b = \left(\frac{1+m}{2}\right)$), then the pattern of underlying correlations will be as in figure 14. However, since the measures we observe are just an indictor of the underlying genotype, the observed occupational status of the parties will be correlated as in figure 15.

Figure 14: Underlying Correlation Pattern in Marriage

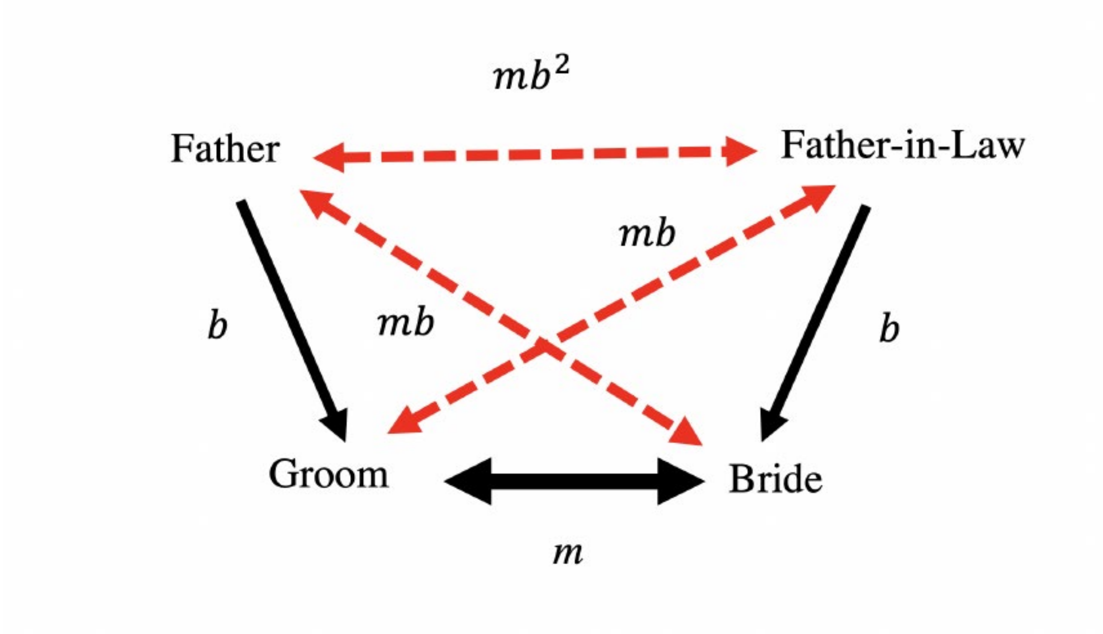


Figure 15: Observed Correlation Pattern in Marriage

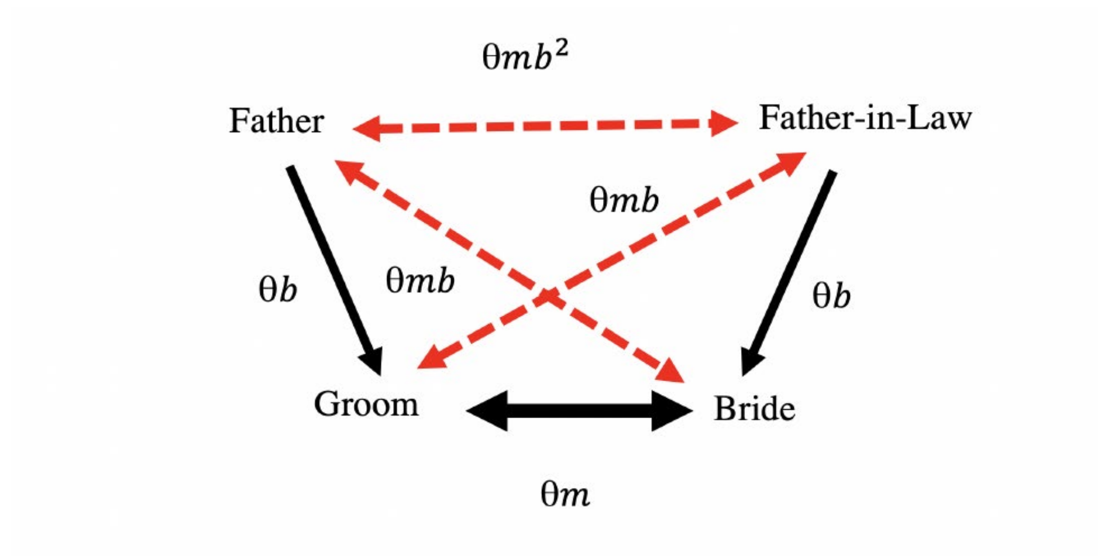


Figure 15 implies that we can estimate the true underlying correlation between bride and groom in social abilities just by dividing the groom's correlation in observed status with his father in law by the his correlation in observed status with his own father. Table 8 displays those correlations by 40 years periods 1837-59, 1860-99, 1900-39, 1940-79, and 1980-2020. The key point is that the correlation of grooms' occupational status with their own fathers' is only modestly higher than the correlation of their status with the father of their bride. By implication the groom and bride are matching very closely on some underlying social status – and potentially on genetics that determine social outcomes.

Table 8: Correlations in Occupational Status in Marriage of Groom and Fathers

Period	ρ_{sf}	se	ρ_{sl}	se	ρ_{fl}	se	N
1820-59	0.640	0.001	0.516	0.001	0.481	0.001	367,198
1860-99	0.618	0.001	0.503	0.002	0.467	0.002	311,029
1900-39	0.521	0.003	0.417	0.004	0.388	0.004	62,886
1940-79	0.391	0.009	0.332	0.010	0.336	0.010	9,482
1980-2020	0.263	0.034	0.175	0.035	0.199	0.035	788

Notes: Only calculated where the occupations of groom, father and father-in-law are simultaneously observed. FOE Occupation Rank. $_s$: Son (groom), $_f$: Father of groom, $_l$: Father-in-Law.

Table 9 shows the implied estimates of underlying marital assortment on social abilities between bride and groom for each period. These estimates are very high, implying a correlation of 0.8 or higher throughout the years 1837-1979 in the underlying social abilities of bride and groom. There is some chance that the correlation declined in the last 40 years, but the small numbers of observations in this period make that conclusion uncertain. The estimate of the underlying father-son and father-daughter correlations of status are also very high being closer to 0.9. These estimates may be distorted by the fact that the fathers and sons occupational status are measured at different points in the life cycle (in a way that does not influence the estimate of m). But they are actually quite consistent with the implication of the genetic model that $b = \left(\frac{1+m}{2}\right)$. If m = 0.8, then the expected value of b is 0.9.

But the important point here is that the pattern of correlations between groom, father, and father-in-law on marriage certificates imply that the wedding records for England 1837-2020 imply that grooms and brides are actually matching very closely on underlying social capabilities all through the period 1837-2020. Thus there is nothing in the data that contradicts the possibility that marriage involved a close genetic match of partners in the genetics relevant to social outcomes.

Table 9: Implied correlation in underlying characteristics in marriage, 1837-2020

Period	m	se	b	se	N
1820-59	0.806	0.003	0.933	0.004	367,198
1860-99	0.815	0.003	0.929	0.004	311,029
1900-39	0.801	0.009	0.930	0.012	62,886
1940-79	0.849	0.032	1.015	0.042	9,482
1980-2020	0.667	0.168	1.137	0.352	788

Notes: Only calculated where the occupations of groom, father and father-in-law are simultaneously observed. FOE Occupation Rank.

Conclusions

It is generally assumed that the elements that define social status – occupational status, educational attainment, wealth, and even health – are transmitted across generations in important ways by the family environment. Above we show that the patterns of correlation of social status attributes in an extended lineage of 402,000 people in England are mainly those that would be predicted by simple additive genetic inheritance of social status in the presence of highly assortative mating around status genetics. Parent-child correlations for a trait equal those of siblings, and the patterns of correlation of relatives of different degrees of genetic affinity is mainly consistent with that predicted by additive genetics. Further family size and birth order, elements that would significantly affect the family environment for children, have modest effects on adult outcomes. The underlying persistence of traits is such that people who have likely never interacted socially, such as second to fifth cousins, remain surprisingly strongly correlated in terms of occupational status and wealth. The patterns observed imply that marital sorting must be strong in terms of the underlying genetics.

If this interpretation is correct then aspirations that by appropriate social design, rates of social mobility can be substantially increased will prove futile. We have to be resigned to living in a world where social outcomes are substantially determined at birth. Personally I would argue that this should push us towards compressing differences in income and wealth that are the product of such inherited characteristics. The Nordic model of the good society looks a lot more attractive than the Texan one.

Data Sources

Births, Deaths, Marriages

General Register Office. England and Wales Civil Registration Indexes. London, England: General Register Office.

FreeBMD. England & Wales, FreeBMD Death Index: 1837-1915 [database on-line]. Provo, UT, USA: Ancestry.com Operations Inc, 2006. Available at www.ancestry.com (last accessed: 01 Apr 2016).

FreeBMD. England & Wales, FreeBMD Birth Index, 1837-1915 [database on-line]. Provo, UT, USA: Ancestry.com Operations Inc, 2006. Available at www.ancestry.com (last accessed: 01 Apr 2016).

Ancestry.com. England & Wales, Birth Index: 1916-2005 [database on-line]. Provo, UT, USA: Ancestry.com Operations Inc, 2008. Available at www.ancestry.com (last accessed: 01 Apr 2016).

Ancestry.com. England & Wales, Death Index: 1916-2006 [database on-line]. Provo, UT, USA: Ancestry.com Operations Inc, 2007. Available at www.ancestry.com (last accessed: 01 Apr 2016).

Ancestry.com. Public Member Trees [database on-line]. Provo, UT, USA: www.ancestry.com Operations. Inc., 2006. Original data: Family trees submitted by Ancestry members.

Social Outcomes

Censuses, 1841-911 and Population Register 1939

England and Wales, Censuses, 1841-1901. Available online at http://www.nationalarchives.gov.uk/records/census-records.htm.

England and Wales, Population Register, 1939. London, UK. 1939 Original Data: 1939 Register, 1939. Kew, Surrey, England: The National Archives of the UK (TNA). Available at http://www.findmypast.com/1939register.

Wealth at Death

England and Wales, Index to Wills and Administrations, 1858-2012. Principal Probate registry, London (available online 1858-1966 at Ancestry.co.uk).

Gov.UK. Wills and Probate 1858-1996 [database on-line]. Original data: Principal Probate Registry. Calendar of the Grants of Probate and Letters of Administration made in the Probate Registries of the High Court of Justice in England. London, England © Crown copyright. Available at https://probatesearch.service.gov.uk/Calendar#calendar (last accessed: 01 Apr

2016).
General Register Office. 1861. Annual Report of the Registrar General 1859. (pp. 173-181) Available online at http://www.histpop.org (last accessed: 01 Apr 2016).
Prerogative Court of Canterbury and Related Probate Jurisdictions: Probate Act Books. Volumes: 1850-57. Held at the National Archives, Kew. (Catalogue Reference: PROB 8/243-250.)

Educational Status

Ancestry.com. UK, Articles of Clerkship, 1756-1874 [database on-line]. Provo, UT, USA: www.ancestry.com Operations, Inc., 2012. Original data: Court of King's Bench: Plea Side: Affidavits of Due Execution of Articles of Clerkship, Series I, II, III (KB 105-107). The National Archives, Kew, Richmond, Surrey. Registers of Articles of Clerkship and Affidavits of Due Execution (CP 71). The National Archives, Kew, Richmond, Surrey. Available at www.ancestry.com(last accessed: 01 Apr 2016).
Ancestry.com. Cambridge University Alumni, 1261-1900 [database on-line]. Provo, UT, USA: Ancestry.com Operations Inc, 1999. Original data: Venn, J. A., comp.. Alumni Cantabrigienses. London, England: Cambridge University Press, 1922-1954. Available at www.ancestry.com (last accessed: 01 Apr 2016).
Ancestry.com. Oxford University Alumni, 1500-1886 [database on-line]. Provo, UT, USA: Ancestry.com Operations Inc, 2007. Original data: Foster, Joseph. Alumni Oxonienses: The Members of the University of Oxford, 1715-1886 and Alumni Oxonienses: The Members of the University of Oxford, 1500-1714. Oxford: Parker and Co., 1888-1892. Available at www.ancestry.com (last accessed: 01 Apr 2016).
Ancestry.com. UK, Crockford's Clerical Directories [database on-line]. Provo, UT, USA: Ancestry.com Operations Inc, 2009. Original data: Crockford Clerical Directory. England: 1868-1932. See title page image of each directory for original source information. Available at www.ancestry.com (last accessed: 01 Apr 2016).
Ancestry.com. UK, Electrical Engineer Lists, 1871-1930 [database on-line]. Provo, UT, USA: Ancestry.com Operations, Inc., 2014. Original data: Society of Telegraph Engineers (later, Institution of Electrical Engineers) Membership Lists, 1887–1930. Institution of Engineering and Technology, Savoy Place, London, England. UK, Civil Engineer Lists, 1818-1930. Available at www.ancestry.com (last accessed: 01 Apr 2016).
Ancestry.com. UK, Mechanical Engineer Records, 1847-1930 [database on-line]. Provo, UT, USA: Ancestry.com Operations, Inc., 2013. Original data: Mechanical Engineering Records, 1847-1930. London, UK: Institution of Mechanical Engineers. Available at v (last accessed: 01 Apr 2016).
Ancestry.com. UK Medical Registers, 1859-1959 [database on-line]. Provo, UT, USA: Ancestry.com Operations Inc, 2008. Original data: General Medical Council, comp.

UK Medical Registers, 1859-1959. London: General Medical Council, 1859-1959. This data is provided in partnership with the General Medical Council. Available at www.ancestry.com (last accessed: 01 Apr 2016).

Royal Military Academy Sandhurst. The Sandhurst Collection [database on-line]. Available at http://archive.sandhurstcollection.co.uk/ (last accessed: 01 Apr 2016).

Senate House Library, University of London. 2016. University of London Students 1836-1934 [database on-line]. Available at http://www.senatehouselibrary.ac.uk/our-collections/special-collections/archives-manuscripts/university-of-london-students-1836-1934 (last accessed: 01 Apr 2016).

Bibliography

Abrevaya, J and H. Tang. 2011. "Body mass index in families: spousal correlation, endogeneity, and intergenerational transmission." *Empirical Economics, 41*: 841–864.

Becker, Gary S., and Tomes, Nigel. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *J.P.E.* 87 (1979): 1153–89.

Becker, Gary S., and Tomes, Nigel. "Human Capital and the Rise and Fall of Families." *J. LaborEcon.* 4, no.3, pt.2 (1986): S1–S39.

Clark, Gregory. 2017. "Estimating Social Mobility Rates from Surnames: Social Group versus Family Effects." Working Paper, March.

Clark, Gregory and Neil Cummins. 2017. "The People, not the Place. The Decline of the North of England 1918-2017: A Surname Investigation." Working Paper, February.

Clark, Gregory and Neil Cummins. 2016. "Family Matters? Do Relatives other than Parents Matter to Social Outcomes, England 1780-2016?" Working Paper, November.

Clark, Gregory and Neil Cummins. 2016. "The Child Quality-Quantity Tradeoff, England, 1780-1880: A Fundamental Component of the Economic Theory of Growth is Missing." Working Paper, August.

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

Clark, Gregory and Neil Cummins. 2014. "Surnames and Social Mobility: England, 1170-2012." 2014. *Human Nature, 25*(4), 517-537.

Clark, Gregory, Cummins, Neil et al. 2014 *The Son Also Rises: Surnames and the History of Social Mobility.* Princeton: Princeton University Press.

Clemons, Traci. 2000. "A Look at the Inheritance of Height Using Regression toward the Mean." *Human Biology*, *72*(3): 447-454.

Crow, James F and Joseph Felsenstein. 1968. "The effect of assortative mating on the genetic composition of a population." *Eugenics Quarterly*, 15:2, 85-97, DOI: 10.1080/19485565.1968.9987760

Falconer, D. S. 1981. *Introduction to Quantitative Genetics*. 2nd ed. London: Longman.

Fisher, R. A. 1918. "The Correlation between Relatives on the Supposition of Mendelian Inheritance." *Transactions of the Royal Society of Edinburgh, 52:* 399-433.

Galton, Francis, "Typical laws of heredity," *Proceedings of the Royal Institution of Great Britain*, 1877, 8, 282–301.

Galton, Francis. 1886. "Regression towards mediocrity in hereditary stature," *The Journal of the Anthropological Institute of Great Britain and Ireland*, *15*, 246–263.

Goldberger, Arthur, "Economic and mechanical models of intergenerational transmission," *American Economic Review*, 1989, 79 (3), 504–513.

Hill, William G., Michael E. Goddard, and Peter M. Visscher. 2008. "Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits." Plos Genetics, 4(2), e1000008.

Holt, Sarah B. 1961. "Quantitative Genetics of Fingerprint Patterns." *British Medical Bulletin*, 17: 247-50.

Mascie-Taylor, C. G. Nicholas. 1987. "Assortative mating in a contemporary British population." *Annals of Human Biology*, *14*:1: 59-68. DOI: 10.1080/03014468700008841

Mascie-Taylor, C. G. N. 1989. "Spouse Similarity for IQ and Personality and Convergence." *Behavior Genetics, 19(2)*: 223-7.

McManus, I. C. and C. G. N. Mascie-Taylor. 1984. "Human Assortative Mating for Height: Non-Linearity and Heteroscedasticity." *Human Biology, 56(4)*: 617-623.

Mulligan, Casey, "Galton versus the human capital approach to inheritance," *Journal of Political Economy*, 1999, 107 (S6), S184–S224.

Nagylaki, Thomas. 1978. "The correlation between relatives with assortative mating." *Annals of Human Genetics, 42*: 131.

Plomin, R 2011. "Commentary: Why are children in the same family so different? Non-shared environment three decades later." *International Journal of Epidemiology* 40: 582–592.

Plomin, R., Fulker, D.W., Corley, R., and DeFries, J. C. 1997. "Nature, nurture, and cognitive development from 1-16 years: A parent-offspring adoption study." *Psychological Science*, *8*: 442-447.

Robinson, Matthew R., et al., 2017 "Genetic evidence of assortative mating in humans," *Nature Human Behaviour*, 2016, 1. Article number: 0016 (2017).

Sacerdote, Bruce. 2007. "How Large Are the Effects from Changes in Family Environment? A Study of Korean American Adoptees." *Quarterly Journal of Economics, 122*(1): 119-157.

Silventoinen, Karri, Jaakko Kaprio, Eero Lahelma, Richard J. Viken, and Richard J. Rose. 2003. "Assortative Mating by Body Height and BMI: Finnish Twins and Their Spouses." *American Journal of Human Biology 15*:620–627.

Solon, Gary. "Intergenerational mobility in the labor market," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics* Volume 3A, 1999, pp. 1761–1800.

Tambs, Kristian, Torbjørn Moum, Lindon J. Eaves, Mike C. Neale, Kristian Midthjell, Per G. Lund-Larsen and Siri Næss. 1992. "Genetic and Environmental Contributions to the Variance of Body Height in a Sample of First and Second Degree Relatives." *American Journal of Physical Anthropology 88: 285-294.*

Watkins MP, Meredith W. 1981. "Spouse similarity in newlyweds with respect to specific cognitive abilities, socioeconomic status, and education." *Behavior Genetics, 11(1)*: 1-2.