To Have and Have Not: The Determinants of Social Status in England, 1600-2026

Draft 6/5/2026

Gregory Clark

Danish National Research Foundation Professor, Southern Denmark University

Economics Chair, Danish Institute for Advanced Study, SDU

Visiting Professor, Economic History Department, LSE

Preface

This book was intended to complete a trilogy of Hemingway punned books from Princeton University Press, with a title "For Whom the Bell Curve Tolls? Genetics and Social Life, England, 1600-2024." But it embraces what is regarded as not just an intellectual failure, but also a moral affront. This is that genetics, at least in a society like England 1600-2026, plays a large role in directly determining people's social outcomes, while social influences are surprisingly weak.

A recent Substack essay by Brad DeLong, titled, *Big Trouble in Little Human Genetic Difference-Worshipper Land,* illustrates nicely the hostility and contempt felt by the academy for anyone who would suggest an important role for direct genetic determination of social outcomes: "[A]ttempts to carve out a large, causal, manipulable "main effect" of genes on complex social outcomes keep reducing, on closer inspection, to approximately zero" (DeLong, 2026).

Thus an article in the *Proceedings of the National Academy of Sciences*, which contains the essence of the argument of the book, has been attacked by such prominent behavioral geneticists as Marcus Feldman and Kevin Lala as "scientific racism" (Lala and Feldman, 2024), and "racist pseudoscience" (Lala, Brown, Twyman and Feldman, 2025). Similarly a presentation of the core arguments of the book at *Glasgow University* in 2021 was indefinitely "postponed" by the Dean in response to a letter signed by 110 Glasgow University Faculty condemning the work as pseudo-science and eugenics.

A majority of the reviewers at Princeton fell in solidly with the received wisdom on the unimportance of genetics, without seriously engaging with the copious empirical evidence presented that cultural and social causation could not explain the patterns of inheritance observed in England. Chicago University Press considered the manuscript, but ultimately reported that they could not find reviewers willing to review it!

The databases that are the empirical center of this book were constructed jointly with Neil Cummins of LSE. Though my earlier book on surnames and social mobility (Clark et al., 2014) was funded in part by an NSF grant, it proved similarly impossible to get grant funding for a project to examine the role of genetics in social mobility. One NSF referee even stated that NSF grants should support only proposals which showed how to increase social mobility. Another noted "I failed to see any really 'nuts-and-bolts' discussion of how these findings might aid in enhancing social mobility or reducing inequality."

Consequently much of that data compilation was funded by generous grants from two private sources. First from Mary Katherine McComb, and second from Michael Dearing in support of Economic History at the University of California, Davis. Indeed without their support many of the chapters in the text would not have been possible.

In the last two years generous funding from the *Danish National Research Foundation* allowed for completion of the project, and also for the chapter comparing England, Denmark and Sweden. I am also very grateful to the members of the Guild of One-Name Studies who shared carefully researched genealogies, and to the members of the FreeReg organization, who transcribed and made available nearly two million marriage records.

The book summarizes work done jointly with Neil Cummins most extensively, but also with my former student Matthew Curtis, and current students Martin Kristensen and Christian Nielsen at the University of Southern Denmark. This does not imply they endorse any of the conclusions I have drawn from the data. Jonathan P. Beauchamp of George Mason University provided very useful critical comments on a draft of this manuscript.

Parts of the manuscript have been published in the following articles:

Clark, Gregory. 2023. "The Inheritance of Social Status; England, 1600-2022." *Proceedings of the National Academy of Sciences.* 120(27) e2300926120.

Clark, Gregory, Neil Cummins, Matthew Curtis. 2019."Twins Support the Absence of Parity-dependent Fertility Control in Pretransition Populations." *Demography*, 57(4): 1571-1595.

Clark, Gregory, Neil Cummins, Matthew Curtis. 2024. "Three new occupational status indices for England and Wales, 1800–1939." *Historical Methods*, 57(1): 41-66.

Clark, Gregory and Neil Cummins. 2025. Hypergamy reconsidered: marriage in England, 1837-2021. *Plos One,* 20(2), e0316769.

Clark, Gregory and Martin Kristensen. 2026. "The Myth of Nordic Mobility: Social Mobility Rates in Modern Denmark and Sweden." *Comparative Economic Studies, 68(1): 6*.

Clark, Gregory and Christian Nielsen. 2026. "The Returns to Education: A Meta-Study." *Kyklos*. https://doi.org/10.1111/kykl.70041.

Clark, Gregory and Neil Cummins. 2027. "How long do wealth shocks persist? Less than three generations in England, 1700-2026." Forthcoming, *Review of Income and Wealth*.

CONTENTS

Postscript

1. Introduction

To what extent are social outcomes such as occupations, income, wealth, longevity, and health socially versus genetically determined? And to what degree can life outcomes be changed by social interventions? Despite generations of research, and armies of social science researchers, these questions remain unresolved. But it is an orthodoxy of modern social science that while there are genetic influences on social outcomes, this influence is modest relative to social factors. Another orthodoxy is that genetics affects social outcomes nearly as much through influencing the child social environment as through direct effects on social outcomes.

This book explores these issues in depth, employing a large genealogical data base of 449,000 English people 1600-2026, as well as the detailed records of more than 1.7 million marriages in England 1837-2022. This rich new data suggests strong and surprising conclusions.

Part I: THE NATURE OF STATUS INHERITANCE

Social status is much more persistent across generations than conventionally measured. The underlying correlation in status across generations is close to 0.8, implying that even relatives as distant as nine generations apart still show correlation in social status.

Moreover, the strength of status persistence remains unchanged 1600-2026. The vast social changes of the Industrial Revolution, the Demographic Transition, emancipation of women, universal voting, and the growth of the modern welfare state, have not increased social mobility rates in any degree. The underlying structures of inheritance are highly stable in England. The vast resources poured into increasing rates of social mobility in modern society have achieved nothing.

The observed patterns of inheritance in England fit extraordinarily well a model where status is mainly determined directly by genetics, and where parents match strongly in the genetics associated with social status. This is not genetic determinism, since chance also plays an important role in social outcomes. But social influences matter little. The orthodox belief in the primacy of social factors and the effectiveness of social interventions is wildly mistaken. This result is laid out in detail in an article in the *Proceedings of the National Academy of Sciences* (Clark, 2023). Here I give a more intuitive interpretation of that result.

Denmark, Norway and Sweden are societies regarded as exemplars of the Nordic welfare state, and as having achieved much higher rates of social mobility than less interventionist economies. But if we correct for measurement errors, the underlying rates of social mobility in these Nordic societies, measured in terms of years of education attained, are not much less than those of modern England. Their social mobility rates are also little higher than in the early nineteenth century England of child labor, dark Satanic mills, and the workhouse. The same model of genetic transmission that so well describes English experience also fits well in the Nordic case. England is not an isolate when it comes to the rate and character of social mobility. Again the more technical aspects of this analysis are in an article in *Comparative Economic Studies* (Clark and Kristiansen, 2025). Here I give a more intuitive interpretation of these results.

If we were to create a society in which there was fundamental equality of opportunity, a true meritocracy, then we would find that genetic endowments, along with random accidents, would be the sole forces determining social outcomes. Parents and family would not matter, except as a signal to the genetic inheritance of individuals.

The finding both in modern and historical England, and also in modern Denmark and Sweden, that genetic transmission can explain observed social outcomes, is indeed evidence of the meritocratic nature of these societies. If governments remove significant social barriers to advancement of children of disadvantaged families, then genetic inheritance will be the only determinant of who has high social status will. England was a meritocracy even before the Industrial Revolution.

Part II: EVIDENCE FOR DIRECT GENETIC INHERITANCE

Parental genetics, however, can determine child outcomes in two very different ways. The first is through the direct effect on child abilities from their genetic inheritance. The second is through indirect effects, through parental nurture and investment linked to their genotype. Social causation. Behavioral geneticists have labelled this mechanism *genetic nurture*. If genetic nurture predominates in determining child outcomes, society can intervene to improve social outcomes, regardless of children's genetic endowment.

Such an indirect effect of parental genetics, however, would be detected in adoption studies. Adopted children would correlate strongly with adoptive parents. They do not, where adoptions are nonselective, and we look at adult outcomes. The indirect role of genetics would also be found in twin studies. Common environment effects would explain a substantial amount of the correlation of siblings. But such environment effects for siblings are not found once we recognize the strong genetic correlation of parents.

An inheritance mechanism with only direct genetic transmission also provides a set of clear predictions of outcomes that are tested in the second part of the book, using the two major datasets. If direct genetic transmission dominates, for example, then mothers and fathers should play equal roles in predicting child outcomes independent of the social status of women. Mothers should matter exactly as much as fathers in the very different social worlds of 1800 and of 2026. This is indeed what is found for all aspects of status, except for wealth where social mechanisms that favored sons predominated in earlier years.

Additional evidence for direct genetic transmission includes:

- In marriages formed before 1880 in England there was no attempt at fertility control. Family size was determined only by accidents of biology, and by the age at which women married. It was largely a random variable. Larger families reduce the amount of parental attention children can receive. They also reduce the material living standards of children. Yet family size had no significant effect on child outcomes.
- Birth order makes no difference to outcomes. Since your genetic inheritance is a random draw independent of birth order, being the first child should confer no advantage. Oldest sons in rich families in the nineteenth century did gain an advantage over younger sons. But for the other 98 percent of families there was a complete absence of birth order effects both in the nineteenth century, and now.
- Another accidental social outcome was where in England a child was born. For recent years, social outcomes on average are much better for those born in London and the South East than elsewhere in England. The chances of an 18 year old in London attending Oxford or Cambridge, for example, are three times those of an 18 year old in the North. Did accidents of birth have significant consequences on child outcomes, once we control for the social background of parents? Here we can show in a variety of ways that there were no disadvantaged locations in England in terms of social possibilities. The advantage of the South over the North in social outcomes came only from sorting of the English population where for the past 200 years those of greater ability moved to the South, and those of lesser ability moved North.
- Adult longevity is correlated with socio-economic status. The rich and educated live long, the poor and uneducated die younger. It is again widely believed that there is social causation of longevity, that work and home environments determine health. But I show that childhood environment has no effect on longevity. Years of education also has no effect. Attained status correlates with longevity, but that seems to be because of an inherent correlation of social status with physical robustness.

- Having the misfortune to lose a parent at a young child has little or no influence on child outcomes. The correlation of a son's adult status with that of their father is, for example, independent of the son ever meeting the father.
- Grandfather status helps predicts child outcomes, but does so in the same way whether or not a child ever met the grandfather.
- Accidental wealth gains get dissipated within three generations. In the long run even wealth is largely a function of genetic endowment.
- For continuous measures of social status, regression to the mean is at a constant rate all across the status distribution, as predicted by genetic inheritance. There are no wealth traps and no poverty traps.

All these findings reinforce the view that it is direct genetic factors that shape social outcomes, challenging the prevailing views that emphasize social influences and interventions.

Part III: SOCIAL IMPLICATIONS OF GENETIC INHERITANCE

The first implication of the importance of direct genetic transmission of social outcomes is that modern societies spend too much on education. In modern England such expenditures are ten percent of all output, nearly as much as expenditures on health care. What drives much of this expenditure is a belief that boosting years of education by less advantaged students increases social mobility rates. Yet despite a substantial narrowing of the gap in years of education between high and low status students 1851-2022, social mobility rates are unchanged.

Studies of the effect of increases in compulsory years of education for less able students in 1918-22, 1947, and 1972 find no effect on employment, income or health. There is every indication that corralling students in schools against their will imposes significant social costs and no benefits.

The second implication of direct genetic transmission is that social outcomes will change – true educational attainment, social abilities – only when the genetic endowment of the society changes. This gives societies, for example, an incentive to select high status immigrants, and reject low status ones. Societies that pursue this strategy – Australia, Canada, New Zealand – will gain permanent improvements in social outcomes.

This does not imply that societies should adopt such policies. There are plenty of reasons for societies to favor low status immigrants fleeing conflict or oppressive social regimes. But it does emphasize that if social causation is limited, the relative fiscal costs of different types of immigrants will endure across all future generations.

Another thing that will affect social outcomes is the relative fecundity of people at the top and the bottom of the social abilities. Here most people would rightly decry any attempt to restrict fertility by social class. But we now live in a world where there is fear that all societies are doomed for extinction from a wealth of material goods and a poverty of children. The popular obsession with social causation of outcomes makes parenting seem an exhausting and endlessly demanding task. A reminder that good parenting does not involve very much effort might be a suitable tonic to encourage fecundity.

PART I: THE NATURE OF STATUS INHERITANCE

Chapter 2 begins by discussing whether twin and adoption studies demonstrate the importance of social causation in status transmission. I show there twin studies have tended to systematically assume a zero genetic correlation between parents, and that this leads to a systematic tendency to overestimate social causation of outcomes. If we allow for what is observed as the high degree of assortment in status between parents then there is no implication from such studies of significant social causation. Adoption studies show adoptive parents have significant effects on early life outcomes, but not so for later life status.

Chapter 3 describes and evaluates the databases used in this study for England.

Chapter 4 shows that in England 1600-2026 we observe a stable pattern of inheritance of a variety of social characteristics – literacy, higher education, occupational status, dwelling value, local social environment, company directorships – that is completely consistent with additive genetic transmission of status with a high degree of genetic assortment by parents, even for relatives as distant as fourth cousins. The high degree of marital assortment means we observe very strong intergenerational persistence of status.

Chapter 5 shows that in recent years in Denmark and Sweden for years of education the same additive genetic model, with very similar parameters, explains well correlations between people very distant in relationship, including those in extended affine relationships. Despite the reputation of Nordic countries for high rates of social mobility, Denmark and Sweden are estimated to be as immobile even in recent decades as England now, or England before 1800.

Chapter 6 considers recent attempts to demonstrate the importance of *genetic nurture*, which is the action of genetic inheritance indirectly through the childhood environment. It argues that these attempts fail, and that if there was significant *genetic nurture* then twin and adoption studies would detect these effects.

2. Social versus Genetic Inheritance

I claim below from extensive lineage information for England that genetic inheritance played the crucial role in social outcomes. If this is obvious from the English data assembled here, why was this not immediately shown when the mechanisms of genetic inheritance were discovered in the early twentieth century?

Once the mechanics of genetics was understood, it became possible to estimate the effects of genetic inheritance on outcomes in a number of ways. The first is to compare the correlation in outcomes of monozygotic (identical) twins who share 100% of their genes with dizygotic (fraternal) twins who share as little as 50%.[1] If, for example, monozygotic twins

[1] With assortative mating by parents dizygotic twins can share significantly more than 50% of the relevant genes. In that case the contribution of genetics to the correlations will be greater, and the contribution of the shared family environment smaller. But the twins literature, for want of a means of estimating those underlying genetic correlations, has assumed it to be 0.

have a correlation on a trait such as occupational status of 0.6, and dizygotic twins 0.4, then genetics explains at least 0.2×2 = 0.4 of the 0.6 correlation of monozygotic twins, and their shared family environment at maximum 0.2 of the correlation.[2]

The fact that even monozygotic twins with the same genetics and family environment are still only correlated 0.6 suggests that chance is explaining 0.4 of the variance of outcomes. Seemingly twin studies should have long ago settled the nature-nurture debate.

When we apply this logic to the inheritance of height in modern societies, for example, we can see that in rich societies where nutrition is abundant for all, height is overwhelmingly determined by genetics. In one study, for example, the height correlation of monozygotic twins was 0.89, and that of dizygotic twins 0.51. That implies genetics explains at least 0.76 of the observed correlation of 0.89 for monozygotic twins, and shared environment at most only 0.13.[3] It also implies that for any child, genes make a contribution to determining height that is at least six times as large as the systematic, shared, elements of the home environment.

Since there is some correlation between parents in the genes determining heights, this share overestimates the role of the environment. As is shown in table 2.1 above, this genetic correlation in the genes determining height could be as high as 0.17 between spouses. If the parental genetic height correlation was just 0.15, for example, then as we will see below genes would cause all the correlation of monozygotic twins, and shared environment none of the correlation. Height in high income societies is overwhelmingly genetically determined.

---

[2] The share of outcome variance determined by genetics is estimated as $h^2 = 2(r_{MZ} - r_{DZ})$ and the share from shared environment $c^2 = r_{MZ} - h^2$. The remaining variance in outcomes $e^2 = 1 - h^2 - c^2$ is called "non-shared environment." If the parents had a correlation of .2 on the relevant genetics, then the dizygotic twins would share 20% of the relevant genes, so genes would now explain 0.5 of the correlation, and the shared environment 0.1.

[3] Silventoinen, et al., 2003, table 3, p. 403.

Table 2.1: Phenotype and Implied Genotype Correlations of Traits Across Partners

Trait	Phenotype Correlation R	h^2	Expected genotype correlation m
Height	0.20	0.85	0.17
BMI	0.14	0.68	0.10
Years of education	0.45	0.58	0.26
Income	0.08	0.34	0.03
Extraversion	0.05	0.46	0.02
Persistence	0.05	0.33	0.02
Novelty Seeking	0.06	0.38	0.02

Source: Zietsch et al., 2011, table 3. The heritability of traits is estimated taking account of assortment in marriage.

Note, however, that while height it now largely genetically determined, in societies where incomes are lower, height also depends significantly on access to food as a child. Thus in poorer societies there are significant social class differences in height. In such a setting the difference in the correlations of monozygotic twins in height compared to dizygotic twins would be less, and the environmental influence on heights greater. Thus how hereditable a trait is depends also on the range of variation in family environments. The smaller is this range of variation the more heritable will traits appear. Indeed paradoxically the aim of the good society, which through social mechanisms gives maximum opportunity to all, is to make genetic inheritance more important than family environment.

Another trait that we can measure in humans is the Total Ridge Count, which is the number of ridges, measured in a standardized way, counted across all 10 digits. For this trait the correlation between monozygotic twins is 0.95, and between dizygotic twins 0.49.[4] Here also the genetic correlation between parents can safely be assumed to be 0, since no one even knows this measure for their partner. Thus we know that the genetic difference between monozygotic and dizygotic twins is 50%. So heritability can be estimated precisely at 0.92. Here the genotype explains almost all variation in phenotype, and there can be no doubt of the importance of genetic inheritance, and the unimportance of environmental influences.

However, the same twin method applied to social outcomes yields much murkier conclusions on the importance of genetics. Consider years of education attained. A meta-study of 34 twin studies across 10 countries estimated the share of variance of outcomes from genetics at 0.39, and that from shared environment 0.36.[5] This makes family background almost as important as genetics in determining outcomes. Genetics is important, but a minority of the story of education determination.

But, as we shall see below, there is good evidence for considerable assortment by parents in terms of the genetics that help determine social outcomes such as education. If *m* is the parental correlation in the relevant genetics, then the heritability of educational attainment will be

$$h^2 = \frac{2(r_{MZ} - r_{DZ})}{(1 - m)}$$

Suppose, for example, that *m* was as high as 0.48. Then in the above meta-study on years of education the share of variance of outcomes from genetics would be 0.75, and that from shared environment 0.

---

[4] Holt, 1961.

[5] Branigan, McCallum, and Freese, 2013, table 5. The correlation for monozygotic twins averaged 0.747, and for dizygotic 0.551 (table 4). The other 0.25 share of variance is a combination of measurement error, and randomness caused by the imperfect measure of educational attainment.

Researchers reasoned they had good evidence that genetic assortment in mating was actually low. Table 1 above shows, for example, the phenotype assortment of spouses from a large survey of up to 8,315 partners of twins in Australia. As can be seen these phenotype correlations are generally modest: 0.2 or less for all except years of education. The implied genotype correlations for these traits will be even smaller if the match is being made based on the phenotype. Thus if people are matching on height and the match is a correlation of 0.2, then in terms of the genetics that helps determine height the correlation will be only h^2r. In this case h^2, the heritability of height, is 0.85, so the underlying genetic correlation will be only 0.17.

For years of education, where the phenotype correlation is 0.45, the genotype correlation is predicted as just 0.26. Consequently in the example above, the heritability of years of education would be only 0.54, and the share of the variance explained by common environment would be 0.20, and by random elements 0.26.

However the assumption that the genotype correlation for a trait such as educational attainment must we weaker than the phenotype correlation is suspect. In the case of height matching is likely driven purely by the phenotypic height, such as by social conventions where men and women prefer not to have a match with the man shorter than the woman. Then the phenotypic correlation will indeed exceed the genotypic. But in the case of education people match in marriage on the basis of a whole set of characteristics relating to education – general knowledge, linguistic complexity, sense of humor, organization – that actually match more closely with the genotype that predicts educational attainment than does years of education. We will show below evidence that matching on social characteristics in marriage correlates more strongly than matching on any particular phenotype measure.

Support for the idea that educational assortment was even stronger for the genotype than for the phenotype comes from a study that looked at couples in the UK Biobank. The correlation for years of education completed, the phenotype, was in the typical range, being 0.41 (standard error 0.01). However, the correlation for "trait-associated loci" for educational attainment, the genotype measure, was 0.65 (standard error 0.014) (Robinson et al., 2017, p.

4). Thus here the genotype correlation seems to greatly exceed that of the phenotype, and to be in the range where genotypes would explain most social outcomes.

In this book I will show that the evidence from England over the past four hundred years is consistent with an interpretation where there is strong marital assortment in the genetics that determines social outcomes. This assortment implies a much stronger role for genetic inheritance in explaining social outcomes than previously estimated.

A second issue that has cast doubt on the role of genetics in social outcomes has been the low correlation between monozygotic twins on some of these outcomes. If we take income, for example, then a typical correlation for monozygotic twins would be only 0.6 or lower.[6] This implies that *at maximum* genetics explains 60% of adult outcomes, and other environmental factors 40%.[7] Even in the case of years of education the average correlation of monozygotic twins who share genes and family environment is still only 0.74.

However, the environmental factors that matter to outcomes are not typically the observable shared environment twins experience – the home environment, the education of parents, their parenting styles, school and neighborhood characteristics. These as we saw above tend to be of minor importance. The environment that matters are subtle differences in conditions across pairs of twins which somehow have profound effects on outcomes.

A second method that can be used to assess the importance of genetics versus systematic family environments is to look at the correlation of unrelated adopted children within a family. If these children truly have no genetic correlation, then any correlation in their outcomes will reflect common environmental influences on siblings. Since adoption placements are generally not random, with children placed in suitable families based on their backgrounds, which will be correlated with their genetics, this test will bias upwards the observed correlation from shared childhood environments. But Plomin here reports again

[6] Bowles and Gintis, 2001, figure 4.

[7] Robert Plomin, a leader in the field of twin and adoption studies in behavioral genetics, argues that across many studies this gene-environment split of 50:50 is found on average (Plomin, 2018).

that for traits measured in adulthood – IQ, completed education, occupation – the correlation of unrelated adopted children is minimal.[8]

Instead the environmental factors that matter, and that typically explain about half the differences in outcomes across monozygotic twins are, as Plomin notes, "unsystematic, idiosyncratic, serendipitous events."[9] Attempts through interviews with identical twins with divergent outcomes in educational attainment, and with their parents, to uncover the nature of these idiosyncratic environmental influences failed. There was no systematic pattern. There is thus no hope of setting up programs to provide positive environmental elements that facilitate more favorable social outcomes. No-one knows what these elements are. To all practical purposes the environment here is just a random shock that helps determine the phenotype from the genotype.

While the focus on the twins literature has been on the subtleties of idiosyncratic non-shared home environments, differences in monozygotic twin phenotypes may mainly come from subtle accumulated differences at the cellular level. The DNA in cells in the body is chemically altered by the attachment of molecules that inhibit the action of segments of the DNA in that cell. This process, methylation, is annulled when DNA is transferred to an embryo, so that each embryo starts with the DNA blueprint as a *tabula rasa*. However, even for MZ twins there are differences in the rates and patterns of methylation. The correlation in rates of methylation of the DNA is only in the order of 0.93-0.98. Thus there is the possibility that random processes at the cell level are changing gene expression even in MZ twins.[10] This would imply that social outcomes are typically 50% determined by genetics, less than 10% by family and social environments, and more than 40% by molecular level randomness.

---

[8] Munsinger, 1975, similarly concluded that, even for IQ measured in childhood, where adoption involved non-selective placements, childrens' IQ correlation with parents was three times greater for biological children than for adoptees.

[9] Plomin, 2018, p. 80.

[10] Since DZ twins have nearly as high a correlation of methylation rates as MZ twins, the estimates of the heritability of methylation, the percentage of the variation explained by additive genetic variation, is only in the order of 5-15%.

These random effects mean that monozygotic twins are not a good control for ability in estimating the effect of schooling on earnings and other outcomes. For example, the literature that estimates the causal effects of years of education on adult earnings, mostly written in the 1990s, wrongly assumed that monozygotic twins were perfect controls for innate talent.[11] This implied that such twins who received different numbers of years of education could be used to estimate the causal effects of education, absent from sorting on ability. These estimates show robust earnings gains from education, and thus the possibilities of remaking outcomes across the population through social interventions. But this literature had to assume that then the variation in years of education across twins had to be random! But since even identical twins are different in talents and abilities this literature shows nothing about the causal effects of education.

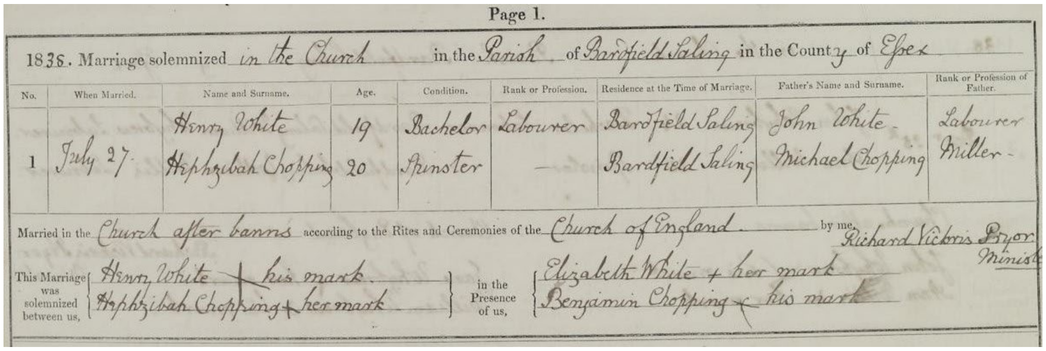
However, despite such evidence the overwhelming sentiment across all the social sciences – economics and sociology in particular – is that social outcomes are largely socially determined. In economics the focus has been on the role of human capital – the investment in education and in "soft skills" – that parents and society make in children. Since economics can seemingly demonstrate that an additional year of education is causally associated with a significant gain in earnings, then by implication economic outcomes for individuals must be changeable just by more investment.[12] In sociology there is consideration of wider socialization processes connected to social class and social networks. But, as with economics, with the implications that social institutions can significantly change social outcomes.

This book thus offers additional powerful evidence against the dominant social science narrative where social outcomes are the elastic results of social forces, in favor of one where societies have much more limited ability to influence social outcomes, and need to construct social policies with that constraint in mind.

---

[11] Card, 1999, summarizes this literature.

[12] I consider below, however, how convincing the demonstrations of the causal role of education are.



Page 1.

1838. Marriage solemnized in the Church in the Parish of Bardfield Saling in the County of Essex

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.
1	July 27	Henry White	19	Bachelor	Labourer	Bardfield Saling	John White	Labourer
		Hephzibah Chopping	20	Spinster	—	Bardfield Saling	Michael Chopping	Miller

Married in the Church after banns according to the Rites and Ceremonies of the Church of England by me, Richard Victoris Pryor Minister

This Marriage was solemnized between us, Henry White + his mark, Hephzibah Chopping + her mark

in the Presence of us, Elizabeth White + her mark, Benjamin Chopping + his mark

2. Families of England Genealogy

The empirical basis of this book consists of two major data collections. The first of these is a genealogy of a set of English families, 1600-2026, who carried one of 499 rare surnames, where we track every holder of the surname, no matter what their location. This database has 436,185 individuals.

These surnames fall in two groups. The first is a set of 233 rare surnames, 61,190 people, which had significantly higher wealth than average in the mid nineteenth century that was used to measure the persistence of wealth in England 1858-2026. This set of surnames includes some well-known in English history including Courtauld, Pepys, Cornwallis, Lane-Fox, Sebag-Montefiore, de Havilland, Bazalgette, Champion de Crespigny, Rothschild, Baring, Pigou, Mitford, and Rusbridger.

The second, larger, group is a set of 268 rarer surnames, 374,995 people, that had close to average social status. This group is representative of the general population of England.

Using census records 1841-1921, the population survey of 1939, marriage records 1837-2022, probate records 1858-2026, voting rolls 1999-2026, and other ancillary sources we can measure for many of these persons their educational attainment, their occupational status, their dwelling value, their wealth at death, their reproductive success, and their adult longevity.

The second database is a set of complete records on church marriages 1837-2022, which contains nearly 1.7 million marriages. Throughout these years these marriage records record husband and wife literacy, as well as the occupations of husbands, wives, fathers, and fathers—in-law.

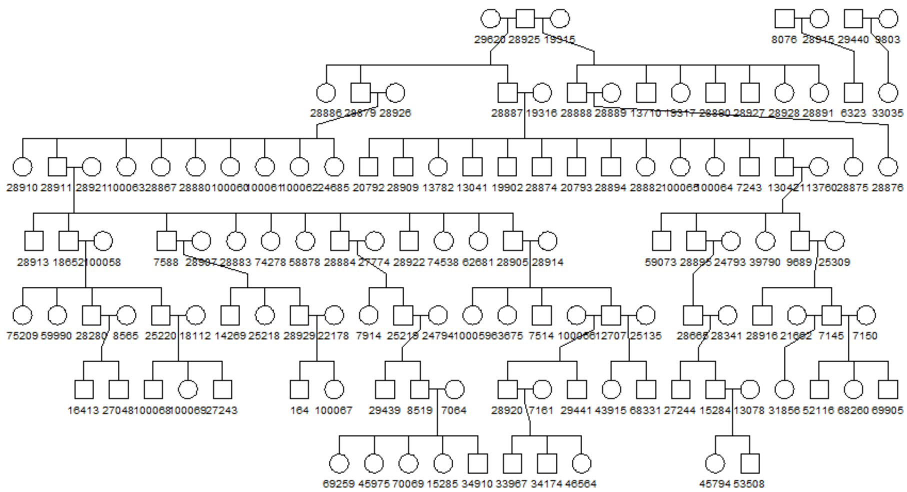
Families of England (FOE) Database

In constructing the FOE database we directly compiled 424 of the 499 surname lineages (with 118,735 persons). The other 75 lineages (317,450 people) were obtained mainly from an appeal to members of the *Guild of One-Name Studies*, an organization devoted to tracing the history of particular rare surnames. These lineages incorporate everyone with a rare surname of interest, wherever they reside, as well as spelling variants of the surname. The Mitchelmore lineage, for example, incorporates the surnames Michelmore, Mitchelmore, Mitchamore, Mitchmore, Mouchemore, Muchamore, and Muchmore.[20] The Auty lineage encompasses Auty, Autey, Awty, Otty, and Ottey.[21] In cases where we only had access to published lineages, these did not typically contain details of living holders of the surname. Here we added that information ourselves from public records of births, marriages and addresses. Lineages were chosen for inclusion based on their completeness, and either the public posting of the lineages, or their creators' willingness to share the data with us for inclusion in the study.

The decision to incorporate families with rarer surnames into the genealogy was in part adventitious. For an earlier study of social mobility using surnames to link generations we had assembled data on elite and underclass wealth in England by rare surname 1858-2012. But focusing on rare surname individuals allows a high degree of linkage across generations. In the FOE database for 84% of adults born with a rare surname we can identify the father and mother. For those born 1800-1939 the percentage linked is even higher at 92%.

Figure 3.1 shows an illustrative fragment of the FOE database, showing linkages across 7 generations. Table 3.1 shows the outline of the source of the data, and its distribution across time, and between general and elite lineages. Table 3.2 shows the numbers of relationship pairs in the data, again by lineage type. The reason for the extraordinarily large numbers of pairs of 2nd-4th cousins in table 3.2 can be seen in figure 3.1, which shows an illustrative fragment of the genealogy database. Average completed family size in England in the nineteenth century was around 3 adult children, but this varied enormously across families, and the bulk of adults in each generation came from larger than average families, so that average sibship size then was 6. Such demographic processes ensured large numbers of cousins, 2nd cousins, etc. for adults in each subsequent generation.

Figure 3.1: Sample of the FOE database



Notes: A sample section of the FOE database, showing linkages across 7 generations. The squares denote men, the circles women.

Table 3.1: Families of England (FOE) Data Outline

Birth Period	All	General Lineages	Elite Lineages
1600-99	5,870	5,477	393
1700-99	28,587	23,457	5,130
1800-49	65,586	55,044	10,542
1850-99	115,988	100,513	15,475
1900-49	79,245	70,056	9,189
1950-2026	48,166	41,765	6,401
All	436,185	374,995	61,190

Source: FOE database.

Notes: Significant numbers of people have no birth date observed.

Table 3.2: Families of England, Numbers of Relationship Pairs

Relationship Pairs	All	General Lineages	Elite Lineages
Spouse	135,751	115,085	20,666
Child-Parent	512,983	445,565	67,418
Sibling (full)	525,973	455,765	70,208
Nephew/Niece	1,079,630	945,026	134,604
Grandchild	520,624	450,914	69,710
Cousin	850,359	768,176	82,183
2nd cousin	1,229,410	1,135,913	93,497
3rd cousin	1,505,612	1,436,407	69,205
4th cousin	1,530,689	1,484,961	45,728

Source: FOE database

Ten Social Outcomes

Using sources such as the censuses 1841-1921, the population register of 1939, marriage records 1754-2021, the probate registers 1858-2024, the electoral registers of 1999-2024, the Companies House registers of company directors 2000-2026, and the death registers of 1837-2026 we construct ten social outcome measures for the FOE members. These ten are listed below.

1. **Literacy at Marriage**. For marriages 1754–1889 we infer literacy from whether the bride or groom signed the marriage register. These signature records extend more recently than 1889, but by then signature rates were high, making the status information content of the measure low. Only a subset of county record offices in England have made images of marriage registers available on-line. So literacy is available only for a subset of men and women marrying in these years.

2. **Adult Longevity.** This we calculate for those born 1700-1929. Adult longevity, even in the nineteenth century, was correlated with social status. Since longevity was trending over time we also construct a standardized longevity measures, by gender, for each birth decade which is years lived minus average years lived divided by the standard deviation of years lived. This measure for each gender in each decade has a mean of 0 and standard deviation of 1.

3. **At work/school aged 10-18**. The censuses of 1851-1921 and the 1939 population register record whether children are at school, or in some type of work. Thus controlling for child age and the year of observation we have an indicator of when a person left education and entered the working population.

4. **Higher education**. This is an indicator variable with a value 1 if the person achieved a higher educational status. Since before 1920 most higher education was closed to women we construct this measure only for men born 1600-1919. Complete records are available for attendees Oxford and Cambridge Universities (1600-2018), Durham University (1837-1939), University of London (1837-1939), the Royal Military Academy Woolwich (1790-1839) 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.

5. **Occupational Status**. Occupations are recorded in the censuses of 1841-1921, the population register of 1939, marriage records 1837-2022, probate records, 1858-1939, ship passenger lists, 1870-1959, army enlistment records 1914-1918, and in the professional directories listed above.[23] Because most married women are listed without occupations, and even large numbers of unmarried women have no listed occupation, we again construct this indicator only for men.

Where a person had multiple statements of occupation at different times ages, the one closest to age 40 was employed. For those born 1780-1919 the very large set of occupation description strings were first assigned to one of 442 categories. Using a large set of 1.6 million marriage records 1837-1939 which give occupations for grooms, their fathers, and the father-in-law we derive an occupational status score 0–100 using Goodman's Association methodology. The occupational status index derived in this way shows strong father-son correlations 1780-1919.[21]

6. **House Value**. We estimate this from the addresses recorded for people alive 1999 and later in the electoral roll. Since the Roll released 2002 and later records only those who consented to the release of their address, there are potential issues of selectivity. However, we see in table 3.4 above that the average house value recorded using the electoral register addresses closely approximates to that of England as a whole. There is no sign that higher status individuals are less likely to permit publication of their addresses in the Electoral Roll.

The Land registry shows house prices for all property sales 1995 and later. From this we construct an average dwelling value, normed to 2017 prices, for each Post Code. British Post codes on average cover only 40 houses. So this gives a good estimate of local house values for the person. Where there was no sale recorded for a postcode, we use the Council Tax Band to estimate the property value. Empirically the log of average house values produces higher correlations between relatives, so we use this measure.

We employ this measure only for men and women aged 24 or above, and living at a different address than their parents. Property values differ substantially by region in England and Wales, as table 3.4 shows. London house values, for example, are more than 4 times those in the North of England. Since people show strong persistence by region across generations, we normalize all house values to remove regional effects, dividing England and Wales into 6 regions for this purpose.

7. **Index of Multiple Deprivation**. This index is a ranking of Lower Layer Super Output Areas (LSOAs), typically with a population of around 1500, in terms of a weighted average of measures of social deprivation. The index is available by post code. The 2019 index, used here, is a weighted average, with weights indicated, of measures of: Income (22.5%), Employment (22.5%), Education (13.5%), Health (13.5%), Crime (9.3%), Barriers to Housing and Services (9.3%), Living Environment (9.3%). Since the index for Wales is constructed in a non-comparable way, we fix the level for that as the average for the English IMD for the North of England.

8. **Company Director**. Companies House in the UK maintains a register of the directors of limited companies. Limited companies include commercial enterprises, but also management companies for housing associations, as well as medical and legal practices. The register also includes people who subsequently resigned the position. We classify anyone alive 2002 and later, and aged 24 and above in 2024, as either a company director or not.

9. **Probated at Death.** Once a national probate system was introduced in 1858, we get a complete record of who was probated at death in England and Wales 1858-2024. Only those above a minimum level of assets were required to undergo probate, so being probated is an indicator of having some wealth at death. The fraction probated was only 17 percent in the 1860s, but increased over time to 46 percent by the 1950s, and has stayed around this level since.

10. **Wealth at Death.** The Principal Probate Registry, 1858-2024, records for each testator the value of the estate. For those not probated, whose estate value fell below the value requiring probate, we attribute a value which is half the minimum value for probate. Thus for 1950 when the minimum value for probate was £500, we assume each person not probated had an estate of value £250. To allow for changing price levels and average wealth, we normalize each estate value by dividing by the average estate value in that decade. Thus the average normalized wealth at death should be around 1 in all periods.

Table 3.3: Correlation of Status Measures, modern period

	Probated	Log Wealth	Occ. Status	Log House Value	Index of Multiple Deprivation	Company Director
Probated	-					
Log Wealth	0.87	-				
Occupational Status	0.37	0.44	-			
Log House Value	0.34	0.44	0.48	-		
Index of Multiple Deprivation	0.20	0.25	0.30	0.50	-	
Company Director	0.23	0.33	0.39	0.30	0.05	-

Table 3.4: Families of England, Social Outcomes

Outcome	Gender	All	General Lineages	Elite Lineages
Literacy	Both	26,164	24,254	1,810
Higher Education	Male	68,732	51,582	17,150
Occupational Status	Male	63,973	51,748	12,225
At Work/School 10-18	Both	36,821	31,577	5,244
Probated	Both	89,759	65,590	24,169
Wealth at Death	Both	84,691	62,067	22,624
House Value	Both	33,577	27,168	6,409
Index of Multiple Deprivation	Both	32,128	25,789	6,339
Company Director	Both	38,927	31,702	7,225
Death age (21+)	Both	142,113	112,630	29,483

Source: FOE database.

The different measures of social outcomes all tend to show highly statistically significant correlations. Thus for people dying 1999 and later, table 3.3 shows the correlation in the four measures of log house value, occupational status, index of multiple deprivation, and company director.

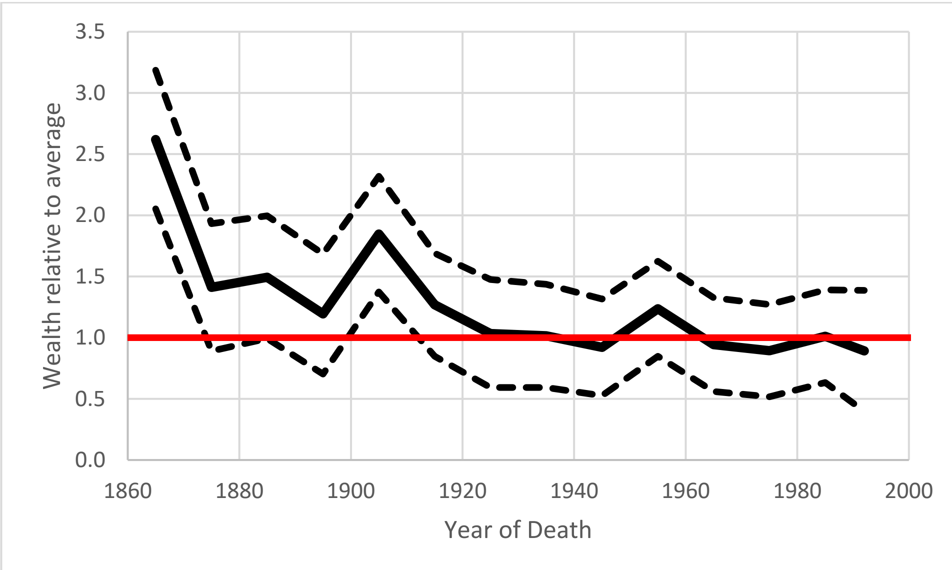
Table 3.4 shows the social outcomes that are available by gender and lineage type. The numbers of any social outcome are much less than the numbers of people in the database because: (1) some outcomes are available for men only, (2) before 1914 a significant number of children die before reaching age 21, (3) for births before 1780 and after 1929 many social outcomes are not observable.

But, as table 3.4 shows, when we consider the numerous pairs of relatives in table 3.2, there are many pairs observed with the same social outcome, that can be used to estimate the underlying nature of social mobility. Thus for male occupational status, for men born 1860-1919 the database contains 241,941 pairs of observations.

How representative are the lineages in the FOE database of the general population in England and Wales? One test is average wealth at death in the general lineage 1858-1996 compared to average wealth of all deaths in England and Wales in these same years. Figure 3.2 shows this ratio by decade 1860-1990.

As can be seen for the death decades 1920 and later, and thus the birth decades 1860 and later, the Families of England average lineages seem representative of the general population in terms of wealth, and thus also in terms of other aspects of social status. For deaths before 1920, and thus births before 1860, average wealth in the general lineages is typically 50% higher than for the general population.

Figure 3.2: Wealth at Death, Average FOE Lineages relative to England



Notes: Average wealth at death by decade of members of the average FOE lineages relative to England and Wales as a whole. Dashed lines show 95% confidence intervals.

The most likely explanation for this is that the processes that generated the rarer surnames used in these lineages were associated with somewhat higher status families in earlier centuries, but that over time slow but steady social mobility has brought these a possibility that lower status holders of the lineage surnames are less likely to appear in the records with a surname recognizable as belonging to these lineages.

Thus for the estimates of familial correlations using the average lineages for births 1860-1919 the sample used is then representative of the general population of England and Wales in its social outcomes. This conclusion is buttressed when we compare the average occupational status of men in the general lineages in the FOE database with the general population. For 16,639 men born in the FOE database who married 1880-1939, the average occupational status for "average" lineages was 37.2 on a scale 0-100, for those marrying outside the major city London. For grooms in a random sample of 591,000 Church marriages

1880-1939, transcribed by volunteers from the FreeReg organization, but drawn almost entirely from parishes outside London, the average occupation status on the same ranking was 35.5. (We discuss this second data source below). Thus again for births after 1860 the FOE sample looks very similar in status to the general population.

For men born 1780-1859 we can compare those in the FOE general lineages with the Freereg sample in terms both of occupational status, and of literacy. Men in the FOE general lineages marrying outside London 1837-79 had an average occupational score of 36.9. In contrast 771,000 grooms in the Freereg sample in these same years had an occupational score of only 31.1. Thus in line with the average wealth evidence, men in the "average" lineages scored nearly 20% higher than the average man in this period. The marriage records, also record if men and women can sign their names at marriage. In the Freereg sample of 484,888 marriages 1837-79 with records on whether brides and grooms signed, 68% of men and 58% of women were literate. For men and women in the FOE general lineages marrying outside London 1837-79, 73% of men and 61% of women were literate. Here the members of the FOE average lineages show only a very modest elevation in social status compared to the average person in England.

Will the modestly elite status of the FOE average lineages for births 1780-1859 bias the estimates of social mobility rates in this period? The answer is that if we draw a sample from the population where the variance of outcomes is different than for the population as a whole we would potentially get a biased estimate of the heritability of traits, h^2. If the variance of the sample outcomes is higher, we will also get a more precise estimate of persistence. But whatever sample of the population we start from, estimates of the level of persistence should be unaffected.

For the recent period 1999-2024, where we observe house values, the Index of Multiple Deprivation, and Company Directorships table 3.5 shows the distribution of the observations used here, in terms of geography and house values versus for the general population. The FOE dataset for individuals observed 1999-2024 has a geographic distribution that largely echoes that of the general population distribution, except for being less frequent in London.

Table 3.5: FOE Observations 1999-2024 versus General Population

Region	National Population Share 2020[26]	FOE Address Data	FOE Address Share	National Average Dwelling Value 2017[27]	Average FOE Dwelling Value
London	0.151	1,914	0.078	£481,556	£576,957
East and South East	0.257	7,872	0.321	£306,534	£307,754
South West	0.096	2,883	0.118	£246,519	£254,057
Midlands	0.181	4,722	0.193	£183,773	£171,814
North	0.262	6,413	0.262	£152,318	£163,603
England	0.946	23,804	0.972	£240,325	£254,118
Wales	0.054	688	0.028	£151,672	£149,206

Source: FOE database.

But the FOE dataset is composed, by design, of long-established English family lineages. London is the area of England with the largest proportion of the population of more recent immigrants. So it is expected that the frequency of the FOE families will be lower in London than for the general population.

The estimated house values observed in the FOE database, adjusted to 2017 prices, again are close to the average across region observed nationally in sales in 2017. The only location with a substantial difference is London, where the FOE house values are higher. But as noted London is the city with the largest share of population of recent immigration to England. Thus there is no reason to expect the FOE house values here to be similar to those of the London population as a whole. Overall house values in the FOE database are 6.6% higher than for England and Wales as a whole in 2017. This is a modest difference.

Marriages of England (MOE) Database

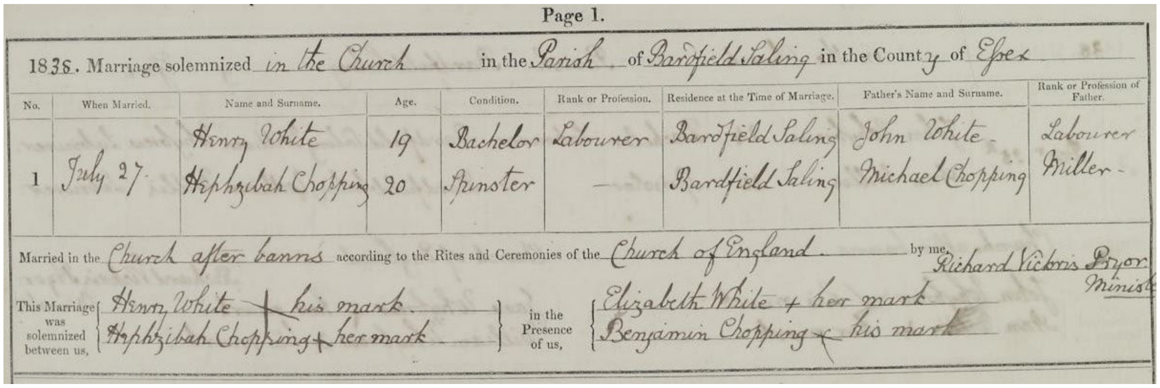
The second database underlying this book is Marriages of England (MOE), a set of nearly 1.7 million marriage certificates for church marriages in England and Wales 1837-2022. The marriage certificate introduced in England and Wales in 1837 has continued since then unchanged down to the present day. Thus figure 3.3 shows a certificate from 1838 compared to one from 2020. The marriage certificate 1837 and later gives a wealth of information on the social status of the parties marrying. First it reveals the literacy of the groom and bride who are required to sign the certificate. For the years before 1890 when there are significant numbers of brides and grooms who are unable to sign the marriage register, and instead record just an "x", this is an excellent measure of educational status.

Marriage certificates also record, however, the "rank or profession" of the groom, bride, father and father-in-law. Thus we shall see that they are informative also about the degree of assortment in marriage, and about the strength of intergenerational persistence in social status.

The UK government now has such records of around 106 million marriages 1837-2024 from England and Wales, with the associated details. However, it costs by statute £11 to obtain a copy of any marriage certificate from the government, and the copy is delivered by mail as a paper reproduction of the government's copy of the marriage certificate. Since copies of the marriage certificates were kept in church registers, and many of these registers have since been deposited in local record offices, these provide an alternative source for marriage records.

The marriage certificates available in local record offices exclude Civil Marriages. But though Civil marriage was introduced in England in 1837, such marriages remained a minority of all weddings before 1914. In 1841 Civil marriages were 1.7% of all marriages, and in 1914, still only 24%. Thereafter there were increasing numbers of civil weddings, as church attendance declined, but also as divorce rates increased. Until recently divorcees were rarely granted permission to be remarried in the Church of England. So 31% of weddings were civil by 1952, 49% by 1982 and 68% by 2012. However for first marriages by 1995 only 40% were civil, and for 2012 63% civil (Haskey, 2015).

Figure 3.3 Sample Marriage Certificates, 1838 and 2020



Page 1.

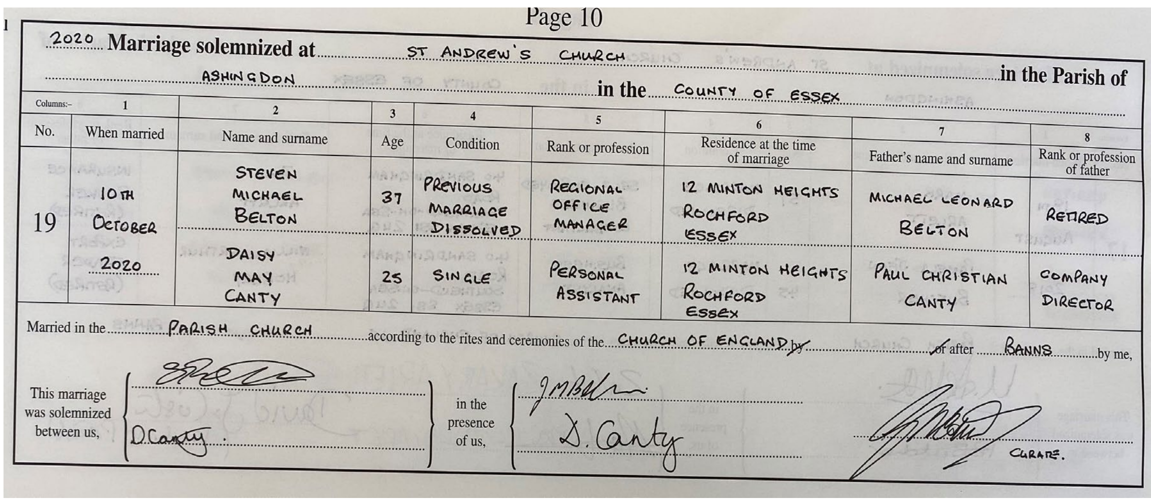
1838. Marriage solemnized in the Church in the Parish of Bardfield Saling in the County of Essex

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.
1	July 27.	Henry White	19	Bachelor	Labourer	Bardfield Saling	John White	Labourer
		Hephzibah Chopping	20	Spinster	—	Bardfield Saling	Michael Chopping	Miller

Married in the Church after banns according to the Rites and Ceremonies of the Church of England by me, Richard Vickris Pryor Minister

This Marriage was solemnized between us, Henry White + his mark; Hephzibah Chopping + her mark

in the Presence of us, Elizabeth White + her mark; Benjamin Chopping + his mark



Page 10

2020 Marriage solemnized at ST ANDREW'S CHURCH ASHINGDON in the Parish of ASHINGDON in the COUNTY OF ESSEX

No.	When married (1)	Name and surname (2)	Age (3)	Condition (4)	Rank or profession (5)	Residence at the time of marriage (6)	Father's name and surname (7)	Rank or profession of father (8)
19	10TH OCTOBER 2020	STEVEN MICHAEL BELTON	37	PREVIOUS MARRIAGE DISSOLVED	REGIONAL OFFICE MANAGER	12 MINTON HEIGHTS ROCHFORD ESSEX	MICHAEL LEONARD BELTON	RETIRED
		DAISY MAY CANTY	25	SINGLE	PERSONAL ASSISTANT	12 MINTON HEIGHTS ROCHFORD ESSEX	PAUL CHRISTIAN CANTY	COMPANY DIRECTOR

Married in the PARISH CHURCH according to the rites and ceremonies of the CHURCH OF ENGLAND by ~~or after~~ BANNS by me,

This marriage was solemnized between us, [signature]; D Canty

in the presence of us, J M Belton; D. Canty

[signature] CURATE.

Table 3.6: Church Register Marriage Data, 1837-2021

Marriage Period	All	Groom Occupation	Bride Occupation	Father Occupation	Father-in-law Occupation
1837-59	540,614	450,905	70,032	413,638	411,789
1860-79	365,463	310,321	42,146	294,935	295,259
1880-99	336,125	285,405	42,870	253,004	273,058
1900-39	343,344	283,040	63,397	242,408	273,831
1940-79	66,636	61,454	39,380	52,986	54,405
1980-2021	15,788	15,601	14,683	14,636	14,640
All	1,667,970	1,401,806	198,446	1,266,052	1,317,687

Source: FreeReg, https://www.freereg.org.uk/. Supplemented 1980-2021 with transcriptions of Essex church registers.

Table 3.7: Church Register Literacy and Age Data, 1837-2021

Marriage Period	All	Groom Literacy	Bride Literacy	Groom Age	Bride Age
1837-59	540,614	289,453	289,459	153,199	162,440
1860-79	365,463	195,428	195,478	230,475	235,971
1880-99	336,125	*	*	291,032	291,778
1900-39	343,344	*	*	322,135	321,606
1940-79	66,636	*	*	64,183	64,723
1980-2021	15,788	*	*	15,744	15,719
All	1,667,970	484,881	484,937	1,076,768	1,092,237

Source: FreeReg, https://www.freereg.org.uk/. Supplemented 1980-2021 with transcriptions of Essex church registers.

Table 3.6 shows the numbers of observations in the MOE database by marriage period. These data were mainly obtained from the website Freereg.org, which posts transcriptions by volunteer genealogists of marriage registers held in county archives. However, the interests of the genealogists tend to be around marriages in earlier years, as table 3.6 illustrates. The table also shows the amount of information available on the occupations of grooms, brides, fathers and fathers-in-law. To supplement the records for the years 1980-2023 11,030 marriage records from various churches in Essex Record Office were transcribed. Essex as a county on the edge of London has a mix of both urban and rural parishes.

Table 3.7 shows for the years 1837-79 the amount of data available on literacy at marriage by bride and groom, and also on bridge and groom ages. While almost all the contributors of marriage transcripts to FreeReg transcribed elements such as occupation, only a minority recorded literacy. However, there is little difference in the rate at which literacy is recorded for the marital partners for brides with high or low occupational status fathers. At the high end of occupational status literacy is reported for 53% of marriages, while at the low end of occupational status it is recorded for 58% of marriages. Thus the literacy evidence will provide an unbiased estimate of the degree of marital assortment. Age in earlier years was often recorded just as either "a minor" or "full age" meaning less than 21 or 21 or older. This was because marriages involving a minor required the consent of their father.

The coverage of the FreeReg data depends on local volunteers, and so is not uniform across England and Wales. Four counties are heavily overrepresented: Lancashire, Kent, Lincolnshire, and Staffordshire. But as figure 3.4 shows there is in general a wide coverage of all geographic areas of England and Wales, with records drawn from 5,184 parishes out of a total of 15,000 across England and Wales.

Figure 3.4: Geographic Coverage of FreeReg Marriage Records

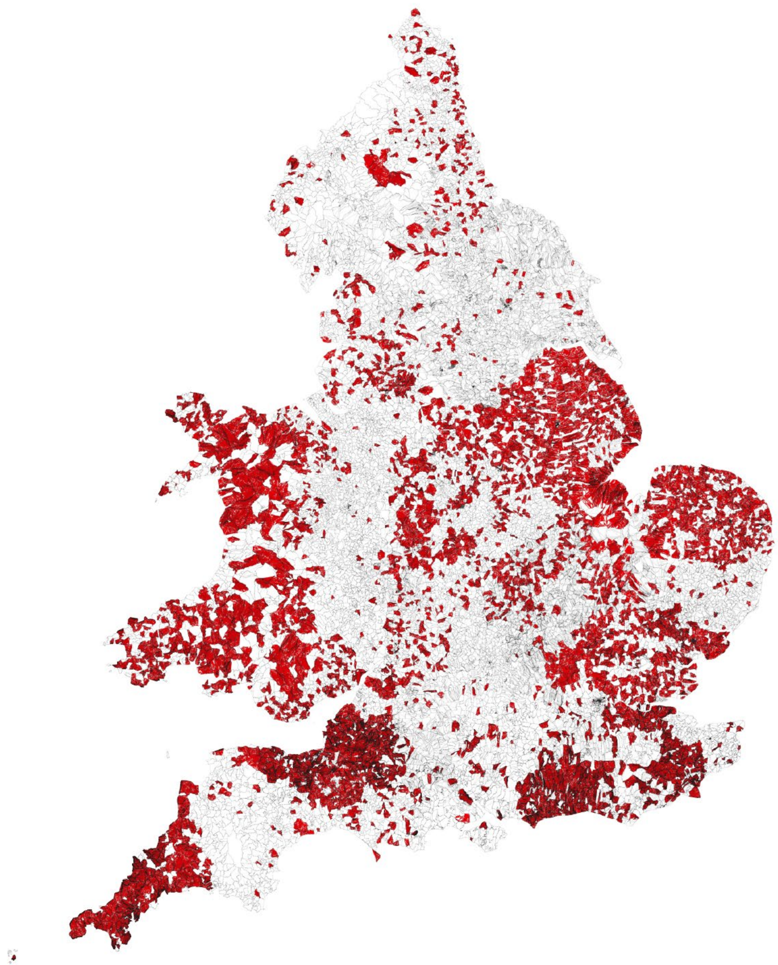
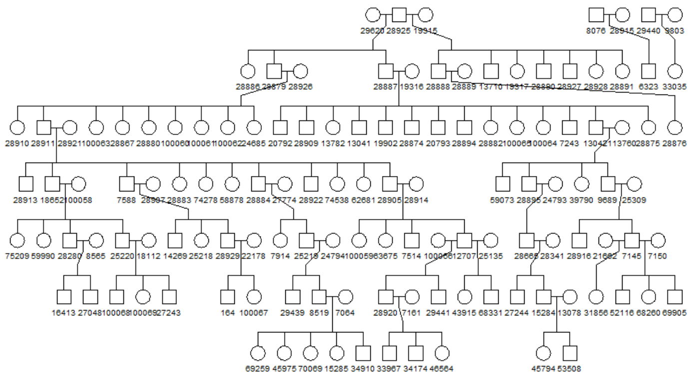


Figure 7: Map of the 5,184 Sample Parishes



4. The Inheritance of Social Status: England 1600-2026[13]

This chapter establishes four important elements about the determination of social status in England.

First, there is much greater persistence of social status across generations than is conventionally estimated. Inheritance, indeed, explains two thirds of the variance in social status.

Second, the very significant social changes in England between 1600 and 2026 have done nothing to change the persistence of status across generations. The enormous modern social investments in education and childcare have failed to increase social mobility rates compared to the laissez-faire world of eighteenth and nineteenth century England where such provision was minimal.

Third, one cause of this strong persistence is very strong status assortment in marriage. Also the strength of this assortment has also been unchanged from 1754-2026, despite very significant changes in the educational and occupational status of women.

Fourth, and finally, the patterns of status inheritance in England 1600-2026 are completely compatible with genetics being the main determinant of social status.

---

[13] This chapter summarizes in a more intuitive way the argument presented in an article in the *Proceedings of the National Academy of Sciences* (Clark, 2023).

Status Persistence and Measurement Error

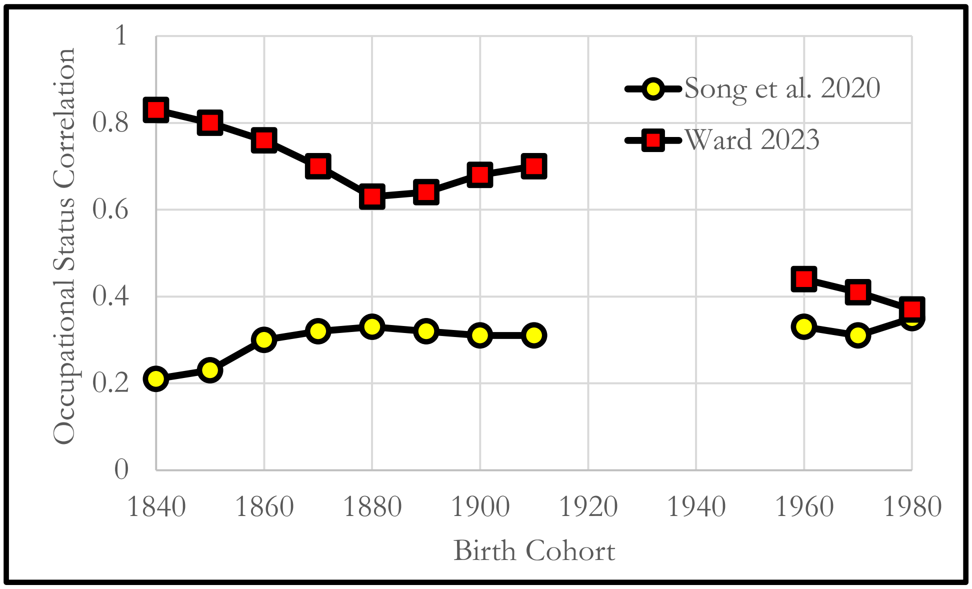
In measuring social mobility, measurement errors in the crude status indicators employed, such as years of education, play a crucial role. These errors always lead to an underestimate of the degree of persistence of status. But such errors vary across time, place, and status indicator.

The most widely used estimate of educational attainment, for example, is just years of education completed. While this measure is undoubtedly correlated with educational attainment, it measures weakly the actual educational status of persons. With this measure a first-class degree in Physics from Cambridge University is counted as the same educational attainment as a third-class degree in Mass Communications from the University of Bedfordshire.

The errors in this years of education measure, when 36 percent of 18 year olds in the UK entered higher education, will be much greater than in 1800 when less than 2 percent of 18 year olds attended universities. Years of education will also convey much more information about educational status in nineteenth century England when, depending on the family, education was completed at ages 5-22, compared with modern England where schooling is required for all up until age 18.

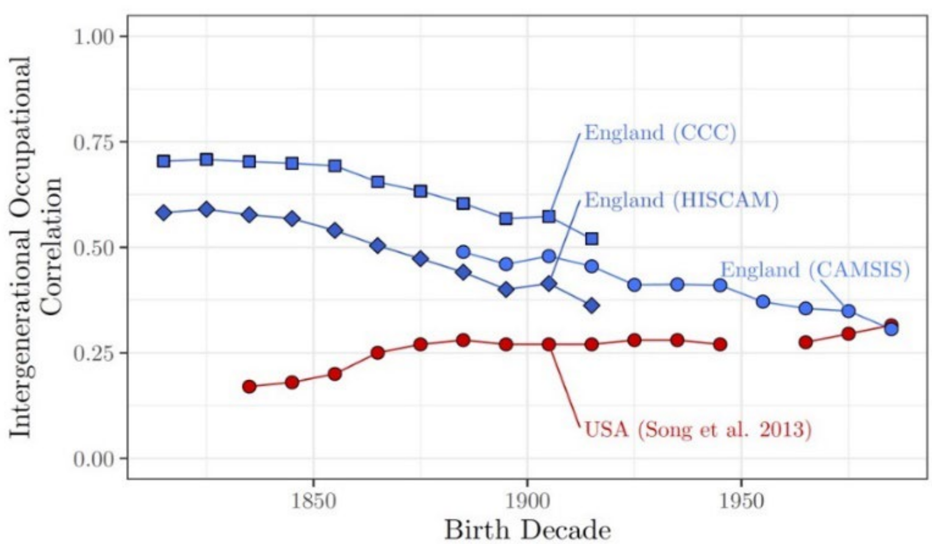
To illustrate the problem of measurement error in the study of social mobility consider some recent studies of social mobility rates over time in the USA and in England. For the USA we have two attempts to estimate intergenerational occupational mobility for men born 1830 to 1980. The first, published in PNAS in 2020 (Proceedings of the National Academy of Sciences) shows high rates of social mobility in the USA all the way from 1830 to 1980, but with a decline in mobility in recent decades. The second, published in the *American Economic Review* in 2023 shows very low rates of social mobility throughput, but increasing mobility in more recent years. Figure 4.1 shows these contrasting series, and their very different pictures of US history.

Figure 4.1: Two Views of Social Mobility in the USA



Sources: Song et al., 2020, Ward, 2023

Figure 4.2: Occupational Status Correlations USA Compared with England, 1810s-1980s



Notes: CCC, HISCAM, and CAMSIS are independent occupational status indices.
Source: Clark, Cummins, and Curtis, 2023

Another illustration of the problems of measurement comes from England. Figure 4.2 shows measured occupational mobility in England on three different indices of occupational status from 1800 to 1980. Just by changing the index in the earlier decades the measured rate of social mobility changes substantially. On all the English indices measured social mobility rates increase over time. They start out much lower than US social mobility rates as measured by Song et al. (2020). But by the 1980s the rates are similar.

However, there is ample evidence in England that the increase in measured mobility rates is just an artifact of the changing nature of jobs, and the terms used to label occupations. The nineteenth century terms such as *laborer* or *carpenter* or *clerk* were more exact statements of the status of the occupation than modern terms such as *manager* or *IT worker* which cover a wide range of job statuses. As errors in the connection between job descriptions and underlying occupational status become more important, measured persistence of status across generations will decline.

It is important to note here that there are three different types of measurement error in any estimation of social mobility rates. For occupational status these potential errors are: the worker may have the wrong occupational label, the occupational status index can assign the occupation the wrong average status level, and within each occupational label workers will vary in the level of skill they possess and their associated status.

The same measurement issues that appear across generations also appear where we try to measure the degree of assortment in marriage, or we try to measure how strongly siblings or cousins are correlated in status, these being alternative measures of social mobility rates.

Measuring Social Mobility without Error Attenuation

Where we have information on the correlation in outcomes across different types of relatives, the impact of these measurement errors, however, can be eliminated by taking the appropriate ratio of correlations between different sets of relative. To show this consider figure 4.3, which shows the observed correlations between siblings, children and cousins on some noisy measure of social status.[14]

Suppose *b* is the underlying parent-child correlation, and *s* the underlying sibling correlation. The observed correlations for each relationship will be θb, θs, where θ is the attenuation factor.[15] The observed correlation between uncle/aunt and child will be θbs, assuming the child-parent correlation is independent of the sibling correlation. This is because in any correlation between relatives, no matter how distant, the error attenuation occurs only once. In turn the observed cousin correlation will $\theta b^2 s$.

From figure 4.3, we see that b and s can be estimated without error attenuation from

$$b = \frac{Corr(aunt/uncle, niece/nephew)}{Corr(siblings)} = \frac{\theta bs}{\theta s}$$

also $$b^2 = \frac{Corr(cousins)}{Corr(siblings)} = \frac{\theta b^2 s}{\theta s}$$

and $$s = \frac{Corr(aunt/uncle, niece/nephew)}{Corr(parent, child)} = \frac{\theta bs}{\theta b}$$

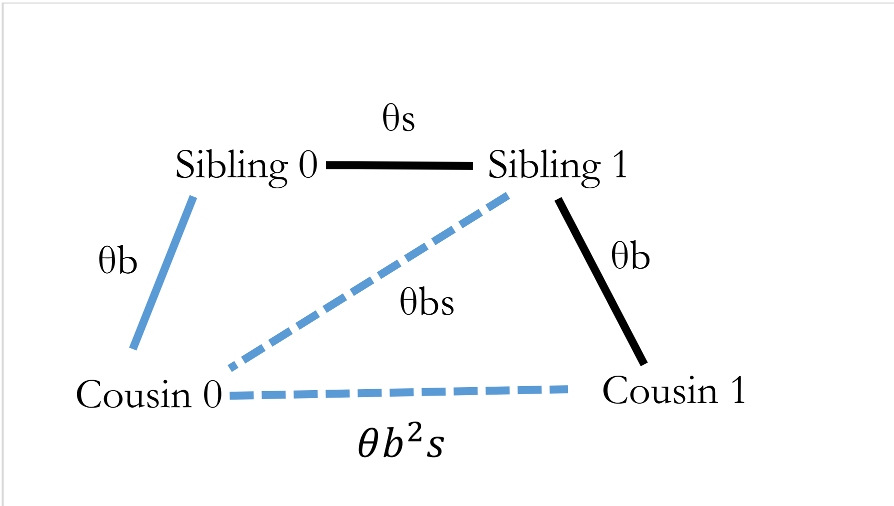
Note that while these ratio estimates will give us the true underlying correlations, since they involve a ratio, the standard errors of this estimate will be large. Thus implementing such estimates will require large amounts of data on familial correlations. Fortunately we have such data.

---

[14] Curtis, 2026, was the first to apply this technique to measuring underlying correlations in marriage in occupational status.

[15] Assuming the parent-child and sibling underlying correlations are independent of gender, and measurement error is independent of gender. I show below that even if not be averaging across all relationships with different gender compositions the result will hold.

Figure 4.3: Observed Correlations between Relatives – Siblings versus Cousins



Notes: Assuming the parent-child and sibling underlying correlations are independent of gender, and measurement error is independent of gender.

Note also that figure 4.3, and also figures 4.4 and 4.5 below, assumes the same underlying correlation between parent and child, siblings, and cousins irrespective of gender. I show below that there is generally good evidence for this simplifying assumption.

Table 4.1 reports the correlations between full siblings, parent and child, aunt/uncle to niece/nephew and between cousins for nine different social outcomes in England and Wales for people born 1725 to 1996. These outcomes for the recent years are house value (normed by region), the index of multiple deprivation for the local area of residence, and an indicator of whether someone had been a company director. For the late nineteenth and early twentieth century there is occupational status for men, and also for men, attainment of higher education. These measures also exist for men born 1780-1859. Finally for both men and women are measures of literacy at marriage, for marriages 1754-1889, corresponding to people born 1725-1869.

Table 4.1: Status Correlations and the Implied True Intergenerational Correlation

Outcome	Birth Period	Full Sibling	Parent Child	Sibling-rem	Cousin	b
House value	1910-1996	0.334	0.352	0.252	0.213	0.80
Imd	1910-1996	0.271	0.319	0.181	0.155	0.76
Comp Dir	1910-1996	0.175	0.141	0.061	0.064	0.67
Occstat	1860-1919	0.496	0.513	0.369	0.277	0.75
Higher Ed	1860-1919	0.359	0.305	0.213	0.146	0.64
Occstat	1780-1859	0.584	0.585	0.502	0.465	0.89
Higher Ed	1780-1859	0.558	0.500	0.398	0.366	0.81
Literacy	1725-1869	0.407	0.436	0.339	0.270	0.81
Mean		**0.40**	**0.39**	**0.29**	**0.24**	**0.77**

Source: FOE database

From the four correlations listed in table 4.1 we get three estimates for each outcome of the true intergenerational correlation of status between parent and child. The last column of the table gives the average of these three measures. Figure 4.4 shows the raw observed error attenuated parent-child correlation for each measure along with the error corrected correlation for each measure.

Table 4.4: Observed and Error-Corrected Parent-Child Correlations, England 1725-2024

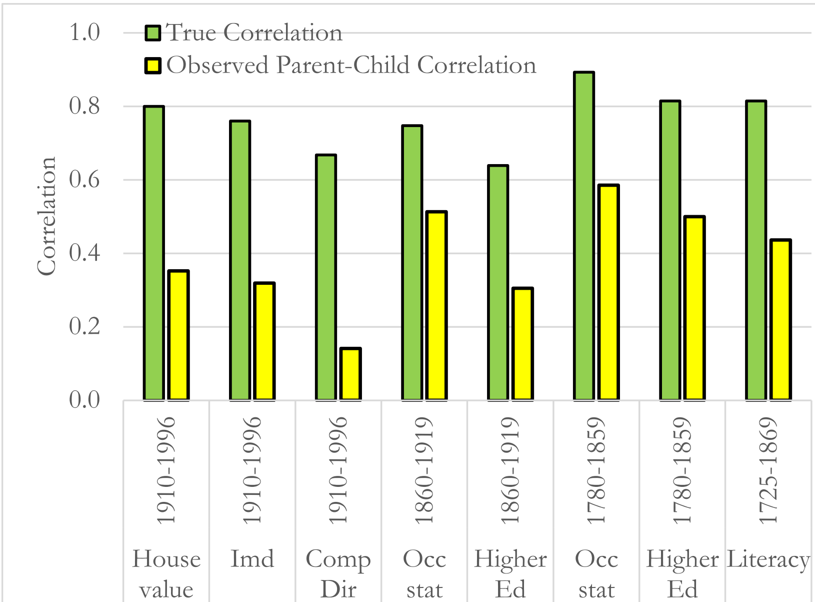
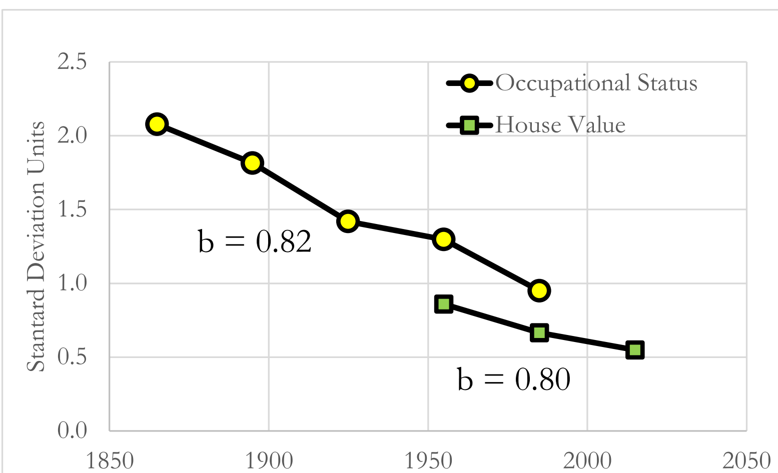


Figure 4.5: Average Status in Elite Surname Lineages, Cohorts Born 1810-1989



Source: FOE database

Notes: Status measured at age 40. Status is measured as average distance from the mean in standard deviation units. Intergenerational correlation for the occupational series 0.82, and for house value 0.80.

In each case the error corrected correlation substantially exceeds the raw observed correlation. The average raw correlation is 0.39, while the average error corrected correlation is 0.77, roughly double the magnitude. Further the error corrected correlation shows little sign of a long run trend. It averages 0.84 for those born around 1820, 0.70 for those born around 1890, and 0.74 for those born around 1955. This might imply some downwards trend. But if we look at figure 4.4 there is enough variation within different measures in each period that it is not evident that there is any trend. We shall see below other evidence that suggests little sign of any trend.

Confirmation that the ratio method above correctly captures the true underlying intergenerational correlation of status is shown in figure 4.5. This figure shows for a group of rare surnames identified as elite by their wealth for deaths 1858-1887 average male occupational status by 30 year birth intervals starting in 1810-39. It also shows by 30 year periods average house value 1999-2024 for men and women born 1900. Occupational status and house value are measured in standard deviation units above the mean.

This method, discusses extensively in my earlier book, *The Son Also Rises, Surnames and Social Mobility* (Clark et al., 2014), works to get an estimate of social mobility not attenuated by measurement error by averaging outcomes across pre-determined groups with higher or lower social status. In this case the average measurement error becomes zero for these groups, and we can observe the underlying rates of social mobility.

In this case the intergenerational correlation of status can be inferred as average status in period t+1, divided by average status in period t. The average persistence rate of status on the occupational measure for those born 1810-1959 was 0.82. For the house value measure of status the average persistence rate for those born 1900-1989 was 0.80. As the figure emphasizes social mobility rates are very low even in recent years, and they are surprisingly stable over the long interval 1810-2024.

Another set of data that allows us to estimate both the underlying marital correlation in social status, and again the underlying intergenerational correlation, is the MOE database of 1.7 million marriages in England and Wales 1837-2022. This reports the occupations of groom, bride, father and father-in-law, as well as the literacy of groom and bride. If we assume that marriages were contracted between groom and bride only, an assumption for which there is good evidence, then figure 4.5 shows the implied pattern of correlations between each of the parties on the marriage certificate. Here I allow for the error attenuation being different where there is a correlation between man and women, compared to between two males.

Now the underlying intergenerational correlation will be the father to father-in-law correlation divided by the groom to father-in-law correlation. The underlying marital correlation will be the groom to father-in-law correlation divided by the groom to father correlation. Table 4.2 shows the raw correlations between grooms, fathers and father-in-law observed in the data by 40 year marriage periods. These raw correlations, as noted, imply much greater occupational status mobility in the final period 1980-2021 than in the earliest years. Because of the large numbers of weddings observed in the earliest years

Figure 4.6 shows the implied underlying marital assortment, r, between grooms and brides. The dotted lines show the 95 percent confidence interval of the estimates. As can be seen this is close to 0.8 throughout the period 1837-2021, with no sign of any significant trend. The large number of marriages observed mean we get very tight estimates of r in the earlier years. The implication is that grooms and brides in England were matching very tightly in marriage from 1837 onwards.

Figure 4.5: Correlations between those listed on Marriage Certificates, 1837-2022

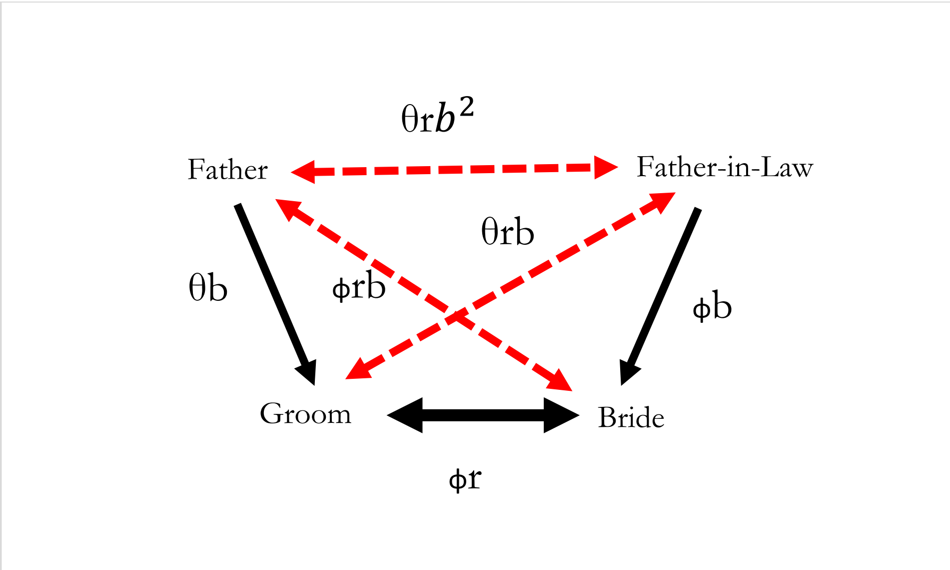
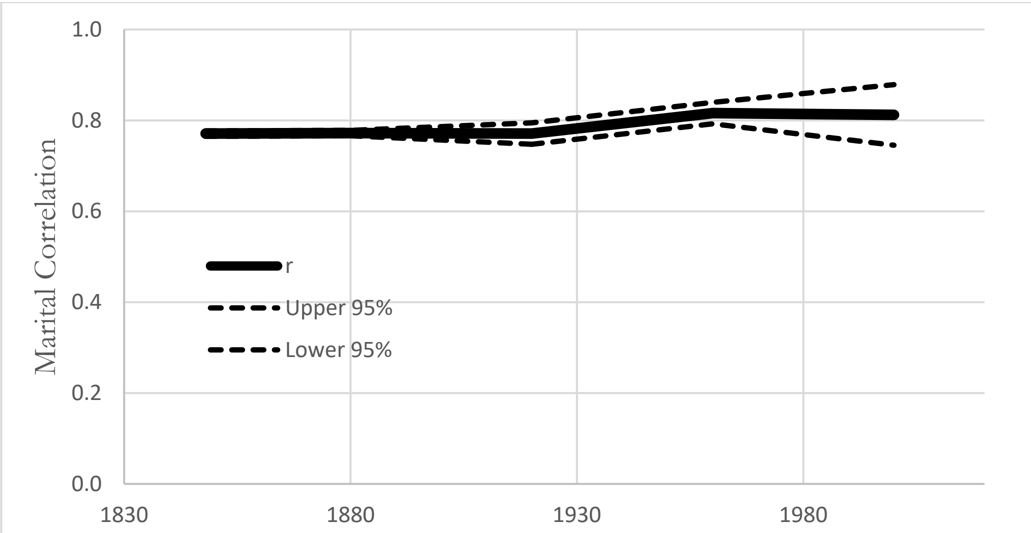


Table 4.2: Occupational Status Correlations, Marriages 1837-2021

Period	Marriages	Groom-Father	Groom-Father-in-law	Father-Father-in-law
1837-59	343,623	0.63	0.48	0.44
1860-99	438,725	0.60	0.46	0.42
1900-39	174,474	0.50	0.38	0.35
1940-79	47,033	0.42	0.35	0.32
1980-2021	10,444	0.34	0.28	0.23

Source: MOE database

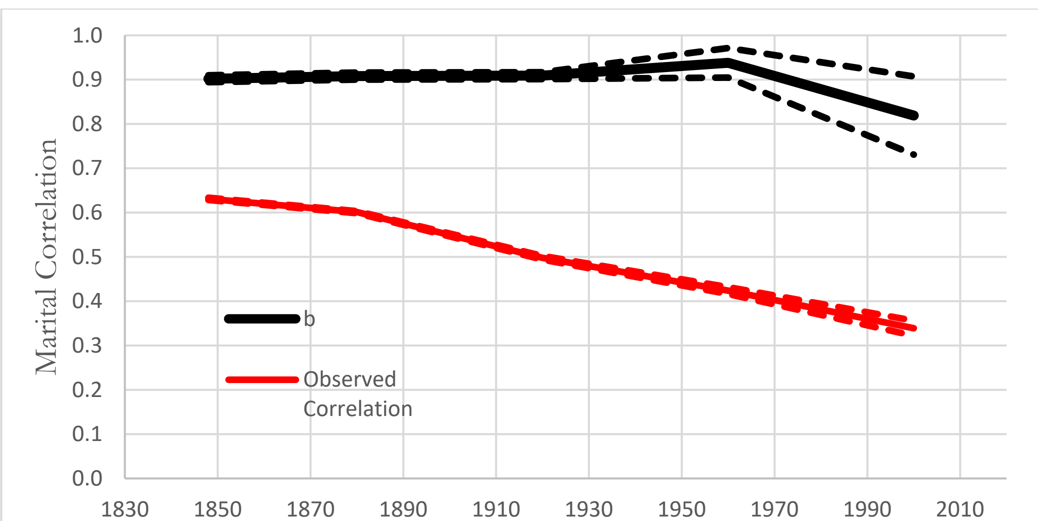
Figure 4.6: Underlying Marital Occupational Status Correlations



Note: the dashed lines show the 5 percent confidence interval.

Source: MOE database

Figure 4.7: Intergenerational Occupational Status Correlations



Note: the dashed lines show the 5 percent confidence intervals.

Source: MOE database.

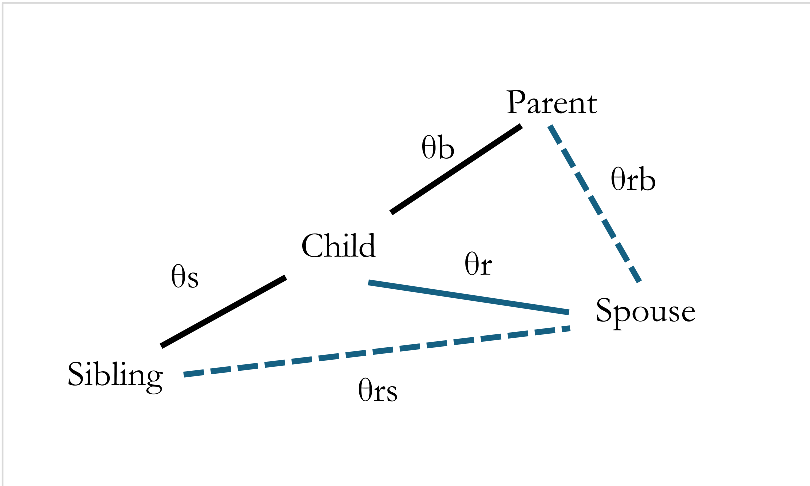
Figure 4.7 shows the corresponding estimate of the true parent-child correlation in occupational status, compared to the measured father-son correlation. The estimated correlation, b, in this case is close to 0.9 for all periods, except the last. But even in that last period the larger 95 percent error bands mean that there is no certainty that the intergenerational correlation declined.

The MOE database establishes the true marital correlation 1838-2021. With the FOE database we can also estimate that true correlation for marriages 1754-1837. To do that consider figure 4.8. This shows the correlations between a spouse, their partner, and their partners parents and siblings, on the assumption that the marriage was formed just between the marital partners.

In the years 1754-1837 grooms and brides were required to sign the marriage register, but there was no systematic recording of occupations of any of the parties to the marriage. Nor was there any recording of their father's or mother's literacy or occupation. However, from the FOE database we can find a substantial number of cases where we observe both the literacy at marriage of bride and groom 1754-1889 as well as the literacy of one or other of their father and mother at their own marriage. Table 4.3 shows a summary of this information for the years 1754-1837 and also for the modern marriage register period 1838-1889.

To maximize the information available, I have calculated for each period the correlation in literacy between the average literacy of the parents and their child, and child-in-law. Table 4.4 shows these correlations. Notice how the parents-child correlation declines substantially between the marriage period 1754-1837 and 1838-1889. This on conventional measures would imply substantial increases in social mobility. But we shall see it is likely driven only by increases in the general literacy rate creating a literacy measure that is less informative of underlying social status.

Figure 4.8: Observed Correlations between Relatives and In-Laws



Notes: Assuming the parent-child and sibling underlying correlations are independent of gender, and measurement error is independent of gender.

Table 4.3: Literacy of Children and Parents, 1754-1889, FOE database

Period	Observations	Literacy Child	Literacy Spouse	Literacy Father	Literacy Mother
1754-1837	1,149	0.67	0.60	0.72	0.50
1838-1889	3,669	0.79	0.78	0.72	0.56

Table 4.4: Literacy Correlations of Children and Parents, 1754-1889, FOE database

Period	Parents-Child Correlation	Parents-in-law Correlation	Implied Marital Correlation
1754-1837	0.518 (0.025)	0.400 (0.027)	0.768 (0.063)
1838-1889	0.383 (0.015)	0.300 (0.016)	0.786 (0.052)

*Notes***:** Standard errors in parentheses. Standard errors for the marital correlation estimate bootstrapped with 10,000 repetitions.

Table 4.4 also shows the implied underlying error free measure of marital correlation in status. That implied correlation is 0.77 for marriages 1754-1837, and 0.79 for marriages 1838-1889. These estimated marital correlations are very close to those found above with the MOE database, but using a different set of information on male occupational status.

Because of the modest numbers of observations, the 95 percent error bound for the 0.77 estimate in 1754-1837 is 0.65-0.89, and for 1838-1889 is 0.69-0.87. So the close coincidence is in part just a matter of luck. But the literacy data is consistent with the possibility that there was no change in the degree of marital assortment for marriages in England 1754-2021, a period of 267 years, and more than eight generations.

In these estimates we do not worry about whether the parents are those of the bride or groom, even though mail and female literacy rates were different. We mix these cases because any differences in measurement error between bride and groom will average out in this case, and the error will thus still be removed by the ratio method.

Marital Assortment and Social Mobility

If we assume fathers and mothers are equally important in determining child outcomes, then there will be a close connection between the status correlation of parents in marriage and the status correlation between each parent and their child. If the underlying status correlation in marriage is zero, then the maximum correlation between a single parent and a child will be 0.5. Conversely if the underlying correlation between parent and child is 0.8, then the marital correlation between parents must be at least 0.6.

Genetics and Social Status

With the extensive genealogical data in the Families of England (FOE) database we observe the pattern of correlations in outcomes between relatives as distant as fourth cousins, for social outcomes in England all the way from 1600 to 2026. This pattern of correlations is consistent with a surprisingly simple mechanism of status inheritance. This is that social status is determined by additive genetic transmission, in the presence of strong genetic assortment of partners in marriage.

Table 4.5 shows correlations in observed status for members of the Families of England lineage, down to the level of fourth cousins. Social status is here measured by six outcomes. The first three - occupational status, higher education status, and literacy at marriage – are measured for people born in the years 1780-1939. The second three - dwelling value, company directorships, and the index of multiple deprivation (IMD) for the residence location – are measured for people alive in the interval 1999-2024. Thus we have both historic and contemporary measures of social status.

What we observe in table 4.5 is that while unrelated individuals show no correlation on any of these outcomes, there is still a significant correlation for relatives as distant as fourth cousins. Fourth cousins have the same genetic share with each other as does a person with a typical single ancestor 8 generations earlier. So fourth cousins are potentially distantly related. Yet for most outcomes such remote relatives remain statistically significantly correlated. In line with the estimates from the first part of this chapter social mobility rates are very low.

Table 4.5: Social Status Correlations by Familial Connection, England, 1600-2022

Outcome	House value	IMD	CoDir	Occstat	Occstat	HighEd	HighEd	Literacy
Birth Period	1910-1996	1910-1996	1910-1996	1780-1859	1860-1919	1780-1859	1860-1919	1725-1869
Pairs observed	98,117	98,711	219,552	104,854	241,941	103,506	219,507	53,047
Correlations								
Full Sibling	0.334	0.271	0.175	0.584	0.496	0.558	0.359	0.407
Child	0.352	0.319	0.141	0.585	0.513	0.500	0.305	0.436
Sibling-rem	0.252	0.181	0.061	0.502	0.369	0.398	0.213	0.339
Grandchild	0.268	0.246	0.104	0.434	0.347	0.426	0.273	0.252
Cousin	0.213	0.155	0.064	0.465	0.277	0.366	0.146	0.270
Cousin-rem	0.158	0.101	0.015	0.325	0.257	0.235	0.181	0.206
Cousin2	0.139	0.074	0.069	0.232	0.151	0.103	0.045	0.225
Cousin2-rem	0.093	0.057	0.025	0.189	0.140	0.098	0.076	0.182
Cousin3	0.095	0.051	0.049	0.114	0.138	0.186	0.099	0.189
Cousin3-rem	0.076	0.017	0.019	0.065	0.139	0.118	0.079	0.202
Cousin4	0.084	0.032	0.024	0.079	0.090	0.020	0.040	0.146
Unrelated	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Source: FOE database

Notes: "-rem" indicates "once removed". Cousin2, 3 and 4 indicates second, third and fourth cousins. "Occstat" is occupational status, "HighEd" an indicator for higher education. "Pairs Observed" is the total number of pairs of relatives for each social outcome.

The second feature we observe in the correlations of relatives is that they decline with genetic distance in just the same way in the twentieth century, the nineteenth century, and the late eighteenth century. Also the rate of decline is the same for each measure of status. The decline in status correlations with each step outward in the lineage is a constant close to 0.8, for different measures of status, and for different epochs from 1600 to 2026.[16]

[16] The estimated correlation is 0.79.

Thus vast social changes in England since the Industrial Revolution, including the extension of democracy to the entire adult population, mass public schooling, and the emancipation of women, have not increased, in any way, underlying rates of social mobility.

The third interesting feature of the correlations are that they conform closely to those predicted by Ronald Fisher in 1918, for familial correlations in a phenotype that is determined by additive genetics, without any environmental input, in the presence of genetic assortment in mating.[17] In genetic models of social outcomes an important distinction is between the phenotype and the genotype. The genotype can be regarded as the plan for the formation of the individual encoded in their DNA. But as each person grows and develops the plan will be more of less successfully implemented. Some of this difference in outcomes will be created by the person's environment. But a significant share of the variance will come from molecular level random shocks that lead to better or worse implementation of the given DNA.

The "additive" specification in the model is the assumption that the social outcomes are determined by a large number of variants in the genome, each of which has a small effect on outcomes, that operates independently of all the other locations on the genome. It is thus the simplest model of genetic determination of outcomes. The total effect of the genome on an outcome like educational attainment is just the sum of all the effects of genome variants at the relevant location. For traits such as educational attainment, large scale whole-genome studies have already identified nearly 4,000 locations where there are variants influencing educational outcomes. But it is believed that the total number of such variants that influence educational attainment is potentially much larger. These whole-genome studies have also failed to find significant interactive effects between variants at different loci.

With such inheritance the phenotype correlations ρ_n between any two members within a family tree typically will be predicted by a formula of the form

[17]Fisher (1918). These results have been extended and clarified by Crow and Felsenstein (1968), Gimelfarb (1981), and Nagylaki (1982).

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

where n is the number of steps apart two people are in the family tree, m is the genetic correlation in parents, and h^2 is the heritability of the trait.[18]

Thus the correlation of siblings is

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

while for cousins it is

$$h^2\left(\frac{1+m}{2}\right)^3 \quad .$$

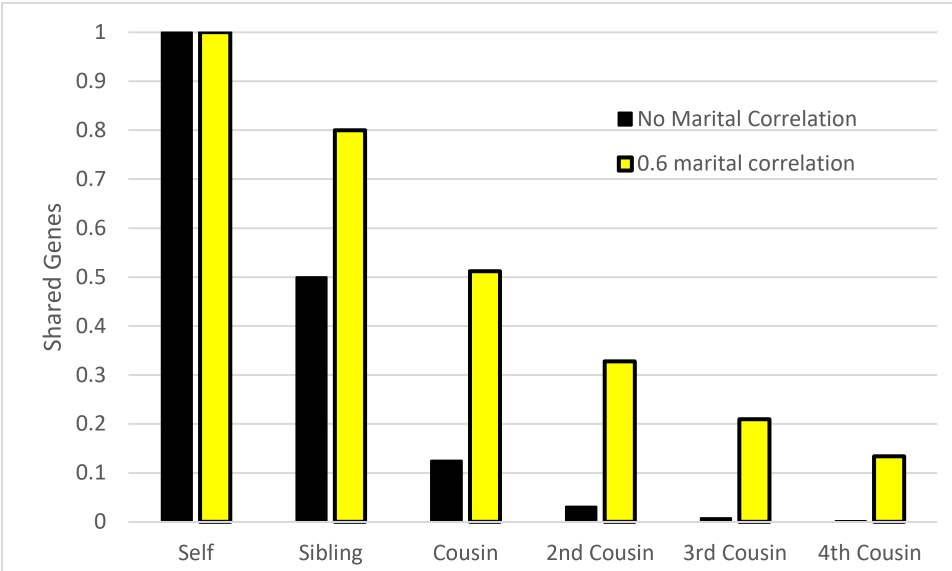
In general, for each step down or across the family tree between, the predicted correlation between relatives will drop by a factor of $\left(\frac{1+m}{2}\right)$, which is the underlying persistence rate.

In the Fisher specification, status persistence across generations is mainly a function just of m, the genetic assortment in marriage. If there is no such assortment, so that m=0, then correlations will decline by 0.5 with every step in the family tree. If m=1, so that people marry their genetic clones, then there would be no decline in status persistence.[19]

Figure 4.9 shows the genetic correlation between relatives at different distances in a family tree where the genetic correlation in marriage is 0, and where it is 0.6. With no assortment then the genetic correlation drops to less than 1 percent by the time we reach third cousins. Distant relatives have effectively no genetic correlation. But with a marital correlation of 0.6 then even fourth cousins share 13 percent of their genes.

[18] h^2 will be the correlation between a child outcome and the average outcome for their parents.
[19] Appendix table A1 gives the formulas for the expected correlations between relatives of different types with assortative mating on a phenotype, and additive genetic transmission.

Figure 4.9: Parental Genetic Correlation and Shared Genotype across Relatives



Note that difference in relative genetic correlation between members of the lineage with and without genetic correlation in marriage is greatest for more distant members of the lineage.

Since the correlations in table 4.5 decline by a factor of 0.785 with each further step in the family tree, this implies that *m*, the genetic correlation in marriage, has to be 0.57 to fit this model.

The genotype of marital partners is, of course, never directly observed. If men and women are matching closely genetically in mating, then that matching has to be based on their observed characteristics, their phenotype. Marital correlation on the genotype will be less than the correlation on the phenotype.[20] This is because the phenotype is determined by a mix of the genotype, social influences, and random effects. When parents match on a phenotype

[20] The genetic correlation in marriage, m, will relate to the phenotype correlation, r, through $m = h^2 r$.

including these social and random influences, they will create much less of a match on the genotype.

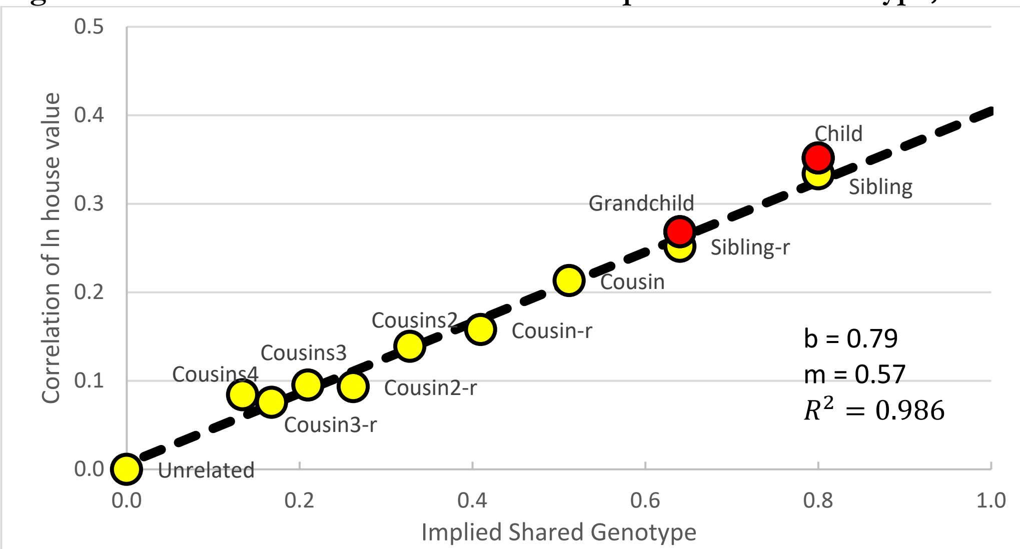
So if there is to be a correlation of 0.57 on the genotype that helps determine social outcomes, then there has to be a correlation between parents on some phenotype that is even stronger than 0.57. We see above in figure 4.6, and table 4.4, that from 1754 to 2021 we observe an underlying status correlation in marriage in England that was close to 0.8. This is strong enough to allow for a genetic correlation of 0.57. The underlying phenotype correlation was also the same across 1754-2021, which is consistent with the stability of intergenerational mobility rates across these years.

Figure 4.10 shows on the horizontal axis the implied genetic correlation between members of the lineage, with the best fitting marital genetic correlation, versus the observed correlation in outcomes, for the log of house values in England 1999-2023.[21] This figure covers men and women born 1920-1996.

There is a remarkably close fit between the estimated genetic correlation and the house value correlation. Notice also that for siblings the correlation is describing the social mobility process over just the years 1920-2023. In contrast the fourth cousins in the sample would have a common ancestor only 5 generations earlier, typically 150 years. So this fourth cousin correlation is summarizing what happened with social mobility across five generations, all the way from 1780 to 2023. Yet these observations fall on the same line as those which cover only one generation. This implies a remarkable stability in the strength of intergenerational persistence across widely different social systems.

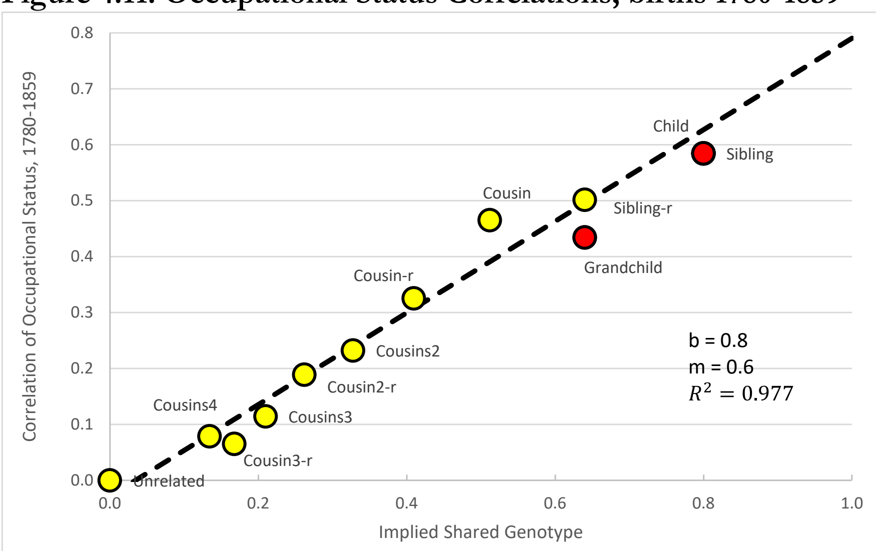
[21] These are house values controlled for regional variation in average house prices across England and Wales. If we do not control for such regional effects then the correlation in house values across relatives is even stronger, and persistence greater.

Figure 4.10: House Value Correlations and Implied Shared Genotype, 1999-2022



Notes: The dashed line shows the best fit to this data. The R^2 reported is for this fitted line. The child and grandchild correlations (in red) will potentially deviate from the fitted line.

Figure 4.11: Occupational Status Correlations, births 1780-1859



Notes: The dashed line shows the OLS fit to this data. The R^2 reported is for this fitted line. The child and grandchild correlations will potentially deviate from the fitted line.

Figure 4.11 shows for occupational status for men born 1780-1859, typically measured around age 40, their correlations with the same extensive set of relatives. The figure displays the same underlying persistence rate, and again conforms remarkably well to the Fisher model. 98 percent of the pattern of correlations is predicted from the model.

Now while brothers would both be born in the years 1780-1859, the common ancestor for fourth cousins could have been born as early as 1650. The fourth cousin correlations are again summarizing what happened in England in terms of social mobility across the long interval 1650-1859. A substantial part of this interval was in the pre Industrial Revolution economy, before substantial urbanization and the shock of new technologies and new industries. Yet social mobility rates show the same stability.

The linearity of the relationship in figures 4.10-4.12 further emphasizes the stability in England of the persistence rate b over generations. If persistence from one generation to the next was much higher in earlier years, then for fourth cousins, where on average in this data the common ancestor was born in 180 years earlier, the measured correlation now would be above the fitted line in the figure for more distant relatives.

There are another seven similar figures we can draw for the other outcomes and time periods: neighborhood characteristics and company directorships, 1999-202, occupational status for men born 1860-1919, higher education of men born 1780-1859 and 1860-1919, and literacy of men and women born 1735-1870. All show the same pattern, the same conformity to the Fisher model, and the same persistence of status with distance in the lineage.[22]

There is one strong regularity in the correlation data in table 4.5 which does not conform to the predictions of the Fisher model. That model holds that if there is parental matching on the phenotype, then the correlation between parent and child should be $\left(\frac{1+r}{2}\right)$, while that between siblings should be the smaller value $\left(\frac{1+m}{2}\right)$. But if we look at figure 4.12 which

[22] See Clark, 2023, Appendix 01, figures S1-S7.

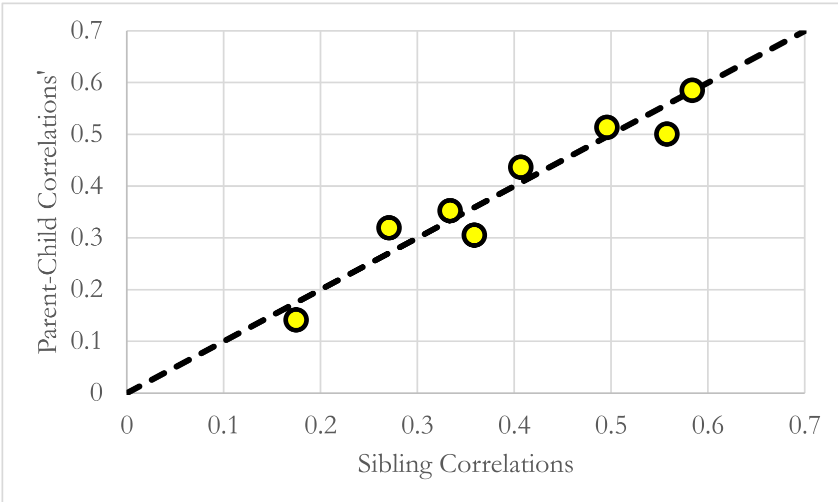
shows the parent-child versus sibling correlations for our nine outcomes we see that in practice these two correlations on average are the same. Such equality would only be predicted on the Fisher model if the parents were actually matching on the genotype itself.

Even this equality in correlations intergenerationally and between siblings is, however, good evidence against social transmission explaining the patterns of correlation found in table 4.5. For to construct a realistic social causation model for status, there has to be both a regression to the mean of family environments across generations, as well as shocks to the family environment to maintain variance. With an underlying persistence of status of 0.8, this implies an average correlation between parent and child environment of only 0.8. But if family environment is the determining factor in social outcomes, then the parent child correlation will be only 0.8 of the sibling correlation. Figure 4.12 decisively rejects this possibility.

In the above estimates of persistence and marital assortment we calculate the genetic link between individuals by their closest genealogical connection. However, if siblings sometimes marry siblings, and cousins sometimes marry each other, then some individuals labelled as first, second, etc. cousins will be more closely connected genetically than the standard cousin.

This will bias upwards the estimates of persistence, and the strength of genetic marital assortment. The question in practice is how significantly in England did the marriage of relatives increase the shared genetics of individuals in a family tree above those assumed in the estimates?

Figure 4.12: Sibling versus Parent-Child Correlations, England, 1754-2022



Notes: The correlations for the nine different outcomes in table 4.5

The biggest contributor to a greater genetic connection than predicted by looking just at the closest common ancestor will be the frequency of cases where siblings marry siblings (this cannot be that uncommon since my own maternal grandparents had siblings who married each other). In this case, for example, the shared genes of cousins (who are now double cousins) will be 0.60 versus 0.48, assuming a genetic correlation in marriage of 0.57.

From the FOE database we can estimate the share of all pairs of siblings who married siblings by use of the birth surname of their marriage partners (which will be the same for siblings). A set of 123,737 such sibling pairs show 1,773 with marital partners with the same surname. A random assignment of these surnames to each other produces only 97 such surname matches. The implication is that as many as 1.35 percent of the sibling pairs married partners who were also siblings, so that 1.35 percent of cousins were actually double cousins.

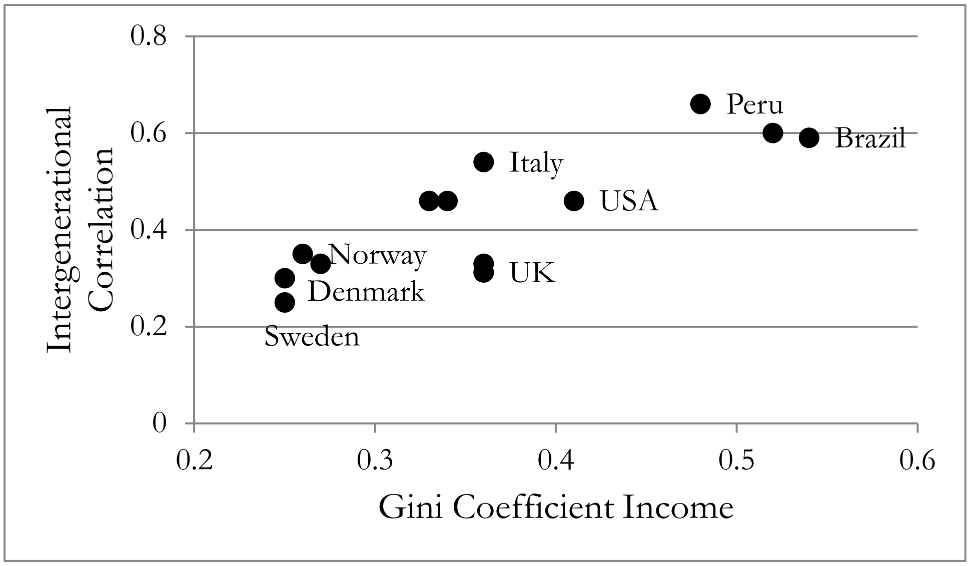
However, this would imply that the estimate of persistence of 0.79 from varieties of cousins would be inflated above the true persistence rate by only a factor of 1.0034 as a result of such extra genetic connections. This is not a significant inflation of persistence.

5. The Determinants of Social Status in Nordic Countries

True social mobility rates in England 1600-2022 were slow, unchanging over time, and with a pattern completely consistent with genetic determination of social outcomes. But how general is that result? Does it reflect just class-bound, sclerotic English society? Are the mechanisms of social mobility very different in the social democratic Nordic societies, where government social and educational supports are greater?

In Nordic societies, such as Denmark, educational provision by the state is notably generous. In Denmark there is free education from age 4 to 24. Further, at later ages the state also provides contributions to living costs to those receiving this education. Limited family resources are not a barrier to educational attainment, and abundant family resources are not a boost to attainment. There is also extensive subsidized pre-school provision ages 1-4, where access is not limited by incomes.

Figure 5.1: Intergenerational Correlation of Years of Education versus Income Inequality



Source: Clark et al. (2014), figure 1.4.

The popular image is that Nordic societies have achieved through such policies both high rates of social mobility, and high levels of income equality. Figure 5.1 shows, for example, the intergenerational correlation in years of education between parents and children for a variety of countries, graphed against the Gini coefficient of income. Educational correlations between parent and child are significantly lower than in societies such as the USA, where students have to finance much more of their own education.

However, as noted in chapter 4, measures such as years of education are only very imperfect proxies for educational attainment. As a result the correlations reported in figure 5.1 will tend to wildly overestimate true rates of educational mobility. In this chapter I show that if we adopt the same ratio strategy for eliminating measurement errors in Denmark and Sweden as in England, then will estimate rates of educational status persistence that are just as high as in modern England, and indeed just as high as in pre-industrial England.

The data for this demonstration comes from the population registers of Denmark and Sweden. From these registers we can measure educational attainment for each adult, where we include only people aged 26 and above to ensure they have largely completed their education.[23] We convert these measures of attainment into the standard measure of years of education.

To estimate rates of educational mobility we want to estimate three key parameters. First is the intergenerational correlation, that between a single parent and child, *b*. Second is the sibling correlation, *s*, which we saw in the case of England tends to equal the intergenerational correlation. Last is the parental correlation, *r*, which plays an important role in determining the other two coefficients.

As in chapter 4, we assume each of these correlations is observed only with a measurement error attenuation, θ. To eliminate this error we consider other relationships where the correlation is some product of *b*, *s*, and *r*. Such relationships are portrayed in the figures in chapter 4 (figures 4.3, 4.5, and 4.8).

Such relationships include cousins ($\theta b^2 s$), sibling-in-law (θrs), child-in-law ((θbr), parent and parent-in-law ($\theta b^2 r$). We can identify such relationships in the register data using the fact that each person in the modern period in the registers is linked to their father and mother. We can thus establish a set of full siblings, as well as a set of parental partners. This leads then also to a set of siblings-in-law. Since the registers encompass three generations since their foundation, we can also establish cousin pairs, as well as niece/nephew links to uncles/aunts.

To ensure the expected correlations in these extended family relationships are the product of the correlation in the individual steps that link people, we have to assume that parental unions were formed based just on matching between the two parties themselves. We also assume that the correlation a sibling and their offspring is independent of the correlation for another sibling and their offspring.

[23] The correlation data for Sweden comes from Collado et al., 2023.

Table 5.1: Years of Education, Correlations with Relatives, Denmark and Sweden

Relationship	Expected Corr.	Denmark Count	Denmark Corr.	Sweden Count	Sweden Corr.
Spouse	θr	1,869,156	0.443	413,062	0.491
Parent	θb	2,993,118	0.366	1,618,853	0.400
Sibling	θs	4,241,096	0.302	1,595,270	0.354
Child-in-law	θbr	3,292,907	0.252	-	-
Aunt/Uncle	θbs	2,545,645	0.283	2,427,198	0.294
Sibling-in-law	θrs	783,564	0.232	2,359,664	0.232
Cousin	$\theta b^2 s$	650,723	0.191	1,198,135	0.191
Sibling of Sibling-in-law	$\theta r s^2$	1,582,569	0.192	1,890,600	0.198
Spouse of Sibling-in-law	$\theta r^2 s$	2,539,664	0.240	943,108	0.231
Parent to Parent-in-law	$\theta b^2 r$	2,709,905	0.193	-	-
Aunt/Uncle-in-law	θbrs	765,909	0.183	1,908,238	0.180

Sources: Danish Population Register, Collado et al., 2023, table 3.

Note: Each row shows correlation of id to all specified relatives, as well as the numbers of such relatives observed. Sweden has a population double that of Denmark, but the Collado et al. paper uses only a subset of these family relationships.

Table 5.1 shows correlations in years of education for a set of such close relatives for both Sweden and Denmark. A strength of the register data is the large numbers of people observed for each relationship pair. By appropriate ratios of the correlations in table 5.1 we derive the estimates in table 5.2 of the underlying correlations *b*, *s*, and *r* for both Denmark and Sweden.

Table 5.2: Implied Underlying Correlations, Denmark and Sweden

Relationship	Denmark			Sweden		
	b	r	s	b	r	s
Child-in-law	0.57	0.83				
Aunt/Uncle	0.62		0.77	0.57		0.66
Sibling-in-law		0.76	0.64		0.72	0.60
Cousin	0.82			0.82		
Sibling of Sibling-in-law			0.68			0.68
Spouse of Sibling-in-law		0.86			0.80	
Parent-parent-in-law	0.77					
Aunt/Uncle-in-law	0.65	0.79	0.73	0.61	0.77	
Average	**0.69**	**0.81**	**0.70**	**0.67**	**0.76**	**0.64**

Source: Table 5.1

Notes: The table shows the implied values of correlations *b*, *r*, and *s* from table 5.1.

The correlations in table 5.1 include all gender compositions of the specific relationship pairing: male to male, male to female, and female to female. The table assumes that the attenuation factor θ is the same for correlations across all these pairings. Even if there is a different measurement attenuation for male-male, male-female and female-female correlations, the results in table 5.2 will still hold. For if the attenuation factor for male-male correlations is φ, for male-female τ, and for female-female ω, then the average attenuation for all relative pairs if calculated separately by gender would now be, for all relationship pairs

$$\theta = \frac{1}{4}(\varphi + \omega + 2\tau)$$

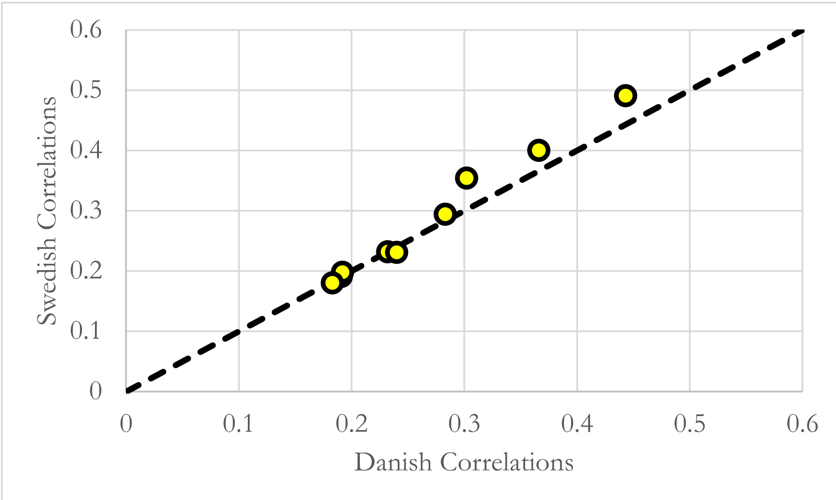
Thus calculating true underlying correlations using ratios will see these disparate attenuation factors appear in the same way in numerator and denominator of each ratio employed.

Figure 5.2 shows that the correlation structure used here is relatively stable across Denmark and Sweden. The only modest divergence is that the single relationship correlations in Sweden (θb, θr, θs) are all somewhat stronger than in Denmark.

Table 5.2 shows the derived underlying values of *b*, *r*, and *s*. There is a fair amount of noise in the individual estimates, though the implied measures without measurement errors are all substantially higher than the observed values of *b*, *r*, and *s*. But the overall estimates for Denmark and Sweden are similar across all three measures.

Table 5.3 shows these Nordic mobility parameters relative to those derived in chapter 4 for England now and in the past. The key element of parental assortment is just as strong in modern Nordic countries as in pre-industrial England. These estimates are remarkably similar for marital assortment. They are a bit lower in the Nordic case for intergenerational mobility and sibling correlations, but only about by 10 percent. These estimates also confirm for the whole population in Sweden the surname estimates of Clark et al. (2014), chapter 2, which also suggested true intergenerational mobility rates in modern Sweden were low, with persistence rates in the range of 0.7-0.8.

Figure 5.2: Education correlation Pattern in Sweden versus Denmark



Source: Table 5.1

Table 5.3: Three Key Mobility Correlations, England, Denmark, Sweden

Population	**Period**	**Marital** r	**Sibling** s	**Parent-Child** b
England	19^{th} c	0.77	0.79	0.80
England	Modern	0.80	0.72	0.77
Denmark	Modern	0.81	0.70	0.69
Sweden	Modern	0.77	0.64	0.67

Sources: Chapter 4, table 5.2.

The strong conclusion from the simple exercise above is that the whole popular notion of Nordic societies as beacons of social opportunity and social mobility is completely mistaken. The vast public expenditures of the Danish or Swedish states on their extensive education system and student support may be justified on other grounds, but certainly not as a promoter of social mobility.

It also emphasizes that attempts to compare social mobility rates in different societies and different time periods using just uncorrected social mobility measures are not just pointless, but also likely to lead to erroneous conclusions.

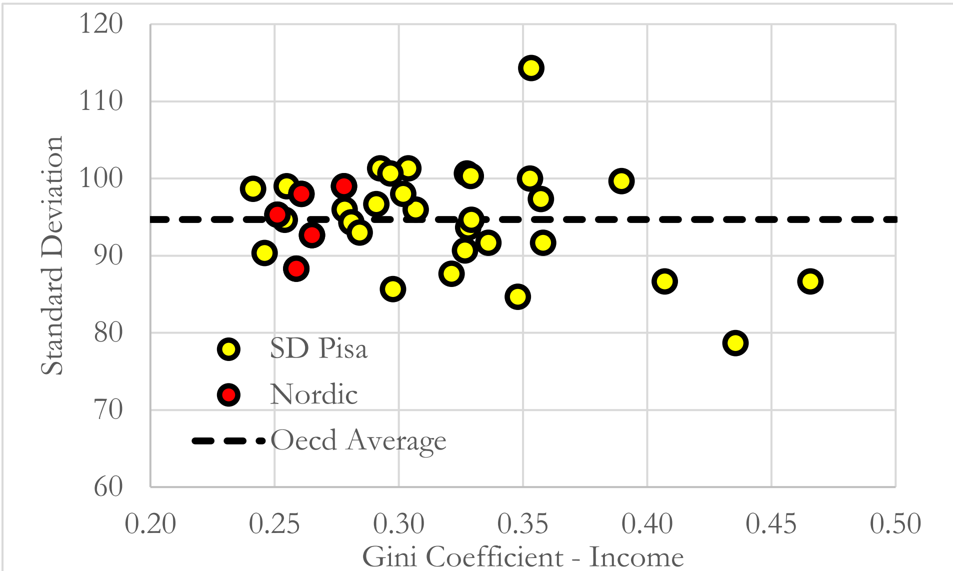
Support for the claim that societies such as Denmark and Sweden have not succeeded in increasing rates of social mobility comes if we look at the distribution of scores in the PISA (Program for International Student Assessment) measures of student reading, mathematical and science knowledge. The aim of this program is to give periodic assessments of the academic achievements of a representative sample of 15-year-olds in each country.

Figure 5.3 shows for each OECD country in the 2018 collection of the measures the average standard deviation by country of the three scores. These are graphed against average Gini coefficient for disposable income in each country for 2012-2022.

What is very clear in the figure is that despite the low Gini coefficients in the Nordic countries, implying more equal home circumstances for 15-year-olds, and despite the generous provision of educational support, Nordic countries have a distribution of student outcomes which is as high as the average for OECD countries.[24] Their extensive social provision has not succeeded in narrowing scholastic performance gaps within each cohort.

[24] The Nordic countries are Denmark, Finland, Iceland, Norway and Sweden.

Figure 5.3: Standard Deviation of PISA scores 2018, by OECD country



Source: OECD, 2019, tables 1.B1.1-6, pp. 216-21.

Notes: The dashed line shows the OECD average standard deviation of 94.7. The Gini coefficient is that for disposable income, averaged for 2012-2022.

In Nordic countries the great majority of children attend publicly subsidized pre-school ages 1-5, where standards of care and education are not dependent on parents' income. There is a substantial recent literature that emphasizes the benefits of such pre-school is particularly strong for children from the least advantaged families, if the care is of high quality (Gormley et al., 2005, Heckman and Lochner, 2000, Waldfogel, 2006). This should compress the educational attainment differences of 15 year olds in Nordic countries. It does not.

The Determination of Social Status, Denmark and Sweden

A recent paper, Collado et al., 2023, claims to be able to establish from the correlations in educational status and income between an extended collection of both consanguine and affine relatives in Sweden, that we can rule out genetic transmission as the source of correlations between relatives.

Collado et al. (2023), were able to identify in Sweden 141 different types of relatives in register data by linking children to each of their parents, parents to grandparents, parents to siblings, and so on through in-laws across as many as five marriages. They then estimate the correlation in years of education and income between each of these relatives. The more remote of these relatives include spouse of sibling-in-law (third degree), which involves linking seven individuals across four marriages and three sibships.

To this correlation data, Collado et al. fitted a social transmission of status model which has 20 free parameters, a combination of latent and observed status, as well as marital matching on both latent and observed status. Then using a selected set of 105 correlations, Collado et al. were able to show a strong fit of this 20-parameter model to the data.[25] That 20-parameter model, however, while showing strong assortment in parenting, and strong persistence across generations, in line with the data discussed above, does not offer much insight into why the explicit and latent variables have the connections estimated.

Collado et al., 2023, also believe they can show that the standard quantitative genetic model will not fit the pattern of correlations observed. They find, for example, that a genetic model where the matching of parties is on the observed phenotype (years of education) will

[25] The baseline model is fitted to only 105 correlations because this same model does not fit well for cousins, or for in-laws to the fourth degree or higher. "Our baseline calibration understates the correlation between cousins by around 25%, and the correlations for in-laws of four or more degrees of separation by an even larger amount (although the deviations are small in absolute value). The gaps close only partially if we directly target these moments in the calibration" (p. 1213).

not fit the observed pattern of correlations well. In figure 5.4 is given one of their attempts to fit the genetic transmission model to the extensive set of correlations, this one just calibrating the model using 15 correlations of siblings, parent-child, and aunt/uncle to niece nephew. The model clearly does not fit with the observed correlations.

Figure 5.4: Model Fit of Genetic Model to Extended Swedish Family Correlations

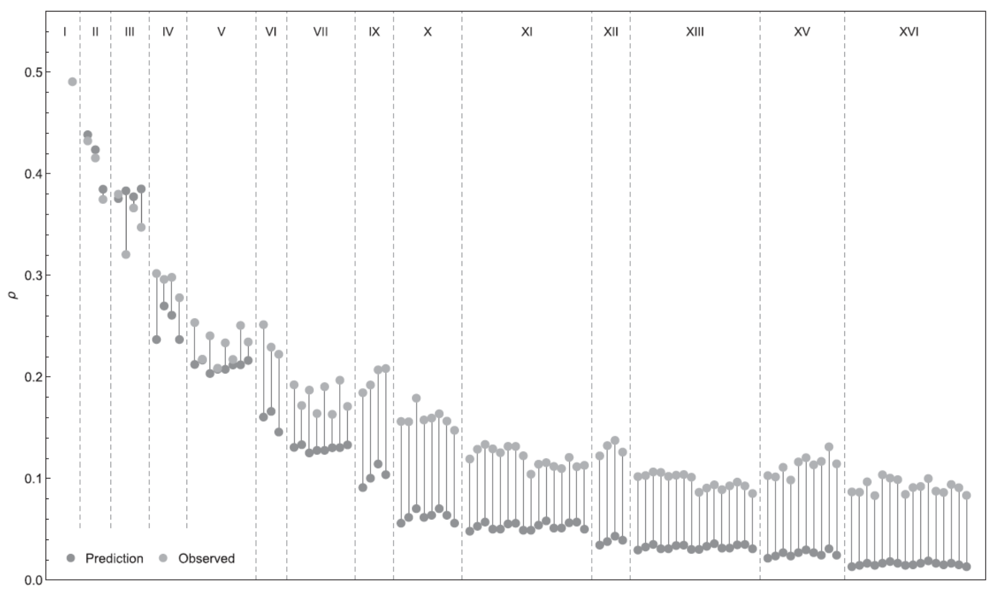


FIGURE 6
Model fit (standard genetic model, education)

Source: Collado et al., 2023, figure 6, p. 1220.

However, in their exercises Collado et al., 2023, do not take into account the measurement errors discussed above. The take "years of education" as measuring without error educational attainment. This leads immediately to the conclusion that the phenotype correlation in education between parents in Sweden was only the observed 0.49. In turn this implies that the genetic correlation could be at a maximum well below this observed phenotype correlation.

As table 5.3 above suggests such values of *r* and *m* will not fit the Swedish data. Both the full attempt to fit the genetic model, and the one using only a subset of relatives then show the characteristic that the correlations decline across more distant relatives too rapidly to fit the strong persistence observed in the data.

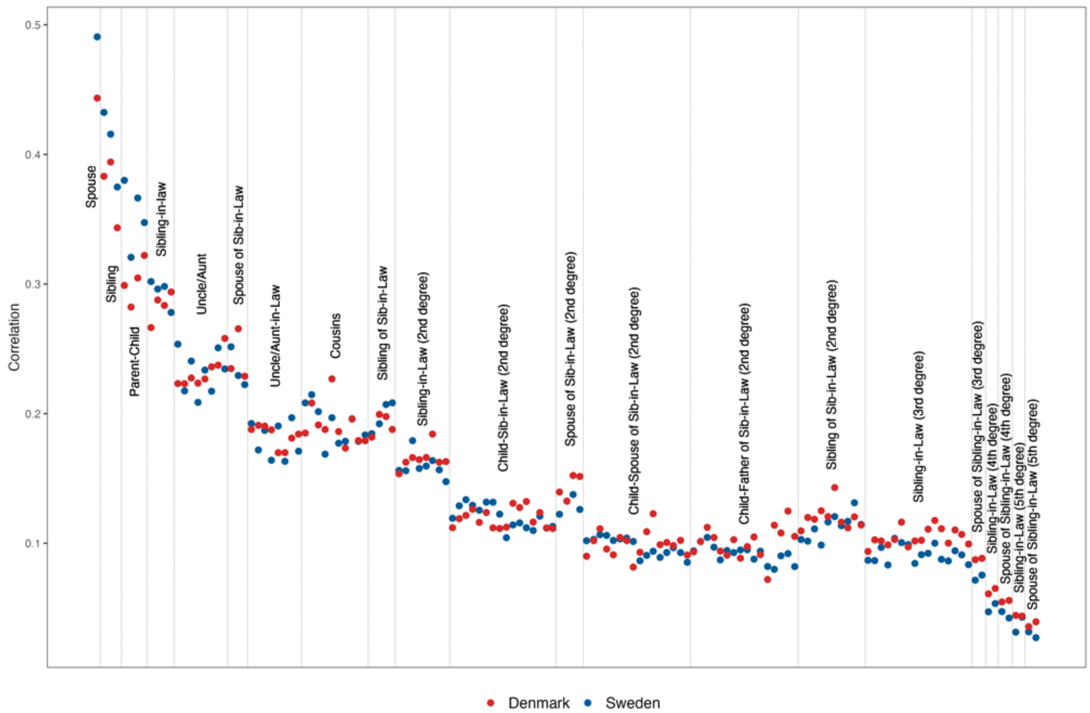
However, if we adopt the same strategy as above and assume that the individuals matching on educational status in Denmark and Norway are matching on a phenotype that "years of education" measures only imperfectly, then the genetic model (with three parameters) will fit the extended educational correlations in both Denmark and Sweden just as well as the 20 parameter model Collada et al. employ. Now the observed phenotype correlation between parents of 0.45-0.49 is consistent with a true phenotype correlation close to 0.8. And this in turn allows for a genotype correlation that can be as high as 0.55. And as we saw in chapter 4, this genotype correlation in the genetic model is the key determinant of social mobility rates across generations.

Table A2 in the technical appendix gives the formulas used for the expected correlation between each set of relatives. The key parameters in these correlations to be estimated are the measurement error, θ, the phenotype correlation between parents, r, the genotype correlation between parents, m, and the heritability of the trait, h^2. Since r, m and h^2 are linked there are actually only three parameters in this model.

Figure 5.5 shows the correlations in years of education between all 141 sets of relatives identified by Collado et al., for both Sweden and also for Denmark. The pattern of correlations is broadly similar in Denmark as in Sweden.

When the genetic model, with measurement error incorporated, is fitted to the correlation data we get the fit shown in figure 5.6 for Sweden. The figure shows both the reported correlations, as well as the predicted ones from the simple genetic model.

Figure 5.5: Comparison of Swedish and Danish correlations in Education



Notes: This figure presents the sample correlations in years of education among Danish and Swedish relatives.

Figure 5.6: Fit of Collado et al. 20-parameter model versus three parameter genetic model, Sweden

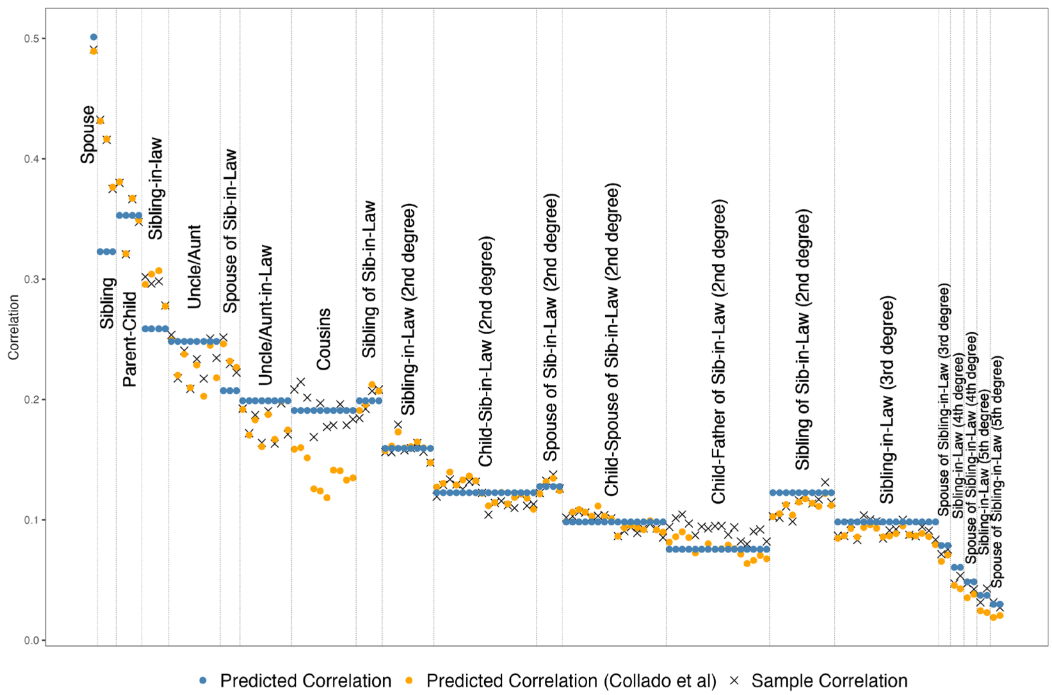


Table 5.4: Fitted Genetic Model, England, Denmark and Sweden

Population	Period	Marital Phenotype r	Marital Genotype m	Heritability h^2
England	19th c	0.77	0.54	0.70
England	Modern	0.80	0.59	0.74
Denmark	Modern	0.83	0.54	0.64
Sweden	Modern	0.80	0.52	0.65

In this case we also have for comparison the fit of the Collado et al. 20-parameter model to the same Swedish correlations. Since the genetic model does not differentiate by gender, all sets of cousins have the same predicted correlation. In contrast Collado et al. have different fits for each gender combination in sets of relatives.

Figure 5.6 suggests, just by visual inspection, that the genetic model fits the data fairly well. It also suggests that it fits as well as the more elaborate Collado et al. model. For some relationships, such as siblings, the fit for the genetic model is less good than for the Collado et al. model. But for others, such as cousins, the genetic model fits better.

Table 5.4 shows for both Denmark and Sweden the key parameters of the best fitting genetic model to the correlations of figures 5.5 and 5.6. Three things stand out. First the fit to the Danish and Swedish data is very similar. There is seemingly a relative stable correlation structure across these two societies. Second there is in all these cases a high implied phenotype correlation across parents, that is close to 0.8. Third there is in all cases a high implied genotype correlation across parents in the range 0.52-0.59. This implies in all these societies low rates of intergenerational mobility, and high familial correlations in educational outcomes.

Conclusions

Chapter 4 showed that for a variety of social outcomes in England we observe the same strong persistence across generations all the way from 1600 to 2022, with intergenerational correlations declining by a factor of only 0.8 across each generation. The chapter also showed that the pattern of correlations across relatives as far apart as fourth cousins is consistent with additive genetic determination of social status, combined with a genetic correlation in marriage of close to 0.6. Vast social changes in England 1800-2022 had no effect on underlying marital assortment or social mobility rates.

Denmark and Sweden are both societies with more even more extensive social welfare systems than in modern England, and less income inequality. Have these social policies changed the nature of status persistence in these societies?

In this chapter we find, using two different methods of estimation, and data on years of education, that parents show the same high underlying correlation in social status as was observed in pre-industrial and modern England. We also find that the implied long run persistence of status is strong. Using just a ratio of correlations between different types of relatives we estimate this at 0.67-0.70. Fitting a fully specified genetic model we estimate persistence at 0.76-0.77, which is very close to the English estimate of 0.78. Fitting the genetic model using years of education we find an implied genetic correlation between parents of 0.52-0.54, compared to 0.54-0.59 for England. Denmark, Sweden and England seem to have essentially the equivalent social structure.

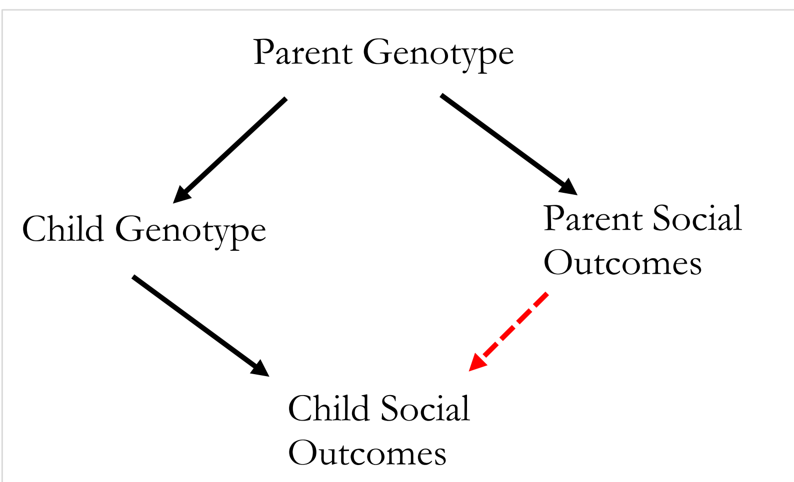
The lessons Denmark and Sweden convey is that if you remove social barriers to educational attainment, so that all across the status spectrum families have equal access to education and health services, then the remaining differences in educational outcomes have to come either from genetic differences, or from the operation of random chance. The Scandinavian meritocracy is a society where the systematic elements determining people's life outcomes are the genes they inherited from their parents.

6. Genetic Nurture

The Nature of Genetic Causation

Suppose in England 1600-2026, and in modern Nordic countries, social status was mainly determined by genetic inheritance. This does not in itself imply that social interventions cannot change social outcomes. There has been much recent discussion of an alternative causal path through *genetic nurture.*[26] Here the parental genotype creates a social environment for the child which favors better child social outcomes. Genetics is correlated with social outcomes, partly directly, but substantially through the indirect pathway of family environment. Figure 6.1 portrays this potential alternative social path of genetic influence.

Figure 6.1: Alternative Causal Paths between Genotype and Social Outcomes



[26] Kong et al., 2018, Selzam et al., 2019, Howe et al., 2022.

There is even recent evidence that parents with higher polygenic scores for educational attainment provided warmer, more sensitive, and more stimulating care to their children (Wertz, J. et al, 2019).

If the social path is an important source of the correlation between the child genotype and the child social outcomes, then social interventions can play an important role in influencing child social outcomes. State interventions associated with the welfare state – income subsidies, public schooling, social housing, medical care - will improve the social environment of children from parents with less favorable genetics.

Demonstrating Genetic Nurture

Through analysis of the connection between single nucleotide polymorphisms in the DNA, and years of education attained, it is possible to construct for people, based on their DNA, polygenic indices that score their DNA in terms of its educational potential. These indices have been improved over time with larger and larger collections of DNA, and there now exist indices EA1, EA2, EA3 and EA4.

Okbay et al., 2022, who constructed the latest and most powerful index, EA4, claim to be able to distinguish between the direct effects of DNA on educational outcomes, and the indirect genetic nurture effects. The way this is done is the first calculate the population connection between EA4 and years of education (EA), where they find

$$\text{EA} = 0.33\ \text{EA4}$$

Then, using data on dizygotic twins, children who share a common environment, Okbay et al., 2021, also finds that if we take the difference in educational attainment between these twins ΔEA and predict that using the difference in their polygenic index $\Delta EA4$, the connection is weaker

$$\Delta EA = 0.18\ \Delta EA4$$

This results echoes similar results found in other studies using the earlier polygenic indices (Liu, 2018, Willoughby et al., 2021). The claim then is that only 56 percent of the observed genetic effect is from direct transmission, with 44 percent flowing through the common family environment parents created for the twins.

A potential problem with this demonstration, though, is that even the latest polygenic index, EA4, still is a very noisy measure of persons' genetic potential for education. We know from twin studies that the share of variance in education explained by genetics is much greater than the share explained by EA4. If we label g as the full genetic score of an individual, then this implies that EA4 only connects to the true genetic potential with a large error component, u.

$$\text{EA4} = \text{g} + \text{u}$$

In this case the connection between years of education (EA) and the index EA4 will be biased downwards compared to the true connection between a persons' genetics and education outcomes. The extent of that bias will be

$$\frac{\sigma_g^2}{\sigma_g^2 + \sigma_u^2}$$

The higher the ratio of the noise, σ_u^2, to the true genetic signal, σ_g^2, the greater the downward bias.

By looking within families at the difference between siblings in educational attainment versus the difference in polygenic scores Okbay et al., 2022, increases the variance of the errors while reducing the variance of the signal.[27] In consequence the connection between

[27] $\Delta EA4 = \Delta g + \Delta u$. Δg will have lower variance than g for the general population. Δu will have four times the variance for u in the general population. Jonathan Beauchamp, one of the authors of

ΔEA and $\Delta EA4$ is biased down more by errors than when the effect is estimated across the whole population.

Thus these within family demonstrations of supposed genetic nurture will seriously overestimate its contribution. The share from the direct effect is much larger than these studies suggest.

A potentially more convincing demonstration of genetic nurture, is Kong et al, 2018. Using data from Iceland the paper calculated for each child their genotype years of education potential score. But also it calculated an equivalent score for the 50 percent of the parental DNA that each child did not inherit. This non-inherited DNA could not play any direct causal role in their educational outcomes. That parental DNA, however, will have an influence on the child's social environment.

Then the paper estimates the effect of the child genotype score, g_c and the non-inherited parental DNA score, g_{-c}, on child years of education, EA_c, by estimating the values of b_0 and b_1 in the expression

$$EA_c = b_0 g_c + b_1 g_{-c}$$

The direct contribution of the child genotype to educational attainment will then be $b_0 - b_1$, while the indirect contribution of parent genetics to child outcomes will be $2b_1$. The estimated values are $b_0 = 0.223$ and $b_1 = 0.067$. This makes the direct contribution of the genotype of 0.156, just slightly more than the indirect effect of 0.134.[28]

---

the Okbay et al., 2022, paper argues in response that for siblings the errors, u, will be correlated, so that the variance of Δu will not increase relative to the variance of Δg. He does not explain the mechanism, however, behind this conclusion.

[28] Young, A. I. et al. 2018 extend this method by imputing parental genotypes from siblings, and again find significant indirect effects of the parental genotype on the child phenotype.

These claimed "genetic nurture" effects face a number of issues, however, which suggest they are a phantom rather than a real effect.

The first problem is that the DNA score that Kong et al., 2018, relies on in the demonstration of genetic nurture above is still a noisy measure of the genotype of the children, and of their genetic predisposition towards more educational attainment. It still predicts, for example, educational attainment less well than just knowing the years of education of each parent. If parents show a correlation in the genetic predisposition for education, then the polygenic score of the non-inherited DNA will be a predictor of the educational potential in the parts of the DNA linked to educational outcomes in the child that have yet to be identified. The non-inherited DNA provides just further information on the genetic educational potential of children. It plays no causal role.

Kong et al. (2018), aware of the issue of assortative mating, made a correction in their estimates for assortative mating. However, Shen and Feldman (2020) argue these corrections involve unrealistic assumptions, such as that the transmitted and non-transmitted genetic scores are uncorrelated in each parent.

And we see above in chapters 4 and 5 ample evidence that parents correlate genetically closely on educational potential. So it is very plausible that the non-inherited DNA polygenic score is just another predictor of child educational potential, and not an indication of genetic nurture.

Another problem with these general demonstrations of genetic nurture is that these effects should have been identified in adoption and twin studies.

If, for example, half of social outcomes are through the indirect path, then adopted children would correlate half as well in outcomes with adoptive parents as do biological children with their parents. But in fact adoptive children, when adoption is random, show very weak correlations with adoptive parent outcomes, or with the outcomes of unrelated adopted children in the same family (see below, chapter 7). As noted in chapter 2, Robert Plomin, who has had a long career in twin and adoption studies, reports that for traits measured in adulthood – IQ, years of education, occupation – the correlation of unrelated adopted children within a family is minimal.[29]

The second problem is that as we saw in chapter 2, once we allow for significant genetic assortment by parents in social characteristics, twin studies for the determinants of years of education find no significant social causation. All the variation comes from genetics or just random effects. If genetic nurture is significant it should appear in the twin studies.

Another claimed demonstration of genetic nurture is that of Selzam, et al., 2019, which uses dizygotic twins. In this case the core expression whose parameters are estimated is

$$EA_c = b_0(g_c - \bar{g}) + b_1\bar{g}$$

where EA is a measure of educational attainment (or in the Selzam et al. paper, exam scores and IQ). $\bar{g}$ is the average genotype score of a child and their twin, and is intended to capture again a family environment effect of parental DNA.

If the genotype influences outcomes only through the direct effect, then the claim is that we will find $b_0 = b_1$. If, however, there is a genetic nurture effect then we will find $b_0 < b_1$. b_0 is the within-family direct effect of polygenic scores on outcomes. b_1 measures the between family effect which incorporates both direct effects, and the indirect genetic nurture effects.[30]

The Selzam, et al., 2019 paper found that for IQ and educational attainment the within-family coefficient was 48-49% less than the between family coefficient. This is consistent with significant genetic nurture effects.[31]

[29] Wang et al., 2021, report that parental educational attainment polygenic scores, for a sample of parents born in Wisconsin, USA, in 1957, are substantial predictors of adopted child educational attainment. But they have no information on how much the adoption process matched child family background with adoptive parents characteristics.

[30] $\bar{g}$ however, will be only a partial estimate of the parental DNA that determines the family environment

[31] Young, A. I. et al. 2018 apply a variant of this procedure using just the sibling genotype.

However if we write out fully the Selzam et al. estimated expression it is

$$EA_c = \frac{b_0}{2}(g_{c_0} - g_{c_1}) + \frac{b_1}{2}(g_{c_0} + g_{c_1})$$

where g_{c_0} *and* g_{c_1} are the polygenic scores of each twin. The means that in the first expression the ratio of the signal to the noise is again reduced in a way it is not in the second expression. Thus the downwards bias in the case of b_0 will be greater than in the case of b_1.

The absence of convincing evidence for genetic nurture still leaves the issue of whether social interventions can be successful in changing outcomes for children in significant ways.

Missing Heritability

"Missing heritability" is a discrepancy that has consistently appeared between the heritability of traits estimated from twin studies or pedigree analysis, and the heritability explained by identified genetic variants through GWAS (genome-wide association studies). While twin studies suggest a strong genetic component to many traits, such as years of education attained, even with very large samples, GWAS typically identifies genetic variants accounting for only a fraction of expected heritability.

Some researchers, such Eric Turkheimer and Sasha Gusev, have concluded that larger sample sizes will not eliminate this heritability discrepancy. Other methods that estimate the maximum possible heritability at the molecular level suggest that for outcomes such as IQ or years of education, the GWAS studies will identify only fractions of the heritability found in twin and pedigree studies.

They thus conclude that twin and pedigree studies overestimate heredity, and underestimate common environment effects. Thus Gusev, in a substack piece entitled "No, Intelligence is not Like Height" argues that even though IQ is as heritable as height based on twin and pedigree studies, it can be shown to be much less heritable at the genome level.

"over the past five years it has become clear that IQ and educational attainment are *not* like height in fundamental and meaningful ways." (Gusev, 2024).

Why do twin and pedigree studies then spuriously identify greater heritability? Given the logic of these studies, which compare the correlations in outcomes between monozygotic and dizygotic twins, the cause has to be that monozygotic twins are treated more equally by parents and teacher than are dizygotic twins. But this conclusion runs into the same problems that we identified above with the idea of genetic nurture. Why in this case are adopted children not more like their adoptive parents, or their adopted siblings? Why, as we will see in the chapters below, do factors that do affect a child's environment such as family size, birth order, and parental death have minimal impacts on child outcomes?

Part II. EVIDENCE FOR DIRECT GENETIC INHERITANCE

The previous section showed that the patterns of correlation in social outcomes, in both England and in Denmark and Sweden, is as would be predicted from additive genetic determination, combined with a high degree of assortment in mating. But since this is only observational data, it does not demonstrate that genetic transmission plays the causal role in explaining social outcomes. It is possible to construct models of cultural transmission which would produce the same patterns – even though these would not be very plausible models.

In this section of the book I show that the patterns of inheritance we observe in England 1600-2026 are all in conformity with the predictions of genetic transmission, but violate what we would expect under social causation.

Men and women had very different social position for most of the past 400 years. Men specialized in market work, women in childcare. If there was important social causation of outcomes, then the characteristics of fathers and mothers would predict child outcomes with different strength. With genetic transmission mother and father characteristics should be equally predictive of child outcomes. Chapter 7 shows that mother and father characteristics are indeed equally predictive of child outcomes, even in eighteenth century England.

Second, family size is an important constraint on the amounts of attention and resources children can receive. For marriages before 1880 in England there was no attempt to control fertility, and family size was a random outcome. With social causation size matters. With genetic causation it does not. Chapter 8 shows that for all outcomes other than wealth, size plays at most a very minor role in determining child outcomes.

With just a genetic lottery determining outcomes, birth order would have no effect. With social causation birth order would have significant effects. First children get more parental attention, and face less competition for resources than later born. Chapter 9 shows birth order is largely unimportant for child outcomes, both now and in the nineteenth century.

In England social outcomes for people born in the North, Wales, and the Midlands are less good than for those born in London and the South. If place of birth matters to life chances, then this is evidence of social causation. Chapter 10, in contrast, shows that the difference in outcomes between regions was the product of generations of selective migration of higher status persons from North to South.

Higher status individuals live longer as adults than do those of lower status. This again is widely believed to be socially caused, where the work and home environment are the determinants of these differences. Chapter 11 shows that what predicts longevity is achieved social status, and not the status of the childhood family. Thus it is quite possible that longevity is correlated with social status because of genetic correlations between social abilities and health robustness.

If social causation is important then parents, grandparents, and other collateral relatives will all help shape social outcomes. Chapter 12 shows, however, that losing a parent while young has inconsequential costs in terms of social outcomes. Also sons are correlated in occupation and education as strongly to fathers they never met as to those they lived with for 18 years. Grandparents also predict child outcomes as well when dead at the time of child birth as when alive.

The randomness of family sizes for marriages before 1880 produced within wealthier families shocks to wealth caused by differences in the numbers of children claiming each inheritance. Chapter 13 shows that these wealth shocks were transitory. While they also appeared in the children of the original generation affected by a wealth shock, they had no effect on the grandchild generation. Though subject to short term shocks, in the long run wealth adapted to family social status, which we saw was largely genetically inherited.

Lastly in this section Chapter 14 shows that for continuous measures of social status, regression to the mean is at a constant rate all across the status distribution, as predicted by genetic inheritance. There are no wealth traps and no poverty traps.

Mother completely concealed beneath blanket, circa 1880

7. Gender Symmetry in Status Transmission

We see evidence above that most social outcomes were largely determined by additive genetic transmission. One important implication of additive genetic transmission is that mothers and fathers should contribute nearly equally to child outcomes. For the fraction of genetic material inherited from mothers versus fathers is close to fifty percent. The actual shares inherited from each parent are:

Daughters	-	mother 50%, father 50%
Sons	-	mother 51%, father 49%

The reason for the difference in sons is that the X chromosome is 3 times the size of the Y chromosome. Men inherit an X chromosome from their mother, and the Y from their father. Thus the greater maternal inheritance. In contrast women inherit one copy of the X chromosome from each parent. [32]

The near equality in genetic contribution of men and women implies a powerful gender symmetry in outcomes, that is independent of the relative social and educational status of men and women over time. Also mothers will always matter the same for child outcomes as fathers whatever the gender of the child, and whatever the social status of the family.

In contrast if social outcomes have mainly social causation, mothers will matter more of less than fathers in different times as the social role of men and women changed over time. Mothers may also matter differently for daughter compared to son outcomes. Mothers in high status families may play less role in child outcomes than those in low status families. As a consequence of these different patterns of inheritance, social mobility rates for men will potentially differ from those for women.

An argument for mothers always playing a more important role than fathers in child outcomes comes from the fact that in England, as in all societies, even today, mothers play a disproportionate role in child nurture. Surveys of time use 1990-2001 found mothers always spent at least twice as much time in child care than fathers, even in the most gender equal societies such as Norway. [33] This parental time differential was even greater in earlier decades.

This implies mother characteristics will have greater importance in predicting child outcomes than father. [34] Arleen Leibowitz, for example, concluded that "since mother's time expenditures on children exceed that of fathers by at least a factor of 4, we would expect the significance and size of the coefficient of mother's education to exceed that of father's education' (Leibowitz, 1974, S116).

---

[32] Mitochondrial DNA is inherited from the mother by both genders, but this constitutes only 16,569 base pairs out of 3.2 billion such pairs for all DNA.
[33] Guryan, Hurst, and Kearney, 2008, table 4. Dotti Sani and Treas, 2016.
[34] Note that Francis Galton in a 1904 paper remarked parenthetically on "A popular notion that ability is mainly transmitted through female lines." Galton, 1904, p. 355.

Though mothers spend more time with children, however, fathers in most earlier societies had disproportionate access to income, wealth, and professional qualifications and careers. In England, for example, up until 1882 husbands had control of women's property after marriage. Fathers also had sole legal authority over children. So fathers earlier likely play a more important role in funding child schooling and training, providing access to career opportunities, and in forming child aspirations and achievement. Conventional approaches to social mobility in England up until the 1980s thus regarded the father's social status as determining the social status of families, and hence child social outcomes (Goldthorpe, 1983, 1984).

There is even the possibility that the relative impact of fathers and mother could depend on the gender of the child, or on the social class of the parents. In England before WWII, for example, in upper class families childcare and education was largely delegated to servants and (for boys) to private residential schools such as the famed Eton, Rugby, and Harrow. The economic resources to hire those servants and pay school fees came mainly from fathers.

In lower class families, in contrast, mothers would play a much more direct role in the lives and education of their children. So perhaps fathers mattered more in upper class homes for sons, and mothers in the homes of the working-class sons.

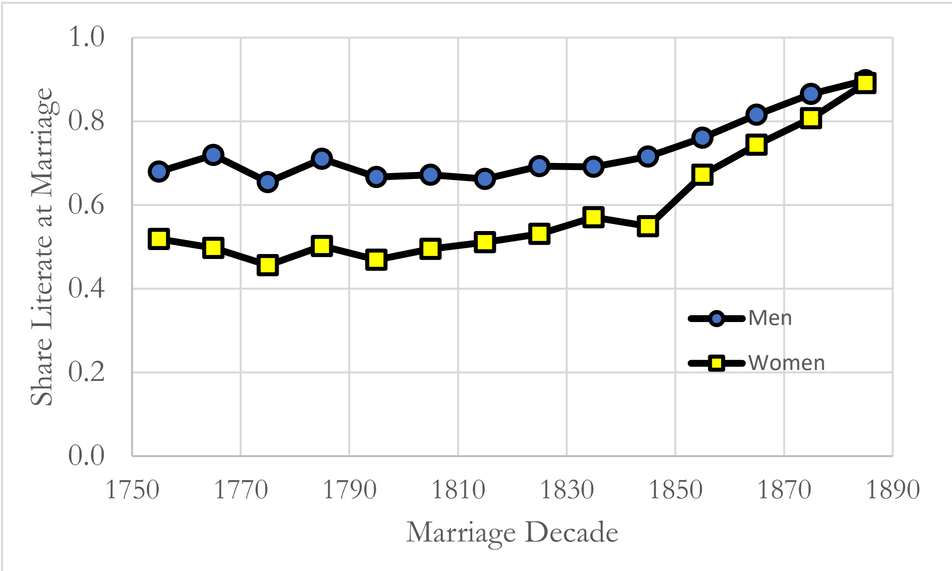
Thus evidence of an imbalance between fathers and mothers in determining child social status, and an imbalance that varied with time and with social class, will be good evidence of social determination of child outcomes. In contrast a constant gender symmetry in determining social outcomes will be powerful evidence for additive genetic determination of outcomes.

If we have equivalent measures of social outcomes for men and women then we could test whether there is indeed symmetry in the determination of status fir children just by measuring the correlations of status between fathers and mothers and sons and daughters. We shall see, however, that for England in the years 1754-1920 the only measure of educational status that is close to equivalent for men and women is literacy at marriage.

This is not a perfect measure of the relative contribution of men and women to child outcomes, however. As figure 7.1 shows, women were less literate than men for most of the period 1754-1889. So the signal of male and female literacy is not necessarily an equivalent signal of male and female educational status. Literacy may convey more information about female educational status than for male.

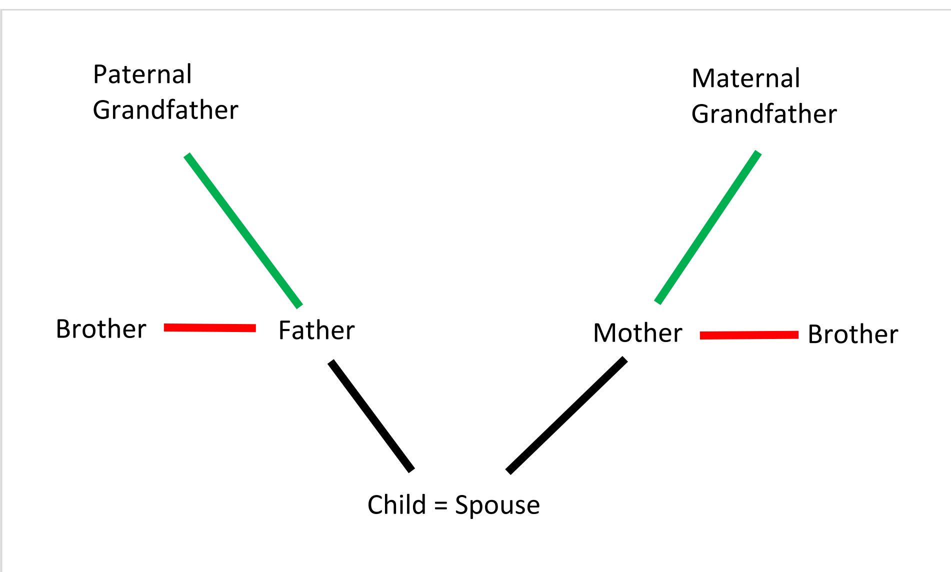
For women born before 1920 there is almost no other direct evidence on educational attainment and occupational status than literacy at marriage. Women do have wealth at death, but the transmission of wealth was much more significant between fathers and sons than to wives and daughters.

Figure 7.1: Literacy at Marriage, England, 1754-1889



Notes: The figure summarizes 24,131 observations of literacy at marriage 1754-1889 in average lineages. *Source:* FOE database, general lineages.

Figure 7.2: Determination of Child Outcomes: Paternal versus Maternal Line



However, in these other cases, we can use men and women's fathers' or brothers' characteristics as a proxy for their status. In that case we can test for gender symmetry using the structure portrayed in figure 7.2. We can test whether paternal grandfathers predict to the same degree as maternal grandfathers social outcomes for grandchildren. If there is complete symmetry in the inheritance of underlying social characteristics from fathers and mothers, then paternal and maternal grandfathers will have equal correlation with child outcomes.

An equivalent test for symmetry will be comparing the predictive powers of paternal versus maternal brothers on their nieces' and nephews' social outcomes. Again, with complete symmetry these uncles will be equally predictive. If transmission along the matriline is weaker than along the patriline this test will reveal that asymmetry.

The results below will show generally that mother's status is an equivalent predictor to father's status for child outcomes. But could it still be the case that father's status has a much more significant causal effect on child outcomes than does mother's status? Could mother's status just be providing more information on the true status of fathers, because of strong assortment in marriage, but the father's status be doing all the causal work?[35] Two considerations show this interpretation is not possible.

First, if mother status served just as additional information on underlying father status, mother's measured status would be more imperfectly correlated to true underlying father status than was measured father status. Assortment in marriage in England was strong, but not perfect. So on this interpretation, the link of mother status to child outcomes would be significantly less than that of father status.

If we simulate outcomes where the child status was determined only by underlying status of the father, but the mother has a status correlated with that underlying father status, then

[35] Again here this is not to imply that education itself creates child outcomes, as opposed to parental abilities which are correlated with education.

we do see clearly that in a joint estimation, father status is a much stronger predictor of child status than is mother.[36]

Second the parents of the fathers in the sample used below to estimate the relative effect of father and mother literacy on child outcomes themselves have literacy that was correlated with that of the father. The literacy of the paternal grandfather had a correlation of 0.34 with father literacy, for example, while their wife had a correlation of 0.43. But if we estimate jointly the connection between child literacy and father and grandfather literacy the coefficient on father literacy, 0.30, is three times as large as that on grandfather literacy, at 0.10. In contrast if we estimate child literacy on this same sample using both father and mother literacy, the coefficient on fathers is 0.18, and on mothers 0.25.[37]

Thus it is very clear that mothers do not predict child outcomes just because they provide more information on the underlying status of fathers. Instead they must play an important causal role in determining child outcomes. Their estimated effect can only be as large as for fathers if their causal influence on children is equivalent to that of fathers.

Previous Studies of Gender Asymmetries

Despite the expectations of greater influence by mothers, the existing empirical literature on the contemporary comparative effect of mothers versus fathers on child outcomes tends to find little difference in effects across parents.

Marks (2008), for example, looked at child test scores at age 15 in the PISA assessment of reading and mathematical performance in 2000, across 30 countries. What was the comparative strength of prediction of child scores from father versus mother educational

---

[36] In a simulation where we have noisy measures of underlying status for father, mother and child, but only a causal link between father underlying status and son underlying status, in a joint regression of child status on father and mother status, the coefficient on father status was nearly 3 times the size of that on mother status.

[37] The difference in these coefficients is not statistically significant.

attainment, measured by years of education? Averaged across the 30 countries, the predictive effect of mother and father education was of near equal magnitude for both reading and mathematics. In contrast, for occupational status the father effect was stronger than that of the mother in both reading and mathematics. But the difference was still less than 20% of the overall father effect (Marks, 2008, tables 1 and 2). And since some women leave outside employment in favor of childcare, the link between social abilities, education and occupational status will likely be noisier for women than for men.

Leibowitz (1974) examined years of education attained by the high IQ Terman sample of 1,528 children of 1921 as a function of mother and father years of education. For sons there was no significant difference in mother and father education as a predictor. For daughters the estimated mother coefficient was higher than for fathers, but the difference was not statistically significant (Leibowitz, 1974, tables 2-3).

Large scale studies by Plug (2004), Björklund, Lindahl and Plug (2004, 2006), (Björklund, Jäntti, and Solon, 2007, table 2), and Holmlund, Lindahl and Plug (2008), mainly using Swedish data, estimated the relative predictive value of both biological parents' years of education on child years of education. They generally found mother and father education to be equally predictive, independent of child gender.

In Björklund, Jäntti, and Solon, 2007, the relative predictive effects of mother and father education did not change when only the father, or only the mother, raised the child. Absent parents predicted child outcomes as well as present ones. Also in Björklund, Jäntti, and Solon, 2007, parental equality of affect was found for both sons and daughters.

The only exception was Holmlund, Lindahl and Plug (2008) where the predictive effect of father years of education in Sweden (0.152) was less than that of the mother (0.198).[38]

For adoptive children, the influence of the adopting parents on child years of education would be limited to social pathways, so based on time inputs by parents, the influence of

[38] Holmlund, Lindahl and Plug (2008), table 1 and table 3, summarizes these studies.

mothers in recent years should be relatively stronger. For Swedish children adopted by Swedish parents the predictive effect of the father's years of education across 3 studies was 0.07, and for the mother's 0.07.[39] For foreign children adopted by Swedish parents the mother and father predictive effects were both insignificantly different from 0 (0.014 versus 0.015).[40]

At least for education, then, we see clear indications in the literature of equal mother and father effects for societies where mother and father education levels tend to be relatively equal. Does this hold for the much more gender unequal societies of the past?

Because of the absence of occupations and education for most women before the last few generations, and the difficulty tracing the maternal line of descent because women change surnames at marriage, there is little evidence on the comparative social mobility of men and women in any earlier period. Paserman and Olivetti, 2015, and Paserman, Olivetti and Salisbury, 2018, estimate for the USA 1850-1940 social mobility of women using as a proxy linkage across generations first names. They find in the earlier USA less intergenerational persistence of female status than for men. In contrast Espín-Sánchez, Ferrie, and Vickers, 2023, using occupational income in the USA 1900-40 find that "The mother's contribution to mobility is almost five times larger than the father's"[41]

Estimating Mother versus Father Effects in a Patriarchical Society

The first problem we have to deal with for England before the modern era is that documentary sources reveal little about the educational attainment and occupational abilities of women. Married women in the censuses 1841-1921, and the population register of 1939, generally have no occupation recorded. Even single women of the upper classes are typically

[39] Holmlund, Lindahl and Plug (2008), table 1 and table 5.
[40] Holmlund, Lindahl and Plug (2008), table 5.
[41] Espín-Sánchez, Ferrie, and Vickers, 2023, p. 3.

listed with no status indicated, or under such terms as "gentlewoman" and "private means". Women were largely excluded from universities, and professional societies, before 1920.

As an example, consider Emma Wedgwood (1809-1896), who married Charles Darwin in 1839, and was mother of 10 children, three of whom became Fellows of the Royal Society. She appears in the census reports of 1851-1891, but only once with an occupational descriptor. In 1851 she is described as a "gentlewoman." Otherwise nothing is recorded of her educational attainments in school records, marriage records, or the censuses.

Yet from other sources we know she was a virtuoso pianist, and that she had gone on a grand tour of Europe in her youth (Healey, 2001). In the 1861 census while the Darwins' son George Howard, 15, was listed as a scholar, no occupation was listed for his sister Henrietta Emma, 17. Henrietta we also know from other sources to be highly educated. She later served as editor for a number of her father's books.

To give another example, Elizabeth Theresa Frances Kelsey (1875-1931) the wife of Samuel Courtauld, the industrialist, was an art collector, and founder of the Courtauld Art Institute in London. In the census of 1901, aged 26 just before her marriage, she is listed living with her widowed mother and with no occupation. In the 1911 census she again has no occupation, and is in a household with her husband, one daughter, and four servants, one of whom is a "child's maid."

Yet Elizabeth Courtauld was a well-known hostess, and a promoter of both modern art and modern music. Along with Malcolm Sargent, the famous conductor, she launched an innovative concert series in London 1927-31 that attracted notable international performers.[42]

The Darwin household records in the census also suggests that if child nurture and education are the crucial paths to success, fathers will be more important to child outcomes

[42] https://www.fondationlouisvuitton.fr/en/events/filippo-gorini

than mothers in richer households. The relative wealth of the Darwins is shown in the considerable numbers of servants who assisted in household tasks and childcare.

Thus in the 1851 census, the Darwin household had 4 children at home ages 1-5, but a total of 8 servants, including two nursery maids. By 1861 there were 7 children at home ages 9-21, and now 11 servants, including a governess and two nurses. If the resources to hire these servants and to pay for high quality education largely derived from the husband's wealth and occupation, then in upper class families husbands would much more strongly determine child outcomes than wives.

The laws of marriage in England and Wales, however, from 1754 on required both parties to sign the marriage register. So we know for a substantial number of parents the literacy at marriage of both father and mother from 1754 on, as indicated by signing ability.[43] This measure will be a good index of the educational status of lower class families, since by the eighteenth century all upper class men and women were literate. Since overall literacy measured at marriage had risen above 90% by 1890 in England, this measure is only useful, however, for marriages before then.

Using literacy we can thus directly measure the predictive content of father and mother education for child outcomes, where these outcomes are being observed in work or schooling ages 10-18, attainment of higher education (sons), literacy at marriage, occupational status around age 40 (sons). This is not a perfect measure of the relative contribution of men and women to child outcomes, however, since as figure 7.1 shows in most of the years of marriage 1754-1850 women were substantially less literate than men. So the signal of male and female literacy is not necessarily an equivalent signal of social status.

For women born before 1920 there is almost no other direct evidence on educational attainment and occupational status than literacy at marriage. Women do have wealth at death,

[43] Schofield, 1973, discusses the use of signature ability as a measure of literacy. He notes that in the years 1750-1850 reading was taught in schools before writing, so that the ability to sign was a good indicator of the ability to read text (p. 440).

but the transmission of wealth was much more significant between fathers and sons than to wives and daughters. For women observed 1999 and later we observe their house value and the social quality of their neighborhood – though this is determined for those married in conjunction with their spouses.

However, as noted, in these other cases, we can use men and women's fathers' or brothers' characteristics as a proxy for their status.

Literacy of Fathers and Mothers, 1754-1889

From 1754 onwards brides and grooms in England and Wales were required to sign the marriage register. Where they could not sign, they made a mark. Table 7.1 shows the data available on parent literacy by decade of marriage 1750-1880, and the outcome variables for children, again by decade of parental marriage. For child literacy the results are from marriages centered around 1820, while for children at work or in schooling aged 10-18 the results are from marriages centered around 1860.

We can thus test for children born in the interval 1754 to 1890 the relative influence of father literacy as opposed to mother literacy at marriage on a variety of child outcomes. Was there a significantly greater connection between child outcomes and their father's education as opposed to their mother's education?

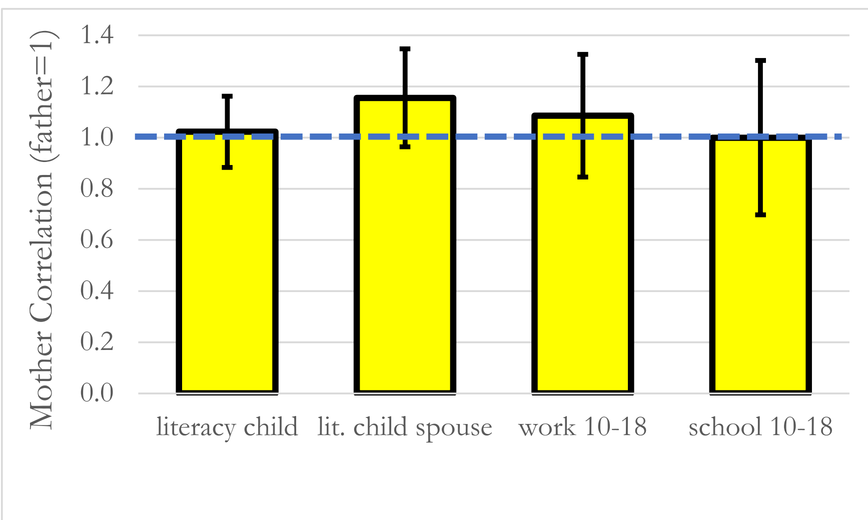
Figure 7.3 shows the estimated correlation between each parent's literacy and child outcomes for both genders combined. The outcomes are child literacy, child spouse's literacy, and children being at work or in education at ages 10-18. For each child outcome the correlations with both parent's literacy is strongly significant statistically. In all four cases the mother correlation is close to that for the father. And in none of the four cases is the mother correlation statistically significantly different from the father correlation. The overall impression is that the correlation of mother literacy with child outcomes is very similar to the correlation for father literacy.

Table 7.1: Summary Statistics on Literacy at Marriage

Marriage Decade	Parent Literacy	Child Literacy	Work/ School 10-18	Occupation Status (sons)	Higher Education (sons)
1750-99	2,474	1,461	0	1,035	1,069
1800-49	8,915	4,701	3,847	5,156	5,188
1850-89	12,742	874	6,725	5,617	5,626
All	24,131	7,036	10,572	11,808	11,883

Note: Observations are classified by the period of the parents' marriage.

Figure 7.3: Relative correlation of mother versus father literacy with child outcomes, sons and daughters, 1754-1889



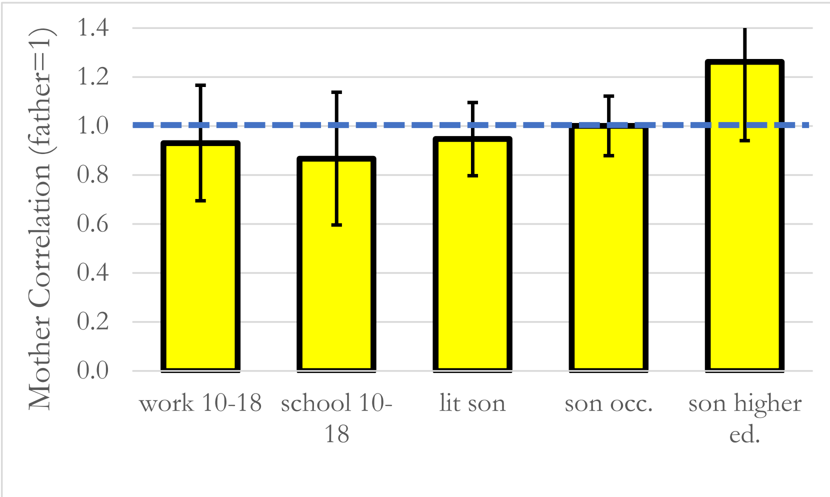
Notes: Error bars show the 5% confidence intervals of the mother correlation equaling the father.

Figures 7.4 and 7.5 show the relative correlations between father and mother literacy and child outcomes for sons and daughters separately. With social causation of child outcomes we could well see parents have stronger effects on children by the gender of the child. Father status could be more important for son outcomes, and mother status for daughter outcomes. For sons the outcomes are for work and schooling 10-18, literacy at marriage, occupational status at age 40, and attainment of higher education. For daughters we just have the outcomes work and schooling 10-18, and literacy at marriage. Unfortunately for daughters in figure 7.5 the error bands for the ratio of mother to father correlations are very wide, so that even large differences are not significant.

For the sons in figure 7.4, the average ratio of these parental correlations was 1.00. None showed significant deviations from one. For daughters the average mother correlation relative to the father correlation was 1.36. But the wide confidence intervals mean that none of these ratios of correlations differed significantly from 1.

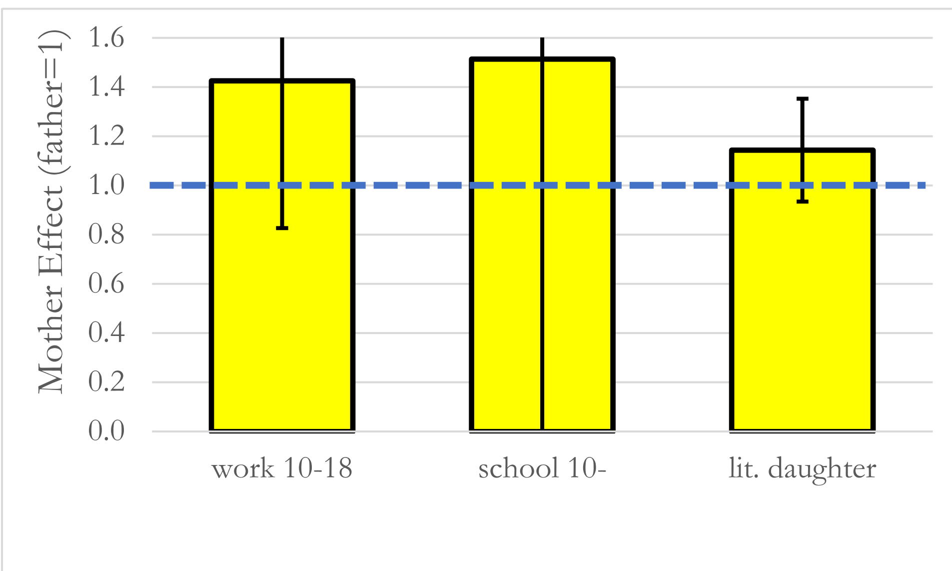
Based on the results reported in figures 7.3-7.5, women in Georgian and Victorian England, despite their low social status, had the same influence as that of men in predicting child outcomes. Despite the much greater exposure of children to mothers as opposed to fathers, there is no sign that mothers systematically had more effect than fathers.

Figure 7.4: Correlation mother v. father literacy with son outcomes



Notes: Error bars show the 5% confidence intervals relative to the mother correlation equaling the father.

Figure 7.5: Correlation mother v. father literacy and child daughter outcomes



Notes: Error bars show the 5% confidence intervals relative to the mother correlation equaling the father.

Grandfathers as a Proxy for Fathers and Mothers

As discussed above, we are missing measures of occupational status and higher education for women born before 1920. We do have measures of wealth at death, but until recently bequests to sons typically exceeded those to daughters. The best measure we have of grandfather social status, in terms of its correlation across generations or between brothers, is occupational status.

Table 7.2 shows the numbers of children where we know their social outcomes and the occupational status of their paternal and maternal grandfathers, by period of marriage of their parents. For house values and the Index of Multiple Deprivation we mainly observe the outcomes for marriages in the first half of the twentieth century. For the other outcomes we mainly observe outcomes for children born to marriages in the late nineteenth century.

Figure 7.6 shows the relative predictive effects of paternal grandfather versus maternal grandfather occupational status on grandchild social outcomes. For all the outcomes in the figure – house value and index of multiple deprivation 1999-2023, company director 2000-2024, at work or in school aged 10-18, occupation and higher education status (males) – paternal and maternal grandfathers have statistically indistinguishable predictive effects on grandchild outcomes.

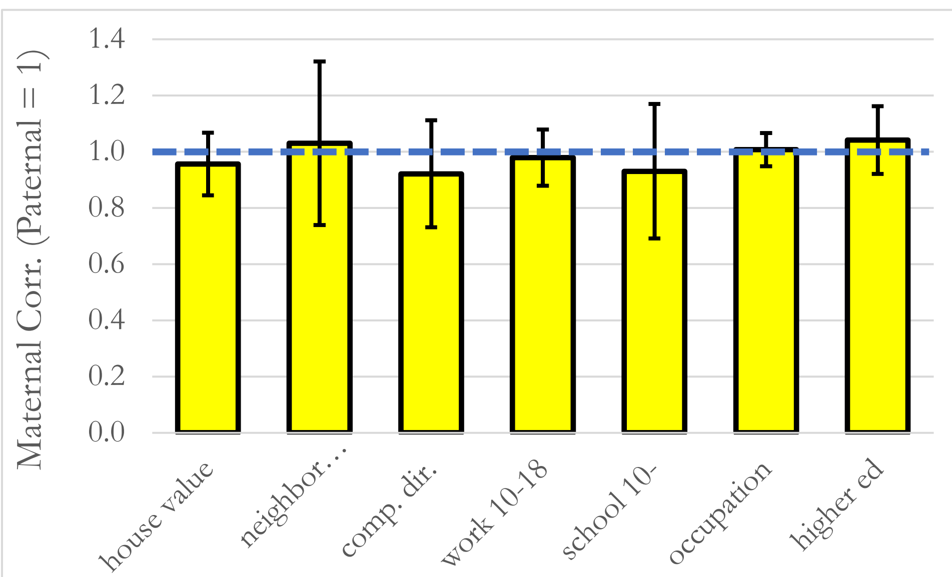
The outcome that is most precisely predicted is the occupational status of grandsons. Here we can be 95% confident that the maternal grandfather correlation of occupational status with grandson occupational status is within 5 percent of the paternal grandfather correlation.

Table 7.2: Observations on Both Grandfather/Grandchild Outcomes

Grandchild Birth Period	House Value/IMD	Occupation (male)	Higher Education (male)	Work/School 10-18
1800-49	0	497	497	454
1850-99	0	4,137	4,183	4,506
1900-49	3,611	1,956	2,226	1,344
1950-99	1,627	113	229	0
All	5,238	6,703	7,135	6,304

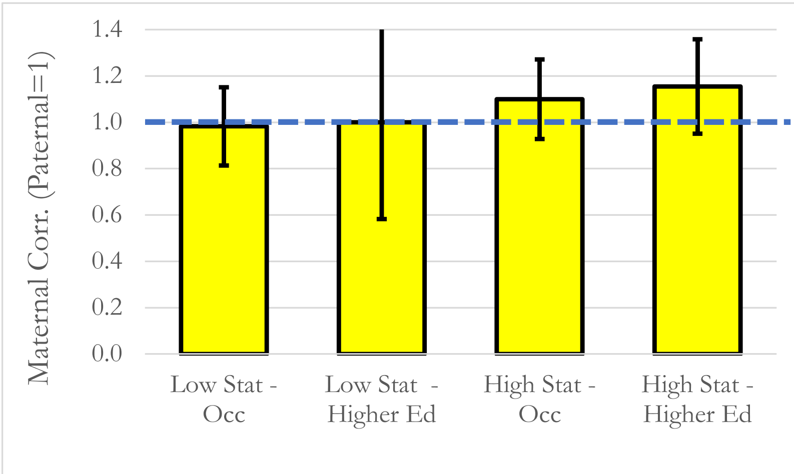
Source: FOE database.

Figure 7.6: Relative Correlation of Maternal Grandfather compared to Paternal Grandfather



Source: FOE database.

Figure 7.7: Relative correlation of maternal versus paternal grandparent occupational status, high versus low status families



Notes: Error bars show the 5% confidence intervals relative to the mother correlation equaling the father. *Source:* FOE database.

Figure 7.7 shows the predictive effect of paternal and maternal grandfather occupational status on grandson outcomes for grandsons born 1756-1919. In this figure, however, grandsons are divided between those from the rich versus the average or poor family lineages. Above we saw that, for upper class families, sons were largely raised by domestic servants and in private boarding schools. What should matter to their success is the funding provided by fathers to provide house space for servants, and the servants themselves, and enrollment in quality schooling.

In lower class families in contrast sons were raised mainly by their mothers, getting formal education mostly through public and charitable schools. But we see in figure 7.7 that for upper class families the occupational status of the maternal grandfather was, if anything, more predictive of grandson status than was the paternal grandfather. And for average and lower class family lineages the paternal and maternal grandfathers had equal weight in predicting grandson outcomes.

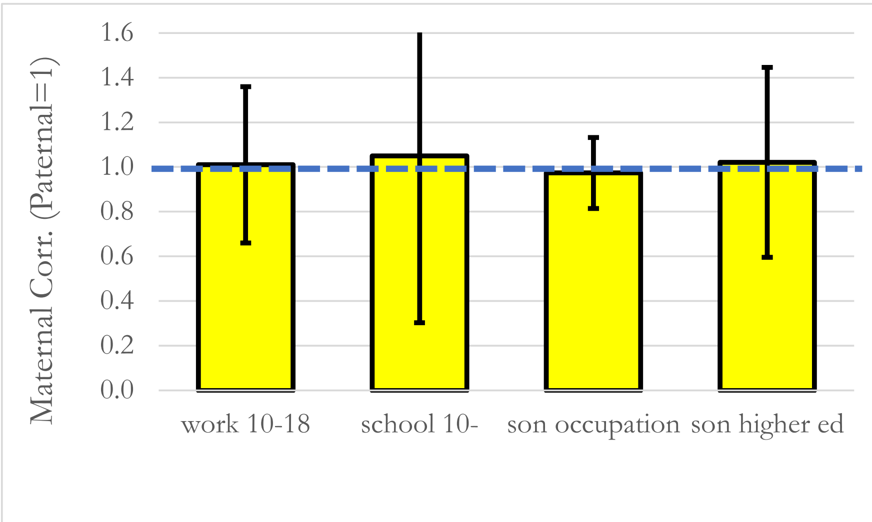
Uncles as Proxies for Parents

We can also proxy the effect of fathers and mothers on child outcomes by use of brothers as proxies for their social status. Figure 7.8 shows the estimated relative correlations of maternal versus paternal brothers, as proxies for the parents. The status measure for these brothers is occupational status. The outcomes here for the children are being at work or at school ages 10-18, attainment of higher education for sons, and occupational status of sons.

The figure shows that for all these outcomes the occupational status of paternal uncles predicts child outcomes in a symmetrical way with that of maternal uncles. Uncle status is always significantly linked with child status. But there is no statistically significant difference between paternal versus maternal uncles in predicting outcomes. The patriline and the matriline are again equally informative.

However, in this case because of smaller numbers of maternal brothers with occupational data compared to maternal fathers, the confidence intervals are relatively wide.

Figure 7.8: Correlations between child outcomes and paternal and maternal uncle occupational rank



Notes: Error bars show the 5% confidence intervals relative to the mother correlation equaling the father. *Source:* FOE database.

Conclusions

If social institutions and conventions matter to status transmission, this would suggest that social status will often be more strongly transmitted between generations on either the patriline or the matriline. The factors favoring stronger transmission on the matriline are the much greater involvement in all societies of mothers in the care and education of children. The greater time investment of mothers in childcare is found in all societies, even those such as in contemporary Nordic countries where gender equality is the most advanced. Thus we would on the human capital interpretation of social outcomes expect a greater maternal than paternal connection in the modern world. We do not see this.

However, a potential countervailing force in earlier times was the greater access of fathers to resources, and professional contacts. Also since in earlier years only fathers had occupations and educational qualifications, the father could be much more of a model for the outcomes of sons. It is thus uncertain whether the paternal or maternal line would better predict social outcomes in any earlier society. But we would expect the paternal effect to be greater in high status groups, and the maternal effect greater in average or lower class families. Again we do not see this.

We also show above that the equal strength of the matriline as the patriline in predicting child outcomes implies that matriline characteristics are playing an equal causal role to patriline characteristics in determining child outcomes. If maternal characteristics were correlated with child outcomes mainly because they were just another signal of underlying father characteristics, then the mother correlations with outcomes would be weaker than those for the father.

The results suggest, however, that the mechanism of transmission is largely independent of parental time interacting with children. The results reported above are thus consistent with the evidence of chapter 4 that the pattern of inheritance of most social outcomes in England 1600-2026 was consistent with direct additive genetic transmission. Such transmission would imply a symmetry of mother and father status as predictors of child outcomes.

Andrew and Elizabeth Clark and their 12 children.

8. Family Size

If direct genetic transmission of social outcomes dominates, family size should have little effect on child outcomes. In contrast social science disciplines - sociology, anthropology, education, demography, psychology, and economics - have all posited a strong negative tradeoff between family size and child social outcomes. In economics this is referred to as the "quality-quantity tradeoff."

Smaller family sizes in Europe after the demographic transition of the late nineteenth century supposedly allowed parents to invest more in their children, and so produce offspring with greater capabilities, spurring modern economic growth. This quality-quantity tradeoff has become a central doctrine of modern economic theorizing about economic growth.

Evidence, however, for a substantial child quality-quantity tradeoff is surprisingly limited. In modern populations family size is chosen, and parents who choose larger family size differ in unmeasurable ways from those who choose smaller sizes. Thus differences in child outcomes between small and large families may come from unobserved differences in parents. For modern populations measuring the effects of family size requires that we can observe random shocks to family size.

The literature on the quality-quantity tradeoff has thus relied on two accidents that produce unplanned increases in family size – twin births, and same gender births. However, since twin births are fewer than 1 in 100 births, and since same gender children for the first two or three births induces only a very small increase in fertility, these studies have generally proved inconclusive. The estimated effects are generally close to zero, but with substantial potential errors. This indeterminacy is discussed in the next section.

Here we use a period in English history already characterized by substantial human capital investment, but no fertility control within marriage, marriages 1780-1879, to measure the effects of family size on child outcomes. In these years even when we control for wife age at marriage, and for husband social status, there was enormous random variation in family size measured either as births, or as children attaining age 21. We can further show that parents were making no attempt to control family size, so that the variation was independent of unobserved differences in parental social abilities or human capital investments.

As we shall see family size had only small effects on child quality, for both prosperous and poor families. The only case with substantial effects was wealth at death for rich families with significant inherited wealth. In families with no parental wealth, wealth at death was unaffected by family size.

There is no significant quality-quantity tradeoff in child attainment even well into the era of modern growth. The inconsequential effects of family size on child outcomes are completely consistent with the dominance of direct genetic transmission in determining social outcomes.

Conventional Estimates of the Quality-Quantity Tradeoff

Modern high income societies have a combination of low fertility levels and high levels of nurture and education for children. Poor societies have high fertility levels, lower levels of nurture for children, and less education. The social sciences, by and large, have interpreted this as a causal association. All else equal, smaller families cause higher child quality.

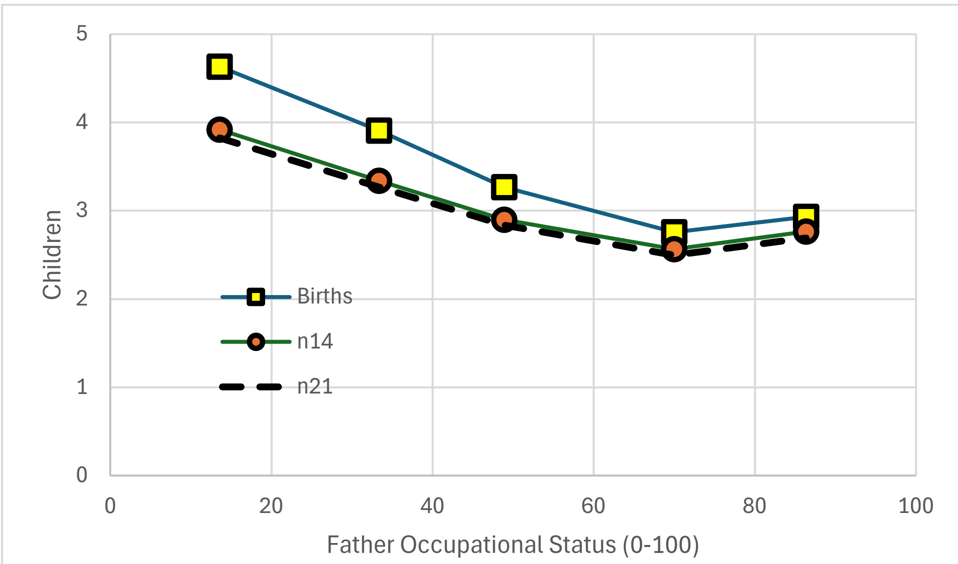
Such arguments can be found in Anthropology (Lawson and Mace, 2010), Education and Psychology (Downey, 2001), and in Sociology (Blake, 1981, 1985). The mechanism most commonly appealed to is "resource dilution." This interpretation is particularly strong in economics, where it underpins most theorizing about the sources of modern economic growth (Willis, 1973, Becker et al., 1990, Galor and Weil, 2000, Galor and Moav, 2002, Lucas, 2002, Galor, 2011).

Yet the empirical evidence that greater sibship sizes cause poorer child outcomes is surprisingly weak. In many modern societies there is a negative correlation between family size and measures of child quality.[44] Thus, for example, in England for men first marrying 1880-1929 we get the relationship between father occupational status and family size shown in figure 8.1. Given that children inherit social status, this implies that even where there is no family size effect, children from larger families will have poorer social outcomes. But

[44] See Grawe (2003) and Lawson and Mace (2010) for Britain, Rosenzweig and Wolpin (1980b) and Kaplan et al. (1995) for the USA, Rosenzweig and Wolpin (1980a) and Jensen (2005) for India, Lee (2008) for Korea, Grawe (2003) for Germany, Desai (1995) for 15 developing countries (using heights as a measure of child quality). But note that for pre-industrial societies, even the raw negative correlation between quality and quantity is either absent or weak. See Carmichael et al. (2015), Figure 2, p. 21.

given that in this era most people were choosing family size, and that those choices correlated with social status, we cannot infer anything about the correlation between size and social outcomes from observational data.

Figure 8.1: Father Occupational Status and Family Size, marriages 1880-1929



Source: FOE database.

Notes: The curves show average family size for fathers with at least one child measured as births, children reaching age 14 (n14), and children reaching age 21 (n21).

If we estimate son occupational status using these marriages 1880-1929, we will find that even controlling for father occupational status, there is still a negative association between family size and son occupational status. As long as there is a negative association between father status and family size, family size provides information about the true status of the father, and hence will be negatively associated with child social outcomes.[45]

[45] The remaining quality-quantity tradeoff, however, is modest. Once we control for father status going from 2 to 3 children reduces occupational status of child by 0.7 points on a 0-100 scale.

Researchers have thus sought to find cases where family size varies for reasons exogenous to the characteristics of the parents. For high income societies in the modern era one such exogenous shock to family size is the accident of a twin birth for the second or later parity. A second exogenous shock to family size, where there is a preference for having a child of both genders, is the accident of the first two children or first three children having the same gender.

These exogenous variation studies in general have produced no consensus on the size of the quality-quantity tradeoff, or even on whether it exists at all. There is indeed no convincing empirical evidence for this central tenet of economic orthodoxy.

A systematic problem these random shock estimates face is that the shocks produced in family size by the accident of twinning or same gender are very modest: typically a 0.1 to 0.5 increase in births.[46] Twinning produces the largest effects, but twinning occurs in less than 1 percent of deliveries. Also since twins are disadvantaged relative to singleton births, we can only measure the size effects induced by twinning by looking at children born before the twin birth. Thus the effects of this exogenous variation is measured very imprecisely, with an error typically 50-100 times that where all children and all size variation is used.[47]

A recent survey of the twin and same gender literature by Damian Clarke found that across ten estimates of the effect of an additional child on years of education, the average estimated effect was -0.10 years of education (Clarke, 2018). But this estimate is both a very small effect of family size, and also insignificantly different from zero.

Mogstad and Wiswall (2016), argue that twin studies fail to detect size effects because size effects are non-linear, being positive going from 1 to 2, and 2 to 3 children, but then becoming

[46] In India, for example, using female births for the first and second births as an exogenous source of larger families (because of son preferences) the first child being a girl increased completed family size by 0.22, while the second also being a girl increased family size by 0.45 (Kugler and Kumar, 2017).

[47] See, for example, Angrist, Lavy, and Schlosser, 2010, table ---.

negative for larger family sizes. For the sample of families examined here another advantage is that, with an average sibship size of 6.6, they are concentrated in the area where Mogstad and Wiswall find negative size effects are large in contemporary societies.

The twin estimates of family size effects thus depend on getting large data sets to be able to get sufficient power to discriminate effects. But they also are no longer possible in higher income societies where assisted fertility methods were introduced in the 1980s. Such methods both produce more twin births, and are associated with parent social status.

Another complication with the twin estimates is recent evidence that twining is actually more common with higher status mothers. This, it is argued, will mask the negative effects of larger family size in the twin estimates (Bhalotra and Clarke, 2019, 2020).

Testing for Family Size Effects, England, marriages 1780-1879

Three elements in marriage and fertility behavior in this period allow us to use the observed variation in family size to estimate the effects of size on social outcomes for children.

The first is that there was no attempt by parents to control fertility within marriage.

The second is that even after controls for parent age and social class, there was very substantial variation in numbers of children born. For wives aged 23 at marriage, where both parents attained age 50, we can observe in the English data marriages with from 0 to 9 children attaining age 21. Those who entered marriage could control very modestly the numbers of children they produced through the ages at marriage of the partners. This explained 12-15 percent of the variance in child numbers. After this family size was random.

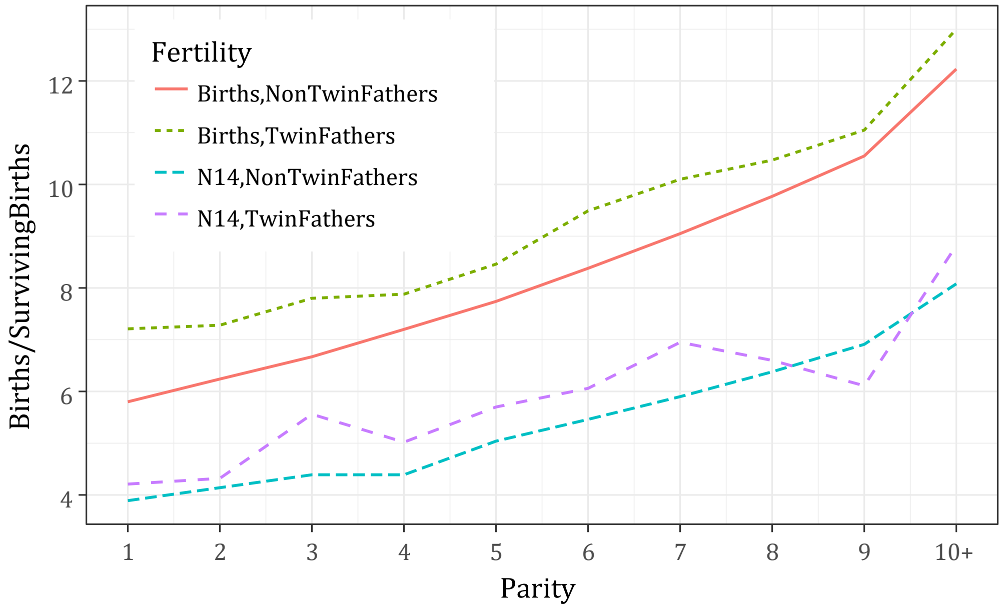
The third element for the period 1780-1879 that allows use of the observational data is that average net fertility, children reaching age 14 or 21, did not differ by the social class of parents.

Strong evidence of the absence of any fertility control within marriage in England, for marriages 1780-1879, comes from the accident of twin births. Consider a population where there is no control of fertility within marriage. In this case, whenever and however, the marriage terminates, the expected number of births will be increased by 1 by a twin birth, assuming the twin birth has no effect on the length of the subsequent birth interval. Also the increase in the final number of births will be the same whatever is parity at the time of the twin birth.

With control, and a target family size for each family, in contrast the increase in births will be less than 1, as families reduce subsequent births to achieve their target family size. Taking account of the higher infant mortality rate of twin as opposed to singleton births, with control a twin in England before 1879 on average would induce only 0.43 additional births (Clark et al. (2019)).

To estimate the effects of a twin birth on total births, we need to control for the parity at which the birth occurs, as well as the age of the mother at the birth, since both higher parity and mother age increase the frequency of twinning. From FOE for marriages 1780-1879 we have 48,877 births with complete family size known, 447 of which were twin deliveries.

Figure 8.2: Observed Births by Parity at Twin Birth, 1780-1879



Source: Clark, Cummins and Curtis, 2019.

Notes:

Figure 8.2 shows the estimated effect of a twin birth at each parity on total births per father for marriages 1750-1879. The average effect of a twin birth was 0.99 additional births (with a standard error of 0.12). Had fertility been controlled, with families having a target family size, the effect of twins on total births would have been an increase of only 0.43 births.[48]

Figure 8.2 also shows at each parity the estimated effect on the number of children reaching age 14 when a twin birth occurred. Overall there was also a substantial increase of 0.72 (standard error 0.14) in the total number of children surviving to age 14 when a twin birth occurred. The expected rise in children reaching age 14 if families controlling fertility would be 0.08.

[48] Clark et al. (2019) details the calculation here. That paper also shows that the frequency of twinning had no correlation with fecundity, or with social status, in England before 1879.

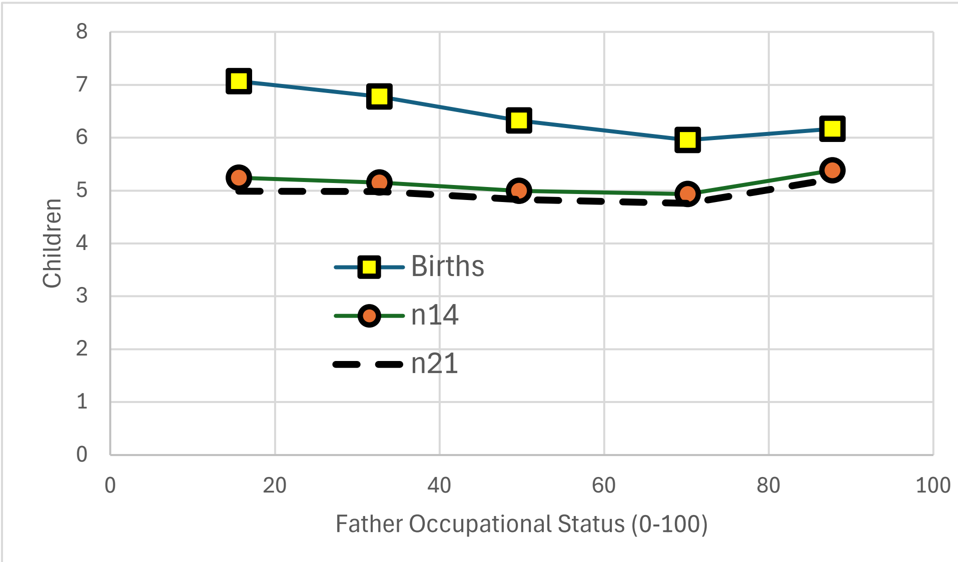
Another way we can test for deliberate fertility control is by looking at the effect of the gender composition of earlier births – how many male versus female – at parity n, on the probability of this being the last birth. Specifically in a society with still a strong preference for boys, if all births up to this parity were female did that reduce the probability that this would be the last birth? And if there was also some taste for mixed genders in children did a set of births all male also reduce the chance of births stopping at parity n?

If we go to England for marriages 1920-1994, where there was extensive fertility control, we do see clear signs that the gender composition of the first two births changed the probability of the family having a third birth. In this period 46 percent of marriages had 3 or more births. But if the first two children were girls the probability of a third birth increased by 5.8 percent, and this effect was highly significant statistically. If the first two children were boys, then the probability of a third birth increased by 3.5 percent, and the effect was again statistically significant.[49]

For marriages 1780-1879 there is clearly no effect on further births if all the births had been boys up to parity 3, 4 or 5. The estimated probability of having at least one more child at each of these parities, where all previous births were male, changed by -1.0 percent, +0.3 percent, and +1.1 percent. None of these effects were statistically significant.[50] For girls the effect is more ambiguous. The estimated probability of having at least one more child at each of these parities, where all previous births were female, changed by +2.7 percent, +1.9 percent, and -0.6 percent, but only one of these estimates was statistically significantly different from 0.

[49] These estimates are from 11,913 men in the FOE database who married 1920-94, and had only one marriage.
[50] The probability of at least one further birth at parity 3, 4, and 5 was 89.8 percent, 86.7 percent, and 83.1 percent.

Figure 8.3: Father Occupational Status and Family Size, marriages 1780-1879

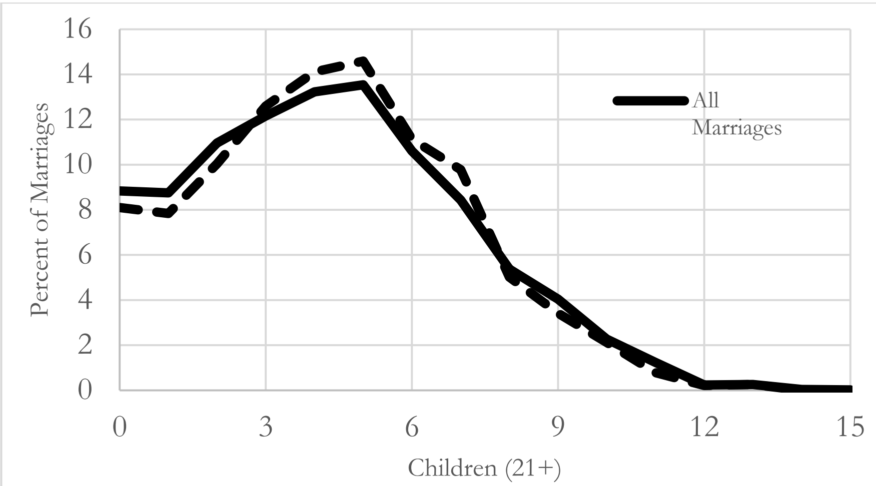


Source: FOE database.

Notes: The curves show average family size for fathers with at least one child measured as births, children reaching age 14 (n14), and children reaching age 21 (n21).

The second important feature of marriages in the years 1780-1869 was the absence of association between fertility and parent "quality". This absence is illustrated in figure 8.3, which shows for marriages 1780-1869 the association between father occupational status and numbers of children born, reaching age 14, and reaching age 21. There is a modest negative association between births and occupational status, but for the more important measures of numbers of children aged 14 and 21, because of higher infant and child death rates for lower status families, there is no connection with status.

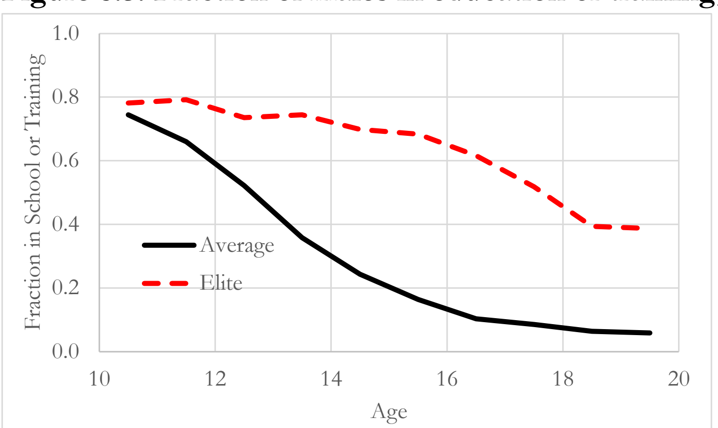
Figure 8.4: Family Size Distribution, marriages 1780-1879



Source: FOE database.

Notes: Family size as measured by children per father. The dashed line shows the distribution of family size for marriages where the bride was 21-25 at marriage, both husband and wife lived to at least 50, and there was only one marriage.

Figure 8.5: Fraction of males in education or training, ages 10-19, marriages pre-1880



Source: FOE database.

Notes: Status as reported in the censuses 1851-1891. This underestimates the fraction in schooling or training for both groups, since occupational status in the census is often left blank for those aged 10-19.

The third important character of fertility in these years was the vast range in family sizes under this natural fertility regime. Figure 8.4 shows the distribution of family sizes, at age 21, measured from the perspective of children. Average family size was 6.6 children, but family size varied between 1 and 18. Sibship sizes in nineteenth century England at the time of the Industrial Revolution were thus among the largest observed across all societies with well recorded demography. As a reflection of the huge variance in average family size in this period. 10 percent of children were in families of 2 or less, 13 percent of children in families of 9 or more. Even if we restrict ourselves to men and women married just once, who each survived to age 50, and where the wife was 21-25 at marriage, the range of family sizes is still great, as figure 8.4 shows.

The period we utilize here, marriages 1780-1869 is already one where there were considerable investments in education and training. Figure 8.5 shows the fraction of males in education or training in the censuses of 1851-1891 by age 10-19 for children from marriages before 1880. This number is a lower bound since frequently, even for rich families where we know education would be provided at younger ages, the census enumerators simply left blank the "occupation" column for children without a job or trade, without specifying explicitly "scholar."

The families are grouped into those from "average" surname lineages, and those from the elite surname lineages. For families from richer lineages at least 68 percent of sons were still in education or training at age 15. Even for the average families at least 16 percent of sons were still in education or training at age 15. The age we are looking at here is already the modern one of significant human capital investments.

In this chapter we use the families of 2,320 fathers from the FOE database who married 1780-1879, had one marriage, lived to at least age 50, had a wife aged 21-25 at marriage, the wife also living to age 50, and who had at least one child. In these 2,320 marriages there were 15,014 births, an average of 6.5 per marriage. The number of births ranged from 1 to 18.

We focus on this subset of marriages so that all the variation in family size was random – not caused by factors such as early parental death, or by men having multiple marriages, or women marrying young or old. These are children mostly growing in intact households until they reach age 14, cared for by their birth mothers and fathers.

Family size is measured in three different ways. First is the number of births per family (*N*0). But a child who dies immediately after birth, as would most of the children dying in childhood in this period, makes few claims on parental time and attention. So another measure is either children surviving to age 14 (N14) or to age 21 (*N*21). Adult children (N21) has the advantage of also being measured with the least amount of error of any of the family size measures.

For this sample of marriages, as table 8.1 shows, at least 96 percent of the variance in numbers of children is random – not explained by age of marriage, of family social status. Mother and father age still influence numbers of children, but this explains a tiny fraction of the overall variance.

What about potential positive correlation between numbers of children and unmeasured aspects of quality that offset a quality-quantity tradeoff? These unmeasured aspects, however, would have to be unmeasured aspects of quality that were not correlated with measured aspects such as occupational status or wealth. For we know the measured aspects of status are uncorrelated with fertility in this period. Characteristics such as thrift, industry, and low rates of time preference would also correlate with wealth at death, education, and occupational status. But these unmeasured aspects of quality in parents would have to then translate into measured aspects of quality in children. This possibility cannot be ruled out logically, but is empirically implausible.

Table 8.1: Children per Father, marriages 1780-1879, women marrying 21-25

Determinant	Births per family	Children (21+) per family
Occupation Rank of father	-0.006	0.001
	(0.005)	(0.004)
Age of Marriage (husband)	-0.034*	-0.016**
	(0.018)	(0.015)
Age of Marriage (wife)	-0.174**	-0.121**
	(0.057)	(0.047)
Number	1,615	1,615
Share of Variance Explained	0.04	0.03

Source: FOE database.

Notes: The table shows for fathers and mothers who both lived to at least age 50 and married only once the numbers of births and children (21+) they had as a function of father occupation, and parent ages at marriage, mothers aged 21-25 at marriage. *significant at 5 percent, **significant at 1 percent.

Family Size and Child Outcomes

The best measure we have of child social outcomes, as judged by the parent-child correlation, is the occupational status of sons. Figure 8.6 shows son occupational status, measured at age 40, as a function of family size (N21). There are 4,904 sons in this sample (with 1,732 fathers), who all came from marriages where the wife was 21-25 at marriage, both parents lived to 50 and above, and the father had only one marriage. Son status is estimated for family sizes 1-2, 3-4, 5-6, 7-8, 9-10, and 11+, controlling for father occupational status.

The figure shows estimated average status by family size for father occupational status of 40 (on a scale of 0-100).

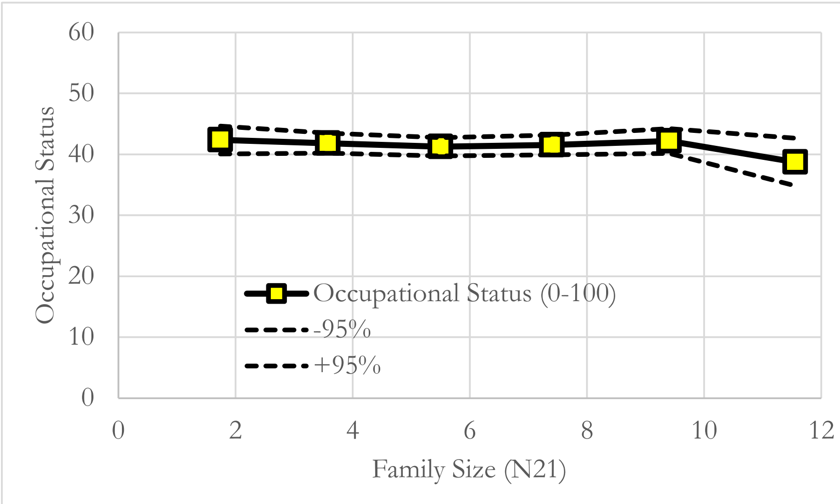
There is no sign of any significant size effect across this wide range of family sizes. In the range of sizes 0-10 there is clearly no size effect. For sizes 11-18 there is possibly some slight loss in son status, but the confidence interval here includes the estimates for the 1-10 range. Also if there is any decline in son outcomes 11+ it is quantitatively insignificant.

The same results hold if we measure family size by births or by numbers of children reaching age 14. If we estimate the effect of 1 additional child on occupational status, then the estimate for N21 is a decline of -0.001 on a scale of 0-100, or for births a decline of -0.002. Family size does not make any difference to son outcomes.

Mogstad and Wiswall (2016), argue that Black et al. (2005), fail to detect size effects on child outcomes because such effects are non-linear, with increases in size benefiting children going from 1 to 2 children, and hurting them only at larger family sizes. Figure 8.6 also clearly rejects any suggestion of non-linear size effects. Across family sizes 3-10 there is no hint of any negative effects of size.

A second outcome we have measures for is literacy for children, measured at marriage by their ability to sign the register. We take this measure just for child marriages 1780-1889, since in later decades almost everyone is literate. Since we also get a measure of their spouses' literacy, we take the average of child and spouse literacy as a better measure of child educational attainment. We have this measure of literacy for 1,007 children (381 fathers).

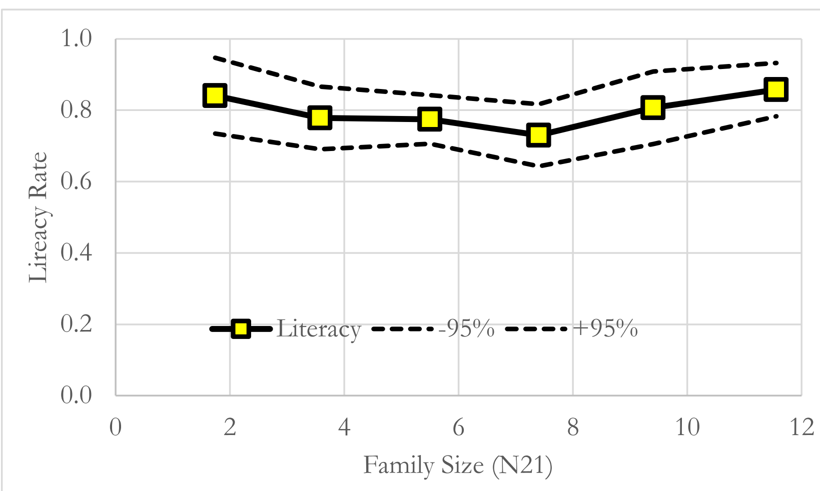
Figure 8.6: Family Size and Son Occupational Status, marriages 1780-1879



Source: FOE database.

Notes: Dashed lines show the 95 percent confidence interval of the estimates.

Figure 8.7: Family Size and Child Literacy, marriages 1780-1879



Source: FOE database.

Notes: Dashed lines show the 95 percent confidence interval of the estimates.

Figure 8.7 shows the estimated average literacy by family size, controlling for mother and father literacy.[51] With this measure because of less data, and the outcome being just a 0-1 variable, the estimates are noisier than for occupational status. But they very clearly imply no decline in literacy for children with family size. Indeed if we estimate the effect of one additional child on literacy rates then the estimate is a decline of 0.2 percent, which is inconsequential relative to the near average 80 percent literacy rate.

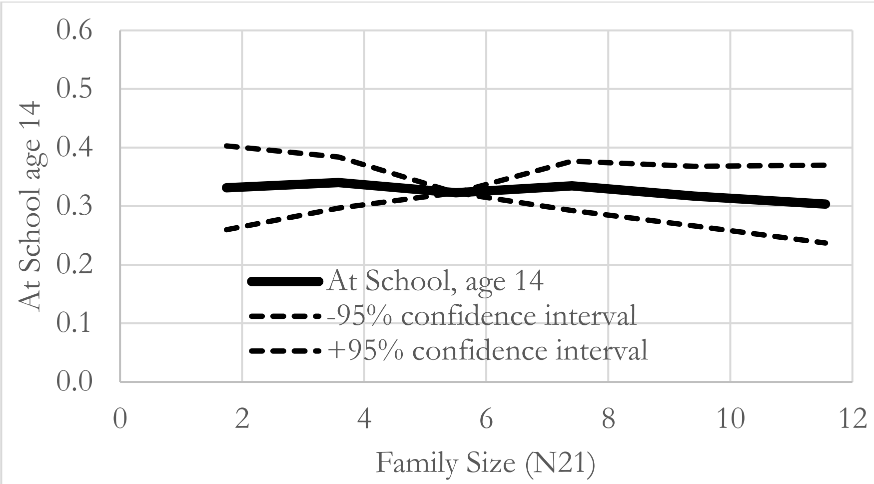
We also have evidence from the census reports of 1851-1911 of whether children ages 10-18 were in education or training. In the period in England there was no compulsory schooling age until 1880. From 1880-93 the school leaving age was 10, 1893-1899 11, and 1899-1918 12. Only in 1891 did primary education become free for all students at state schools. Thus 85 percent of our sample of children reached age 11 in the period where there were still school fees, and where education after age ten was voluntary. This was a period with costs for education, and where beyond age ten education was entirely at parental discretion. For our sample families, 48 percent of children 10-18 were in school or in training.

Figure 8.8 shows the estimated probability a child will be in schooling age 14 as a function of children surviving to age 21 in each family, with controls for the father occupation, the census year the child was observed in, and the child age. The confidence interval for other age groups is shown relative to ages 5-6. None of the other size groups show any significant difference in schooling probabilities from the size group 5-6.

These results also imply that declining family sizes on their own will explain nothing of the secular increase in years of education per child since the Industrial Revolution.

[51]This estimation uses OLS. Since the dependent variable has a value of 0 or 1, statistically the correct estimation would be a logit. But the OLS results are always similar and much easier to interpret than the corresponding logistic estimates.

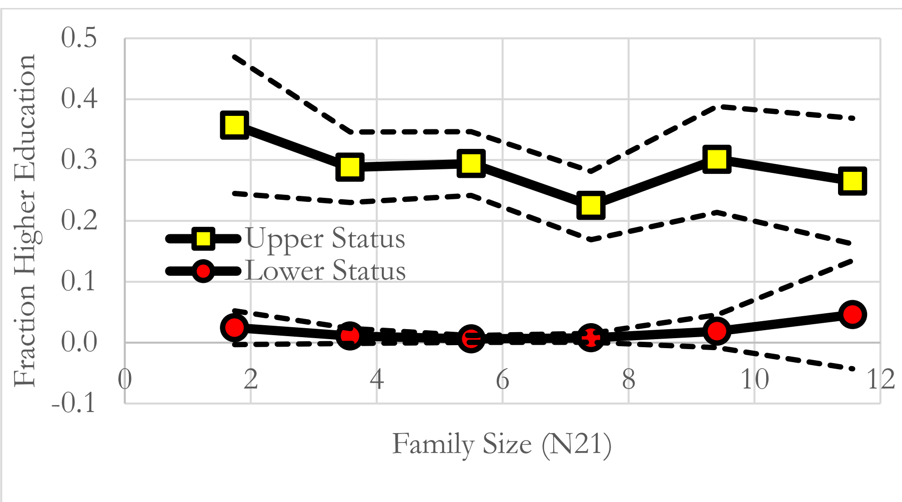
Figure 8.8: Family Size and Schooling, England, marriages 1780-1879



Source: FOE database.

Notes: Dashed lines show the 95 percent confidence interval of the estimates.

Figure 8.9: Family Size and Higher Education, England, marriages 1780-1879



Source: FOE database.

Notes: Families grouped by sizes 1-2, 3-4, 5-6, 7-8, 9-10, and 11+. 95 percent Confidence Intervals indicated by dashed lines.

Another measure of educational attainment is whether a son attained some form of higher education: attending university, ordination in the Church of England, military academy, medical or legal training, or membership in an engineering society. The likelihood of such qualification was heavily connected with father occupational status. For fathers in the upper half of the status ranking 28 percent of sons reaching age 21 so qualified, while for those in the lower half, 1 percent of sons. The upper half of the occupational status distribution included in the lower reaches such occupations as farmers, commercial clerks, pawnbrokers, postmasters, and salesmen.

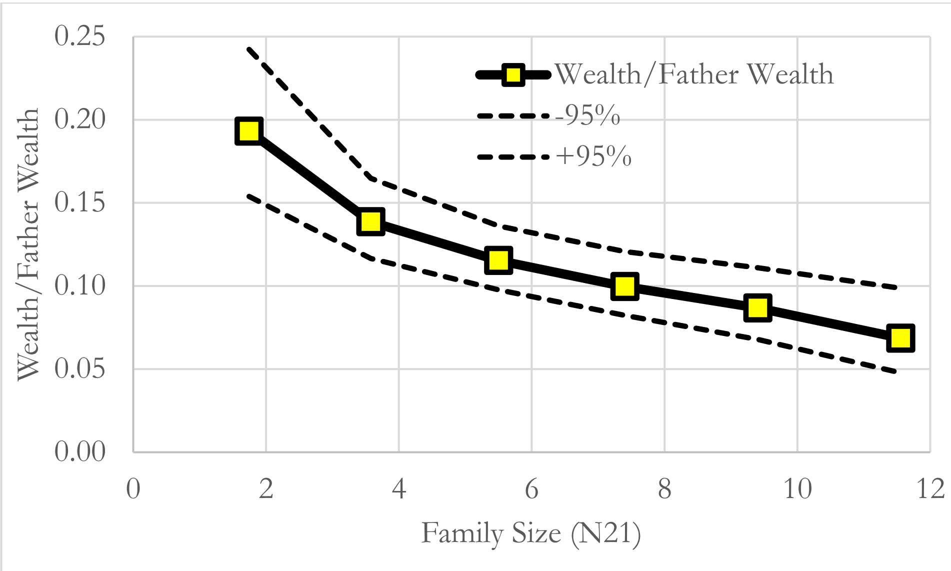
Family Size and Child Wealth

We have estimates of wealth at death for all people dying in England 1858 and later. This comes from the Principal Probate Registry, and is from 1858-1893 a statement just of the personalty of the deceased (assets aside from real estate), and after 1894 a statement of all assets. For those not probated we attribute a probate value which is half the estate value requiring probate.[52]

In the years 1799-1857 there is more selective information on the value of personalty available mainly for wills probated in the highest of the ecclesiastical probate courts, the Prerogative Courts of the Archbishop of Canterbury and of York. However, only about 5 percent of men were probated in these courts, and wealthy men might be probated elsewhere.

[52] In each period after 1858 there was a minimum estate value at which probate was legally required: £10 (1858-1900), £50 (1901-1930), £50-500 (1931-1965), £500 (1965-1974), £1,500 (1975-1983), and £5,000 (1984-2024) (Turner (2010), 628). We thus took as the value of estate for those not probated as typically half the minimum requiring probate: £5 (1858-1900), £10 (1901-9), £15 (1910-9), £20 (1920-30), £25 (1931-9), £50 (1940-9), £100 (1950-9), £250 (1960-1974), £750 (1975-1983), and £2,500 (1984-2024).

Figure 8.10: Family Size and Child Wealth, Richer Families



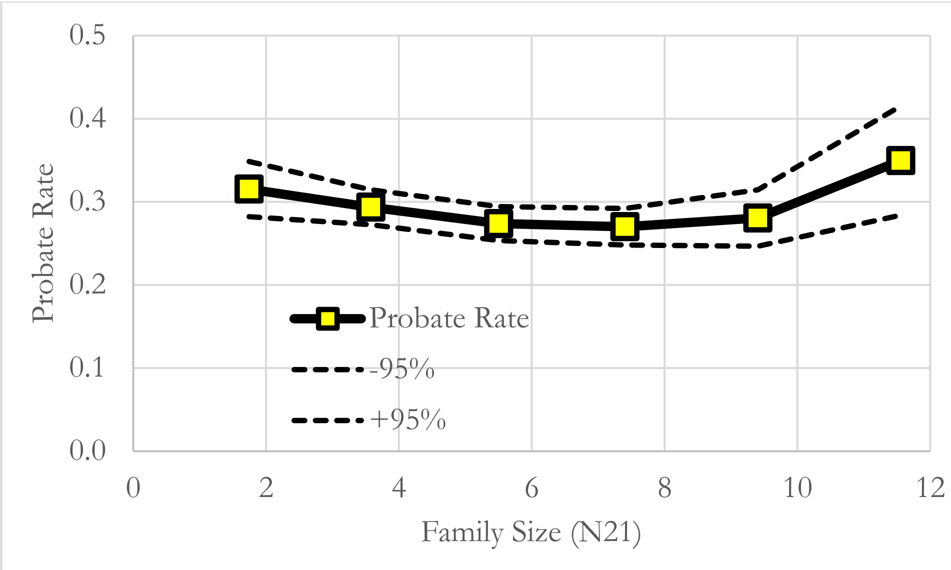
Source: FOE database.

Notes: Families grouped by sizes 1-2, 3-4, 5-6, 7-8, 9-10, and 11+. 95 percent Confidence Intervals indicated by dashed lines. The figure is based on 7,933 adult children of 2,103 fathers.

Since the nominal value of average wealth increased greatly between 1858 and 2012 we normalized by the estimated average wealth at death in each decade. Also since wealth at death has a very skewed distribution, we use the logarithm of normalized estimated wealth to produce a distribution closer to normal.

Figure 8.10 shows average child wealth as a fraction of father wealth for families of different sizes as measured by the numbers of children reaching age 21. Only fathers with above average wealth were included. As can be seem child wealth at death declines substantially relative to father's. But also this decline is greater the larger is the family size. For wealth there is a strong negative size effect. Relative to their fathers children in families of 9 or more children have wealth at death that is less than half that of children in families of 1-2 children.

Figure 8.11: Family Size and Child Probate Rate, Parents with no Wealth



Source: FOE database.

Notes: Families grouped by sizes 1-2, 3-4, 5-6, 7-8, 9-10, and 11+. 95 percent Confidence Intervals indicated by dashed lines. The figure is based on 11,651 adult children of 3,832 un-probated fathers.

In families where the parents were not probated there would be no wealth to be inherited by children. Wealth these children owned at death would have been accumulated in their lifetimes, and would reflect their education and occupational abilities. Are children from larger families disadvantaged in wealth accumulation? Figure 8.11 shows the proportion of children from families of different sizes who accumulated enough wealth in their lifetimes to be probated at death. The figure suggests their might actually be an advantage in wealth accumulation for children from larger families. But if we estimate just the probate rate as a function of number of adult children the estimated coefficient is 0.000, with a standard error of 0.002. Overall family size has no effect on wealth accumulation.

Why is the Quality-Quantity Tradeoff Weak or Even Non-existent?

Why does increased competition for resources and parental attention with larger family sizes fail to lead to poorer outcomes for children, except with the notable exception of wealth in richer families?

For richer families the absence of any substantial quality-quantity tradeoff in child outcomes is not surprising. In nineteenth century England the children in such families were raised and educated at home with the aid of nursemaids and governesses. The annual salary of a governess at £35-50 was modest relative to the annual income of such families. For married men dying in the wealthier lineages before 1914, for example, assets at death averaged more than £55,000.

When older, sons often went to private boarding schools such as Eton, Harrow, Rugby and Winchester. The annual fee for a private boarding school circa 1861 would be £120-£250 per year, and day schools would be considerably cheaper (Turner (2015) p.249). Tuition at Oxford in the 1840s was only £16-£25, though many students paid £40-£50 in addition for private tutoring, again modest relative to the resources of even middle class families (Curthoys (1997) p.50-151).

The absence of significant quality-quantity effects for the average and poorer families is more surprising. In larger families parents would have less time to interact with young children, with more children to care for. More children also implied less ability of women to contribute through market work to family incomes, and less food consumption, space and clothing for other family members.

The costs of formal schooling, even before the era of free provision of compulsory education after 1891, were mainly the foregone wages of children. In the mid nineteenth century the actual weekly costs of school attendance were typically 6d or 12d a week. By age 13 or 14 a son or daughter could expect to earn 36-72d a week at work, depending on the nature of the local industries (Humphries (2010) p.230-33, 316-7).

Apprenticeships, typically entered into around age 14, would provide maintenance for children. But these usually required premiums, though in lieu of that payment the length of apprenticeship could be increased, and the earnings above maintenance in later years reduced.

Why didn't these costs of schooling and training force families with larger numbers of children to set their children to unskilled labor at a young age, limiting their future occupational status and earnings?

We do see, from individual examples in the FOE database, that even poor families with large numbers of children could, if they wished, afford to keep them in schooling.
Alfred Albert and Eliza Wimbleton, for example, who produced 9 children between 1866 and 1888, where Alfred was a general laborer, and Eliza had no occupation. In 1881 6 of those children were at home with the parents, aged 15 to 1. The elder four, respectively 15, 13, 11, and 10, were all still at school.

Large families, even at the bottom of the earnings ladder, could afford to keep children in school. In contrast Charles William Wressell, son of an agricultural laborer, in 1861 was at work age 12, also as a laborer, despite his parents having only one other dependent. Thus in nineteenth century England the choice whether to educate children or not, was driven more by parental attitudes, and the abilities of children, rather than being dictated by any binding budget constraint.

One reason that there was no binding budget constraint on children's care, nutrition and education, even at family sizes larger than 10, was the typical spacing of surviving children at birth intervals of 2 or more years. The modal age of marriage for women was 23, and of men 25, in nineteenth century England. Before marriage older children would often contribute financially, and with child care, to their families. Thus the younger children in large families, those occurring at birth orders 7 and above, often got transfers of care and resources not just from their parents, but from siblings. However, as we see, these transfers from older siblings did not have much measurable impact on outcomes such as occupational status or chances of attaining higher education.

Conclusions

The results above are clear. In England for marriages 1780-1879 parents made no attempt to target fertility, as witnessed by their lack of any behavioral response to twin births or gender imbalances. There was great variance in both births and numbers of surviving children, but uncorrelated with any aspect of observed social status. These shocks resulted in a great range of family sizes, from 1 to 18 for children surviving to age 21, even when we confine ourselves to marriages where the bride was 21-25, and both parents survived beyond age 50. Reflecting the randomness of fertility there was not even a significant correlation of marital fertility between fathers and sons, while sons correlated strongly with fathers on wealth, occupational status, and education.

The consequent random shocks in sibship sizes had significant effects on child outcomes only in the case of wealth among families with significant wealth inheritance. Otherwise for occupational status, education, and wealth accumulation, family size did not matter.

As noted above, estimates from random shocks to family size from the accident of twin births or gender imbalances using modern data fail also fail to find significant family size effects. Thus the economic dogma of the quality-quantity tradeoff is a dogma in the worst sense of the word: a belief unsupported by either logic or empirical demonstration.

Adam Henry Pearson Maclay, National Library of New Zealand

9. Birth Order

With genetic transmission of abilities, birth order should not affect social outcomes. If birth order matters in significant ways for social outcomes such as occupational status or educational attainment, then it shows the importance of social causation of child outcomes.

There has been a long debate in psychology, and also in economics, initiated by Francis Galton in 1874, about whether birth order significantly affects psychological and social outcomes. Galton famously claimed that 48 percent of famous English scientists were the first-born son in their family, while by chance that number should be no more than 33 percent.[53]

There are clear signs that as birth order increases, parental attention per child diminishes. Buckles and Kolka, 2014, for example, show within families that mothers are more likely to seek prenatal care, take prenatal vitamins, and breast feed for children earlier in the birth order.

[53] Galton, 1874. Assuming average pre-industrial family size was 6, the eldest son would on average be one of 3 sons.

Price, 2008, showed for the USA 2003-5, using time-use studies, that within families, children later in the birth order got less parental time than those earlier.[54]

There is ample evidence in the Families of England genealogies of birth order and social outcomes both for people still alive, and for those born in the years 1680-1929. Table 9.1 summarizes the available data. In earlier years because of parental mortality, and later because of divorce, we observe significant numbers of families where later siblings are only half siblings to those earlier in the birth order. In the figures and tables below I thus concentrate on families where fathers and mothers married only once, and all children are full siblings.

There is a large literature in psychology, sociology and economics estimating birth order effects, which is discussed below. The most reliable estimates should be those using information from within families. Each child's outcome is measured relative to the outcomes of their siblings. But large samples of such within-family data is not generally available, except for Nordic countries with comprehensive population registers. Even with within-family studies, the estimate needs to control for gender, potential time trends in the child outcomes, and for parental age at birth. Lower order children will necessarily have older parents, and older parents may generate better or worse outcomes for children.

An alternative is to use cross-family data. But then we need controls for family size, as well as for birth year and parent ages. One transparent way to do this would be to compare outcomes for all birth orders for children from families of size 2, 3, 4 and so on. This, however, will necessarily mean that lower order children are those with older parents. But it will reveal, on net, whether there appear to be significant birth order effects.

Here we first estimate simple net birth order effects for current families of sizes 2-4 using cross-family data. This suggests that unless there is an exactly offsetting positive parental age effect, there are no significant birth order effects. I then consider whether more sophisticated controls for parental age and birth cohort will find birth order effects.

[54] Price, 2008, table 2. Notably the Price, 2008, data suggests that for a given birth order, children in larger families get more parental time. This presumably driven by differences between parents having small versus large families. But that would suggest in the modern era that we might even find a positive family size effect.

Table 9.1: Data Available, FOE Genealogy, Birth Order and Family Size

Outcome	Birth Years	Gender	Number	Average Birth Year
House Value 1999-2026	1910-1999	Both	21,291	1957
Index of Multiple Deprivation	1910-1999	Both	21,291	1957
Company Director	1910-1999	Both	27,271	1958
Survival to Age 14	1910-1999	Both	50,395	1948
Occupational Status	1680-1929	Male	33,257	1865
Higher Education	1680-1929	Male	33,526	1865
Oxbridge Attendance	1680-1929	Male	39,430	1865
Literacy at Marriage	1722-1874	Both	7,730	1828
At School, Work 10-18	1834-1929	Both	27,391	1874
Wealth at Death (21+)	1800-1929	Both	43,244	1874
Longevity (aged 21+)	1680-1929	Both	63,605	1864

Notes: Children from intact single parent families with both parents surviving to age 50.
Source: FOE database.

Contemporary Data

For recent years in the FOE genealogy we have the address of individuals 1999-2026 from the Electoral Register. We include individuals born 1910-1999 where we know their birth order and sibship size, where they were aged 24+ at the time of observation in the Electoral Register, and where they lived at a separate address from their parents.

From the land registry, which records sales, we can then infer their residence value from their address. Since houses sell only periodically we estimate value at the level of the postcode, typically 20-30 dwellings. We also use the log of house value here because the measure itself is highly rightward skewed, which taking the log corrects. It is clear that the children of larger families lived in houses of lower value.

We also know for their local area, typically with about 1,000 inhabitants, the index of multiple deprivation (IMD), a measure of the average social status of the local neighborhood.[55] This status is derived by averaging multiple measures of the quality of life, such as education, crime, employment, and health. The index for England is given as an ordinal ranking of 32,844 areas in terms of most to least deprived. This we have converted into a 0-100 scale, where higher on this scale is less deprived.[56]

Table 9.2 shows the data available for these measures by birth order and sibship size, for birth orders and sibships 1-6+. Birth order and family sizes greater than 6 are set to 6. Figure 9.1 shows the underlying trends in log house value and the index of multiple deprivation by birth year. In both cases there is a modest decline in house values and IMD as births go from 1955 to 1999, but little trend for births before the 1950s. Thus birth cohort effects will be small.

The simplest way to look for birth order effects in the contemporary data is to group children by family size and birth order, and then just take a simple average of their outcomes, their birth years, and their parents' average age. This is done in table 9.3 for house value and for the index of multiple deprivation, for families of sizes 2-4. For families of size 5-6, as table 9.2 shows, the data for each family size and birth order becomes limited. This makes the outcomes by birth order noisier for such sized families.

Table 9.3 suggests no significant effect of birth order on house value, or on neighborhood quality. This can be seen in figures 9.2 and 9.3. On average house values do not decline with birth order, nor does neighborhood quality. Table 9.3 also shows the average year of birth of each child type. Children later in the birth order come from slightly later birth cohorts. But figure 9.1 shows that birth cohort effects are very modest, and for the differences in birth years observed, at maximum two years, of little consequence. But we do see in table 9.3 that parents are 3.5 years older for every step in the birth order.

---

[55] https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019

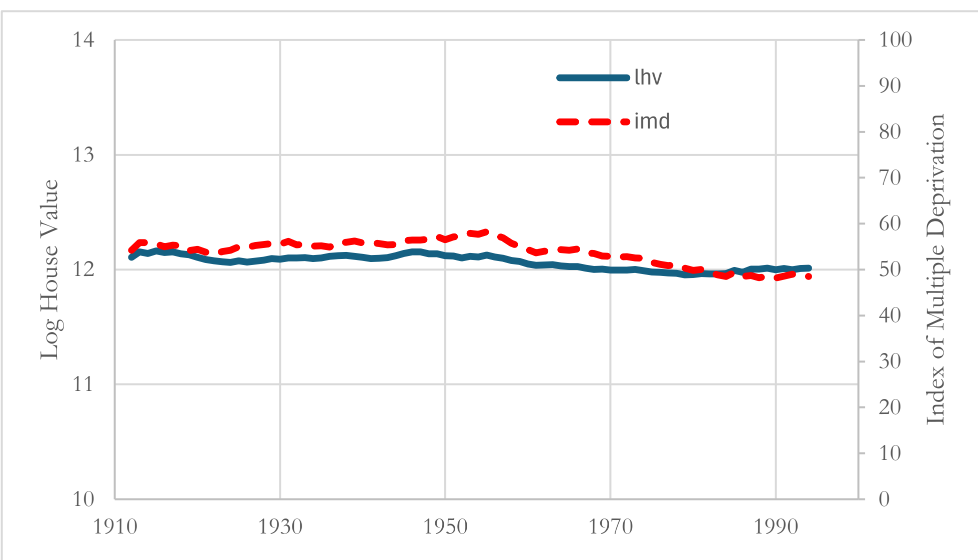
[56] There are separate rankings for Wales and Scotland. We convert those into an English equivalent ranking by assuming the Scottish and English means are the same, and that the Welsh mean is the same as the North of England mean.

Table 9.2: House Value and IMD by birth order and family size, 1910-1999

Size	1	2	3	4	5	6+
1	2,409					
2	3,971	3,942				
3	2.024	2,058	2,125			
4	843	879	928	963		
5	338	345	381	402	437	
6+	285	335	346	388	411	1,305
All	9,870	7,559	3,780	1,753	848	1,305

Notes: Family size is on the vertical axis, birth order on the horizontal. Birth order and family sizes 7 and above are in the 6+ category. Individuals were included only if aged 24+ when observed in the electoral register, and not living with their parents.
Source: FOE database.

Figure 9.1: Time Trends by Birth Year in ln(house value) and IMD



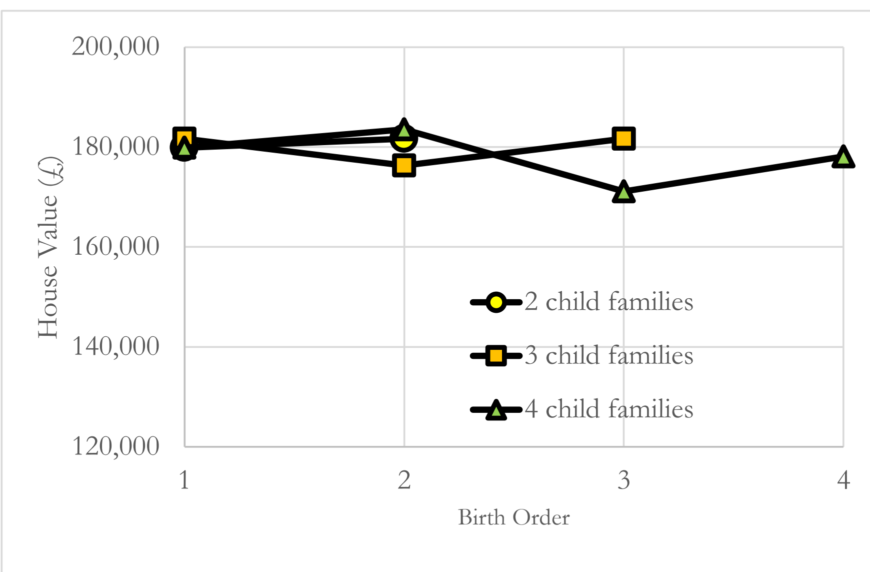
Notes: 5 year moving average of values by birth year.
Source: FOE database.

Table 9.3: Houe value and IMD by birth order and family size

Family Size	Birth order	Count	Birth Year	Parent age	House Value	Index of Multiple Deprivation
2	1	3,354	1957.5	27.7	179,872	58.3
2	2	3,274	1958.1	31.3	181,680	58.2
3	1	1,782	1955.5	26.5	181,680	54.6
3	2	1,803	1956.2	29.5	176,310	56.4
3	3	1,770	1957.0	33.8	181,680	55.3
4	1	780	1951.9	25.7	179,872	53.6
4	2	810	1952.3	28.4	183,506	52.4
4	3	862	1952.3	31.9	171,099	53.2
4	4	847	1954.0	35.8	178,082	53.0

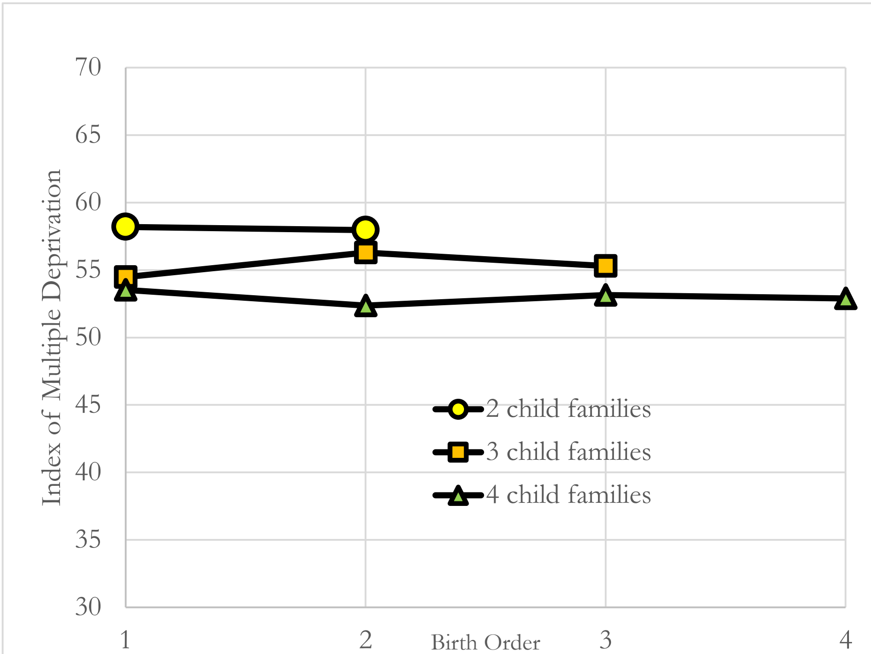
Notes: Average house value calculated from exponentiating average of log house values to reduce influence of outliers. *Source*: FOE database.

Figure 9.2: Birth Order and House Values, 2-4 Child Families, 1999-2006



Source: FOE database.

Figure 9.3: Birth Order and Index of Multiple Deprivation, 2-4 Child Families, 1999-2026

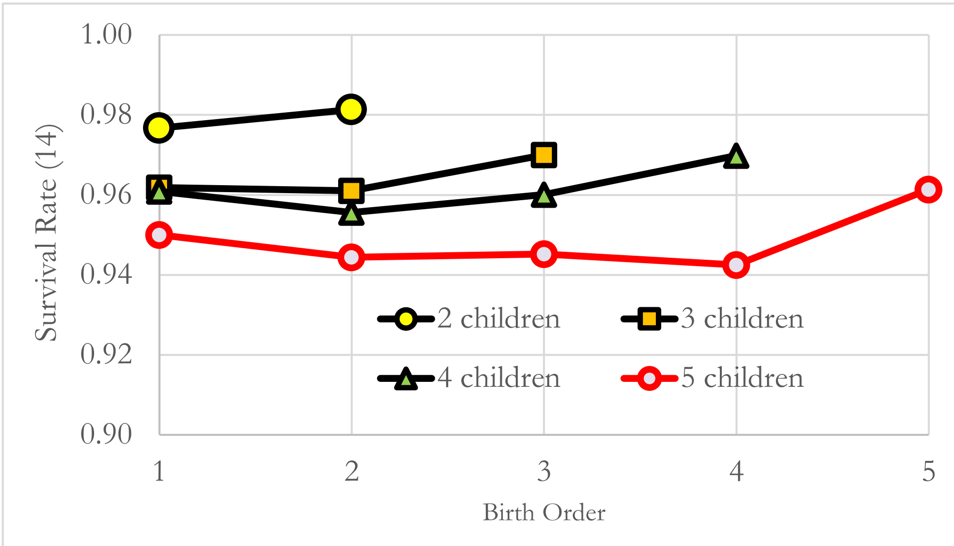


Source: FOE database.

The data in table 9.3 is thus consistent with an absence of both birth order and parent age effects. But it would also be consistent with strong negative birth order effects along with strong positive parent age effects. Younger children in the family will get less parental attention, which may negatively effect their social outcomes. But they also get parental attention from older parents, who may be more effective in that attention.

However, Ockham's Razor suggests the countervailing birth order and parental age effects explanation of the absence of visible birth order effects is implausible. For this requires that two independent effects, those of birth order and parent age, happen to have equal and opposite size both for house values, and for local neighborhood quality. It also requires that the birth order effect is of the same magnitude going from the first to second child, as for going from second to third, and third to fourth child.

Figure 9.4: Birth Order and Survival to Age 14, 2-5 Child Families, births 1910-1999



Notes: Survival rates adjusted for birth cohort time trends in survival rates.
Source: FOE database.

Similar estimates of the effects of birth order on the probability of being a company director are too noisy to reveal much about birth order effects. Since the outcome here is a 0-1 variable where only 14 percent of people are company directors, this is an inherently noisier outcome than house value or index of multiple deprivation. The observed birth order effects here are, however, consistent with an absence of birth order effects.

There is more data, however, on another measure of social outcomes which is survival to age 14. Figure 9.4 shows the net birth order effect for families of size 2-5. Larger families have lower child survival rates in part because they are found more frequently in earlier birth years in the period 1910-99, where child survival rates were lower. So in the figure the survival rates are adjusted to correct for birth cohort effects.

If we look at the average survival rates for first born versus second born children in families of 2-5, then it is 0.962 for first, and 0.961 for second. For families of size 3-5, the average survival rate of second children was 0.954, and for third 0.958. Overall there is again no sign of birth order effects. These could only exist, as before, if there are significant parent age effects of an equal and opposite magnitude to birth order effects.

Birth Order before 1930

As table 9.1 shows, there is significant evidence in the FOE database for social outcomes by birth order for those born before 1930. A complication for this earlier period, however, is the existence of advantages given specifically to first born sons. Such sons can, however, be the third, fourth or even higher in the birth order. Patrick Wallis and Cliff Webb assembled a sample of 645 sons from gentry families born 1550-1699, with records of their education and training.[57] They find the oldest son has a significantly higher chance of attending Oxford or Cambridge, or undertaking legal training at the Inns of Court, than later sons (42 versus 21 percent). They find later sons are more likely to instead be apprenticed (14 versus 4 percent).[58]

There is evidence, however, in the FOE database that the first-born son preference was limited to a very small group of high status families, and would play little role in creating birth order effects overall.

The FOE database has three sets of lineages: those of average status, those of high wealth in the mid-nineteenth century and those of moderate wealth in the mid-nineteenth century. We can thus observe how strong birth order effects were in the population as a whole born before 1930, as well as in moderate and high wealth families. Specifically we can also test if there were significant advantages for first born sons.

Table 9.4 shows the characteristics of the three lineage groups in these years. Wealth is measured relative to the average for deaths in the decade of death. As can be seen the elite groups are distinctive in terms of wealth, and in terms of much greater levels of education for sons. Figure 9.5 shows for each of these social groups the wealth, occupational status, chance of attending Oxford or Cambridge, and chance of any higher education for oldest sons relative to their brothers.

For sons from the average lineages, 98 percent of the population, there is no sign of a specific first son advantage in wealth, occupational status, or higher education. For the sons of the elite lineages there are significant advantages for the eldest son in terms of education and wealth. But, interestingly, there is no advantage in occupational status.

---

[57] Wallis and Webb, 2011, table 1.
[58] Wallis and Webb, 2011, table 3.

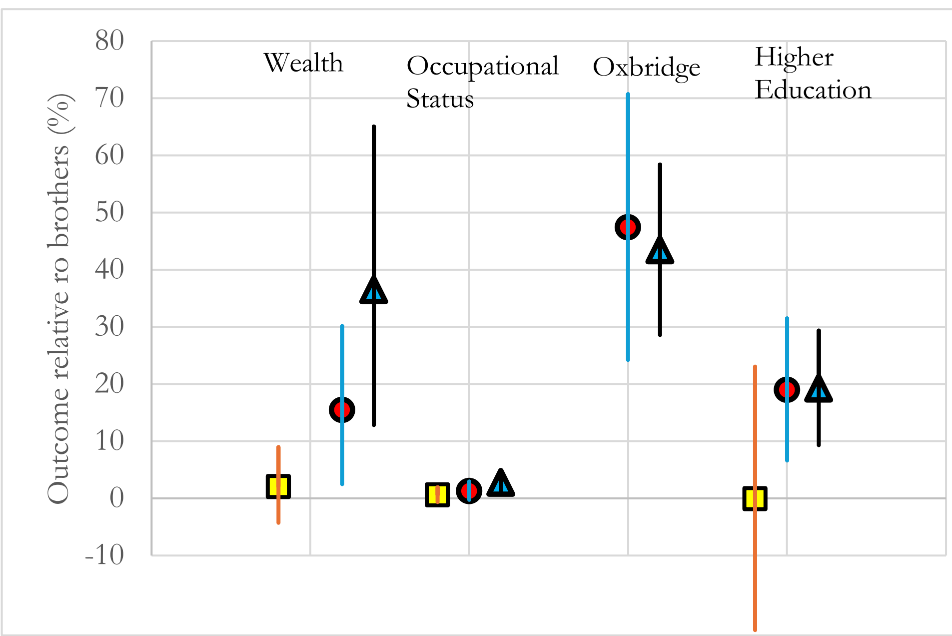
Table 9.4: Characteristics of Average and Elite Lineages, births 1680-1929

Outcome	Genders	Average Status Lineages	Moderate Wealth Lineages	High Wealth Lineages
Wealth at Death (average = 1)	Both	0.8	7.6	62.7
Occupational Status (0-100)	Males	36	46	63
Attended Oxbridge	Males	0.004	0.059	0.223
Higher education	Males	0.017	0.126	0.352

Source: FOE database.

Notes: For the first four outcomes we include only persons surviving to at least age 21.

Figure 9.5: Effect of being the first born son, births 1680-1929



Notes: Percent increase in outcome for first son versus other sons. □ = average lineages, ● = wealthy lineages, Δ = highest wealth lineages. *Source*: FOE database.

Thus the social practice of favoring first born sons seems to have been confined to a small elite share of society in England and Wales in the eighteen and nineteenth centuries, a share that was likely less than two percent of all families. For the vast bulk of the population if birth order effects existed it would not be through this social mechanism of privileging the first born in terms of wealth transfers or educational opportunities.

In estimating birth order effects for the general population I thus include controls for the eldest son. But these controls generally show insignificant gains to eldest sons. While for the general population, first born sons derived no special advantage, there remains the possibility of a more generalized advantage for those low in the birth order, in a world where incomes were much lower than for the last hundred years, and average family sizes much larger.

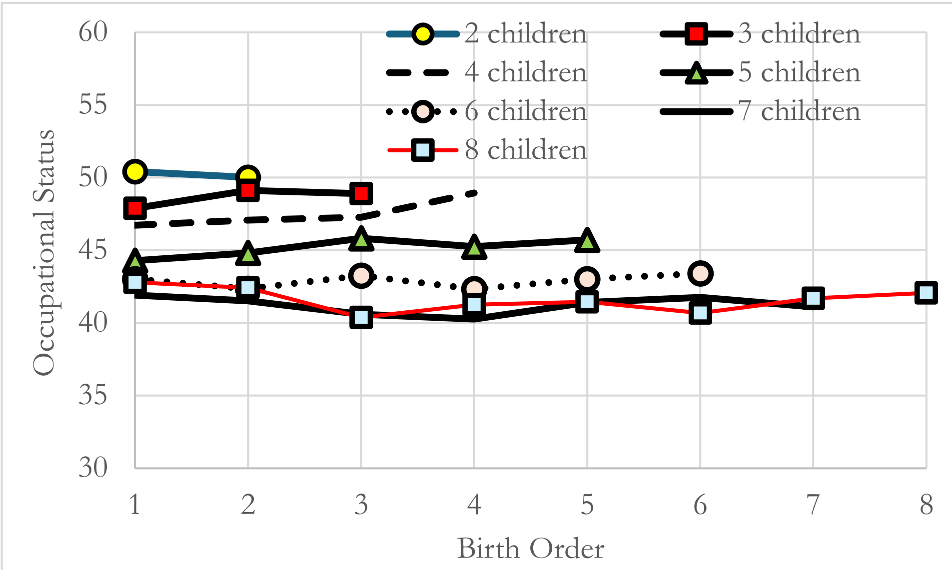
Occupational Status

As figure 9.5 illustrates occupational status for sons is the earlier social outcome that is most accurately measured. We can then follow the same strategy as above, splitting sons into families of sizes 2-8, to examine whether there is any net birth order effect on son occupational status. Figure 9.6 shows the results of that process.

The result, overall, is very clear. On net there is no birth order effect within a whole range of family sizes 2-8. Occupational status is higher for smaller families, but this just reflects the early switch of high status families to fertility control after 1880. But if we compare the first born to the second born across family sizes 2-8, then on average the status of first born is the same as of the second born (44.6 versus 44.5). Similarly for families of sizes 3-8, the status of the second born is the same as for the third born (43.7 versus 43.7). Third versus fourth borns also have the same average status, for families of size 4-8.

Again birth order is correlated with the average age of parents, with higher order children having older parents. Since we are looking at the net effect of birth order, it is possible that there is a substantial negative birth order effect cancelled out by an equivalent positive parent age effect. But even more so Ockham's Razor argues against this possibility.

Figure 9.6: Son Occupational Status by Birth Order, family sizes 2-8, 1680-1929



Notes: Occupational status measured on a scale 0-100. *Source*: FOE database.

For example, if we look at first born versus second born sons, then the parental age gap is 3.8 years for two child families, but just 2.1 years for 8 child families. Why did double the parental age gap for two child families produce the same net effect of zero in terms of child social outcomes as in families of eight?

Thus, in line with additive genetic transmission of social status, there is no evidence of significant birth order effects either in the modern era, or in the nineteenth century. The only exception is that first born sons in elite families in the nineteenth century were given significant advantages in terms of elite education, and of inherited wealth, compared to younger sons. But for 98 percent of families there is no sign of birth order effects even in the nineteenth century.

Alternative Estimates of Birth Order Effects

An alternative way to estimate birth order effects is where we look both within and between families, but control statistically for birth cohort effects, and also for the effects of parental age. The idea is that there should be enough variation in birth intervals, and child dates of birth, that we can also estimate empirically separate birth order, parent age, and birth cohort effects. This method has been widely used in modern estimates of birth order effects (see, for example, Black et al., 2005). These estimates, surprisingly in light of the above, tend to find significant negative birth order effects.

However, these regression methods seem prone to find birth order effects only through fitting implausible birth cohort effects, or through fitting parent age effects that cancel out the birth order effects. As an example, consider figure 9.7, which shows the cross-family estimate of the effects of birth order on years of education in modern Norway from Black et al., 2005.

Figure 9.7 shows significant negative birth order effects, with the 10^{th} child getting a year less schooling than the first. But notice that it also estimates significant and countervailing parent age effects.[59] A 10^{th} child will get one more year of education compared to a first from this parent age effect. The birth order effects, and those of mother's age cancel each other out, so that the net effect of birth order on education is close to zero all through the birth order.

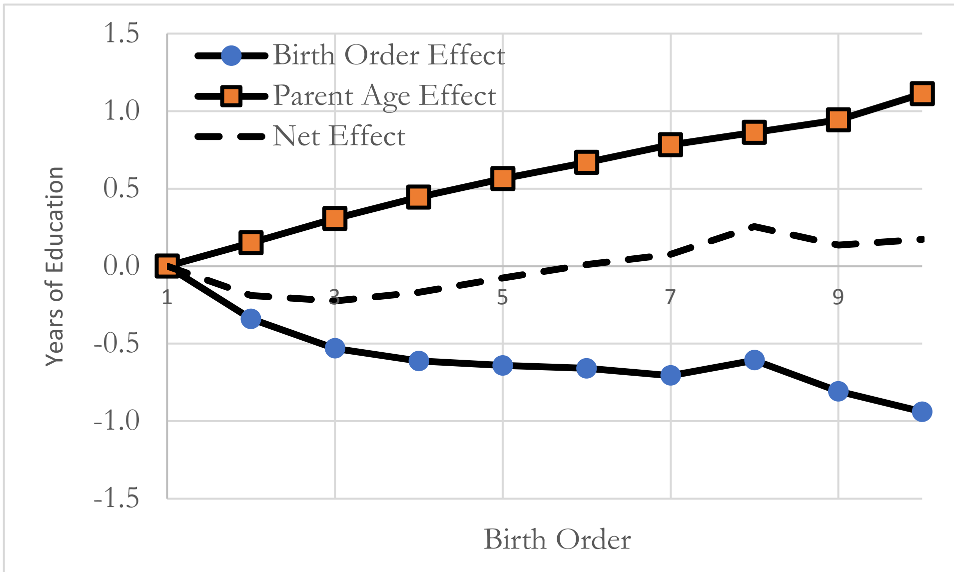
Black et al., 2005, interprets their results as showing a causal effect of birth order, and an independent causal effect of mother age. But the regression allocation of negative birth order and positive parental age effects as the best fit for the data does not imply any causal effect of birth order, keeping parental age constant. These are just the coefficients that best fit this set of data.

Applying these techniques to the modern English data on house values and neighborhood characteristics does yield estimates that have negative birth order effects and positive parent age effects. These estimates, however, on close examination contain pathologies that suggest they are purely artifacts of the estimation method, and not real effects.

---

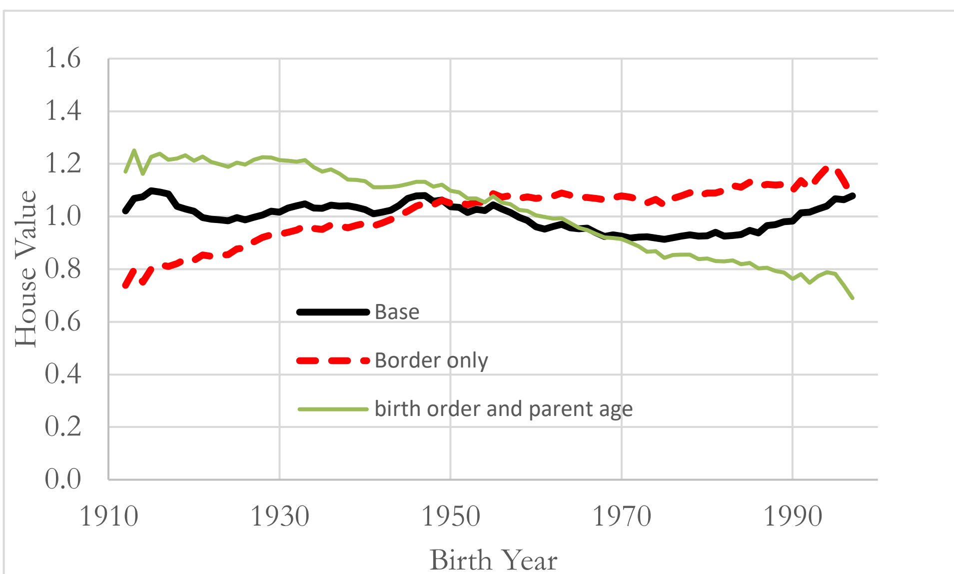
[59] Mother's average age at birth orders 1-10 was estimated from English birth patterns 1910-89 in the FOE database.

Figure 9.7: Birth Order and Mother's Age Effects on Years of Education in Norway



Source: Black et al., 2005

Figure 9.8: Birth Year Effects underlying Birth Order Effects, House Value



Notes: The black line shows the 5-year moving-average house value by birth year for all children. The dashed red line shows estimated house value by birth year where just birth order effects are estimated. The green line shows estimated house value by birth year where birth order and parent age effects are estimated.

In the cross-family estimates of house value and birth order, for example, the negative birth order effect is almost exactly counterbalanced by a positive parental age effect, so that the net effect is zero. Above we have seen that such an exact offset is implausible, and more likely a statistical artifact.

In the within-family estimates, we again get a negative birth order effect, and positive parent age effect. But in this case the parent age effect is stronger than the birth order effect, which implies rising house values for children later in the birth order. But we see in table 9.3 that house values are the same by birth order. To fit the data the estimation thus generates a substantial negative birth cohort effect, where house value by birth cohort falls by more than a third as we move from 1910 to 1990. Yet in the overall data house value by year of birth changes little from 1910 to 1999. This is shown in figure 9.8.

We can show conclusively that these estimated results are pathological, by creating for the families in the FOE database house values which by construction shows no birth order or parent age effect. To do this we use for each family the house value listed for the first child observed. Then each subsequent child gets that value plus a random effect. This implies there is no birth order effect. It also implies no parent age effect. With this artificial data set the estimation procedure, across families, produces an estimate of strong negative birth order effects, counterbalanced by positive parent age effects. Here these effects are clearly spurious. The within family estimates produce a significant positive parent age effect, and to balance this a spurious negative birth year trend.

The statistical imputation methods used in recent birth order effects studies is that unreliable in both cross-family and within-family estimates. It is liable to hallucinate birth order and offsetting parent age effects, even when by definition none exist. Or it will fit birth cohort effects that fail to match the underlying data in the within-family estimate.

This implies that we have to rely on the simpler pictures of outcomes by family size and birth order, as in figures 9.2-9.4 and 9.6 above. These pictures are completely consistent with an absence of birth order effects in both the modern period and the nineteenth century. Significant negative birth order effects could only exist if there is also a significant and offsetting parent age effect. But Ockham's Razor applies in this case. Why would we invoke two unobservable offsetting effects to explain a constancy, without anything to suggest their existence, or why such unrelated effects always counterbalance?

Conclusion

With significant social causation of child social outcomes we would expect substantial birth order effects. The first children in families get significantly more parental attention and interaction than later children. This happens in part because new parents are enthusiastic about the novel experience of childcare. It also happens because children early in the birth order face less competition for parental time and attention in their first years of life than later children.

Yet the data for England in both the modern period and the nineteenth century suggests that, controlling for family size, birth order has no effect on social outcomes. This is consistent again with status being largely determined by additive genetic inheritance.

Oborne Street, Sheffield, 1959

10. Place of Birth and Social Outcomes

Does place of birth exercise an important influence in social outcomes? England, like many modern countries, shows substantial and persistent differences in social and economic outcomes by region, detailed below. In particular output per worker, wages, health, longevity, and education are all better in London and the South-East than in the North of England, and in Wales. Is this a product of the social environment of London and the South-East versus that of the North? If the social environment of different places matters to social outcomes then this shows the possibilities of important social causation of outcomes even within an economy.

The different economic and social outcomes South versus North in England could have two very different sources. First would be where there are strong effects of place. There are local equilibria which either favor, or discourage, economic activity and social success. The same person born in a favored location, such as the South, will have better life outcomes than if born in an unfavored location, such as the North.

A second source would be that there had been sorting of people over time. Those who were more productive, educated and socially successful gravitated to the South. Those less productive, less educated, and less socially able stayed in the North. For a person of a given level of ability the South and North offer equal economic, educational and social opportunities. But for people of higher economic ability the South is a more attractive location.

In this chapter I show that the better outcomes in London and the South-East stem mainly from several hundred years of sorting within England, where there was a persistent net flow of people with greater talent and social abilities towards London and the South-East, and away from northern England, the midlands and Wales. These flows are shown to be sufficient to explain the differences in outcomes in terms of occupational status and wealth between North and South in recent years, if they created persistent differences in ability between North and South.

By implication the poorer performance of the North reflects no economic failure requiring the channeling of government resources to public works, or the building of high speed railway lines. The economic returns to a given level of skill are equal in North and South. It is just the skill endowment that differs.

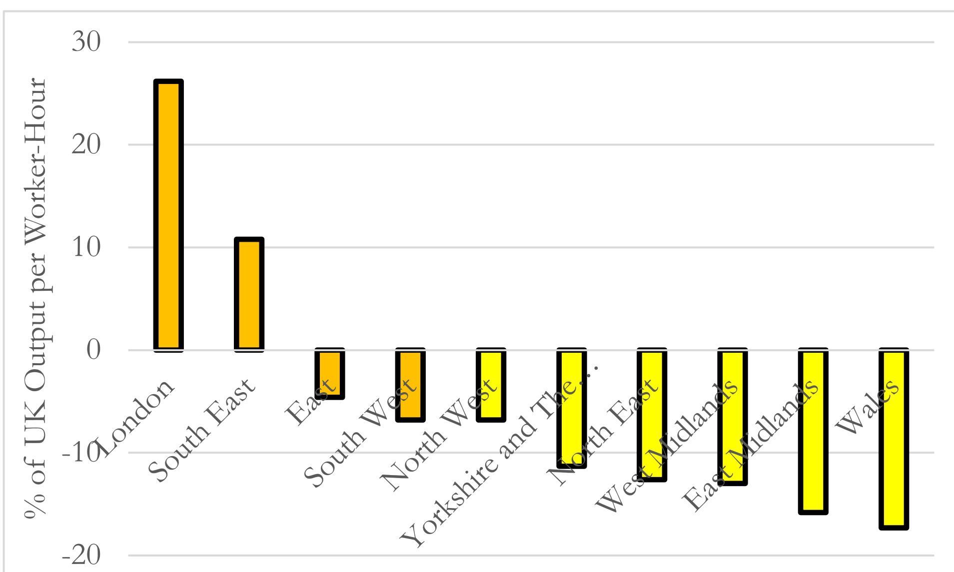
If, as asserted in chapters 4 and 5, social abilities are largely determined by genetics then any migration of higher status individuals North to South will have an enduring effect on relative abilities in North and South. The children of high status migrants will tend to regress to the mean. But the addition the parents bring to the average ability level will be preserved, since the genetics of the Northern migrants will become a permanent share of the genetics of the South, assuming they have the same fertility level as the native Southerners.

Regional Differences in Economic and Social Outcomes

Figure 10.1 shows the ten regions of modern England and Wales. Figure 10.2 shows output per worker-hour in each of these regions in 2022, relative to the average output for the UK. Four regions in the south of England - London, South-East, East and South-West - show higher output per worker than the other regions in the midlands and north of England, and in Wales. London, the best performing region, has an output per worker-hour more than 50 percent greater than in Wales. Below I denote these four southern regions as “the South”, and the other northern regions and Wales as “the North”. In 2023 the North and South each had 50 percent of the population of England and Wales. In 1851 the division was that 54 percent of the population lived in the North, 46 percent in the South.

Figure 10.1: Regions of England and Wales

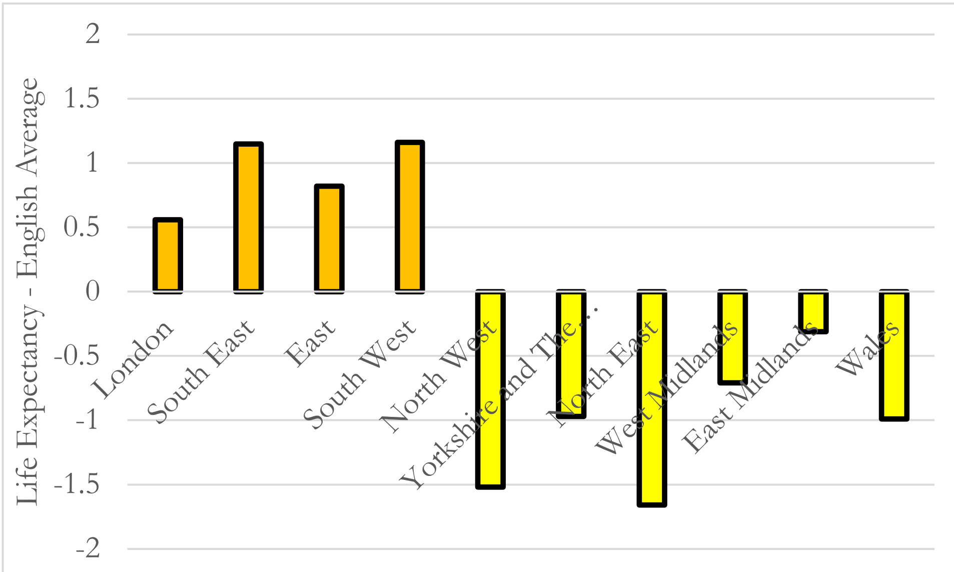
Figure 10.2: Regional Differences in Output per Worker-Hour, England and Wales, 2022



Source: Office of National Statistics

Notes: The regions highlighted in orange constitute the South, the others the North.

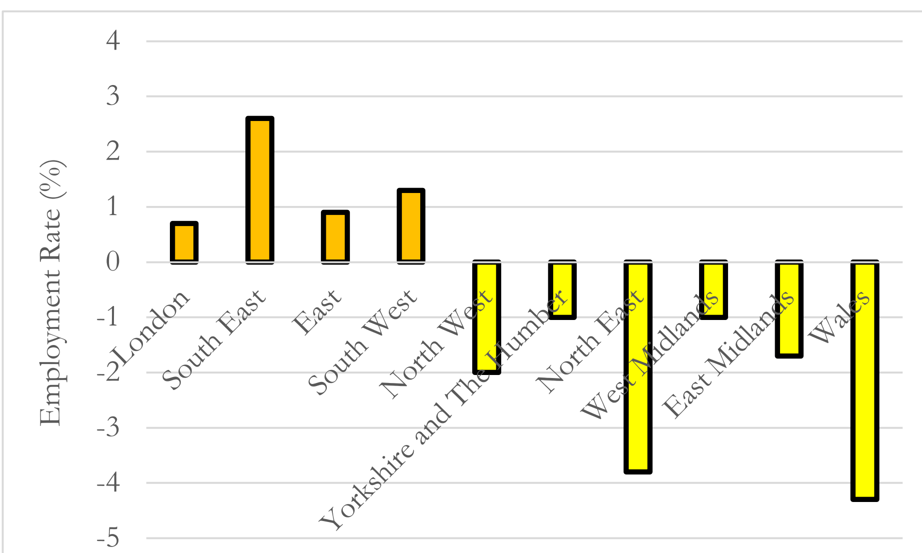
Figure 10.3: Life Expectancy, 2020-22, in regions of England and Wales



Source: Office of National Statistics

Notes: The figure shows for each region the difference in life expectancy between the region and the English average life expectancy of 80.83 years.

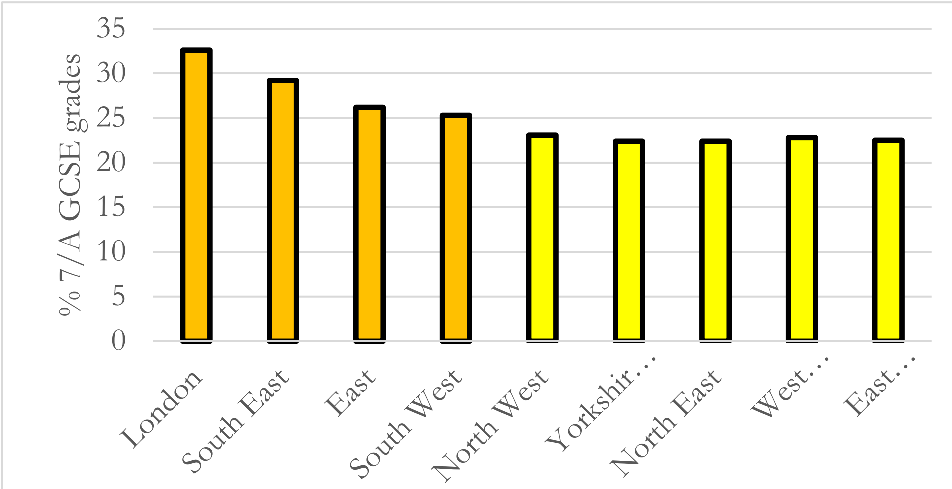
Figure 10.4: Employment Rates 16-64, relative to Average (%)



Source: Office for National Statistics (ONS), 13 February 2024, ONS website, statistical bulletin, [Labour market in the regions of the UK: February 2024]

Note: The average employment rate for England for those 16-64 in 2023 was 75.4%.

Figure 10.5: Educational Attainment, 2022, in regions of England and Wales



Source: Winchester, 2022, table 2.
Notes: The figure shows the share of students getting grades 7-9 (corresponding to the earlier grade of A), on a 9-point scale, in the General Certificate of Secondary Education (GCSE) exams taken ages 15-16.

Table 10.1: Benefit recipients North versus South, England, 2022

Town	16-64 year olds receiving benefits (%)	16-64 incapacity benefits (%)	16-64 out of work benefits (%)
Middlesbrough	22.7	6.4	15.7
Liverpool	20.8	7.2	13.0
Sunderland	110.6	6.3	12.7
Manchester	18.5	5.0	13.1
Oxford	7.4	2.1	5.1
Richmond	6.8	2.0	4.6
Cambridge	6.6	2.2	4.3
Guildford	6.3	2.1	3.9

Source: Office for National Statistics (ONS)

These differences in economic performance are echoed in other social indicators. Figure 10.3, for example, shows life expectancy in 2020-22 in each of the regions compared to England as a whole. The South shows nearly 2 years greater life expectancy than the North.

Employment rates of those aged 16-64 also vary significantly between South and North, as figure 10.4 shows. The South has employment rates typically 1.5 percent above the English average of 75.4 percent. In contrast the North had employment rates that were typically 2 percent below that average. One source of this was that for those in the age range 16-64, 6.8 percent of those in the North in 2022 were out of the labor force because of sickness, while in the South that share was only 4.6 percent.

These regional social differences extend also to educational attainment. Figure 10.5 shows for 2022 the percent of test takers around age 16 who got an A grade in the GCSE exams. Again the South outperforms the North. The percent getting an A grade in London was almost 50 percent higher than in the North-East, Yorkshire or the East Midlands.

Table 10.1 shows the extremes of the differences North versus South, in the form of the share of people in 2022 who received benefits payments in four economically depressed northern cities, versus four prosperous towns near London. These disadvantaged northern cities have substantially more people receiving government benefits. 20 percent of 16-64 year olds were in receipt of such benefits, compared to less than 7 percent in the favored southern towns. The northern population in these depressed cities is less likely to be employed both because a larger percent of this population is incapable of work, and because a larger fraction of those in the labor force are unemployed.

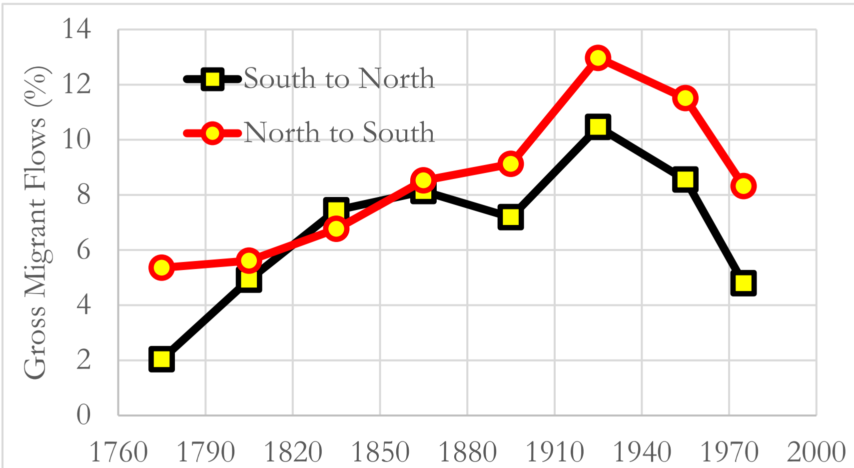
Sorting by Social Abilities, England and Wales, 1770-2024

There has always been movement of people across the regions of England. Figure 10.6 shows the gross flows of individuals by birth generations of 30 years between North and South. A migrant is counted as a person born in one region who dies aged 21 and above in the other region. Flows are measured as the percentage of each generation born North and South which reached aged 21, who die in the other region, measured relative to the population born in the North. It thus shows net flows into and out of the North as a share of the population of the North.

There was a period at the height of the Industrial Revolution, people born 1790-1879, where these flows were equal. But both prior to the Industrial Revolution, and for births after 1880, the net flow within England was consistently from North to South.

But what figure 10.6 emphasizes is that the main effects of these population movements on the characteristics of people North and South will come from differences in the status of the movers to the South versus to the North. And there is good evidence that all the way from 1770 to 2000 the migrants to the South had higher educational attainment and social abilities than those moving North.

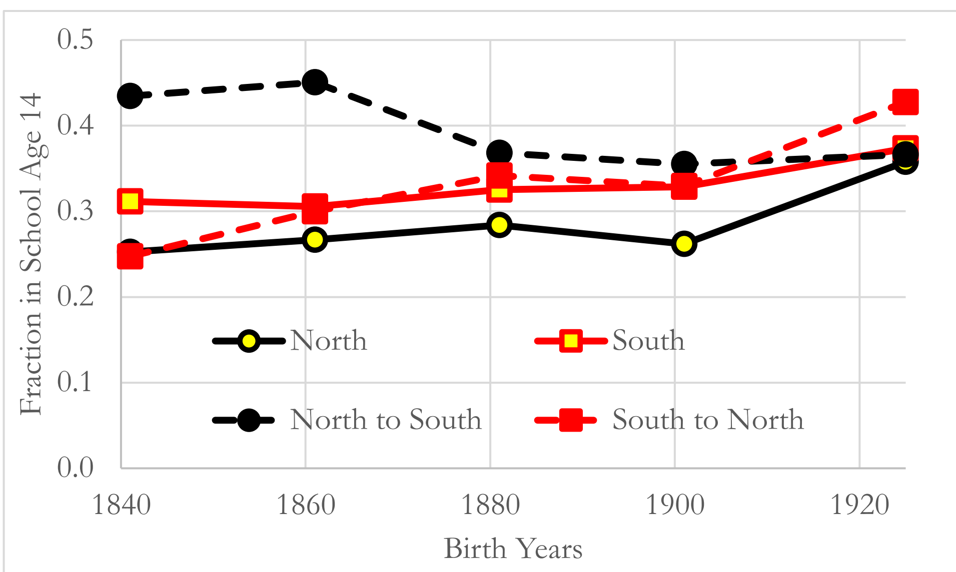
Figure 10.6: Gross Flows between North and South, births 1770-19710.



Source: FOE database

Notes: For births 1910-79, the later location of people is identified from the electoral rolls 1999-2024. The figure shows flows of those aged 21+ North and South as a percent of those born in the North in each period who attained age 21.

Figure 10.7: Fraction in School, Age 14, by Migrant Status, births 1841-1925



Source: FOE database

Notes: People categorized by place of birth and death.

Figure 10.7 shows educational status for children in the general lineages of the FOE database aged 10-18 from the censuses of 1851-1921, and the population register of 19310. The figure shows the fraction of 14-year-olds who were still in schooling. These children fall into four groups: those born and dying in the North, those born and dying in the South, those born in the North but dying in the South, and those born in the South but dying in the North.

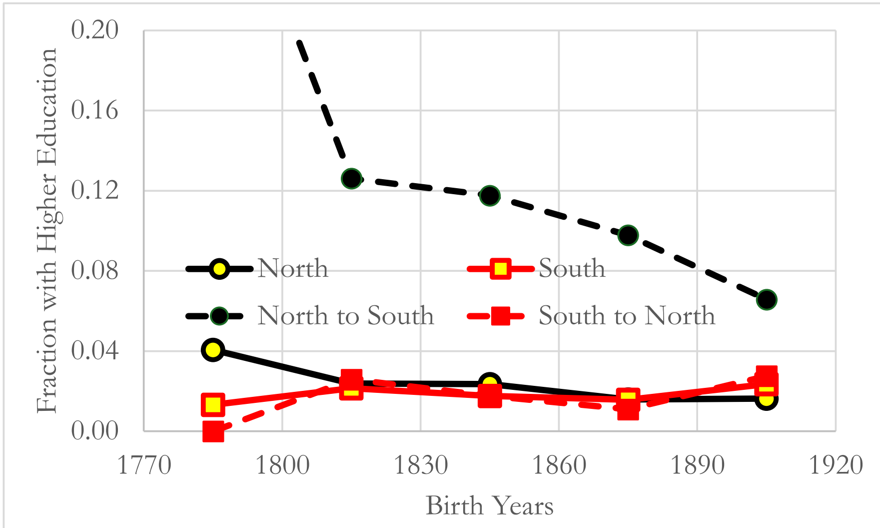
For the cohorts observed 1851-1881 there is a strong positive selection among those migrating North to South. The North was losing more educated people to the South. In 1851, for example, 25 percent of those born in the North who stayed in the north were in school aged 14, compared to 43 percent for those moving to the South. In contrast migrants from the South to the North tended to be similar to the general population in the south. In 1851 31 percent of those staying in the south were in school aged 14, compared to only 25 percent for those moving north.

Thus we see for births in the years 1840-1900 a general pattern of the movement of more educated people from the north to the south. The measure of status from schooling is helpful, since schooling occurs early in life, before most people would have moved between regions. It reflects just their status within the North or South.

This selective movement north to south is confirmed by another source, which is the fraction of men who attained higher education. This attainment would be slightly later in life, typically ages 18-21. Figure 10.8 shows the fraction of men attaining higher education again classified as Northern or Southern stayers and movers. The figure again covers birth years 1770-1919, by 30-year periods.

In general only 2 percent of men achieving such higher education. Here we see an even more dramatic effect of men migrating from north to south being much more educated than both those staying in the north, but also than those staying in the south, or migrating south to north. In the first birth period, 1770-1799, nearly one third of men moving north to south have some kind of higher educational attainment. The higher levels of education for men moving from north to south is still prominent for the last group observed, those born 1890-19110.

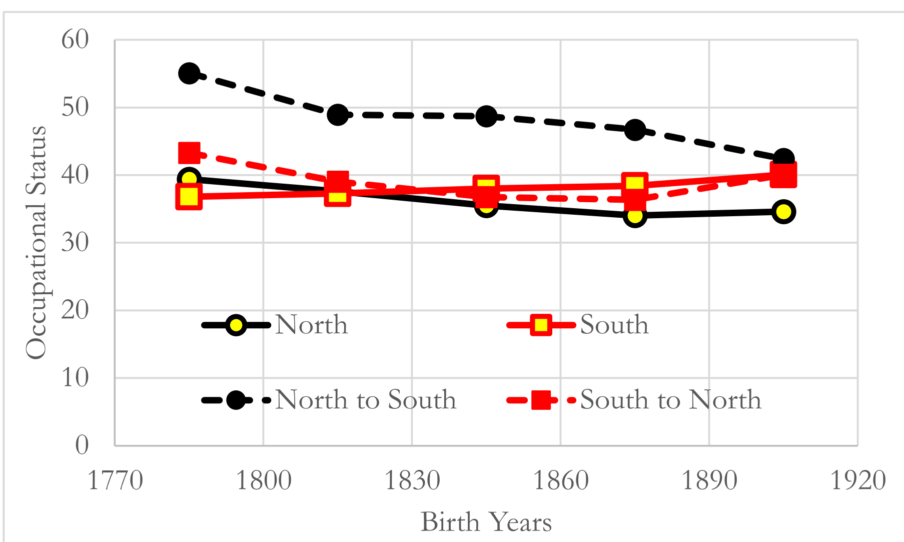
Figure 10.8: Fraction of Men with Higher Education, by Migrant Status, births 1770-1919



Source: FOE database

Notes: The fraction of men with higher education born in the North 1770-1799, and dying in the south was 0.31.

Figure 10.9: Occupational Status of Men, by Migrant Status, births 1770-1919



Source: FOE database

Notes: Occupational status at age 40, on scale 0-100.

Another measure we have for men is occupational status attained around age 40. This is not such a good measure of the intrinsic abilities of men migrating versus staying. For if location matters to achieved occupational status, many of the men moving will be exposed to the new location before their occupational status at 40 is measured. The movers north to south may thus gain an occupational benefit from the move. But if average occupational levels are similar for south versus north, then this potential gain will be modest.

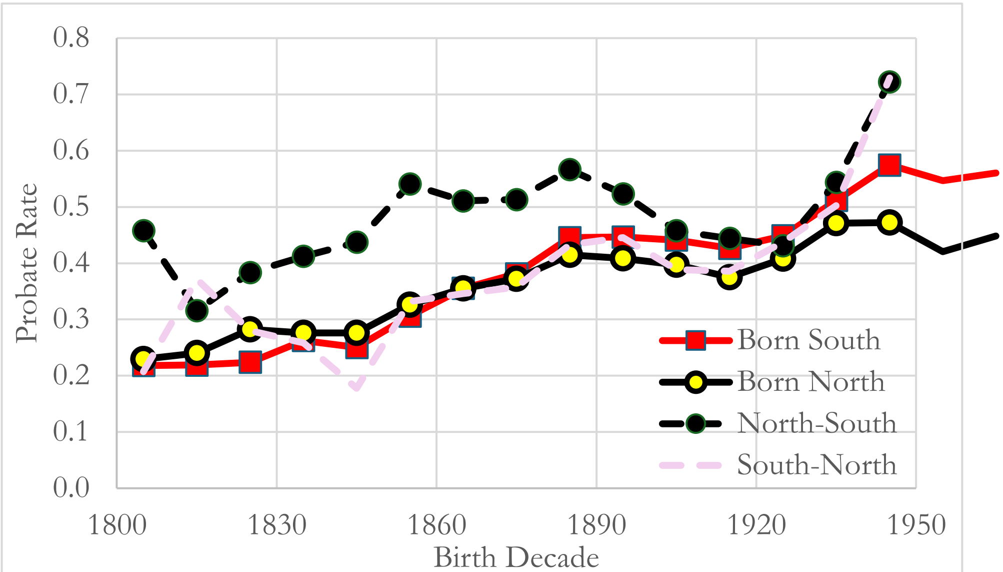
Figure 10.9 shows the results. Men leaving the North for the South showed substantially higher occupational status than those staying in the North, for all periods 1770-19110. At the same time there was little difference in occupational status for men staying in the South versus those moving North.

Another indicator of the selectivity of migration between North and South is the fraction of people who were probated at death. Only those with assets above a certain minimum were subject to probate. Thus the fraction of each group probated is an indicator of the economic status of individuals. For the FOE database we have, for general lineages, measures of these probate rates for those dying 1858 to 20110.

Figure 10.10 shows the probate rates for four groups of persons by birth decade, from 1800 to 19610. The first are those born and dying in the North, the second those born and dying in the South, the third those born in the North but dying in the South, and the last those born in the South but dying in the North. Since for those born in the 1930s to 1960s we only observe those dying at younger ages, I control in the estimation for age at death. The probability of being probated rises with age at death until a person reaches ages in the 90s.

Being probated is an indicator of having wealth at death, since all those whose wealth exceeded a varying minimum were required to be probated. Since 1985 the minimum wealth requiring probate has been only £5,000, but still less than half the population is probated. In figure 10.10 in the early nineteenth century the probability of being probated was slightly higher for those born and dying in the North, compared to those born and dying in the South. But thereafter the probability of being probated in the South has risen relative to the North, and for those born in the 1960s the predicted probate rate in the South is 56 percent, compared to only 42 percent in the North.

Figure 10.10: Probate Rates, births 1800-1949



Source: FOE database

Notes: Groups classified by place of birth and death.

Those moving North to South born 1800-1899 showed much higher probate rates than stayers in the North and South. Those born 1900-1969 show higher probate rates than those staying in the North, but rates similar to stayers in the South. Finally those moving South to North generally showed probate rates much lower than those moving in the opposite direction.

The evolving gap in probate rates between North and South thus is plausibly just a reflection of the shifting economic abilities of the populations of North and South created by persistent elite migration out of the North. This migration in terms of the evidence of probate rates extended back as far as those born 1780 and later.

Can Selective Migration Explain Social Gaps between North and South?

As noted above, if social status is genetically determined, then each transfer of higher status population North to South will have a permanent effect on economic and social performance of North versus South. Were the transfers observed of sufficient magnitude to explain the difference observed now?

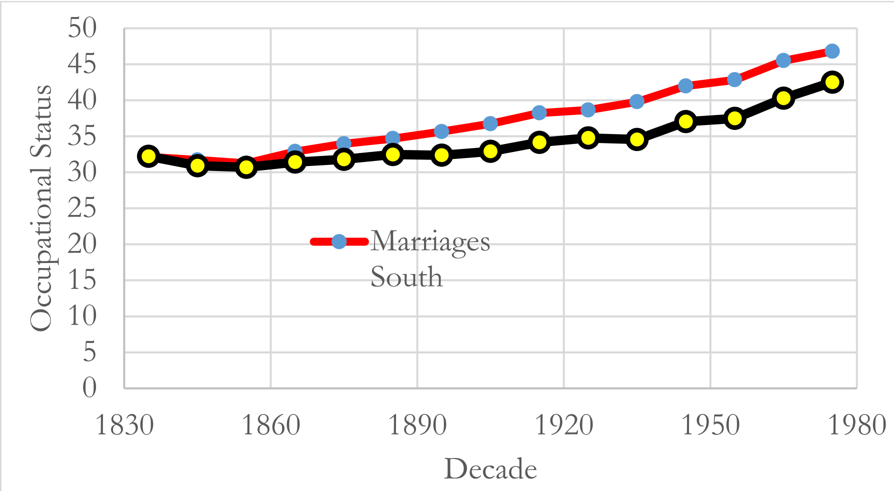
Figure 10.11 shows the average occupational status of grooms marrying 1837-1979 in the North versus the South of England, based on a sample of 1.6 million marriages transcribed by volunteers of the FreeReg organization (the MOE database). While these statuses were equal North and South for marriages before 1860, thereafter there has been a steady divergence in favor of the South.

Figure 10.9 above shows a consistent migration of men of higher occupational status from North to South for men born 1770-19110. Was this migration of sufficient quantity and selectiveness to explain the change in relative occupational status between North and South? The numbers to assess this are found in table 10.2, which shows occupational status for men born in England 1770-1919 from the FOE database, as well as the numbers in each group. From this we can estimate the expected effect on average occupational status in the next generation for those born North versus born South.

In the birth interval 1860-89, representing one generation, for example, migration North to South is estimated to decrease occupational status in the North in the next generation by 1.1 points, and increase it in the South by 0.75 points. Similarly migration South to North should have increased Northern occupational status by 0.2 points, and increased southern by also 0.2 points. This gives a net decline for the North of 0.9 points, and gain for the South of also 0.9 points.

Figure 10.12 plots out for each birth period 1800-29,…, 1920-49 the projected change in relative status, South – North, from the migrations observed in table 10.2. Over these five generations there is a projected gain of 5.5 points, on a base of 32 points, a substantial projected gain.

Figure 10.11: Groom occupational Status North Versus South, 1837-1979



Source: MOE database

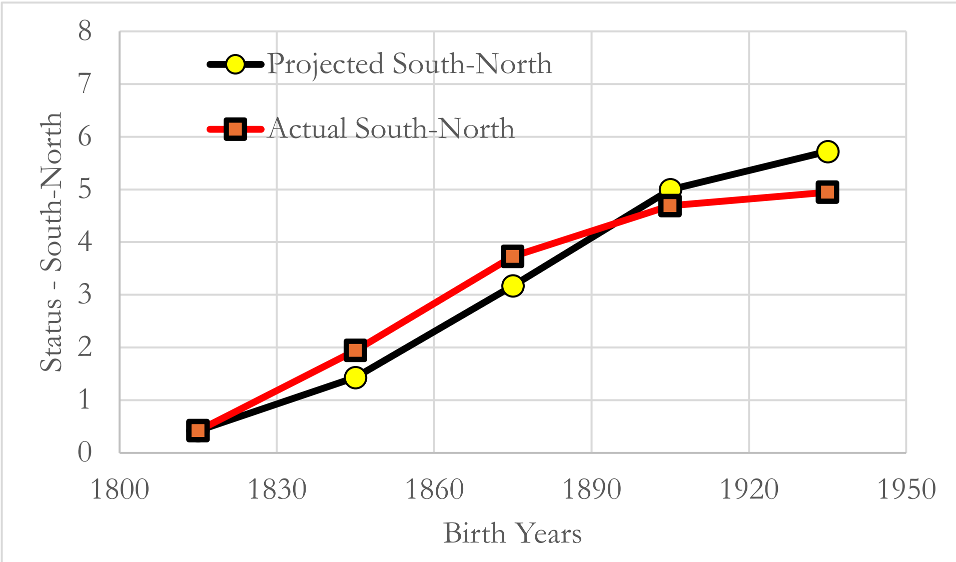
Notes: Occupational status on a scale of 0-100.

Table 10.2: Occupational Status by Migrant Status, 1770-1919

Period	Births North Occ. Stat.	Births South Occ. Stat.	North to South Occ. Stat.	South to North Occ. Stat.
1770-99	31.4	36.8	55.1	43.3
1800-29	37.6	37.3	48.9	31.0
1830-59	35.5	38.0	48.7	36.8
1860-89	34.0	38.4	46.7	36.4
1890-1919	34.6	40.1	42.4	40.1
	Number	Number	Number	Number
1770-99	816	928	50	24
1800-29	2,191	2,305	115	116
1830-59	3,318	3,324	242	278
1860-89	4,492	4,235	384	369
1890-1919	3,969	3,461	410	259

Source: FOE database.

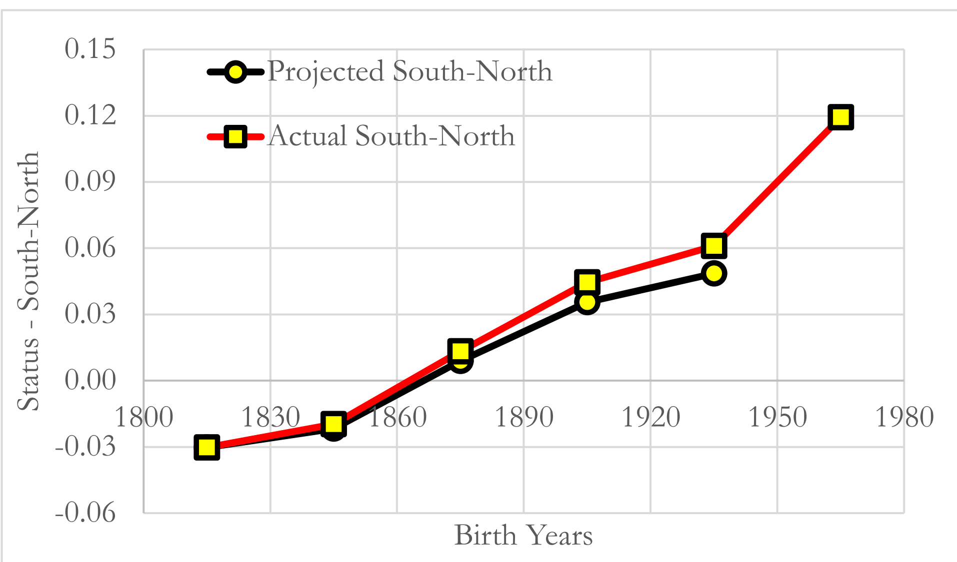
Figure 10.12: Actual versus Projected South-North Occupational Status Diffeence



Source: FOE database, MOE database.

Notes: Difference in average groom occupational status on scale of 0-100, South minus North, by birth period, versus projected difference from selective migration.

Figure 10.13: Actual versus Projected South-North Probate Rate



Source: FOE database.

Notes: Difference in average probate rate, South minus North, by birth period versus projected difference from selective migration.

Figure 10.12 also plots the actual gap in male occupational status by birth generation observed in the MOE database. The movement of the projected gap is close to that of the actual gap. The gap between the regions in figure 10.9 based on the FOE database moves in the same way.

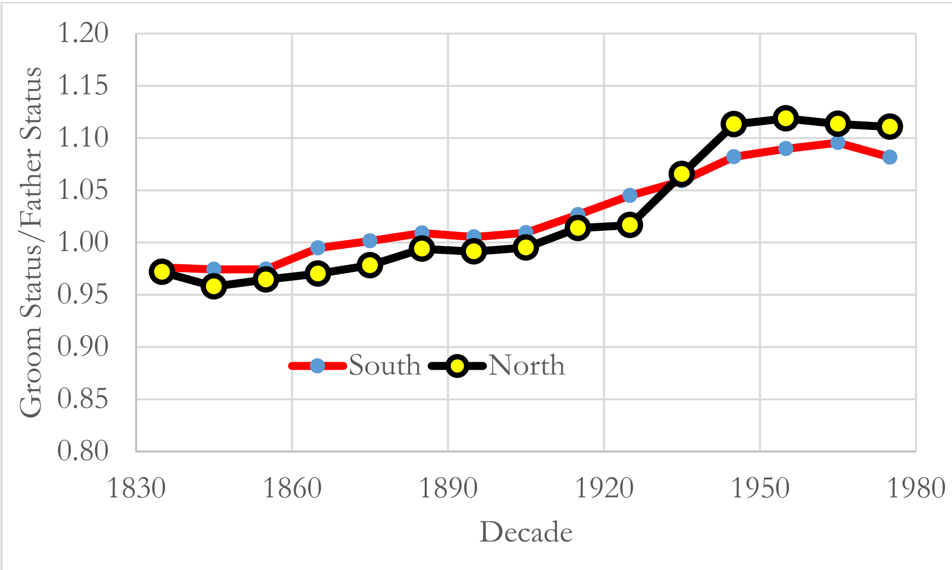
We can do a similar exercise using the records of probate rates by migrant status, as in figure 10.10. Again we can project, from the probate rates of movers versus all people born North or South, what the projected probate rates will be the next generation, assuming complete group inheritance of these characteristics. Figure 10.13 shows the results. There is also here a remarkably good fit between the projected probate rates from selective migration, versus the actual rates by region.

Figures 10.12 and 10.13, for occupational status and probate status, suggest that all of the North/South differences in social and economic outcomes could be just the product of the long history of selective migration from the North to the South. Further evidence that the differences between North and South are the product of selective migration versus better opportunities in the South comes from considering the relationship between father occupational status and son status. If average occupational status rose in the South relative to the North because of better opportunities in the South, then we would expect to see sons rise in status relative to their fathers more in the South than in the North.

Figure 10.14 shows for grooms marrying North and South 1837-1979 their average occupational status relative to their father. As can be seen, in the South from the 1920s on we see son status rising 5-10 percent relative to father status. This helps explain the rise of average occupational status for men in the South across time. But interestingly the occupational status of grooms in the North shows the same average rise 1837-1979 as in the South.

In the period 1837-1929 grooms in the North fall slightly in occupational status relative to their fathers, but by an average of only 1 percent. Grooms in the South have the same average status as their fathers. For 1940-1979 grooms in the North rise in status by 11 percent relative to their fathers, and in the South by 9 percent. The greater rise in status in the North may be explained by the decline in coal mining. Coal mining was 8 percent of fathers in the North, but only 4 percent of grooms. Coal mining also had a very low estimated occupational status.

Figure 10.14: Groom Occupational Status Relative to Father, 1837-1979



Source: MOE database

Thus there is no indication that the South offered better occupational status opportunities than the North in a way that would explain the rise of occupational status in the South relative to the North after 1860. Instead selective migration seems to be the key.

If migration is driven by people's attained phenotype, then the children of migrants will regress to the mean to some degree, and this will mean that the exercise above will overestimate the contribution of selective migration to the growing status gaps between North and South. However, we saw in chapter 4 that social status is highly heritable, so that the average difference from the average of social status of children will be 0.75-0.80 that of their parents. Thus selective migration will still explain a most of the occupational and wealth status gaps between North and South.[60]

[60] There was also migration from Scotland and Ireland into the North and South, and also migration mainly to the South from outside the UK.

Confirmation of Genetic Differences between North and South

It is well established that genetic differences across individuals predict at least some of their differences in educational attainment, measured as years of education. If the differences in social outcomes between North and South in England were and are attributable to differences in opportunities for education in each region, then we should observe that relative to their genetic potential people in the North acquire less education than those in the South.

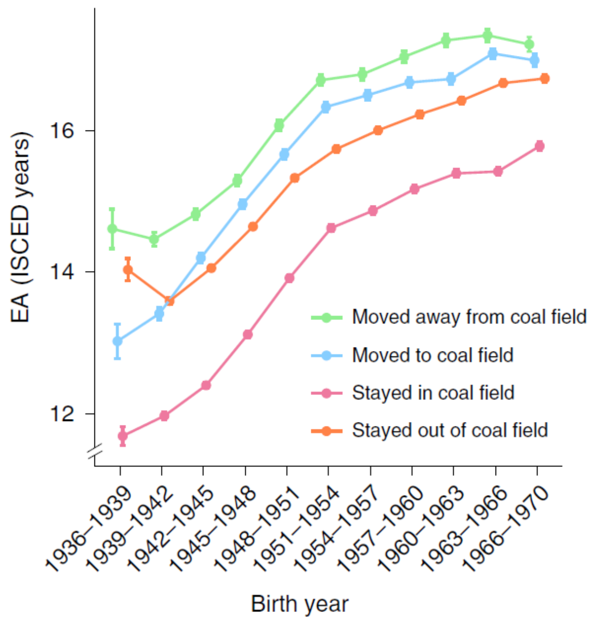
A recent study, Abdellaoui, Hugh-Jones, *et al.* 2019, provides support to the idea that regional migration in Britain led to differences regionally in people's genetic potential for social status. Using data from the 450,000 individuals UK Biobank, the authors divided the country into the areas of the coalfields (mainly in the North and in Scotland), and the non-coalfield areas. Looking at people born 1936-1970, who had completed their educations, they divided the sample into those born and still resident in a coalfield, those born in the coalfields, but now resident elsewhere, those born and still resident elsewhere, and those born elsewhere now resident in a coalfield.

The coalfields are areas of relative deprivation, with an average difference of about a year of education in favor of those born in the non-coalfield areas. But we see the same pattern we saw with migration between North and South, where migrants out of the more deprived coalfields are more strongly positively selected than migrants out of the non-coalfield areas. Figure 10.15 shows average years of education of each of the four groups. Those staying in the coalfield areas have about a year less education than those staying in the non-coalfield areas. Figure 10.16 shows their corresponding genetic score, GPI, for predicted educational attainment, measured with a standard deviation of 1. The non-coalfield stayers have an polygenic education scores nearly 0.2 of a standard deviation higher than the coalfield stayers.

Those migrating or staying in the region of origin showed the expected difference in polygenic scores for educational attainment, in line with their differences in educational attainment of each group. There is no sign that relative to genetic potential, those remaining in the North were achieving less education than would be expected.[61]

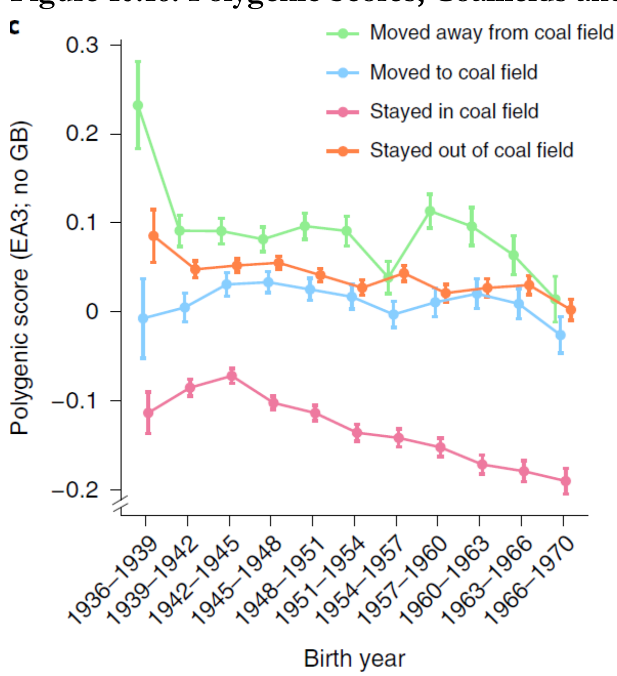
[61] Papageorge and Thom, 2020, and Willoughby et al., 2021 both gives the connection between the genetic educational attainment predictor AE3, and years of schooling. From this we can estimate the expected differences in polygenic score from a given difference in education.

Figure 10.15: Educational Attainment, Coalfields and Elsewhere



Source: Abdellaoui et al., 2019, Figure 5.

Figure 10.16: Polygenic Scores, Coalfields and elsewhere



Source: Abdellaoui et al., 2019, Figure 5.

Conclusions

The lineage and marriage data both suggest that the modern differences in social outcomes between southern England and the north and Wales are substantially explained by selective migration of individuals with greater social abilities from North to South. Estimates from the volume and selectivity of migration suggest that this migration can explain at least 80 percent of differences in average occupational status, or in probate rates. Migration can also explain the widening of these differences over time, from 1800 to 1950.

These conclusions imply that the North-South divide does not represent an economic problem. There would, for instance, be no gain from encouraging more people from the North to move South. There would similarly be no gains from greater public expenditures on education, or on infrastructure, in the North versus the South.

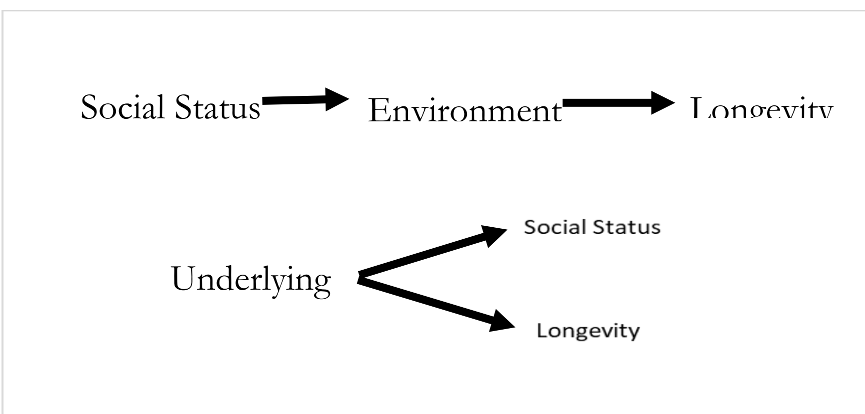
11. Adult Longevity

Child mortality is clearly determined by social factors, and can be significantly changed by social interventions. But what about adult longevity? This was, and is, significantly greater for people of higher social status. But is that because social status creates an environment that confers longevity benefits? Or is it because the genetics that generates higher social status is also associated with robust health?

It is rare to observe people whose social circumstances are changed by some random, exogenous intervention. So it has been hard to assess whether there is significant social causation of adult longevity. It can be shown, however, in a number of ways that a child's family environment does not influence their adult longevity. It can also be shown that if adult environment matters, then it is home environment, and not the conditions of work life. Thus while it cannot be demonstrated that adult social circumstances do not play a causal role, it seems most plausible that adult longevity is associated with social status because the genetics generate higher social status are also associated with robust health.

Figure 11.1 shows there are two very different ways in which social status and longevity can be linked.

Figure 11.1: Possible Links between Social Status and Longevity



When we think of the link between environment and longevity, we would expect that if environment has causal effects, childhood environment would be the most important. What were the conditions when children's bodies were growing, immune systems developing, and organs being formed? But we shall see below that the FOE data strongly suggests that if social environments matter to adult longevity, it is only the attained social environment of peoples' adult years.

The Inheritance of Adult Longevity

To investigate the inheritance of adult longevity I employ data on the lifespan of all people in the FOE database born 1750-1929, who lived to at least age 21. The long lifespan of humans means that we can only discuss longevity inheritance for people born before 1930. To control for trends in longevity over time, age was normalized by gender and birth decade to have always a mean of zero and standard deviation of one.[62]

Table 11.1 shows the longevity statistics for the collection of parents and children where we know the ages of all three, 54,031 observations. Children were included here only if they attained at least age 21. Note that the average parent has above average longevity for the decade they were born in. Their children, in contrast, have close to average longevity. This explains why the average age of death of the parents exceeds that of their children.

Table 11.2 shows the correlation between normalized longevity for parents and children. The first thing that stands out is the extraordinarily low correlation between the longevity of each parent, father and mothers, and the longevity of their children. On average this is just 0.073. Even if we take the average longevity of both parents, the correlation with child longevity increases to just 0.100. This 0.1 is the heritability of longevity.

This low heritability implies that a person, both of whose parents live 20 years longer than average, can expect themselves to live just 2 years longer than average. This weak inheritance of longevity stands in stark contrast to the very strong inheritance of occupational status in the same years, where the observed correlation would be in the range 0.6-0.7, and the underlying correlation was around 0.8.

[62] The normalized age for each decade t was constructed as $Agenorm_{it} = \frac{(Age_{it} - \mu_t)}{\sigma_t}$, where μ_t is mean longevity per birth decade and σ_t the standard deviation of longevity.

Table 11.1: Characteristics of Parent-Child Sample, births 1750-1929

Group	Average Death Age	Standard Deviation Age	Normalized Death Age	Standard Deviation Normalized Age
Fathers	68.91	13.86	0.280	0.802
Mothers	69.45	15.75	0.153	0.862
Average of Parents	69.18	11.87	0.217	0.609
Children	67.40	18.43	-0.011	1.026

Source: FOE database, births 1750-1929

Table 11.2: Correlations in Longevity Between Father, Mother, and Children

Pairing	Number	Correlation	Standard Error
Father-Mother	54,013	0.0728	0.0043
Father-Child	54,013	0.0756	0.0043
Mother-Child	54,013	0.0712	0.0043
Parents-Child	54,013	0.1000	0.0043
Father-Son	31,995	0.0795	0.0056
Father-Daughter	22,018	0.0697	0.0067
Mother-Son	31,995	0.0669	0.0056
Mother-Daughter	22,018	0.0765	0.0067
Parents-Son	31,995	0.0992	0.0055
Parents-Daughter	22,018	0.1002	0.0067

Source: FOE database, births 1750-1929

Notes: 95% of the time, the true correlation of longevity will lie within two standard errors of the estimated correlation.

This low parent-child correlation in longevity reveals an important element in all our lives. Though there are now important social class differences in average longevity, the low heritability of longevity implies that nevertheless the age each of us will die is largely random at the individual level. Reflecting that, the correlation in longevity between husband and wife was only 0.073, despite their shared social class, diet, and family environment.

A nice illustration of the randomness of adult longevity comes from the case of the world's current oldest person, Ethel Caterham, born in England in 1909, and now 116 years old.[63] Her husband died in 1976, aged only 71. Both her two daughters are dead, respectively at ages 71 and 82. Whatever environmental factors contributed to her extraordinary longevity did not communicate to her husband or children.

When we look at the correlation in longevity between fathers and sons, mothers and daughters, and across genders, we do see that in each case the same gender correlation is higher than the cross-gender correlation. However, this effect while suggestive, is of tiny quantitative importance. It is not a statistically significant difference in either case. Finally the correlation between average parental longevity and child longevity is the same for both sons and daughters.

As table 11.3 shows the correlation in longevity between siblings is very similar to the correlation between parent and child. But that correlation is stronger for same gender siblings than for opposite gender. This sibling gender effect is quantitatively more significant, and also statistically significant.

Table 11.4 shows the relative contribution of fathers versus mothers in predicting child and grandchild longevity. In general fathers and mothers make equal contributions, but again with indications of a slightly greater weight for the parent of the same gender in predicting child longevity.

Patterns of Inheritance

The patterns of inheritance of longevity across various members of the family tree support the conclusion that longevity is largely genetically inherited. One striking feature is the equal importance of the men and women in predicting child and grandchild longevity, as in table 11.4. This is consistent with additive genetic transmission. But it is unexpected on theories of social causation, where men would be regarded as more important in determining the social status of families than women.

[63] Now is 22/9/2025.

Table 11.3: Correlation in Longevity, Siblings

Type	Count	Correlation	Standard Error
All	120,333	0.0717	0.0028
Brothers	41,234	0.0832	0.0047
Sisters	22,067	0.0946	0.0067
Mixed Gender	56,832	0.0545	0.0042

Table 4: Relative Contributions of Fathers and Mothers to Child/Grandchild Longevity

Child	Number	Father	Mother
All	54,013	0.0708	0.0660
		(0.0043)	(0.0043)
Son	31,995	0.0748	0.0612
		(0.0056)	(0.0056)
Daughter	22,018	0.0649	0.0721
		(0.0067)	(0.0067)
Grandchild	45,589	0.0230	0.0239
		(0.0047)	(0.0047)

Notes: Standard errors in parentheses.

Source: FOE database, births 1750-1929.

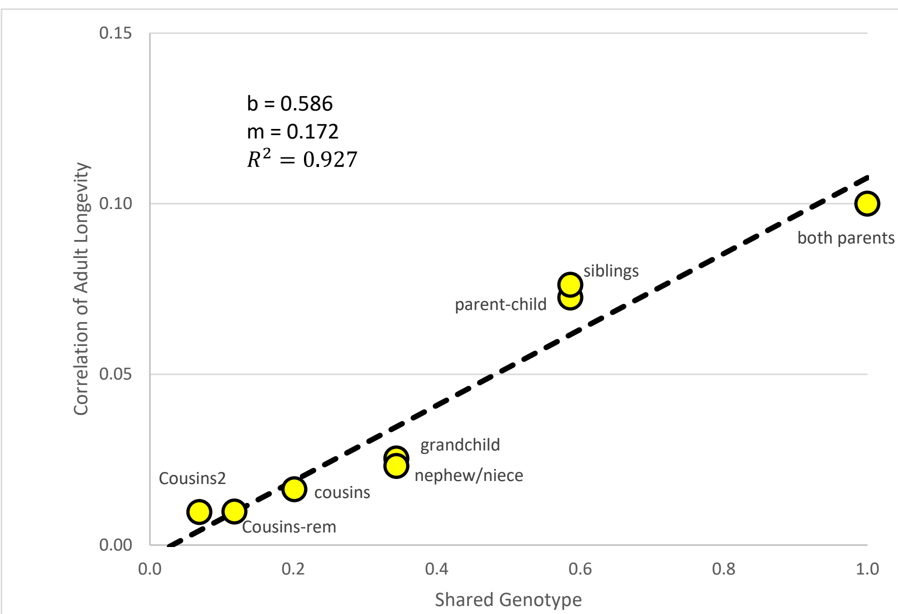
Table 11.5: Patterns of Inheritance of Longevity

Relative	Relative Genetic Distance	Observations	Average birth year	Correlation	Standard Error
Average of Parents	0	54,013	1869	0.100	0.004
One parent	1	117,587	1869	0.073	0.004
Sibling	1	114,912	1861	0.076	0.004
Grandparent	2	110,231	1872	0.025	0.004
Uncle/aunt	2	215,869	1874	0.023	0.003
Cousin	3	172,469	1867	0.016	0.004
Cousin Once Removed	4	320,682	1879	0.010	0.003
Second Cousin	5	225,508	1874	0.010	0.003

Notes: Standard Errors are clustered by the most recent common ancestor.

Source: FOE database, births 1750-1929.

Figure 11.2: Additive Genetic Model of Longevity Inheritance



Source: FOE database, births 1750-1929.

Table 11.5 shows the longevity correlations between various relatives down to those as distant as second cousins. If we fit this data to a model of additive genetic determination with parental genetic assortment in longevity, then the implied degree of genetic assortment by parents is 0.17, and the underlying intergenerational persistence of longevity is 0.59. Figure 11.2 plots the longevity correlations against the implied longevity genotype correlations of each related pair. The fitted additive genetic model explains 93 percent of the correlation variations, so it is a well-fitting description of the data.

The conclusion here is that parents match very loosely in terms of the genetics of longevity (0.17) compared to their underlying genetic match in terms of social outcomes such as occupational status and education, where the genetic correlation is implied to be close to 0.6. But given that longevity is only very weakly predicted by social status, which is what couples match strongly on, the much looser matching in terms of longevity is entirely explicable. These results match with the larger scale study of longevity in genealogies conducted by Kaplanis et al., 2018.

Social Causation of Longevity

There is strong belief in the social sciences that the social environment we find ourselves – our incomes, housing quality, job characteristics - has a strong causal effect on our health and longevity. This is referred to as "social causation" of health.

The claim is that if we could experimentally assign a group of 18 year olds to either a set of working class jobs, with associated living conditions, or a set of upper class jobs with again associated social environment, then those assigned the upper class jobs would live healthier and longer. On this basis, such august bodies as the American Psychological Association have called for policies to reduce socio-economic inequalities.

A popular proponent of the social causation approach has been the epidemiologist Michael Marmot. In path breaking studies of civil servants in Whitehall, England, who had similar types of desk-bound jobs, Marmot and his collaborators found that the higher the rank of a person within the Civil Service, the lower the incidence of heart disease and of mortality in general. The only seeming difference in their work environment was that those with a high rank gave orders, while those of lower rank took orders.[64]

This led Marmot to posit that a major contributor to socio-economic differences in health was the amount of autonomy a person had in their work. In his popular book, *The Status Syndrome: How your social standing directly affects your health and life expectancy*, he argues that if we make our workplaces and society more participatory and inclusive, it will significantly improve public health.[65]

But again, how do we know that the things that make people rise in the Civil Service do not also make them healthier? A review of studies that examined the correlation between socio-economic status and markers of stress, for example, found lower status was not consistently related to higher cortisol levels (Dowd et al., 2009).

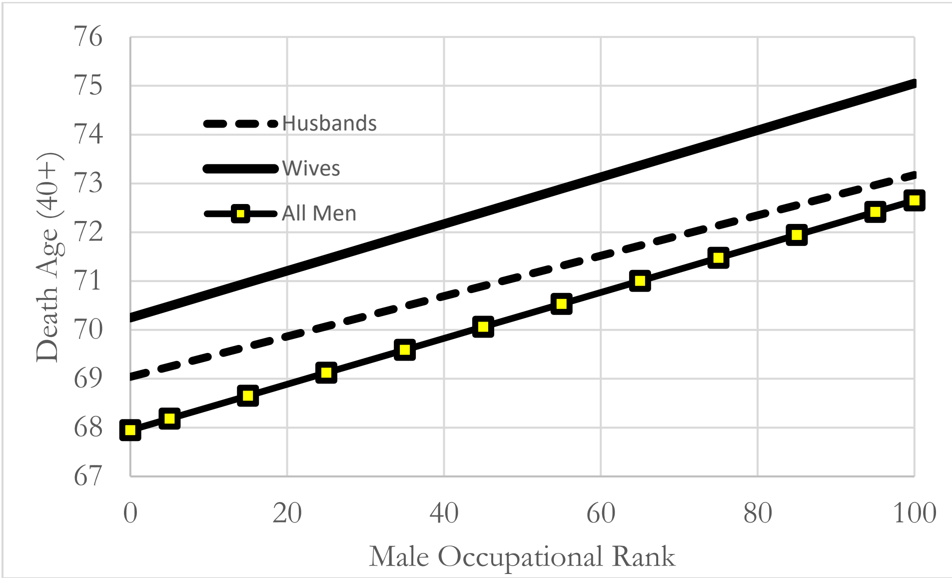
The literature establishing causal links between social status and health and adult longevity is modest and inconclusive, since exogenous variation in social status is rare. What literature there is has relied on such elements as compulsory schooling laws which extend schooling, changes in political boundaries, and quasi-accidental social promotions such as winning Oscars or Nobel Prizes (Lleras-Muney, 2005, Muzumdar, 2008, Arendt, 2005, Redelmeier and Singh, 2001a, 2001b, Rablen and Oswald, 2008, Kemptner et al., 2011).

In the FOE database for the birth interval 1750-1929 there is already a strong correlation between social status and adult longevity. Consider male occupational rank, which is measured at the closest measure to age 40, and is on a scale 0-100. There are some life cycle effects on occupational status. A person who dies at age 25 is unlikely to be an Admiral in the Navy as opposed to a Lieutenant, or unlikely to be a Judge as opposed to a Barrister. To ensure there is no induced correlation between occupational status and longevity, I thus include only men dying aged 40 and above in analyzing the connection between occupational status and longevity.

---

64 Marmot et al., 1978, Marmot, Shipley, and Rose, 1984, Marmot and Shipley, 1996, Marmot, 2004, Marmot, 2006.
65 Marmot, 2004.

Figure 11.3: Adult Longevity and Male Occupational Status, births 1750-1929



Source: FOE database.

Notes: The figure shows longevity for men and women attaining age 40 and above. Husbands are all men married before age 40. Wife status is proxied by that of her husband.

Figure 11.3 shows the fitted linear relationship between occupational status and adult longevity (40+) for men born 1750-1929, controlling for decade of birth. The highest occupational status men are estimated to live 4.7 years more on average after age 40 than the lowest status men.[66]

For women we do not get direct measures of occupational status, but we can proxy the family status for married women using the occupational status of their husbands. Looking just at husbands and wives who both live to at least age 40, we get the connection shown also in figure 11.3. Longevity for wives rises just as strongly with husband status as does longevity for the husbands themselves.

[66] The slope coefficient is estimated at 0.047, with a standard error of 0.003, so the connection is highly statistically significant. There are 40,442 observations.

This immediately casts significant doubt on the Marmot claim that for men it is the stresses of their work life that determines their longevity. For the wives in this birth interval, 1750-1929, would largely have been engaged in domestic occupations. These domestic occupations would have their own stresses, but there is no expectation that these would be perfectly correlated with those of their husbands at work. But given the strong status correlations seen in marriage, it is plausible that the wives showed the same health robustness as their husbands if that robustness was just a byproduct of social abilities.

Figure 11.3 also suggests that for men marriage itself by age 40 is associated with only a very modest longevity premium of about one year.

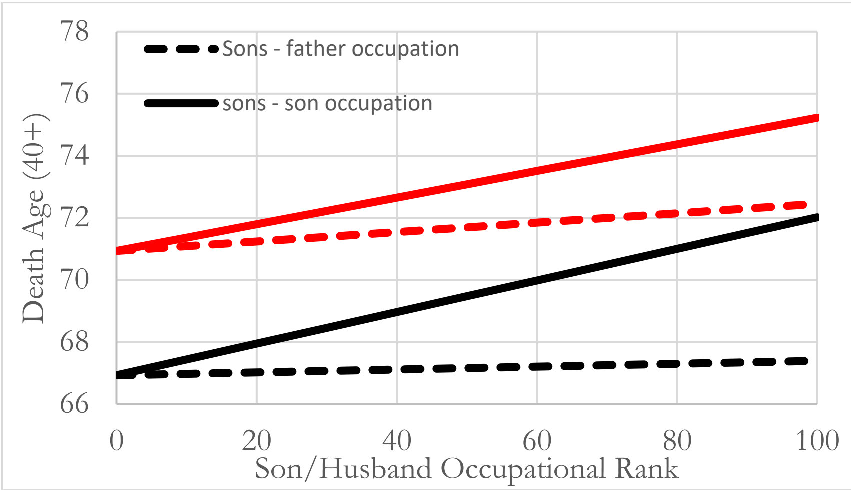
The longevity premium associated with higher occupational status has at least three potential sources.

> Since occupational status is inherited strongly, the gain may come from the better childhood environment of men with high occupational status. Since marriage is strongly assortative their wives will also potentially have benefited from a better childhood environment.
>
> The gains may come from the better shared adult environment experienced by men and their wives of high occupational status households.
>
> The longevity gains may stem from a greater health robustness of men who attained high occupational status, and an associated robustness of their spouses, who match closely in terms of social status.

The first two explanations would imply social causation of health, while the third would suggest just a selection into higher status occupations of healthier individuals.

To investigate the relative potential importance of childhood versus adult environment figure 11.4 portrays the connection between father occupational status, as well as that of sons and daughters, on longevity at age 40. Daughters' occupational status is proxied by their husbands, so the daughters included here are only those who married.

Figure 11.4: Occupation Rank and Longevity, Fathers, Sons, Daughters, births 1750-1929



Source: FOE database.

The lines in figure 11.4 for sons show how their longevity increases as a function of father occupational status, keeping son occupational status constant, and as a function of their own occupational status, keeping father occupational status constant. What these lines imply is that what matters for male adult longevity is the attained occupational status of sons, not their childhood environment linked to father occupational status. Once we control for son status, childhood environment has no impact on adult longevity.

For married daughters there is a very similar pattern, though father status has more effect on longevity and husband status less effect. One interpretation here would be that in both cases what matters to child adult longevity is the characteristic of the child, as proxied by their own occupational status (sons) or the occupational status of their husbands (daughters). Because husband occupation is a less good proxy than son occupation, father occupation in the case of daughters still supplies some addition information as to daughter characteristics, which is why the father coefficient is larger here.

Support for the conclusion that the childhood environment does not matter to longevity comes the case of brothers. In the FOE database we observe 25,787 unique pairs of brothers where we know both lifespan (above 40 years) and occupational status. When we estimate the effect of differences in brother occupational status on years lived by each brother, controlling now for child environment, we get a positive effect of status on longevity that is just as high as when we compare unrelated men. If child environment was an important determinant of adult longevity, then differences in status between brothers would have less effect on longevity than differences in status between unrelated men.

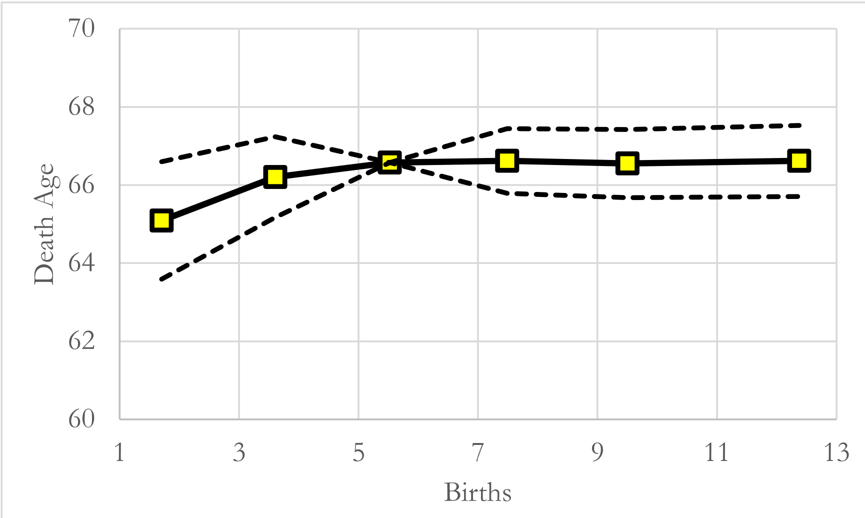
Family Size and Longevity, 1680-1879

As discussed in chapter 8, one random shock to child environment was family size for marriages formed 1780-1879. As discussed in that chapter there was no attempt in this period to limit fertility within marriage, so that the only significant determinant of marital fertility was for completed marriages was the age of the bride.

With larger family sizes children got less parental attention. Their material living conditions would be less good with the family income having to house, cloth and feed more young claimants. Here I test whether sibship sizes had any effect on adult longevity. To do this I concentrate on families where the wife was 21-25 at marriage, both parents lived to 50 and above, and both had only one marriage, so that almost all variation in family size was random. I measure size by births per family. If we measure family size by the number of children reaching age 21 then there is a positive and significant longevity increase with family size. Children from the largest such families, measure as numbers of adult children, live 3 years longer than those of families of 1-2 children.

Child longevity is estimated for family sizes 1-2, 3-4, 5-6, 7-8, 9-10, and 11+, controlling for child gender, birth decade, father occupational status, and the longevity of both parents. Child longevity is significantly and equally predicted by father and mother longevity. It is also significantly predicted by father occupational status. Figure 11.5 shows the results for the effects of family size on longevity. For families of between 3 and 18 births there is no connection between family size and child adult longevity. For families of size only 1 or 2 births the point estimate is of a lower longevity, though this result is not statistically significant.

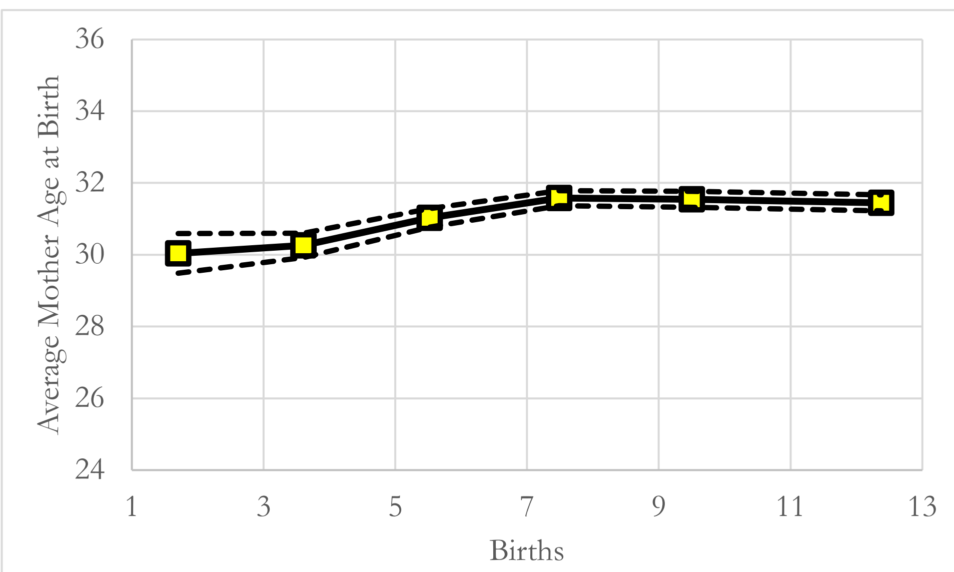
Figure 11.5: Child longevity and Family Size, marriages 1780-1879



Source: FOE database.

Notes: Horizontal axis is average births per family, grouped in cells of 1-2, 3-4, 5-6, 7-8, 9-10, 11+. Dashed lines show 95 percent confidence intervals relative to the group with 5-6 births.

Figure 11.6: Average Mother Age at Birth and Family Size, marriages 1780-1879



Source: FOE database.

Notes: Horizontal axis is average births per family, grouped in cells of 1-2, 3-4, 5-6, 7-8, 9-10, 11+. Dashed lines show 95 percent confidence intervals.

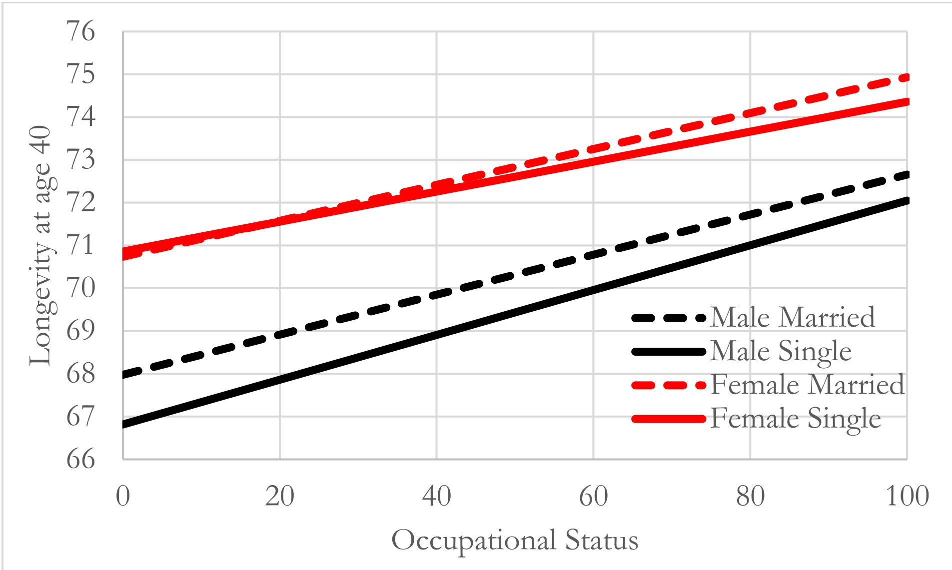
One possibility would be that larger families were created by mothers having longer reproductive spans, so that their children on average were born when the mother was older, and this conferred either social benefits from having older parents, or even biological benefits from having mothers and fathers who were healthy enough to remain fecund at older ages. However, as figure 11.6 shows, the rise in average mother age at childbirth was actually very small as total births increased. There was a rise from age 30 to age 31 as we go from 1-2 births to 5-6 births. But thereafter the age increase is minimal all the way from 5-6 to 11+ births. Thus the stable child longevity for families as we go from 3-4 to 11+ births is not going to be explicable by a parental age effect countering the social costs of larger families.

A number of modern studies confirm that childhood socio-economic status does not contribute to mortality risks in adulthood, once we control for adult socio-economic status. Thus studies in Finland and Sweden which predict either all cause mortality rates for adults, or mortality rates from cardiovascular disease, find that what matters is adult social class, and not social class of origin (Lynch et al., 1994, Tiikkaja and Hemström, 2008, Tiikkaja et al., 2009, Tiikkaja et al. 2012). Even studies which purport to show the importance to adult mortality rates of childhood conditions, on examination, are actually consistent with the claim that once we include adult conditions, childhood conditions no longer matter (Hayward and Gorman, 2004).

Marriage and Longevity

It is widely believed that married people live longer than those who are single. It is, however, debated still whether this is a causal effect of marriage, or just a selection into marriage of healthier people. Figure 11.6 shows for the FOE database the estimated longevity premium for those married before age 40 versus those unmarried at this age, for births 1750-1929. In the figure we control for occupational status, since we see above that this is the most important predictor of longevity. For this control we use father's occupational status, so that we have the same control for men as for women.

Figure 11.6: Longevity Premium from Marriage, births 1750-1929



Source: FOE database.

Notes: Occupation Measured as father's occupation.

For men there is on average a longevity premium of 1.2 years to those marrying before age 40, which is consistent across the occupational status spectrum. For women the premium averaged only 0.3 years, and is statistically insignificantly different from 0. This may reflect, however, that women in much of this period still faced risks of death in childbirth even in their 40s, which would fall mainly on married women.

Conclusion

Overall, there is no significant evidence, either in the FOE database, or the wider literature, for social causation of longevity. We see above that childhood social circumstances, which we would think are the years when health would be most subject to environmental conditions, do not influence longevity. Men who achieve greater occupational status get as much of a longevity gain compared to their brothers, with whom they shared a childhood environment, as compared to men in general. Children from large families, growing up in more crowded conditions, with poorer nutrition and less parental attention, show no less longevity than those from small families. Indeed if we look at family size measured by numbers of adult children there is a robust positive correlation of adult longevity and family size.

Marmot identifies longevity gains with being in control of your work conditions, but the wives of such men show almost as much longevity gain as they do, even in the era where high status women mostly had no occupations. So the longevity gain from status is not attached to work conditions.

Marriage by age 40 for women is associated with no increase in adult longevity. For men it is associated with a gain of less than a year in longevity. This result for men may be, of course, just the result of selection into marriage of healthier men.

This leaves a very narrow path for social causation to operate within. It has to attach just to adult social conditions in the home and community. To observe this having a causal effect we would need to observe adult men or women who for exogenous reasons saw a significant change in socio-economic status. English history since 1600, fortunately of unfortunately, does not provide many such social natural experiments.

12. Dead Relatives

There is an extensive literature in anthropology, economics, human ecology, psychology, and sociology on the role relatives and other social connections play in shaping the lives of children. These relatives include parents, siblings, grandparents, aunts/uncles, and cousins.

Recent studies make clear that in predicting the social outcomes for children, the status of relatives other than the parents is predictive, even when we control for the characteristics of the parents.[67] This is true for the FOE lineages. Controlling for father characteristics, the status of grandfathers, uncles, and cousins all significantly also predict child outcomes. If you have a high-status grandfather, then controlling for your father's status, you will have higher status yourself. The same holds for your uncles and cousins.

Why does the status of relatives other than parents predict outcomes in this way? One prominent school of thought in both anthropology and sociology believes that relatives play a direct causal role in the outcomes for children. In evolutionary anthropology, part of the reasoning behind this is the puzzle as to why women enter menopause long before the end of their natural lifespan.

A proposed solution to this puzzle is the "grandmother hypothesis." Grandmothers forego reproduction at later ages in order to improve the reproductive success of their children (Williams, 1957, Lahdenperä et al., 2004, Hawkes, 2004). Sarah Hrdy argued more generally that humans evolved as cooperative breeders — relatives other than the parents, such as the grandparents, and older siblings – all contribute to the rearing of children (Hrdy, 1999, Mace and Sear, 2004).

All such writings assume the importance of social causation in child outcomes. In this chapter, however, I show that at least in the context of England 1780 and later, there is no evidence that help from grandparents and other relatives was important to child outcomes.

[67] See, for example Chan and Boliver, 2013; Dribe and Helgertz, 2014; Ferrie, Massey and Rothbaum, 2016; Knigge, 2016; Møllegaarda and Jæger, 2015; Zeng and Xie, 2014.

Indeed even the loss of parents or grandparents in childhood was not associated with any significant loss of adult social status.

Below I consider the effects on children of losing parents when they are young, or of losing grandparents before their birth. The conclusion in all cases is that while such relatives provide information on the likely outcomes for children, whatever help and resources they provide does not significantly improve child outcomes. This finding that grandparent help has little effect on child outcomes is in line with a substantial recent literature that finds living grandparents do not predict any better outcome than do dead ones.[68] There is also a literature finds adverse effects of early parental death on adult mortality rates for children (Debiasi et al., 2021, Smith et al., 2014). However, here there is possibly a direct genetic causal link between parent and child.

Parental Death

Children in nineteenth century England often saw one or more parent die before they reached age 21. Thus of children born 1680 to 1929, who lived to adulthood, 19 percent had one or more parents die before they reached age 14. In part this was due to the marriage pattern in England where the average age of first marriage for women was 24, and for men 26. There was no restriction in fertility within marriage, so men and women were still producing children even in their 40s, in a world where life expectancy for adults was often only 65 or less. Further there significant risks to women of death in childbirth.

The other factor we saw in the chapter on longevity is that, even though there was a social gradient, at the individual level mortality was largely random.

If a child lost a father when young they faced growing up in a family with constrained income without a male breadwinner. If a child lost a mother then they lost the care and attention mothers provided.

[68] See Bol and Kalmijn, 2016, Braun and Stuhler, 2018, Ferguson and Ready, 2011, Kroeger and Thompson, 2016, Neidhöfer and Stockhausen, 2019. Only Zeng and Xie, 2014, reports substantially greater association between grandparent education and grandchild outcomes for co-resident grandparents.

To illustrate the disruptions parental death in childhood provoked, consider William Adamthwaite born in Brough, Westmorland, 11 days after his father's death in 1853. In the 1851 census his father was recorded as a tailor, prosperous enough to be employing an apprentice. William and his five siblings aged 0 to 11 had their mother to care for them only until 1858, when she also died.

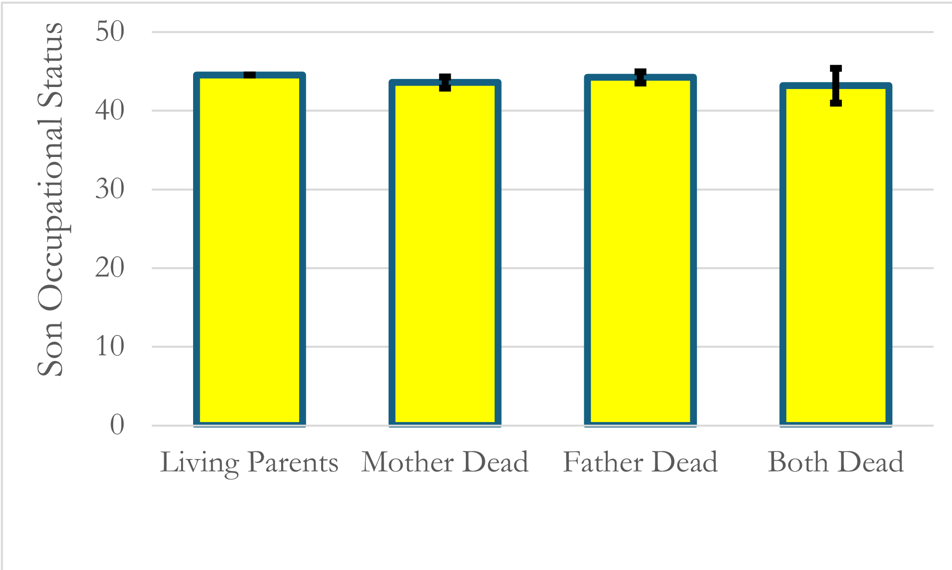
In the 1861 census the two oldest surviving sons were live-in farm servants aged 14 and 16, while the three younger children were living with the paternal grandparents (aged 71 and 73). The grandmother died by 1863, and grandfather by 1868. So by the 1871 census William had moved to Durham with an older brother, a coke burner, and was now, age 17, a coal miner. His other surviving older brother became a railway laborer.

All three sons in their young work years saw significant downgrading in occupational status compared to the father, though William himself eventually emigrated to Australia, where he elevated his status to Florist.

What was the general effect of early parental death on child social outcomes? If direct additive genetic determination was the overwhelming cause of social outcomes, parental death at young ages for children should have little negative impact on child outcomes. Alternative arrangements for child rearing, such as relatives, remarriage, or even the Workhouse, will do the job adequately. If social causation of child outcomes was important in practice, then the disruptions caused by early parental death should leave their imprint on the fortunes of children.

In looking at this with observational evidence we would expect a bias where lower status parents would be more likely to die younger. So if we find little effect of parental loss, then, a fortiori, it implies contact with parents mattered little to outcomes.

Figure 12.1: Dead Parents and Son Occupational Status, 1680-1929



Source: FOE database.

Notes: The figure shows son adult occupational status where both parents were alive at son age 14, only the mother was dead, only the father was dead, and both parents were dead. The bars shows the 95 percent confidence interval for each estimate.

One feature that does appear, however, is that children whose father died before they reached age 14 had fathers older at birth. Thus for children whose father was alive when they reached age 14, father age at their birth averaged 34 years, while for those with a dead father by age 14, their father's age at birth averaged 38 years.[69] However, looking within families we see no net effect of father's age at birth on son occupational status. Younger sons in families gain nothing in occupational status compared to their older brothers.

Figure 12.1 shows the effect of losing either your mother, your father, or both parents before age 14 for the occupational status of sons born 1680-1929. In this estimation father occupational status is controlled for. The first column shows the average adult occupational status of sons with both parents living at age 14. The average such status is 44.5 on a scale of 0-100. The second column shows the status for those losing a mother by age 14. There is here an estimated decline in status of 0.9. While this decline is statistically significant, it is of

[69] For mother's the respective ages were 30 and 31, much less of a gap.

extremely modest magnitude relative to average occupational status. For a dead father, the estimated decline is only 0.3. For both parents dying the estimated decline was 1.3. But neither of these last two estimates are statistically significant.

Using instead us a cutoff age of 10 for parent death, the results are very similar to those shown in the figure. Parental death, even at a very young age, had negligible effects on the adult occupational status of sons.

Perhaps what is happening here is that there were sufficient social safety nets, even in nineteenth century England, that the loss of one or more parents was largely compensated for. Children losing a parent young tended, as noted above, to be the younger members of sibships. Where mothers died, older sisters would often become surrogate mothers to these younger children. Where fathers died, older sons would become the family breadwinners. If both parents died, the married siblings of parents would take in the orphans. In other cases the mother or father would quickly remarry. The median time to remarriage for men with children who lost a spouse when they were under age 45 was just 2.4 years. For families forced into the workhouse by loss of a male breadwinner, the conditions were certainly uncongenial, but there was decent nutrition for children, and also educational provision.

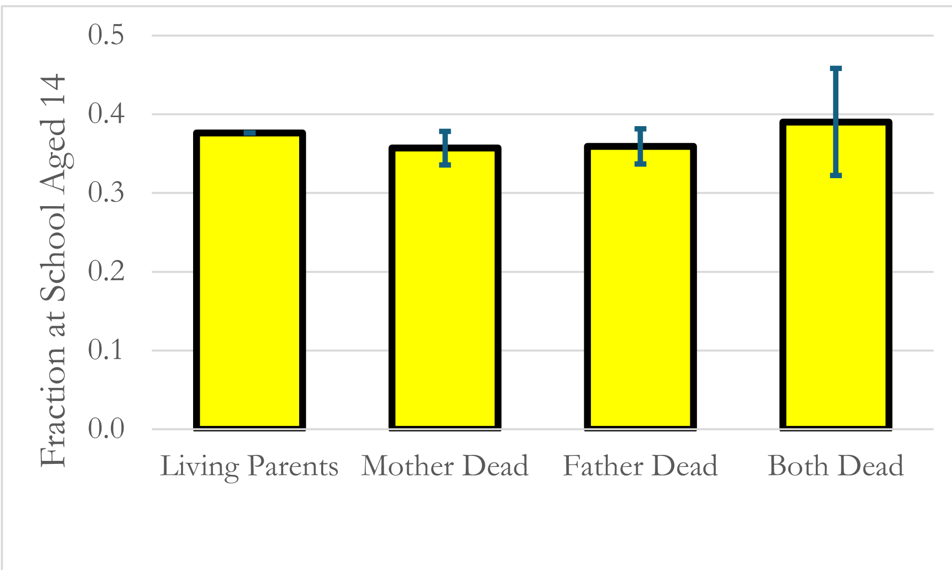
In line with this, there is evidence that parental death created surprisingly modest disruptions in the environment of children. This comes if we look, for example, at the probability a child was still in schooling at age 14, where here we include also girls. Figure 12.2 shows this probability, where there are controls for father occupational status, and for the year of observation.

As can be seen, there are signs of a slight decline in school attendance where a single parent has died. For the case of orphans, the estimate is of no effect on school attendance, though here because of fewer cases, there is a large error band. Thus for the orphans nothing can be said with any confidence on whether there was an effect of their loss on school attendance.

A similar exercise done looking instead at the probability a child was at work age 14 when observed in a census or population register 1851-1939 finds again modest effects. There was a slight increase in the work probability at age 14 for children who had lost a parent before that age.

Studies of the effects of parental death in childhood elsewhere in northern Europe are consistent with the results here. Dribe, Debiasi, and Eriksson (2022), for example, look at the effects of parental deaths in Sweden from the influenza pandemic in 1918-19 on child status in adulthood, measured by occupational status. The hope is that influenza deaths were an exogenous adult mortality shock. They conclude in their abstract "The patterns were similar for sons and daughters and do not support the theory that parental loss had important negative effects on socioeconomic outcomes in adulthood." But they summarize the state of the literature on the effects of parental death in "contemporary and historical Western contexts" as not finding consistent and important negative effects.

Figure 12.2: Dead Parents and Fraction of Children at School Age 14, 1851-1939



Source: FOE database.

Notes: The figure shows child school attendance at age 14 where both parents were alive at child age 14, only the mother was dead, only the father was dead, and both parents were dead. The bars shows the 95 percent confidence interval for each estimate.

Parental loss did not seem to cause too much disruption in the home environment of children. But if fathers had a significant social imprint on children, then their loss should have significantly reduced the correlation in occupational status between son and father. With their deaths early in a child's life fathers were being replaced by a stepfather, by an older male sibling, or by the workhouse. These new parental influences would be imperfectly correlated with the dead father in occupational status.

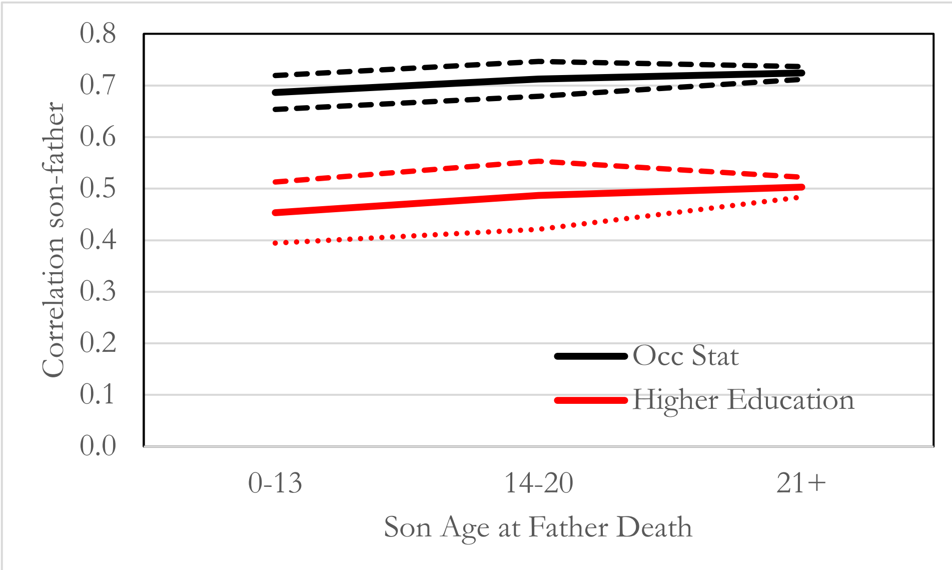
Mothers, for example, correlated strongly in social status in marriage with their male partners. But even with a 0.8 correlation of underlying status between marital partners, the implied social status correlation of a father with a stepfather would be only 0.6. Thus if fathers transmit social status through social means to their children, the early death of a father should significantly reduce the correlation in social status between son and biological father.

Figure 12.3 shows the correlation between son and fathers in occupational status, and in attainment of higher education, as a function of son age when the father died. For occupational status we see potentially a slightly lower son-father status correlation if the father died before the son reached age 14. But this potential effect is very small – an estimated decline in the correlation of 5 percent. The father-son correlation in this case declines from 0.72 to 0.69, but statistically it is possible that there was almost no decline given the confidence intervals.

For attainment of higher education there is again a slightly smaller father-son correlation estimated for the case where the father dies when the son is under age 14. But again the decline estimated is very modest, being a less than 10 percent decline compared to the correlation in cases where the father is still alive when the son attains age 21.

Figure 12.3 is thus consistent again with a society in which genetic transmission of status dominates, so that fathers remain strongly correlated in outcomes with their children, even when they have little or no social contact with these children.

Figure 12.3: Father-Son Correlation versus Son Age at Father Death



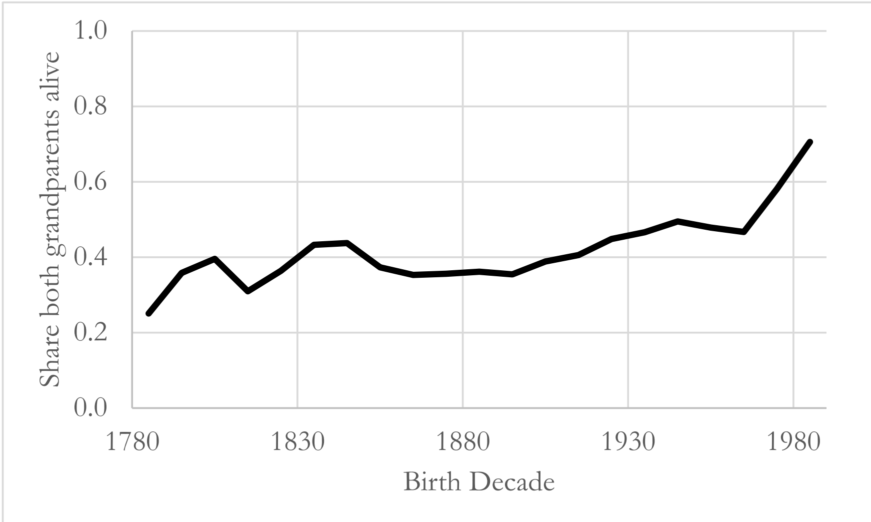
Source: FOE database.

Grandparents

Even controlling for parent status, collateral relatives – grandparents, aunts/uncles, and cousins – always provide more information on child outcomes. This has been widely interpreted as showing that there is a causal connection between collateral relatives and child. As noted this connection has been claimed to explain why women live long after the end of their reproductive careers.

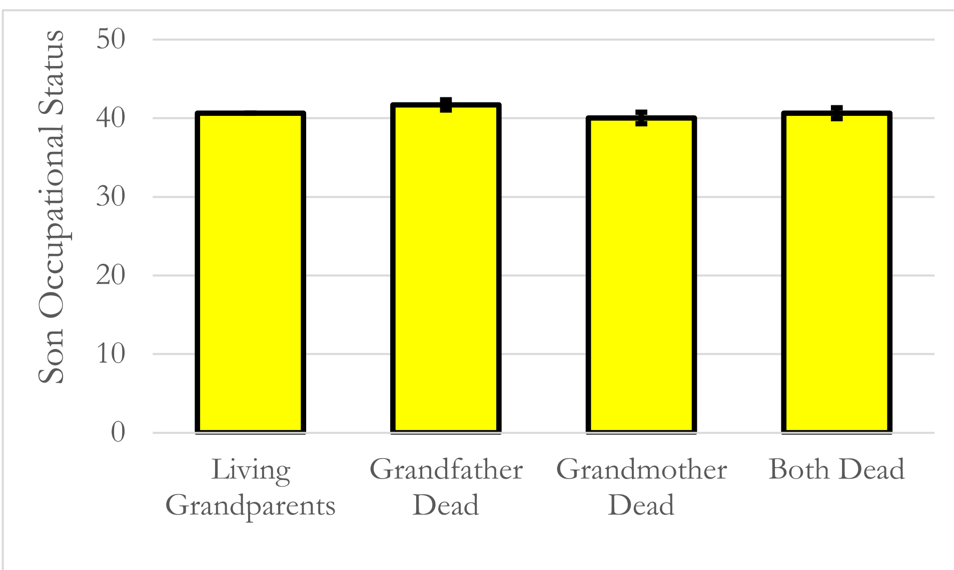
In this section I show two things. First having living grandparents in England does not create any benefit in terms of child outcomes. Second grandparents in England provide as much information on child social outcomes when they are dead at the time the child is born compared to when they are alive.

Figure 12.4: Fraction of Children with both grandparents alive at their birth



Source: FOE database.

Figure 12.5: Son Occupational Status at age 40 as a function of grandparent status



Source: FOE database.

Notes: Occupational status of 15,974 sons who lived to at least age 21, as a function of grandparent status at their birth, controlling for father occupational status. 95 percent confidence intervals shown by bars.

Figure 12.4 shows for the families in the FOE database the fraction of children by each birth decade with both paternal grandparents alive at the time of their birth, for those born 1780-1979. The share with living paternal grandparents is around 40 percent for most of this interval, but rising in the 1920s and later as life expectancies increased.

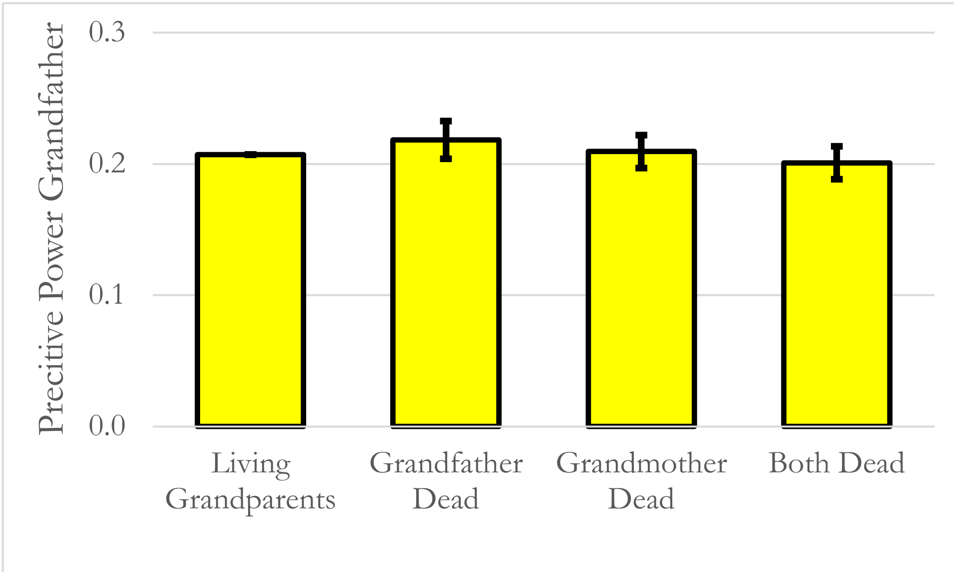
Did living grandparents improve child outcomes? To measure this we first concentrate on sons and their attained occupational status around age 40. We divide sons into four categories, depending on whether their paternal grandfather or grandmother was alive at their birth. We then measure average occupational status for sons for each category, controlling for father occupational status. Figure 12.5 shows the results. There is essentially no effect of living versus dead grandparents on son occupational status. There is actually estimated a slight positive effect from having a dead grandfather (but living grandmother). But the quantitative magnitudes, as seen in the figure, are inconsequential. Having living grandparents provides no benefits to grandsons.

Even when we control for father occupational status, grandfather occupational status has strong predictive power for grandson occupational status. On a genetic interpretation of the transmission of status this just represents the fact that grandson outcomes are dependent on the father genotype, which is only imperfectly revealed by the observed father phenotype.

Grandfather social status, however, provides further information on the father genotype. A father with high status, whose own father also had high status, is more likely to have a high status genotype. A father with high status, whose own father had low status, is more likely to have a lower status genotype than is their phenotype. But importantly, if direct genetic transmission is the key to social status, the predictive power of grandfather status should be independent of any interactions between grandfather, grandmother and grandchild.

Figure 12.6 measures the additional predictive power of grandfathers' status for grandson occupational status, after controlling for father status, for four cases: both grandparents alive at grandchild birth, grandfather dead but grandmother alive, grandmother dead but grandfather alive, and both grandparents dead. As is predicted by direct genetic transmission dominating, grandparent interactions with grandsons makes no effective difference to the predictive power of grandfather status for son occupational outcomes.

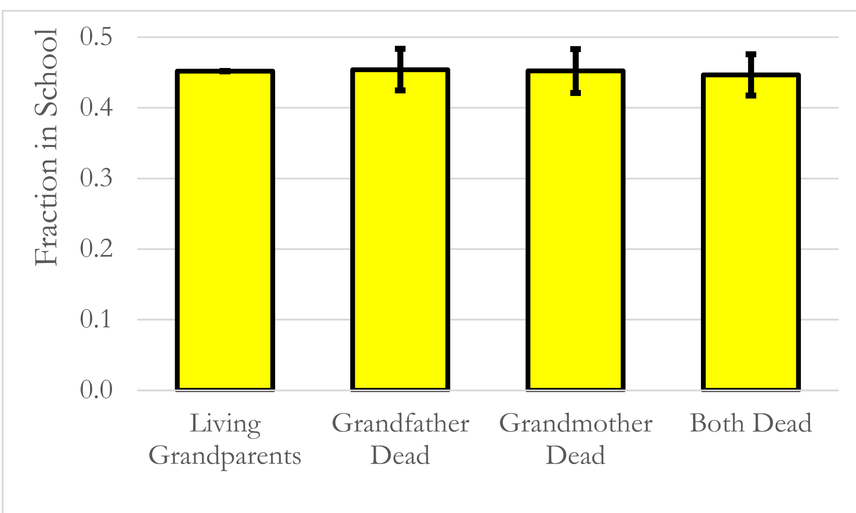
Figure 12.6: Predictive Power of Grandfather Occupational Status for Grandson Status



Source: FOE database.

Notes: Limited to cases where grandparents die in the same region of the country as grandchild born.

Figure 12.7: Fraction at School Age 14 and Grandparent Status



Source: FOE database.

Notes: Schooling observed in the censuses 1851-1921, and the population register of 1939. Controlling for father occupational status.

The figures above concern just male outcomes, but the same irrelevance of grandparent contact to grandchild outcomes appears when we look at schooling ages 10-18 for both sons and daughters, as in figure 12.7. Grandchildren with living paternal grandparents at birth had a 45 percent chance of still being in schooling at age 14 when observed in 1881. Grandchildren with one or more grandparents dead when they were born show no significant decline in school attendance compared to those with living grandparents.

Conclusions

Even controlling for the status of parents, the characteristics of other relatives – grandfathers, uncles, and cousins – all are predictive of social outcomes for children. There is a significant literature in anthropology, sociology, and economics that ascribes a causal role to these associations. In this chapter I test for causal effects by looking at two things. First, controlling for father occupational status, do children do better if they have living grandparents? Second, do children with living grandparents correlate more strongly with the grandparents in social status? Do living grandparents have more influence on their grandchildren? I do not detect either effect. Dead grandparents without a causal pathway to influence grandchildren are associated with just as good grandchild outcomes, and correlate as closely in terms of grandchild outcomes.

These results are consistent with a model of status transmission across generations where the process is actually first-order Markov, but the observed social outcomes in each generation are affected also by transitory elements, so that the status of other relatives conveys information about the likely outcomes for children independent of the status of the parents. One variant of such a process would be the case where status transmission from parents to children is through genetics.

13. Wealth Shocks

We have focused above on the meritocratic elements of English society 1600-2026. The argument is that between 1600 and 2026, England was largely a meritocracy. People were born with social abilities, and were able to exercise these largely independent of the circumstances of their birth.

Wealth, however, complicates this picture. It is both created by entrepreneurial activity, but also passes between generations, not based on merit, but by rules of social connection. Accidents of demography, how many children people had and what gender they were, determine who inherits what. In the FOE database for the years after 1858, those leaving more than average wealth at their death can be estimated to have typically received nearly 80 percent of that wealth as bequests.

In this chapter I look at wealth shocks from demographic accidents that made the children of families of the same parental wealth richer or poorer. How persistent was such wealth acquisition across future generations. In effect, how much of modern wealth holding can be attributed to wealth gains people made many generations in the past? In the form of wealth, does the past live on today in a way that is very different from genetic transmission?

There has been considerable attention recently to the persistent effect of wealth shocks - sudden unexpected gains or losses in wealth – in terms of the wealth and general social status of later generations. Such shocks, even occurring hundreds of years ago, in popular imagination, explain current differences in wealth and other social outcomes.

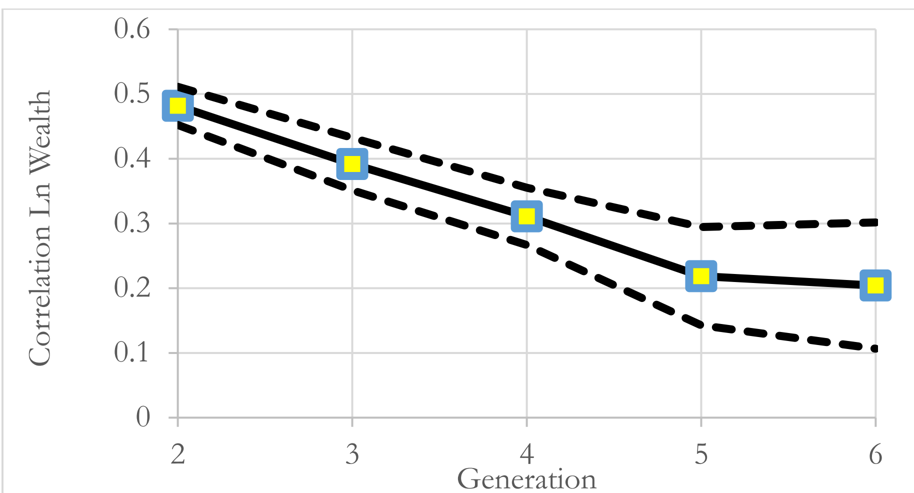
Thus the compensation paid to British slave owners with the emancipation of slaves in the British Empire in 1834 has been argued to be still a foundation of the wealth of many richer Britons now, including David Cameron, the former Conservative Prime Minister (Hall et al., 2014). Even cultural institutions, such as the Booker Prize for literature, were alleged to have been ultimately funded in part by 1834 slave compensation (Creamer, 2024).

But this raises an interesting question. If we in 1834 were to give £100,000 to a randomly chosen British person, and then did an audit of their descendants, 190 years (6 generations) later, would we find them wealthier than the average person in 2026? Do shocks to wealth persist across generations? Or does unexpected, unearned wealth dissipate quickly, so that, for example, none of the Booker Prize money can be attributed to slavery, and none of David Cameron's assets face that same taint?

If we look at the intergenerational persistence of wealth among men in England 1700-2026 then we do observe an even stronger persistence of wealth that for other attributes such as occupational status or education.

Thus if we look at ln wealth at death across six generations in England, using the Families of England (FOE) database, where the average birth year of the first generation was for men born 1769, and for the sixth generation men and women born 1924, then the correlation was still 0.2. Even for a modest sample of 539 such cases, the correlation after five generations was highly significant statistically.[70] Figure 13.1 shows the correlation with the initial wealth in generations 2 to 6.

Figure 13.1: Wealth Correlations across 6 generations, men, England 1700-2026



Notes: Wealth measured as ln wealth at death relative to average for the death decade. The dotted lines show the 95% confidence interval. The initial set of fathers had above average wealth at death, and married before 1880.

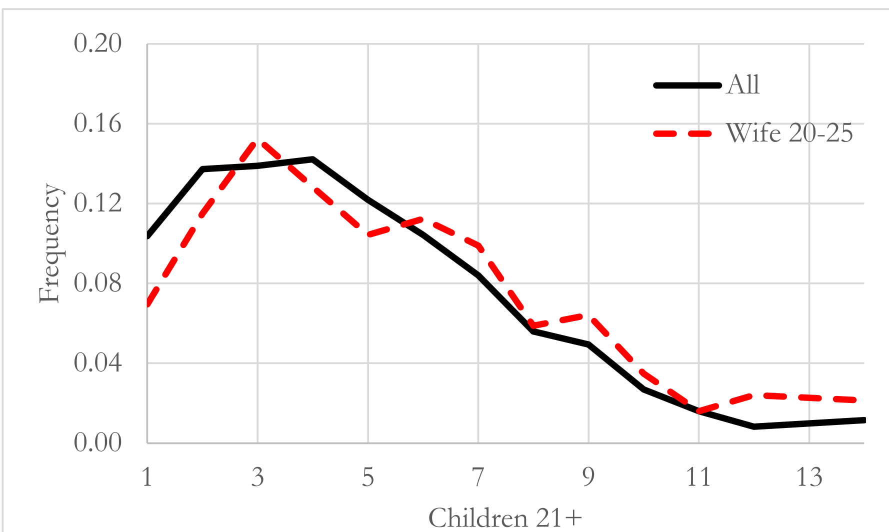
[70] Ln wealth is used since wealth is so skewed to make the wealth measure have a distribution closer to normal.

Wealth Shocks from Family Size

In this chapter I examine the effects of random shocks to wealth that were produced by the accident of sibship size. In chapter 8 I showed that family size was essentially random in England for marriages 1780-1880. The only significant predictor of family size was wife age at marriage. But even that explains a very modest fraction of the variance of family size.

So we can observe what are random wealth shocks in England in the years before 1880, and measure their effects on future generations, again using the Families of England (FOE) database. Figure 13.2 shows for 2,148 marriages before 1880, where the father left an estate at death of above average value, the distribution of family sizes, measured as the number of children who attained at least age 21.[71] Notice the enormous dispersion of family sizes, looking at adult children. In this biological lottery some marriages produced just 1 child, while others with parents who were very similar in wealth and social status, produced 10 or more children. As we shall see below these shocks to family size created substantial shocks to child wealth relative to fathers.

Figure 13.2: Family Sizes among Testators (Wealth ≥ 1), England, marriages pre-1880



Notes: The figure shows the distribution of family sizes, measured as children reaching age 21, for male testators of above average wealth whose first marriage was pre-1880. The dotted line shows the size distribution for men whose wife was aged 20-25 at the marriage. *Source*: FOE database.

[71] Families without any adult children were not included.

Figure 13.2 also shows the distribution of family sizes for men whose first wife was 20-25 at marriage, where the husband was 22-28. As can be seen the range in family size even for this restricted set of marriages in terms of spousal age was nearly as great as the overall range in sizes. Shocks to child wealth for families with wealth were largely random.

Inheritance of family size

Did fathers with large families produce children who also had large families? In that case the effects of family size on wealth in the initial generation would be further magnified in later generations.

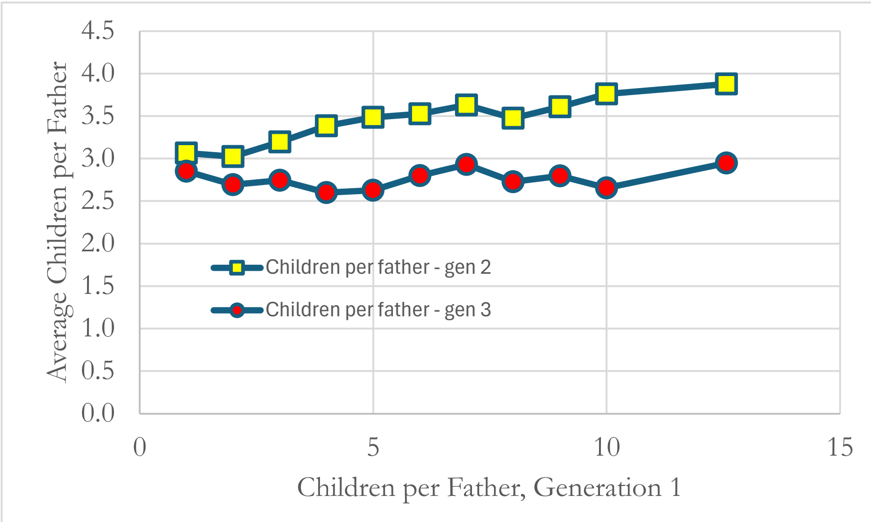
The correlation in family sizes for fathers and sons was actually very low, illustrating the largely random source of family size variations. Thus in the sample of fathers marrying first before 1880, and with wealth at death above average, the correlation between their family size and that of their sons was only 0.032.[72] However, all the fathers in the first generation have at least one adult child, while those in the second can have 0 adult children. The correlation between fathers and sons in numbers of adult children, where only sons with one or more children are included, rises to 0.093, but is still very low.[73] These results are consistent with a meta-study of fertility transmission across generations, which found that pre-transition populations generally had weak fertility inheritance (Murphy, 1999, figures 1-2).

By the third generation, the correlation between grandfathers and their grandsons in numbers of adult children becomes insignificantly different from 0, even when only including grandsons with at least one adult child. Thus there is a slight echo in the second generation of large or small family size, but no echo by the third generation. Family size really is a transient influence on wealth.

[72] Though with a standard error of 0.012, this is statistically significant at the 1% level. Size measured as numbers of children attaining at least age 21.

[73] With a standard error of 0.015 this is again highly statistically significant.

Figure 13.3: Adult Children per Father, by Generation



Notes: The figure shows the number of adult children per father in the patriline in the second and third generations as a function of family size (adult children) in the first generation, counting only fathers with at least one child. The initial set of fathers were those with above average wealth at death, marrying before 1880. *Source:* FOE Database.

Figure 13.3 shows this effect graphically. The horizontal axis shows the number of adult children from each father in the first generation. The first line shows the number of adult children on average from each of their sons who had adult children. This line has a positive slope of 0.073. If the grandfather went from 1 adult child to 11, their sons would be expected to have 0.73 additional adult children. So there is some echo of previous generation fertility in the first generation. But by the third generation, net fertility is the same across all men, independent of what their grandfather's fertility was.

Figure 13.3 implies the negative shock to wealth created by many children in an initial generation will be transmitted across subsequent generations largely unchanged. A man with few children will typically have few grandchildren and few great-grandchildren. A man with many children will have many grandchildren and great grandchildren. If inheritance is the main source of wealth, the grandchildren and even great grandchildren of a wealthy man with few adult children in a marriage before 1880 will have greater wealth, all other things being equal, than the grandchildren or great-grandchildren of a man with many children.

Family Size and Child Wealth

For families with father's bequeathing wealth, family size, the number of adult children, had a significant effect on average child wealth relative to the father. When there were multiple heirs, birth order had only modest effects on the amount bequeathed to each child. The oldest son was typically about ten percent wealthier at death than his younger brothers. Sons, however, seem to have been bequeathed more than daughters. Thus family size, as well as the gender composition of siblings, was an important determinant of wealth inheritance for all children. The huge variation in family size implies that for equally wealthy fathers, their children would inherit very different amounts.

To illustrate this, consider Edward Cazalet (1827-1883), who married Elizabeth Marshall in 1860, when Elizabeth was 23. Despite her youth, the marriage produced just one child, a son. Edward Cazalet died with wealth £0.35 m, and that one son with wealth £0.65 m.[74]

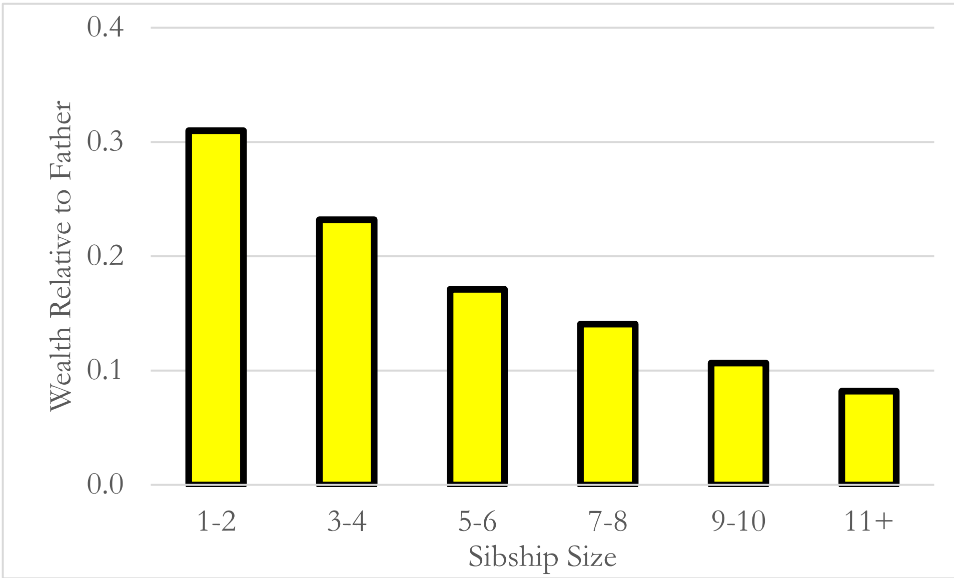
In contrast Richard Thomas Pulteney (1811-1874) married Emma Dalison in 1845, and fathered fourteen children, twelve of whom outlived him. His wealth at death was £0.5 million. Those twelve children left an average of only £0.03 m. each, £0.3 m. in total. For families with wealth, size mattered for wealth per child.

Figure 13.4 shows the median the wealth of children relative to their father versus adult sibship size. Because the ratio of child wealth to father is very asymmetric, and ranges from 0 to 302, the median rather than the mean is used. Clearly sibship size strongly influences median child/father wealth. Relative to the father, child wealth at sibship sizes 11 and above was less than a third that for sibship sizes of one or two.

Daughters overall in families with above average wealth showed 45% less wealth than sons. If we look just at the richer fathers, daughters on average had 60% less wealth at death than sons. In looking at the long run effects of these demography driven wealth shocks we will be following the patriline in families. In the FOE database complete fertility in each generation is only recorded for the patriline. Fortunately, the same effects as in figure 13.4 of family size and child relative wealth appear if we concentrate just on sons as heirs.

[74] These are all the nominal amounts reported.

Figure 13.4: Median Child Wealth at Death relative to Father by Family Size, marriages pre 1880, father wealth ≥ 1

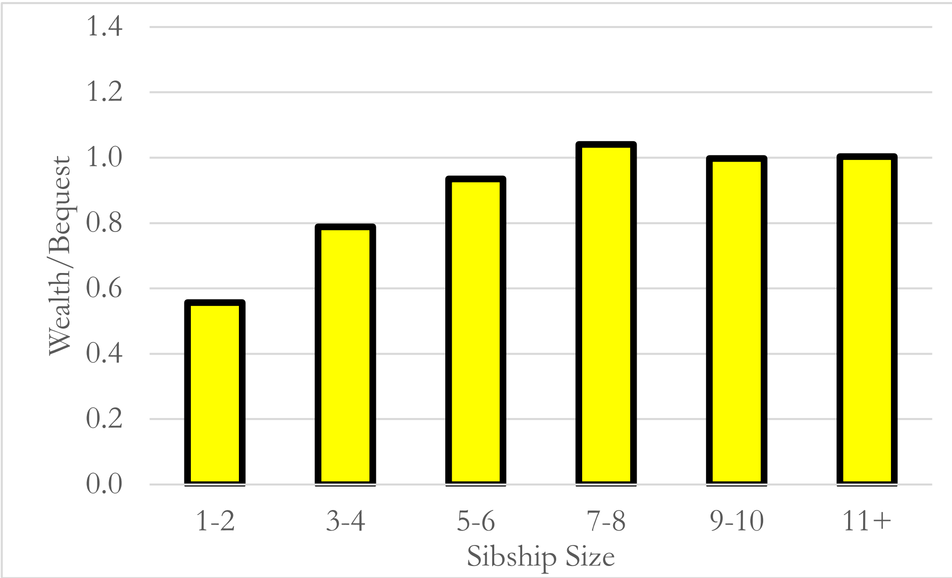


Notes: The horizontal axis shows the sibship size. The vertical axis shows median child wealth at death by sibship size, relative to father wealth, for fathers with wealth greater than average.

Source: FOE database.

Figure 13.5 illustrates why sibship size had less effect than expected on wealth. The figure shows the median ratio of wealth at death of children to their expected inheritance, by sibship size. If this ratio is less than one then the child dies with less wealth than their expected inheritance. If the ratio is greater than one the child dies with more than their expected inheritance. For adult sibship sizes of 1-6, children die with less wealth than their expected inheritance from their father. But for sibships of 7 or more children die with as much or more wealth than inherited from their father. In smaller families, where children inherit more, they consume more of their inheritance. In larger families, since average wealth levels were increasing over time, children were typically accumulating wealth above what they inherited.

Figure 13.5: Median Child Wealth versus Expected Inheritance, by Sibship, Father Wealth ≥1



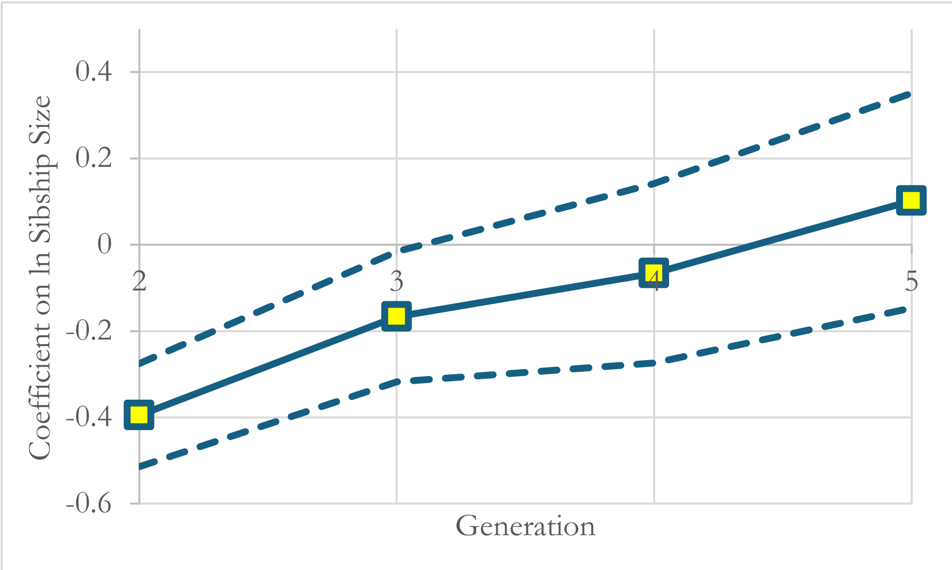
Notes: Median child wealth/expected inheritance, by adult siblings, fathers marrying pre-1880. Children in sibships of 1-6 have less wealth than their expected inheritance. Children in sibships 7 and above have equal or more. *Source*: FOE database.

Wealth Shocks Across Multiple Generations

How long do the shocks to wealth created by family size in the first generation last across subsequent generations. To estimate this I look at wealth of children in generations 2-5 relative to the wealth of the patriarch in the first generation, but also relative to how many children were inheriting in the first generation. What is the effect of that family size shock in generations 2-5? If inheritance is the dominant source of wealth, then that initial family size shock will persist across multiple generations.

Figure 13.6 shows the effect of sibship size shocks on child wealth at death, controlling for father wealth. In the second generation, the children, a move from sibship size four to five was associated with an 8 percent decline in wealth for each sibling. This effect is statistically significant. Note that if child wealth was determined just by their share of a fixed inheritance then we would see a 20 percent decline in child wealth as sibship moved from size 4 to 5.

Figure 13.6: Effects of Sibship Size on Wealth Across 2-5 Generations



Notes: The figure shows the estimated effect of ln(sibship) on ln(wealth) for descendants in generations 2-5 of men marrying pre-1880 and dying with above average wealth. The dotted lines show the 95% confidence intervals.

In the third generation the effect of sibship size in the previous generation declines by a half. Now a move from sibship size four to five was associated with only a 4 percent decline in wealth for each sibling. By the fourth and fifth generations, the estimated effects of sibship size two generations earlier on wealth at death have become zero. There is still some uncertainty about the size of these effects, reflected by the 95 percent confidence intervals in the figure. But the smooth transition from significant effects in generation 2, to modest effects generation 3, to no effect in generations 4-5 suggests that the disappearance of size effects is real.

Figure 13.6 shows what happens with initial fathers having above average wealth. If we were to draw the same figure but include only in the initial generation father's with wealth at least 10 times average wealth, then the results would look very similar. So even in the very rich there is indication that wealth shocks from sibship size dissipate within 3 generations.

Implications

The estimates above suggest that, at least for England post 1880, pure wealth shocks – random wealth changes not associated with family characteristics – dissipated completely within three generations. Shocks to wealth from sibship size, occurring to the children of a testator, were still observable in the grandchildren, though of much smaller size, but disappeared by the time of the great-grandchildren.

There is still a strong correlation of wealth between men and their great-grandchildren (as in figure 1). This implies that the main mechanism of wealth transmission across generations, again in the case of England, is not the actual physical transfer of wealth. If what mattered was just inheritance of wealth, then the demographically induced wealth shocks observed from marriages before 1880 would persist even beyond five generations.

This has two implications. First, for English families with wealth now, even where we can trace that wealth back through inheritance to the nineteenth century, there is no causal connection between their nineteenth century inheritance and current wealth. The wealthy in England are typically distinguished from the rest of the population not just by the accidental creation or inheritance of money. In the famous, but fictive, exchange where F. Scott Fitzgerald notes "The rich are different from you and me" and Ernest Hemingway responds "Yes, Scott, they have more money", Fitzgerald is the truth teller.[75]

The second implication is that, again for England, wealth holdings of individuals stem largely from their social and economic abilities. It is this which links them strongly in figure 13.1 across six generations to the wealth of their great-great-great-grandfathers, not the actual wealth that they have inherited from those forbears.

There has been debate in the literature about what share of wealth in modern society in each generation is inherited, with estimates varying from 8% to 89% (Atkinson, 2018, Kotlikoff and Summers, 1981, Ohlsson, Roine, and Waldenström, 2014, Piketty, Postel-Vinay, and Rosenthal, 2014). The argument here is that these high estimated shares likely

[75] This exchange was mistakenly reported by the literary critic Lionel Trilling, in a review of a collection of Fitzgerald writings in 1945.

depend on the social abilities of inheritors strongly correlating with bequestors. If wealth was randomly bequeathed to a set of inheritors, then much more of that wealth would dissipate between each generation, the share of wealth explained by inheritance would decline, and the overall wealth stock would diminish.

14. Linearity of Inheritance

Continuous traits inherited by additive genetic means show a characteristic linear pattern of inheritance, illustrated well by height. In this pattern there is consistent regression to the mean all across the range of the parent outcomes. This regression towards the mean is the same at the top and bottom of the distribution. A person who has parents 10 cm above the mean falls in height on average 2 cm compared to parents. A person who has parents 10 cm below the mean gains in height on average 2 cm compared to parents. Social mobility looks the same at the bottom and the top of the distribution.

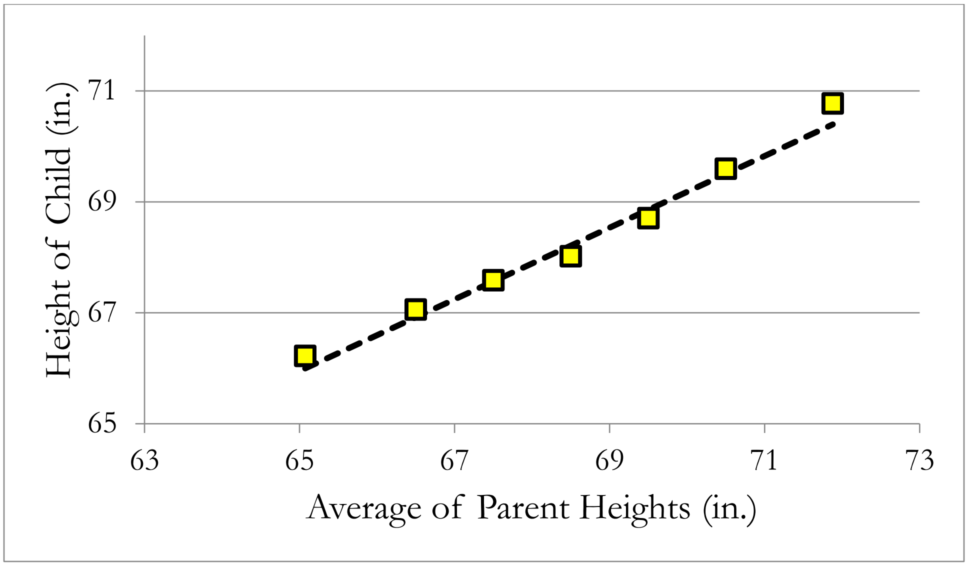
Figure 14.1 shows this pattern for Francis Galton's famous study of the connection between the heights of parents and children. Presented to the Royal Society in 1885, this study introduced the concept of regression to the mean.[76] The figure shows Galton's 928 observations, grouped into seven clusters of average parent heights (at one-inch intervals for the five central groups). Also shown is the best linear fit of the 928 observations. All across the parent height distribution, the observations lie close to this linear fit. Children of short parents approach the mean in the same way as children of tall parents.

With socially mediated forms of inheritance there is no such expectation of linearity across the distribution. There would be the possibility, for example, of poverty traps and privilege traps. Parents at the extreme of the distribution of social status could tend to transmit their poverty or privilege more strongly than those in the broad middle of the distribution. The social groups the poorest parents occupied would isolate their children from educational and other opportunities for upwards mobility. In contrast, for elite parents, their social milieu would reinforce the privilege of their children.

This would lead to the pattern of inheritance shown in figure 14.2. There would be greater persistence at the extremes of the distribution, and greater social mobility in the center.

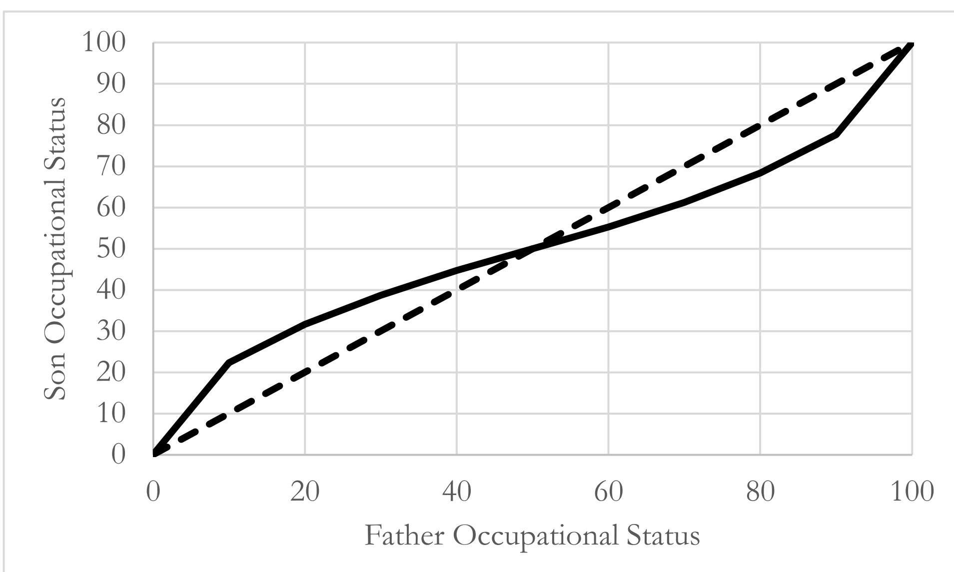
[76] Galton 1886. The data is from http://www.math.uah.edu/stat/data/Galton.html.

Figure 14.1: Height Inheritance, Galton 1886



Source: Clark et al., 2014, figure 7.5.

Figure 14.2: Poverty and Privilege Traps



Notes: Dashed line shows son inherits father status. Solid line shows pattern of inheritance with poverty and privilege traps.

With social causation the symmetry of the movement to the mean from high status and low status families is also surprising. The social conditions of the high status families are very different from those of low status families. So why would the social processes that lead to regression to the mean have the same effect at one end of the distribution as at the other?

In the FOE database we have a number of continuous outcome variables: occupational status and age at death for those born before 1929, as well as house value and the index of multiple deprivation for 1999-2026. Do these show the expected linearity of transmission which would characterize genetic transmission?

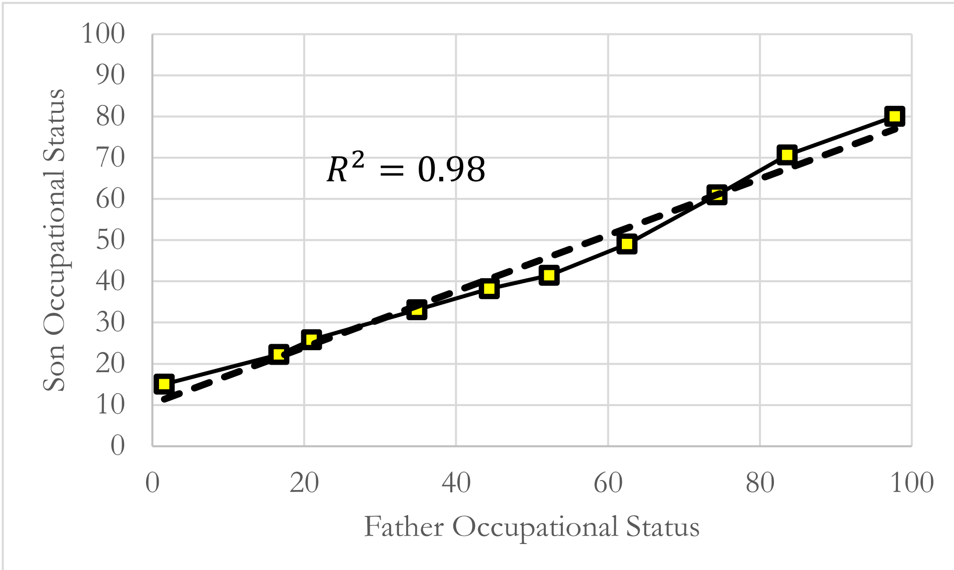
Figures 14.3 and 14.4 show the patterns of inheritance of occupational status by sons compared to their fathers, from wedding records 1837-1879 and 1880-1939. In both cases the fathers have been grouped by each decile of occupational status, which ranges 0-100. Also shown is the best fitting line to these decile averages.

Both figures show a linear relationship between father and son status. There is no sign of any excess persistence at the top of the occupational scale, or at the bottom. The linear fit to the data is so good that it explains respectively 98 percent and 99 percent of the variation in son outcomes in each period.

Figure 14.5 shows the connection between average parental longevity and child longevity, for births in the years 1700-1929. Longevity is measured for each birth decade in standard deviation units difference from the mean of that gender and birth decade. Only those living to age 21 and above are included in this calculation.

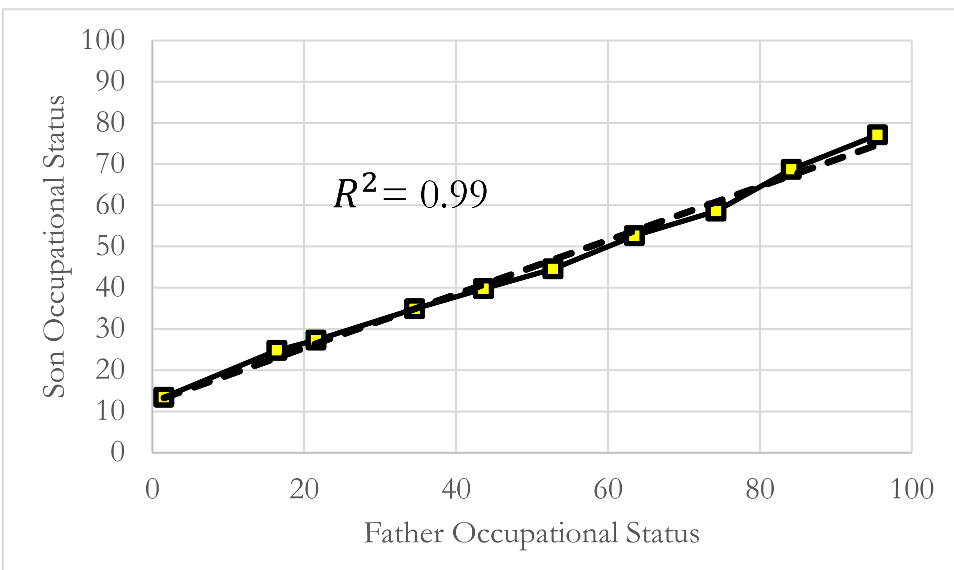
As noted in chapter 11 there is in fact very weak inheritance of longevity. This is reflected in the fact that for every gain of one standard deviation in longevity by the parents, child longevity increases by only 0.16 of a standard deviation. Yet again this relationship is linear. 90 percent of the variation in child longevity, by parent longevity decile, is explained by a linear fit. Again there is no sign of any greater persistence of higher longevity by children from the top parental decile, or of lower longevity from the bottom parental decile.

Figure 14.3: Occupational Status, Father-Son, Marriages 1837-79



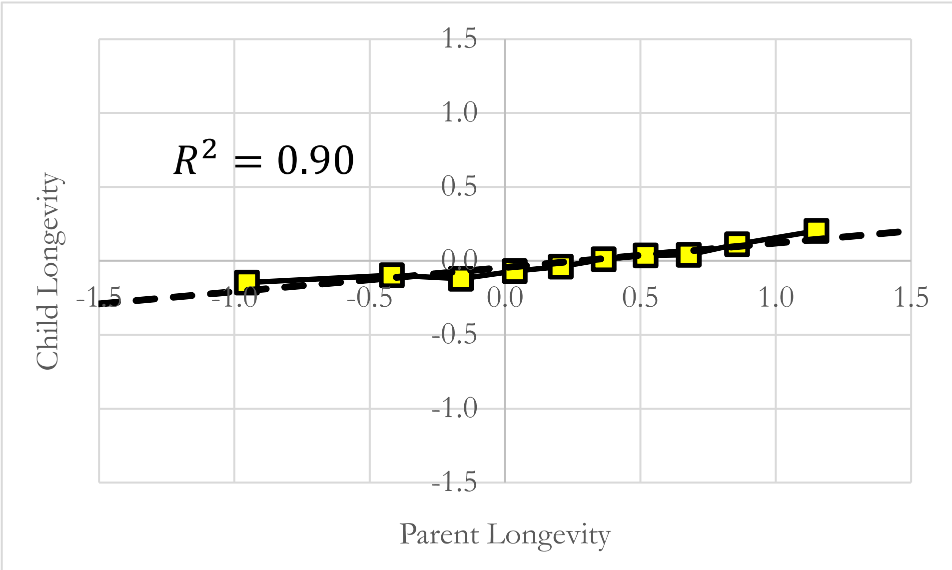
Source: MOE database

Figure 14.4: Occupational Status, Father-Son, Marriages 1880-1939



Source: MOE database

Figure 14.5: Longevity Inheritance, births 1700-1929

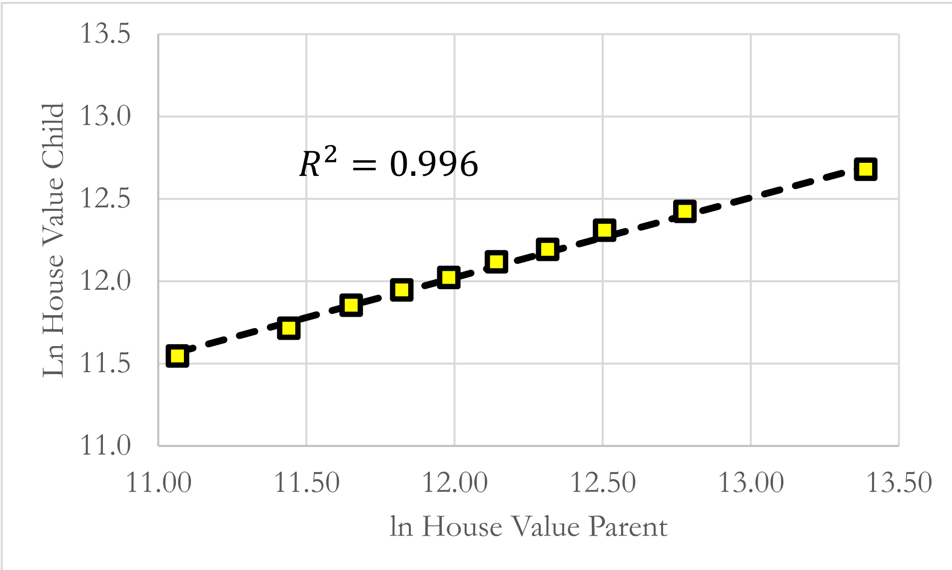


Notes: Longevity of every individual measured in standard deviation units by gender and birth decade. Horizontal axis shows by decile average longevity of parents.

The last two figures in this chapter look at the inheritance of house values and neighborhood quality as observed in contemporary data, 1999-2026. Since the house value data reflects the values of the homes people were dwelling in, rather than values of home ownership, the house value data is better interpreted as a measure of permanent income than of wealth. Since house values are right skewed, the figure ranks individuals by the natural logarithm of house value, so that the outcomes represent now close to a normal distribution. Children's house values were included here only when they were aged 24 and above, and living at a different address than the parent or parents.

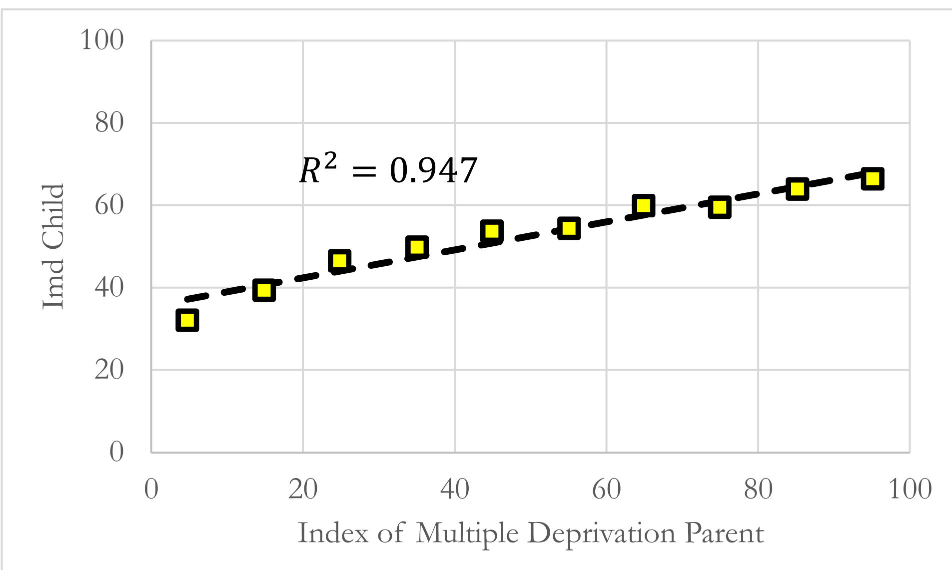
Neighborhood quality is measured by the Index of Multiple Deprivation (IMD) for local areas of around 1,500 people. That index aggregates components such as the percentage of people in the area who receive benefits and tax credits, who are employed, who have had education and training, as well as crime rates. As employed here it is measured as the percentile rank of each area on the index, 0-100, going from the highest deprivation neighborhoods to the lowest.

Figure 14.6: Ln House Value, Children versus Parent, 2002-2026



Source: FOE database.

Figure 14.7: Index of Multiple Deprivation, Children versus Parent, 2002-2026



Source: FOE database.

Figure 14.6 shows the parent log house values by decile, and the corresponding child log house values. Again there is a remarkable linearity to the relationship, with a simple linear relationship in this case explaining 99.6 percent of the variation in the child house values. Since the values are in logarithms the linear fit implies that the children at the bottom experience the same percentage move towards the mean as the children of parents at the top of the distribution.

Figure 14.7 shows the corresponding Index of Multiple Deprivation by parent deciles for the local neighborhood of their address, as well as the average index for their children. Lower numbers mean more deprivation. Linearity in regression to the mean again appears. Children of areas with the least deprivation show increases in the deprivation ranking that match the ranking gains for children from the areas of most deprivation. In this case 95 percent of child variation is explained by the linear fit.

Conclusion

For a set of social outcomes which fall on a continuum - occupational status, longevity, dwelling value, and neighborhood deprivation score – we observe two connected features. The first is a linear connection between parent and child outcomes. There is a uniform movement towards the mean by children throughout parent distribution. The second, an implication of the first, is a symmetry in regression to the mean at the top and the bottom of the distribution. This pattern is not surprising for phenotypes determined by polygenic inheritance such as height.

The pattern is surprising for outcomes determined by social environment. The ends of the parent distribution will correspond to very different social environments: extremes of abundance and deprivation. Why that extreme abundance should produce the same movement downwards towards the mean as the families of extreme deprivation is a puzzle in terms of social influences. It is confirmation of the likely dominance of direct genetic transmission in determining most social outcomes.

III SOCIAL IMPLICATIONS OF GENETIC INHERITANCE

In this third part of the book I consider the implications for economic and social policy of the claim that in a society such as England 1600-2026 genetics largely determines educational and social outcomes.

The stability of social mobility rates 1600-2026, and their failure to be influenced by social interventions suggests that the persistent belief that we live in an unfair world that consistently denies opportunity to children from lower status families is misplaced. A society such as England is, and has been, close to a meritocracy for many generations. There is no social mobility problem. Low rates of social mobility reflect just the high degree of assortment by social abilities in parenting.

The first implication is that the extraordinary large expenditures on education in modern high-income societies will largely have no effect on social and economic outcomes for students of lesser abilities. Raising the relative number of years devoted to education for lower class students since 1880 in England has had no effect on social mobility rates. Further there is good evidence for England that adding years of compulsory education in 1918, 1946, and 1972 resulted in no economic or health gains to the students affected. In chapter 15 I show how government policy reduced the variance of years of education in English society 1851-2026 without increasing social mobility rates.

To improve the social conditions of the least advantaged, it will be much more effective to transfer income directly to this group, or to follow policies that compress the wage distribution. Nordic countries succeeded in creating the good society not because of their very generous subsidies to ever increasing amounts of education, but because they have compressed material inequalities.

The second implication is that the character of emigration and immigration will be powerful forces shaping the fortunes of societies. Losing high status migrants will permanently reduce the social abilities of a society. In particular the Huguenot refugees of

1685 had significant effects on English social abilities from 1685 till now, even though completely assimilated into English society culturally within a few generations of their arrival. Similarly gaining high status immigrants will permanently improve the abilities of a society. Immigrants will through slow processes of social mobility eventually be absorbed into average status in any society. But these immigrants will leave an indelible imprint in that absorption process on the average level of abilities.

This means that the costs and benefits of different immigrant groups will extend across multiple generations. It does not imply that all countries should emulate the selective immigration policies of countries such as Australia, Canada and New Zealand. There are many considerations about immigration that go well beyond economic costs and benefits. But it does imply that calculations such as in van de Beek et al. (2023) on the costs and benefits of different types of immigrants by education level will substantially understate the gains from high educational status immigration, and the costs from low educational status immigrants.

In this respect England has been the beneficiary in recent years of a selective immigration which has raised social abilities. Now the white British are one of the lower classes of modern England, and a variety of immigrant groups form a new upper class in the society.

The third implication is that the relative fertility of different social classes can potentially have a large impact on both the average level of abilities, and also of the variance of abilities within a society. From 1250 to 1780 England had a pattern of social class fertility where higher status was associated with substantially greater net fertility. That has left a permanent genetic endowment for future generations of higher social abilities.

The fourth implication is that the degree of assortment in social abilities in marriage will be a substantial determinant of the distribution of abilities within a society. The more assortative is marriage, the greater the distribution of abilities. Changes in the degree of assortment in marriage in north west Europe between the middle ages and 1750 were potentially a factor leading to the Industrial Revolution.

15. Education and Social Mobility: 1851-2026

Social mobility rates were slow and unchanged in England 1600-2026. In addition, social position seems largely genetically determined. The fact of genetic determination matters little, however, if social interventions can significantly change social outcomes. In this chapter I consider further the evidence that social interventions – primarily provision of universal schooling, and child support – can indeed significantly change social outcomes.

Remarkably for England specifically, and for education in general, there is no good evidence that expanding educational provision to lower ability students has produced gains in earnings or health. Despite the central role that investment in human capital plays in economic doctrine, the move towards prolonged educational provision for all has demonstrably no effect on outcomes for less able students.

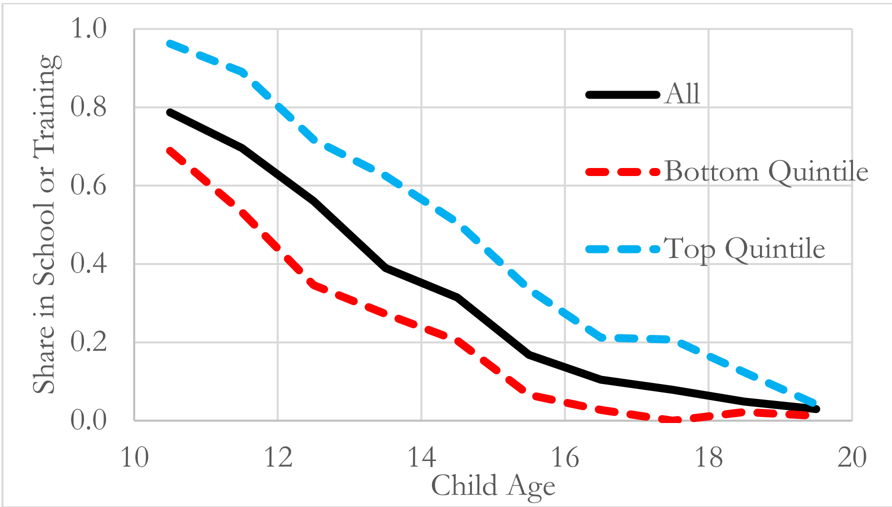
The evidence indeed is that much of the 10 percent of national income typically spent on education by high income economies is largely a social waste. In exaggerated hope of economic gains, and in pursuit of reduced inequality, and enhanced social mobility, modern high-income societies supply too much useless education.

State Education Expansion, England, 1851-2026

In 1851 educational choices by parents in England were completely unconstrained by government. Parents freely chose the quantity of education for their children. And those choices strongly reflected the social status of families. Thus figure 15.1 shows for children 10-19 years old at the census in 1851 the fraction in school or in apprenticeships for children as a whole, and for those whose father was in the top or bottom quintile of occupational status.[77]

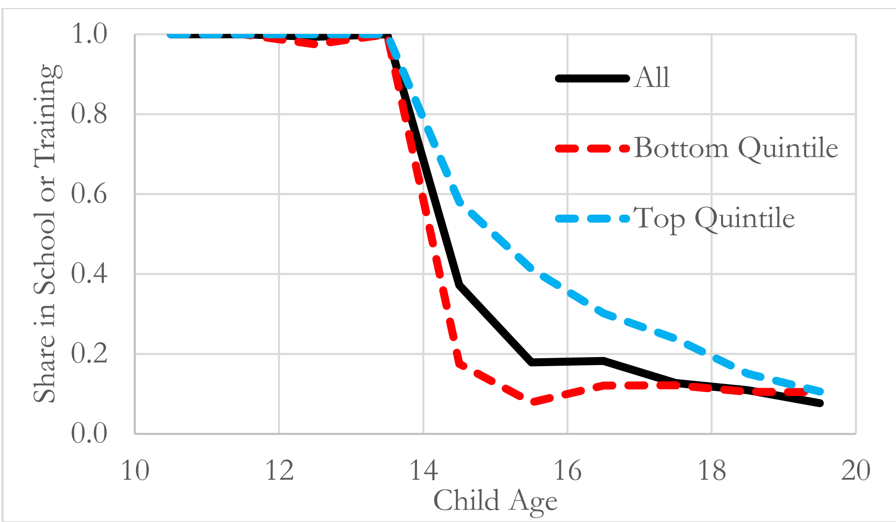
[77] Low status fathers had occupational status less than 28 on a 0-100 scale, high status fathers 48 or above.

Figure 15.1: Share of Children 10-19 in Schooling, 1851/61 census



Source: FOE database.

Figure 15.2: Share of Children 10-19 in Schooling, 1939 Household Survey



Source: FOE database.

The educational regime of 1851 had substantial inequalities in schooling by social class. Children of the lowest quintile of status on average got 5.6 years of schooling before age 20. Top quintile children got 9.4 years, 68% more education.[78]

Between 1880 and 2026, the government in England mandated ever increasing years of compulsory education. The school leaving age increased as follows: 1880, 10; 1892, 11; 1899, 12; 1922, 14; 1947, 15; 1972, 16; 2013, 17; 2015, 18. Along with increases in compulsory schooling went free provision of education, and eventually other child supports.

The household survey of 1939 shows the effect of earlier government interventions. Schooling until age 14 was compulsory, and there was near universal school enrollment of children under age 14. Figure 15.2 shows overall enrollment in schooling and apprenticeships by age, along with enrollment for bottom and top quintile family status children. In 1939 school enrollment for lower and upper-class children was equal until age 14.

As a result while total years of schooling by age 20 increased between 1851 and 1939 by 4.1 years for lower-class children (9.7 versus 5.6), it increased by only 1.7 years for upper class children (10.8 versus 9.1).[79] Government mandates in education 1880-1918 expanding the compulsory years of education narrowed the social class education gap. That gap has continued to be narrower to the present day.

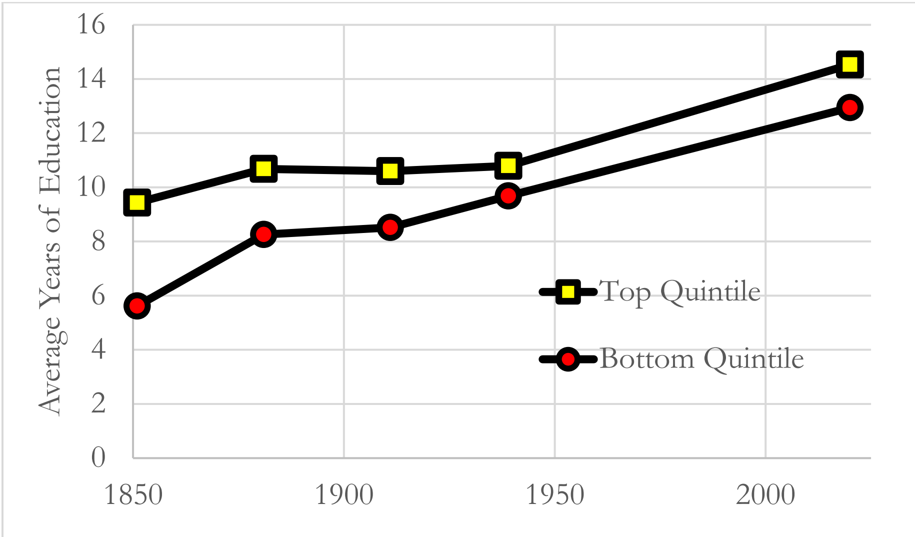
Figure 15.3 shows for 1851, 1881, 1911, 1939 and 2020 years of education for the first and fifth quintile of social status. From 1851 to 1939 we see a significant narrowing in social class differences in years of education. But that narrowing, significantly influenced by government policy, had no effects on social mobility rates. Children from lower class backgrounds did not experience enhanced rates of upward mobility from the 1920s onward.

---

[78] Assuming educational participation ages 5-9 was the same as for age 10.

[79] This estimate looks just at education completed by age 20. Some upper-class children would still be in education even at ages 21 and 22 in 1939.

Figure 15.3: Years of Education by Status Quintile, England, 1851-2020



Sources: FOE Database, Espinoza et al. 2020, Table A1.4.
Notes: Espinoza et al. reports distribution of schooling socio-economic status and by qualification levels, which have been converted to years of education.

Effects of Compulsory Education Expansions

The reason the expansion of education for lower class students did not affect their upwards social mobility is that for England there is good evidence that the expansions of compulsory education in 1918-22 from 12 to 14 years, in 1947 from 14 to 15 years, and in 1972 from 15 to 16 years had no effect on the economic and social outcomes of the students who received more education as a result of these government measures

The 1947 educational expansion is the one that has been most intensively studied. Students born in 1932 and later had to stay in school until age 15, while those born 1931 and earlier could leave at age 14. Those born in 1932 and later got an extra 0.6 of a year of schooling compared to those born in 1931 and earlier. The children affected by this administrative change were mainly those of families of lower social status who would have transitioned from school to work at age 14 (Clark, 2023).

Despite earlier claims of substantial gains in affected cohorts from such extra years of education in England, a recent study by Damon Clark suggests convincingly that there were no gains from the schooling extension for either men or women. This shows in several ways. First using an administrative sample that tracks earnings by 1 percent of the population, Damon Clark shows the movement of average earnings aged 38-60 of those turning 14 three years before and after the schooling extension. This is shown in figure 15.4. There is no sign of any gains associated with the extended years of education. This is confirmed by statistical analysis which estimates a gain of 0 percent in wages for both men and women.[80]

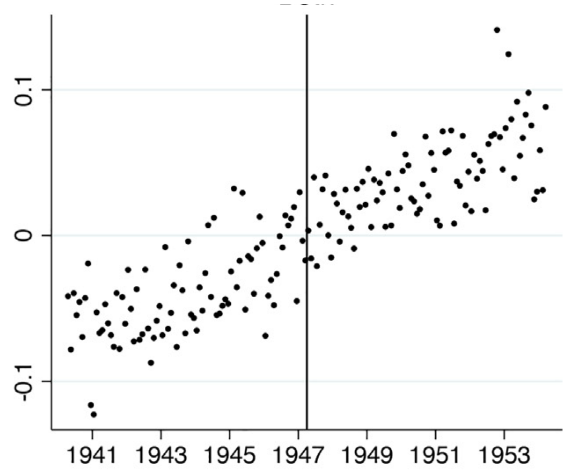
The lack of gain in wages is echoed in the failure to observe any higher labor force participation of those affected by the schooling extension, when observed in the 1991 census, when those first affected by the extension would be 59. This is shown in figure 15.5. Another measure of the effects of the schooling extension is the fraction of men and women affected who were reported as home owners in the census of 1991. That fraction is unchanged by the extension for both men and women.[81]

Further, while there is a robust positive association between years of education and both health and longevity, in an earlier article Damon Clark and Heather Royer document that there was no gain from lower mortality in subsequent years for those affected by the schooling extensions in both 1947 and 1972 (Clark and Royer, 2013).

Damon Clark explains the failure of extra education to create any gains in social mobility in 1947 from the character of this particular reform. At this time students in England were tracked by ability. The students affected by the reform were almost all in the lowest track of school, the Secondary Moderns, which focused on practical skills as opposed to academic qualifications. Data from the 2011 census in England and Wales, which collected information on academic qualifications, vocational qualifications, and apprenticeships, shows no gains in any of these three qualifications from the 1947 schooling extension.[82] Clark does not dispute that education is in general productivity enhancing, just that this particular schooling extension notably failed to deliver any more education.

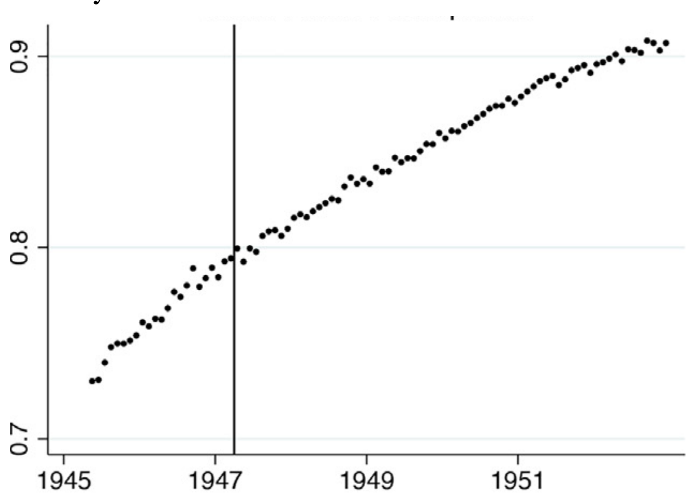
[80]If we take a 4 year bandwidth, the estimate of 0.008 of the wage gain for men and women together, plus standard error of .0093 implies that we can reject with 95% confidence that the wage gain exceeded 2.3% (Damon Clark, 2023, table 2).
[81] Damon Clark, 2023, Appendix figure F6.
[82] Damon Clark, 2023, Appendix, figure E1.

Figure 15.4: Impacts of 1947 compulsory schooling reform on log weekly earnings



Notes: Vertical axis shows log earnings per week. Horizontal axis shows date of reaching age 14.
Source: Damon Clark, 2023.

Figure 15.5: Impacts of compulsory schooling reform 1947 on male labor market activity in 1991



Source: Damon Clark, 2023.
Notes: Vertical axis shows fraction of men active in labor market 1991. Horizontal axis shows date of reaching age 14.

There is convincing evidence, however, that the extensions of schooling in 1918-21 (12 to 14 years), and 1972 (15 to 16 years) also failed to generate gains in income, wealth, or health for the students affected. The 1972 extension to 16 years of schooling, for example, was conducted in an era where educational tracking by school was being replaced by comprehensive secondary education. The impetus was a desire to democratize education, and make academic qualifications available to all students regardless of ability. The Labor Government of 1964-70 had pushed Local Education Authorities to plan for the reorganization of secondary education on comprehensive lines, and large numbers of these Authorities by 1972 had created comprehensive secondary schools. Thus by 1972 nearly half of secondary students were in comprehensives.[83]

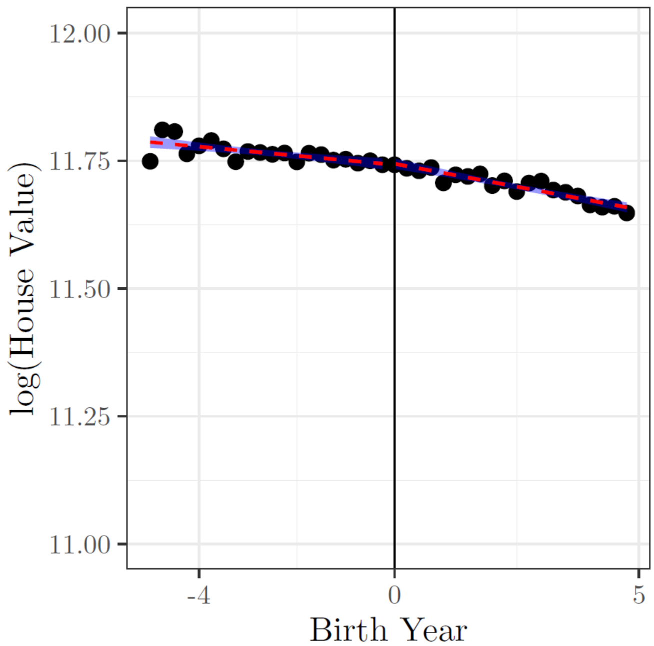
We can get evidence on the effect of the 1972 leaving age extension from the 1999 Electoral Register for England and Wales. The 1999 Register gives the address of all registered voters. These addresses can be linked to estimated average house values by postcode, from 12 million property sales in the Land Registry 1995-2005. There are 1,758,31 postcodes in the UK, so this is a highly specific estimate of average local house values covering only 20-30 houses. Using individuals with unique names born 1952-62 we can link them to the value of the house they reside in 1999. The value of housing people were living in is a good indicator of their permanent income and social status. Thus for a sample of men from the FOE database, where we get an occupational rank score, that score correlates strongly with their postcode house value.

Indeed as we move from the bottom to the top of the occupational rank, predicted house value increases by a factor of 10.[84] Average house value by postcode in 2017 prices varied from £8,000 to £24,000,000, with a pronounced right skew in the distribution. Therefore in the estimations of the effects of additional schooling we use the logarithm of house values to have an outcome that has a closer to normal distribution.

---

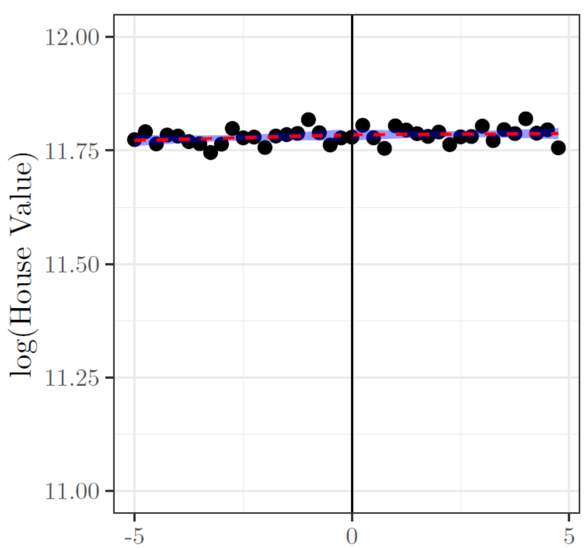
[83] Bolton, 2012, p. 8.
[84] Clark and Cummins, 2024, table 8.

Figure 15.6: House Values 1999, 1972 schooling extension



Notes: 95% Confidence interval displayed. Birth years 1952-1962.
Source: Clark and Cummins, 2024, figure 4.

Figure 7: House Values 1999, 1947 schooling extension



Notes: 95% Confidence interval displayed. Birth years 1928-38.
Source: Clark and Cummins, 2024, figure 4.

Figure 15.6 shows average log house values by year and quarter for those born 1952-62, five years before and after the compulsory school extension in 1972, as well as the fitted trend before and after the break. There clearly were no substantial gains in house values for the cohort affected by the schooling extension.[85] Figure 15.7 shows the same house value estimates for year and quarter for those born 1928-38, who reached age 14 before and after 1947. As can be seen there is no sign of any upwards movement associated with the increase in average years of education.[86]

The 1918 Education Act, which extended compulsory schooling from age 12 to 14, had a more drawn-out introduction than the later educational extensions. Administrative data on school enrollment from the Reports of the Board of Education for those aged 12-16 in 1916-1926 show that the leaving age requirements of this Act were in fact introduced over 4 years.

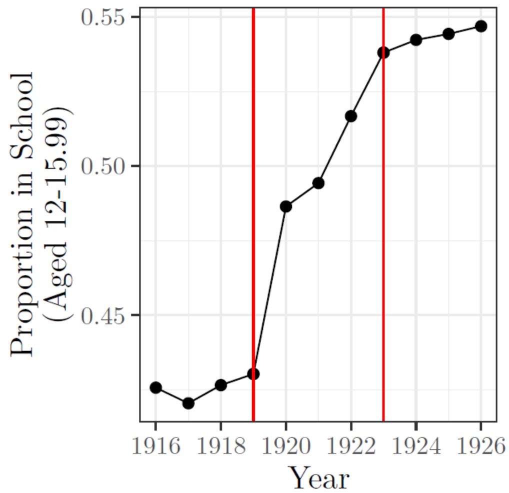
Figure 15.8 shows there was a sharp rise in school attendance by 12-15 year olds in the school year 1919-20, but then three more modest increases all the way to full implementation of the leaving age of 14 in the 1922-3 school year. The overall effect was an increase of 0.43 years of schooling, with 0.22 of that increase occurring in the school year 1919-20. Therefore in the estimates below the treatment cohorts are taken as births 1910(Q4) to 1915(Q3), and the pre-treatment cohorts as births 1902(Q4) to 1907(Q3). The children now attending school until age 14 were largely taught still in elementary schools.

---

[85] Clark and Cummins, 2023, table 10 shows that the gain in house values post the 1972 extension is estimated at 0. With 95% confidence the effect of an additional year of schooling is less than a 4% gain in house values.

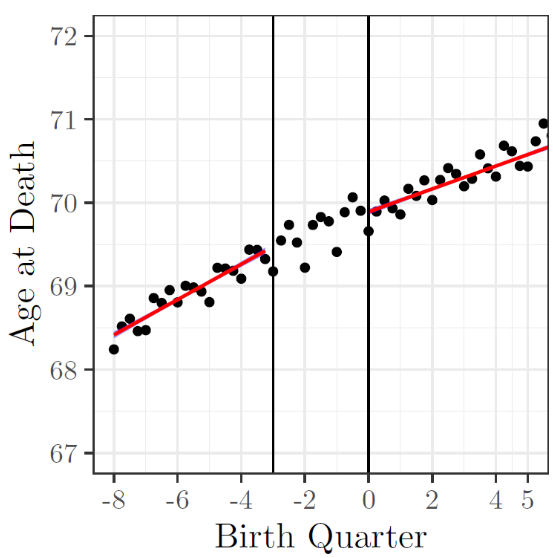
[86] Clark and Cummins, 2023, table 9 shows that the gain in house values post the 1947 extension is again estimated at 0, and with 95% confidence the gain from a year of education is less than 3.5% in house value.

Figure 15.8: England, School Attendance ages 12-15, 1916-1926



Source: Clark and Cummins, 2024, Appendix table 8b.

Figure 15.9: Average Age at Death by Birth Year, 1903-1915



Notes: The black dots plot averages by birth quarter. The fitted line is estimated from a linear regression. Year 0 = 1910(Q4).

Source: Clark and Cummins, 2024, figure 1.

While the drawn-out implementation of the 1918 Schooling Extension is a drawback, for this earlier education extension we do get a record of both place and age of death of the entire cohorts immediately before the extension, and the entire cohort immediately after.

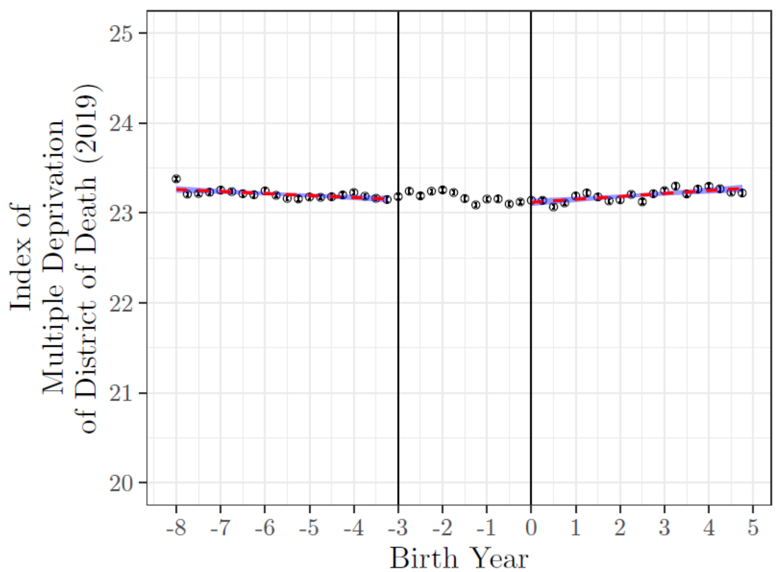
Education has a well-documented association with reductions in mortality. A recent meta-study concluded that for those aged 18-49 an additional year of education was associated with a 2.9% reduction in mortality, and for those aged 70+ still a reduction of 0.8% (Balaj et al. 2024, p. e160). There has also been claimed by Lleras-Muney (2005), using state level variations in compulsory education in the USA, that education has a strong causal effect on longevity, where an additional year of schooling increased adult longevity by 1.7 years.

For 4.3 million men and women born in the five years on either side of the 1919-22 schooling extension we have measures of average age at death. Figure 9 shows by birth quarter average age at death on either side of the compulsory schooling extension. Also shown is a best linear fit for both sides of the leaving age change, omitting the transition years. We would expect on the Lleras-Muney (2005) estimate a rise consequent on the increase in years of education 1922 and later of 0.8 years of longevity. In fact the estimate is of no significant gain in longevity at the discontinuity.[87]

Thus despite the observed link between years of education and adult longevity, in none of the three extensions of compulsory education in England in the twentieth century – 1922, 1947 and 1972 – do we observe any reduction in mortality consequent on the increased years of education.

[87] Clark and Cummins, 2024, table 4.

Figure 15.10: Impact of the 1919-22 Educational Extensions on the Index of Multiple Deprivation of Place of Death



Note: Neighborhood characteristics of 6.5 m people born 1902-1915.
Source: Clark and Cummins, 2024, figure 5.

Data from the FOE database shows that there is a strong association, for men dying 1920 and later, between their occupational status and locations where the Index of Multiple Deprivation indicates less deprivation in 2019. [88] We can thus use the Index of Multiple Deprivation by death location as an indirect measure of the social status of those affected by the 1922 extension of required schooling compared to the previous cohort. There should be enhanced movement among those whose schooling was increased by compulsory schooling extensions to residential districts with more educated people, wealthier people, healthier people, and districts with fewer social problems such as crime.

Figure 15.10 shows the results for 6.5 million people born in England and Wales 1902-1915, who died age 21 and above, in terms of the socio-economic characteristics of their place

[88] Clark and Cummins, 2024, table 3. The strong association between occupational status for people dying in the 1920s and the social status of places in 2019 shows that social status by place is a very persistent feature of locations.

of death. Those born 1910(Q4) and later, who all stayed in school until at least age 14, show no sign of any advance in social status as reflected in their place of death. They show no advantage in the socio-economic status of their place of death compared to the cohorts born before 1907(Q4), who received 0.43 years less schooling on average.

Thus, remarkably, for all three extensions of compulsory education in England 1918-1972, going from 12 to 16 years, there is no evidence of any economic or social gains from these extra years of education for the lower status students who were affected by the extension.

Returns to Schooling in General

The general problem in the estimation of the causal effects of schooling is that people of higher ability choose more schooling, so that the observed earnings gains from schooling likely overestimate the actual returns. A number of strategies have thus been devised to control for ability in estimating the return to schooling, as discussed in Card (1999, 2008).

One popular method to control for ability differences was using monozygotic (MZ) twins who differed in years of education. A slew of such studies were published in high profile journals between 1994 and 2011. They suggested returns to education observed between twins that were as high as the raw cross-sectional returns to education, implying an absence of significant ability selection into education.

However, these MZ twin estimates do not in fact control for ability. For despite their genetics being the same, monozygotic twins can differ significantly in academic abilities, which helps explain their different years of schooling. Even within the same family environment there is significant variation in MZ twins' phenotypes, despite the same genotype.

On average MZ twins in a British sample, for example, at age 16 differed in academic performance by nearly half a grade level (von Stumm and Plomin, 2018, p. 5). Since these performance differences between the MZ twins at ages 12, 14 and 16 were correlated, this was not just measurement error, but represented persistent academic achievement differences. Robert Plomin, the distinguished behavioral geneticist who has made extensive exploration of the cause of these "non-shared environment" effects, has concluded that whatever these effects are they have no identifiable systematic source (Plomin, 2011, 2024, Plomin and

Kawakami, 2023). But they imply that for MZ twins, as for the general population, differences in schooling are associated with difference in academic abilities.

Consequently the majority of studies of education returns that control for ability employ administrative changes in the minimum school leaving age. Clark and Nielsen, 2024, surveys the literature and finds 79 such estimates. Remarkably, while the majority of these studies find substantial gains from education, a number of well-grounded studies find no effect. The average return from these studies still implies substantial average gains from an extra year of education: an average of 8.2%. But the pattern of reported returns shows clear evidence of publication biases: omission of studies where the return was not statistically significantly above 0, and where the estimated return was negative. Correcting for these omitted studies, the implied average causal returns to an extra year of schooling is only in the range 0-3%**.**

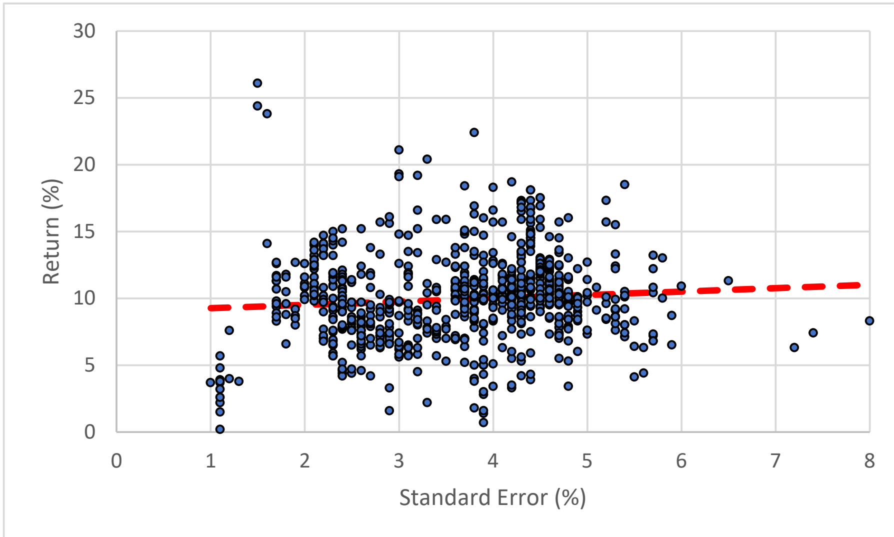
If studies fail to be published unless they find an estimated return to education significantly greater than zero, then the average estimated return will be biased upwards. Absent publication biases there should be no connection between the average effect size estimated and the standard error of that estimate.

A meta-study of the gross returns to an additional year of schooling, without controls for ability, for example, reported estimated returns and the standard error of these estimates for 819 estimates (Montenegro and Patrinos, 2014). The average return was 10.1%. Figure 11 shows these estimated returns against the standard error of the estimates, along with the fitted regression line.[89] As can be seen there is very slight evidence of publication bias in the form of a positive slope of the regression line. However, the slope of this line was not significantly different from zero. Thus there is no clear evidence of publication bias in this literature.

In contrast, figure 12 shows the estimated returns to a year of additional education, from compulsory schooling extensions. Here an increase in standard error of 1% is associated with a 1.42% increase of reported returns to an extra year of education. The linear fit to this data suggests that at a standard error of zero, the measured rate of return would be 3%, as opposed to the 8.2% observed on average for these studies. There is substantial publication bias.

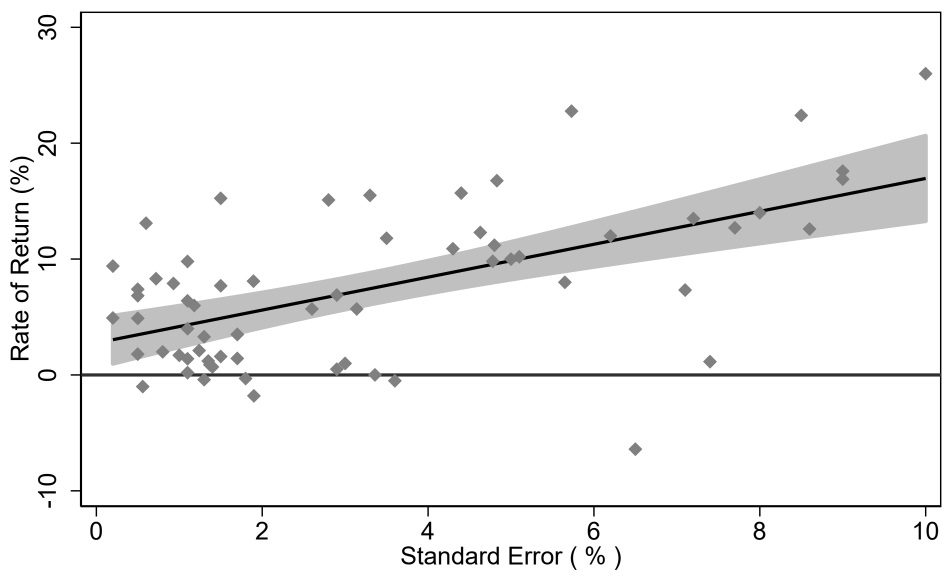
[89] The regression line was fitted including an indicator for returns in high income countries.

Figure 15.11: Gross Returns to Schooling versus Standard Error



Source: Montenegro and Patrinos, 2014, Annex Table 1, columns A and B.

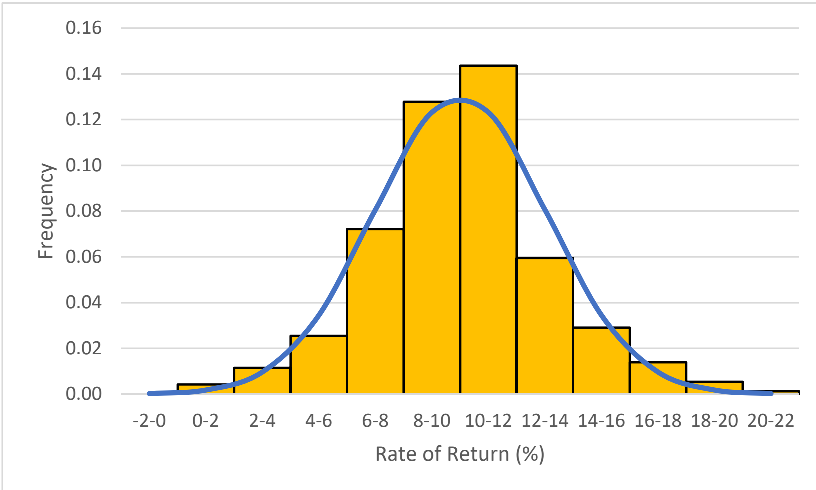
Figure 15.12: Compulsory Education, Rate of Return versus Standard Error



Notes: Six outliers with standard errors above 10% are omitted. Grey shading indicates the 95% confidence intervals of the regression line.

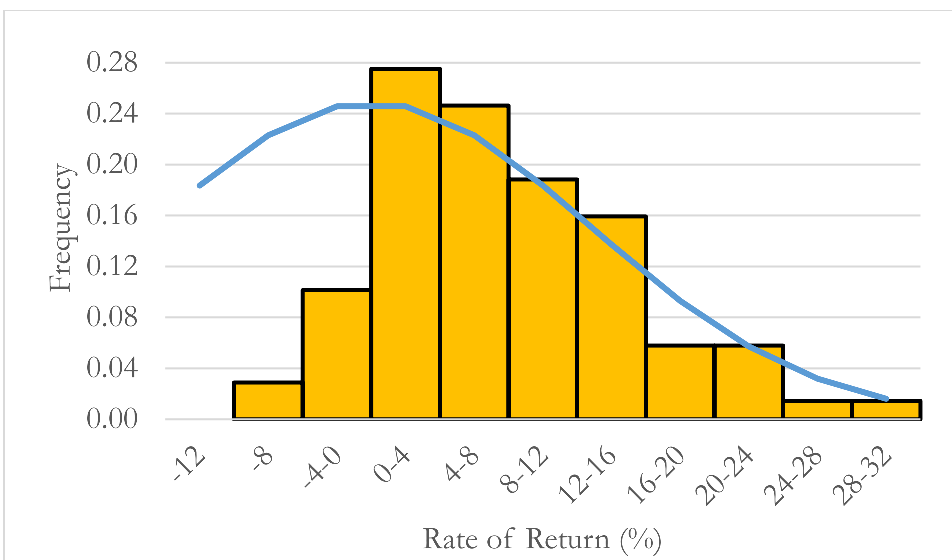
Source: Clark and Nielsen, 2024, figure 7.

Figure 15.13: Distribution of Returns to a Year of School, 1970-2012



Notes: The normal curve shown is that with the mean and standard deviation of the reported returns.
Source: Montenegro and Patrinos, 2014, Annex Table 1, column A.

Figure 15.14: Estimated Rates of Return to Compulsory Schooling Versus a Normal Distribution with mean 0



Notes: Estimated rates of return to a year of education in the meta-study in range -12% - 20%, versus a normal distribution with mean 0, and same variance for positive returns.
Source: Clark and Nielsen, 2024, figure 9.

The second evidence of publication bias appears in the non-normal distribution of the compulsory schooling estimates in the meta-study. The meta-study of gross returns to schooling produced just such a normal distribution, shown in figure 13.[90] This normal distribution will be the product of the sampling error in the underlying studies, as well as heterogeneity in actual underlying returns to education.

Based on the case of reported gross returns to education, we would expect, absent publication biases, that the estimated causal returns to education would also follow a normal distribution. The average return in the non-causal estimates (10.1%) is, after all, not much above the average of 8.2% for the causal estimates. This is assuming the two sources of variation in reported returns to education are a normally distributed underlying rate of return, combined with sampling error. Simulation using the reported standard errors of estimates suggests sampling error will explain half the reported variance in returns.

We instead find in figure 14 a causal returns distribution that cannot be explained without publication biases. Estimated returns to education have a highly asymmetric distribution with a strong peak (24% of all returns) in measured returns in the 0-4% range. There is a distinct discontinuity in the distribution of reported returns at 0%. Given sampling errors reported in these studies, this peak in the 0-4% range implies there should be many more studies reporting negative returns than we observe.

One possible interpretation of figure 14 is that because of a strong belief that the true returns to education cannot be negative, there is omission from published studies of those that find negative returns. As can be seen in the figure, the reported studies with a positive return to education in fact approximate well to the upper half of a normal distribution of returns to education centering close to 0, and with the same variance as these positive returns studies.[91]

[90] There is sign, however, even in this case of some publication bias. There were no estimates of returns which were negative. If there was symmetry with the upper tail of the estimates, there should have been 6 such estimates. Correcting for this likely omission the average return was 10%.

[91] It has been suggested that the heavy concentration of returns in the range just above 0% could be the product of the underlying returns having a log normal distribution, which would imply no true returns falling in the range below 0. You would expect, however, that since the average causal return estimate (8.2%) is close to the average gross return estimate (10.1%), then gross returns would also display a log normal distribution. As seen in figure 13, those estimates show a normal distribution.

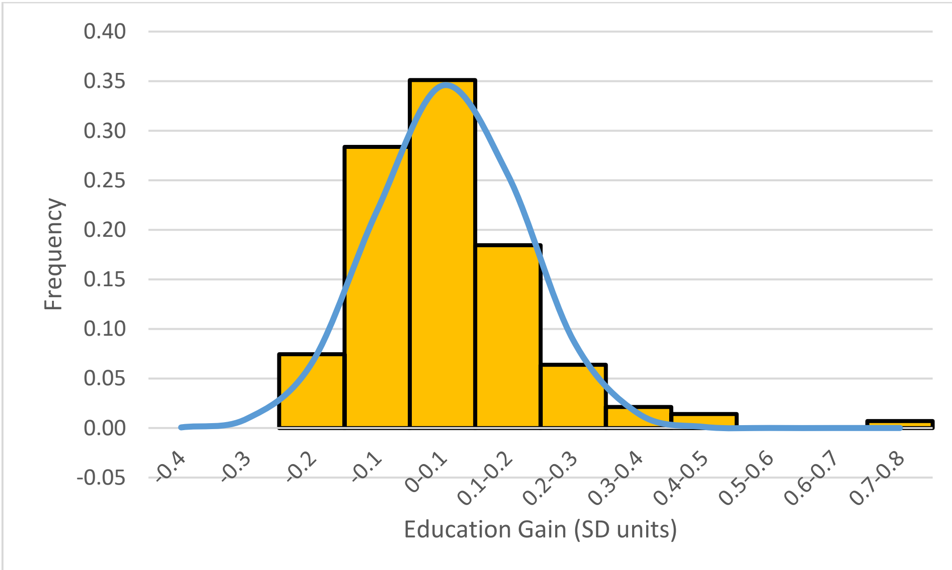
Thus the published literature on the returns to schooling seems to be shaped by two biases. First an inflation of returns caused by publication bias towards returns that are statistically significantly above 0. But second an omission of studies that find returns below 0. Taking the first bias into account reduces estimated returns to only 3%. But allowing also for the second bias then plausibly the returns could be as low as 0. Thus the oddity reported above, where returns to increased compulsory education in England 1922, 1947 and 1972 all are estimated at zero, may not be an oddity at all, but instead a general feature of the expansion of education. Zero net gains from additional years of education for students of lower ability would help explain why the substantial narrowing of gaps in years of education between upper class and lower class students in England 1880-2026 had no effect on rates of social mobility.

Note that published research on the effectiveness of educational interventions seems in general to have been plagued by publication biases. Hattie (2009) in a meta-analysis of more than 800 educational interventions finds an average reported effect size of 0.4 standard deviations among students of the targeted outcome. These are enormous average gains, given that normal learning gain in a year for students is 0.23 of a standard deviation.

An alternative meta-study looked at the effect of educational interventions aimed at improving academic achievement in kindergarten, primary and secondary education in England and the USA. In the studies included the allocation of students to the intervention and control groups was random, the trial was conducted by researchers independent from those promoting the intervention, and the results of all trials were published (Lortie-Forgues and Inglis, 2019). This resulted in 141 trials of interventions, with a total of 1,222,024 students.

The mean effect size with the controls above was surprisingly small – a gain in performance of 0.06 standard deviation units on average. Figure 15 shows the distribution of the results. Here the fit with the normal distribution is not as clear as in figure 13, but note the substantial number of trials which had reported negative effects on student outcomes. Since the typical confidence interval for these outcomes was 0.3 standard deviations, for almost all of the individual interventions, it was not clear if there was any positive effect allowing for sampling error.

Figure 15.15: Distribution of Effect Sizes from Educational Trials



Source: Lortie-Forgues and Inglis, 2019, figure 1.

Conclusions

The above all illustrates the inability of socially provided education to substantially change social and economic outcomes for poorer families. Despite modern high income societies typically spending 10% or more of national income on education, there is no sign these expenditures have changed slow rates of social mobility. Extensions of compulsory education in 1922, 1947, and 1972 in England did not improve outcomes for the 40-60% of the affected cohorts who now received an additional year of schooling. Nor is there evidence that such extensions of education in general increased earnings for the recipients.

The inability of state provided education and other family resources to change social outcomes for poorer families is what explains the unchanging rates of social mobility in England 1600-2026.

16. Migration

John Houblon, first Governor of the Bank of England (1694), on the Bank's £50. Houblon descended from Huguenot refugees who fled France for England in the sixteenth century. Even now Huguenot surnames remained four times as common at Oxford and Cambridge as the average English surname.

Migration is of interest for several reasons. In a world where social outcomes depend on genetic endowments, migration in and out of countries will potentially have long lasting impact on the average level of social abilities. The American ideal of U.S. society as a "melting pot", where within one or two generations immigrants assimilate to the dominant culture so that the history and characteristics of migrants makes no difference to social outcomes, will not hold. If those leaving are above average status, they permanently lower the social abilities of the home population. If those arriving are above average status, they permanently raise the social abilities of the host society.

In this chapter I consider the effect of migration on social abilities in England 1600-2026. In particular what was the effect of the significant emigration of people to the New World countries – Australia, New Zealand, Canada and the USA – on the ability levels of the population remaining in England? But alongside migration from England, there was also significant immigration from Scotland, Ireland, and Eastern Europe 1760-1940, and later significant flows of migrants from India, Pakistan, Bangladesh, the Caribbean, Poland, Romania, and China. How did the earlier and later in-migration flows affect average social abilities in England?

A second issue that arises with migration, however, is whether people can improve their social position through migration. Was the very strong persistence of social status across generations in England the product of restrictions on mobility peculiar to English society, restrictions that disappeared in the more open New World economies?

Above I argue that migration within England did not have this effect, instead mainly redistributing the population with greater social abilities to the South. But do we see significant gains in occupational status by migrants who leave England for the New World societies? Here I show that while migration likely increased migrants' incomes, it had no effect on social position.

Using the FOE database I find that migration out of England, as measured by father occupational status, was modestly positively selected. However, comparing migrants abroad to those moving within England, there were no benefits in terms of social and educational status, either for the migrants themselves, or for their children. Indeed men who migrated abroad saw their occupational status fall compared to brothers who moved within England. Their sons, however, ended up with an occupational status equal to that of their cousins who were the children of internal migrants. The determinants of social status were very similar within England as for the major receiving countries of Australia/New Zealand, Canada, and the USA.

Emigration from England

The FOE database also allows us to look at the effects of foreign migration by the English and Welsh on domestic social abilities. I categorize external migrants as children whose fathers died in England, who themselves died aged 21 and above in another country. Stayers were children who died aged 21 and above in England, and whose fathers also died in England. I categorize the second generation of migrants as the children born abroad to the first generation of migrants.

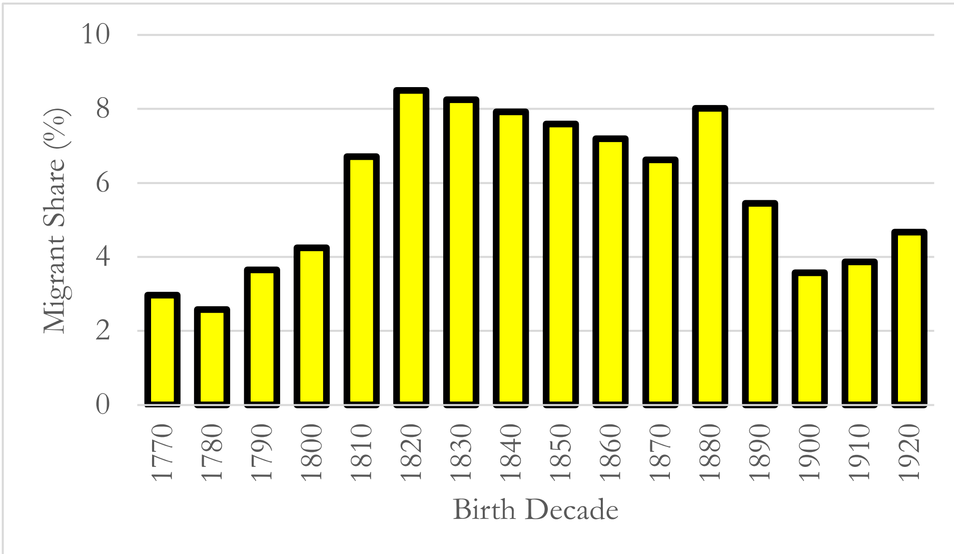
Figure 16.1 shows the numbers of migrants from England and Wales in each decade. The period of substantial migration was for births from 1810-1900, where around 7 percent of each generation left England. Overall, of children in England reaching age 21 born before 1960, 5 percent died abroad.[92] The losses of population in each generation 1700-1959 amounted to about 30 percent of English population in this period overall by 1960. If this loss was highly selective it could affect the characteristics of the remaining population. Was that loss selective enough to change the character of the remaining English population?

There has been extensive study of the composition of migrants from England to the USA and British colonies in the nineteenth century (Cohn, 1992, Erickson, 1972, 1981, 1986, 1989, 1990, Green, Mackinnon and Minns, 2002). However, the literature remains inconclusive on the question of emigrant selectivity, or the relative performance of immigrants in the countries of settlement compared to stayers in the homeland.

Though the FOE database contains modest numbers of migrants, it offers good information on the selectivity of migrants, and on the performance of migrants and their children in the countries of destination compared to their siblings and cousins in the home country.

[92] Deaths abroad in wars, such as WWI, have been counted as domestic.

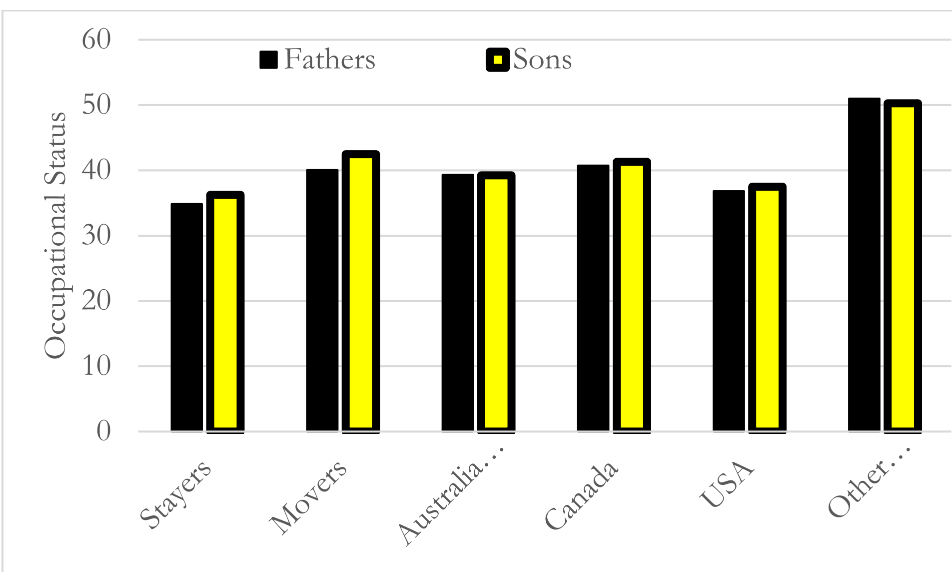
Figure 16.1: Share Migrants by Birth Decade, 1770-1929



Source: FOE database.

Note: External migration rate for general lineages. Migration rates estimated as share of children born in England and Wales dying abroad aged 21+.

Figure 16.2: Migrant Occupational Status versus their Fathers



Note: Stayers and movers indicates domestic stayers and movers.

Source: FOE database.

The chief destination of English migrants was the New World – Australia, New Zealand, Canada and the USA absorbed 82 percent of all English and Welsh migrants. Figure 16.2 shows evidence for the selectivity of external migration in the FOE database. The black columns show average father occupational status 1770-1959 for those in the general lineages who stayed within their region in England, those who moved internally to another region in England, and those moving to the major foreign destinations.

Measuring status by father status, those moving abroad had a father occupational status 13 percent higher than the father occupational status of those remaining in England. So the overall effect of emigration from England 1700-1959 would be a modest reduction of around 4 percent in average English social abilities.

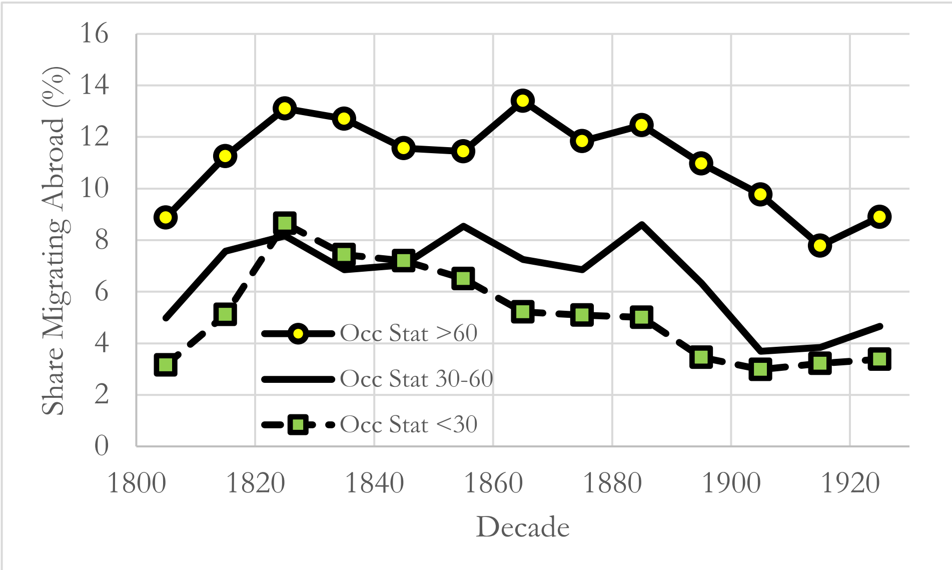
Interestingly the least selective foreign migration was that to the USA, where migrant father status was close the average father status in England. Migration to "other" destinations, principally Ireland, Scotland, and India, was the most selective. Figure 16.2 divides stayers in England into those who stayed in their home region, and those moving between regions.[93] This is to define a reference group, domestic movers, with whom we can compare the migrants externally in terms of their and their children's outcomes.

We can also measure the propensity to migrate abroad by father status, dividing up fathers into statuses 60 and above, 30-60, and less than 30. Figure 16.3 shows these propensities by decade, 1800-1930. Children of professional fathers were in all decades much more likely to emigrate. Overall nearly 11 percent of these cohorts migrated, compared to 6 percent for the intermediate cohorts, and only 5 percent for the lowest status cohorts. Over the years 1760-1930 this would imply over half of this upper social class emigrated from England.

Figure 16.2 also shows the occupational status attained by stayers and movers within England, and migrants abroad compared to the occupational status of their fathers in England. There is sign here that external migration was associated with less gain in occupational status relative to fathers than happened with internal migrants.

[93] The regions are the North, Wales, Midlands, South West, London and South East.

Figure 16.3: Propensity to Migrate Abroad by Father Occupational Status, 1800-1930



Source: FOE database.

Notes: Migration rates estimated as share of children born in England and Wales dying abroad aged 21+.

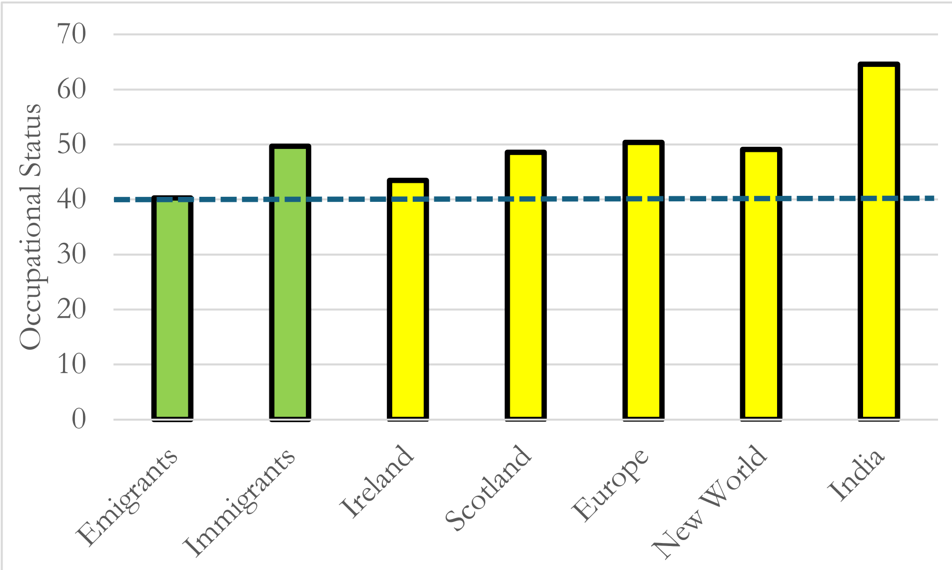
Immigration

We see above that England 1760-1929 there was some loss of social talent as a result of selective migration abroad of those with more social abilities. But at the same time, England also saw immigration. These migrants before 1930 mainly came from Ireland, Scotland, and Europe, including Jewish immigration from the Russian Empire. Overall 1760-1929 in the FOE database, immigrants averaged 4 percent of the domestic population, while emigrants in the same period averaged 5 percent.

Figure 16.4 shows the average status of emigrants and immigrants to England and Wales for births 1760-1929.[94] Immigrants substantially exceeded emigrants in social status, by a margin of nearly 10 points, where average occupational status in England was around 37.

[94] For emigrants the status is that of their father, which as figure 16.2 shows will be close to that of the sons. For immigrant groups the status is that of the immigrant themselves.

Figure 16.4: Immigrant Occupational Status, births 1760-1929



Note: Average occupational status of male immigrants to England and Wales born 1760-1929, by region of origin.

Source: FOE database.

Figure 16.4 also shows the average occupational status of the most important subgroups of immigrants. The first three are each about 1 percent of each cohort in England, and the last two each 0.5 percent. The lowest status group if immigrants were those from Ireland, but even they register as being well above average status. A recent paper by Neil Cummins and Cormac O'Grada suggests that the Irish in England have had persistently low social status (Cummins and O'Grada, 2024). But they identify the Irish as those as having distinctive Irish surnames. Those from Ireland in figure 16.4, however, includes anyone born in Ireland including the higher status Anglo-Irish and Scots-Irish, which can explain the discrepancy.

The other major immigrant groups all had status in the 50-65 range, substantially above the social average in England of around 37. The highest status group were those born in India, but this echoes the emigration figures which shows India was the destination of relatively high-status English migrants.

This data on migration flows and migrant status suggests that England and Wales 1760-1929 experienced modest gains in average social abilities. The overall gain would be, assuming genetic determination of social status, about 3.2 points on the 0-100 occupational status scale, where the average score was 37.

Post 1945 there is evidence that immigrant flows have likely improved social abilities in modern England, so that now the ancestral English are among the lowest social status groups in English society. Though immigration has recently been a major political issue in England, and unhappiness about free labor movement within the EU led to the positive vote on the Brexit Referendum of 2016, the evidence is that English society has seen social and economic gains from immigration in the last 50 years.

The scale of immigration in the UK has greatly increased since the pre-WWII years, with the immigrant share of the UK population 16 percent by the time of the 2021 census. This contrasts with an immigrant population in England and Wales of less than 4 percent in the 1911 census. But because immigrants are younger, and have higher fertility, the share of next generation which will have parents who are not "white British" will be much higher. Thus in 2022 the share of 16-year-olds in England who were classified as not "white British" was 35 percent, as can be seen in table 16.1

As can also be seen in the table, children of "white British" background are the second lowest scoring of the major ethnic groups, with only Black Caribbean scoring lower. Scores in the Chinese, Indian, and "Asian Other" ethnic groups are notably well above the society average. Were all immigrants removed from English society average educational attainment at age 16 would decline, though not by a dramatic amount.

To consider the likely long run impact of the elite Chinese and Indian immigrants consider the influence of the earlier Huguenot immigrants on English society. The Huguenots were a minority Protestant group, followers of John Calvin, in Catholic France in the sixteenth and seventeenth centuries. They were periodically tolerated and persecuted in France 1555-1685. The Royal Edict of Nantes of 1598 granted them civil rights.

Table 16.1: GCSE Results by Ethnicity, England, 2022-23

Group	Average Score 2022-3	Students
Chinese	65.5	3,383
Indian	59.4	19,601
Asian Other	54.3	12,152
Bangladeshi	51.9	10,981
Black African	48.7	26,334
White other	48.4	38,104
Pakistani	47.0	27,242
White British	44.9	392,497
Black Caribbean	40.0	6,787
All	46.3	606,915

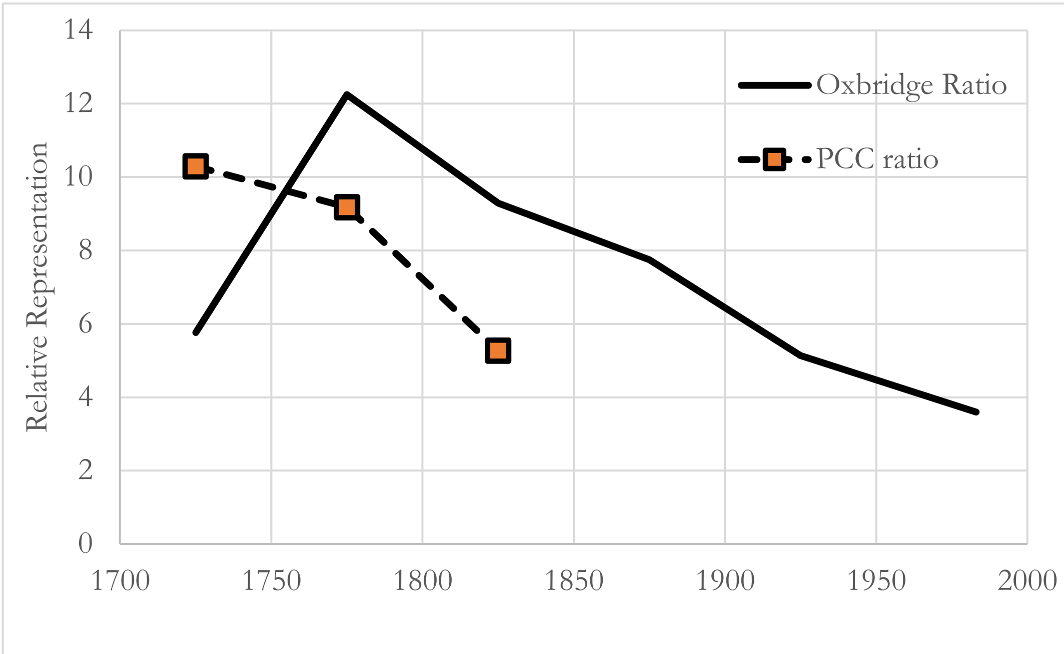
Source: UK, Department of Education, 2024.

Notes: The "Other White" category is mainly Polish and Romanian. Scores range from 0-90.

But in 1685 Louis XIV revoked the Edict of Nantes. The Huguenots had to either convert to Catholicism, or leave France. Of around 900,000 most converted to Catholicism, but 200,000 chose exile to retain their faith, with their main destinations the Netherlands, Germany, England, and South Africa.

50,000 Huguenots migrated to England, where they constituted close to 1 percent of the population at their arrival. This was when the term "refugee" was introduced to English language. Since the Huguenots had a distinctive set of surnames – such as Cazalet, Bulteel, Pigou, and Courtauld – we can follow their fortunes in England by tracking the status of their characteristic surnames.

Figure 16.5: Huguenot Representation in Two English Elites, 1700-2013



Notes:

The Huguenots assimilated quickly culturally into English society. Initially they had their own churches, schools, and charities. But though of Calvinist origin, they soon began to worship in the Church of England, intermarried with partners whose English origin surnames showed they were not Huguenots, and gave their children English first names. First names such as Pierre and Jean replaced by Peter, John, George and Edward within one or two generations. An example of their quick assimilation and social success was that the first Governor of Bank of England, in 1694, was John Houblon, a Huguenot.

Despite their complete social and cultural assimilation, descendants of the Huguenots remain a social and educational elite in England, even to the present day. We can track their status 1700 to 2013 by looking at the relative representation of Huguenot surnames at Oxford and Cambridge 1700-2013, and among wealthier testators 1700-1858, compared to their share in the population. The average surname in England will have a relative representation of one throughout.

At their peak in the late eighteenth century Huguenot surnames were 12 times as likely as the average surname to be found as students at Oxford and Cambridge. Even in the period 1950-2013 these surnames were still four times as likely as the average name to be found at these universities. Thus 300 years after they arrived in England, 10 generations, the descendants of the Huguenots still have not completely assimilated in terms of social status to the English average.

In terms of wealth, Huguenots appear disproportionately among high status wills, probated at the court of the Archbishop of Canterbury 1700-1858. Already 1700-49 they were 10 times as frequent as their share of the population. Still in 1800-58 they were five times as frequent.

While figure 16.5 makes it seem that the contribution of the Huguenots to educational attainment in England has gradually faded over time, note that by now the large majority of Huguenot descendants, 340 years (11 generations) after their arrival in England will have non-Huguenot surnames. So while with regression to the mean, individual Huguenot descendants have become less socially distinctive, this is compensated by their rising share of the population. With genetic transmission of social status dominating the Huguenots will have left an imprint on English society that never fades.

This consideration implies that when polities consider the benefits and costs of immigration within modern welfare state societies such as in northern Europe, they would have to consider different types of immigrants as creating long-lasting fiscal benefits or costs. A recent study in the Netherlands of the fiscal costs of immigration was able to do a detailed estimate of net fiscal contribution by highest educational attainment, shown in table 16.2. There are enormous differences in lifetime fiscal costs and benefits depending on education level, where the difference from highest to lowest is over €1 million. The fiscal benefits associated with highly educated immigrant communities will continue undiluted across the generations. Similarly the fiscal costs associated with less educated immigrants will persist across all future generations.

Table 16.2: Expected Lifetime Net Fiscal Cost, 2016, by Education, Netherlands

Highest attained education	Net contribution (€)	Number
Primary education	-545,000	782,182
Lower secondary - vocational	-394,000	646,182
Lower secondary – theoretical	-368,000	266,326
Upper secondary - vocational	-232,000	2,117,060
Upper secondary - theoretical	-49,000	396,324
BA	128,000	1,443,886
MA	477,000	910,534

Source: van de Beek, 2023, table 9.2, p. 184.
Notes: Undiscounted net fiscal costs, in 2016 euros.

This in no way implies any recommendation that countries should adopt the policies of Australia, Canada and New Zealand, which prioritize skilled immigration. There are reasons for admitting immigrants, those fleeing conflict for example, that are independent of their future fiscal contributions. But it does emphasize that all immigrants have much more substantial future social costs and benefits than is normally taken into account. Also given the fiscal strain modern welfare systems are under, the maintenance of these systems in the future may require governments to be careful not to extend too far the number of potential future beneficiaries.

Migration and Social Mobility

Did migration abroad show substantial benefits in terms of the social status of migrants? To test this I take advantage of the genealogical nature of the FOE database and compare the outcomes for three groups of brothers: those who stay in place in England, those who move to other parts of England than that their father dies in, and those who migrated abroad. The outcome variable I focus on is occupational status for men, measured as close to age 40 as possible, and so typically representing occupational status after migration. For this exercise I include people from both groups of lineages in the FOE database.

Given the modest amounts of data available on foreign migrants we look at migrants to each of the "New World" destinations of Australia/New Zealand, Canada, and the USA, and those to all other countries. This gives us six categories of sons of fathers dying in England and Wales, with numbers as shown in table 16.3. I then estimate son occupational status by son place of death, with father fixed effects. I do this in the table for the pooled population in the elite and "average" lineages, and also for the average lineages alone.

Table 16.3: Migration and Social Status, First Generation

Death Location	All N	All Occupational Status	All s. e.	Average Occupational Status	Average s. e.
Internal Movers	6,469	44.7	-	39.1	-
Stayers	17,972	43.2**	0.26	37.5**	0.30
Canada	391	41.1**	0.84	37.0*	0.93
Australia/NZ	650	40.0**	0.68	37.1*	0.79
USA	607	42.0**	0.71	37.7	0.76
Abroad-other	599	43.7	0.68	38.8	1.04

Notes: ** significantly different from Internal Movers at 1% level, * significantly different from Internal Movers at 5% level. "All" includes also people from the elite lineages. Father fixed effects.

Source: FOE database.

As table 16.3 shows, brothers who moved regions within England and Wales show a modest but statistically significant occupational status premium compared to their siblings who were stayers. This difference shows both for the average lineages and for the more elite group of families. This status difference could be the product of higher ability sons being more likely to move, or of moving generating better social prospects. I investigate this further below.

Note, however, that migrants to the New World countries – Australia/NZ, Canada, and the USA - did not experience any migration status premium relative to their stayer brothers, for either the broader or narrower sample. Indeed they mostly had statistically significantly lower occupational status than brothers who were internal migrants within England and Wales, though the quantitative effect was modest: a decline of 2 points ranking in a scale of 0-100.

Migrant brothers to other countries similarly experienced less gain in status compared to stayer brothers than did migrant brothers within England and Wales. Thus foreign destinations for migrants did not offer any better status outcomes in the first generation than found within England. The first generation migrants to the New World countries, though, likely made substantial gains in terms of their labor income.

The Second Generation

First-generation migrants to the major receiving countries - Canada, USA and Australia/New Zealand – had modestly lower occupational status relative to those who migrated within England and Wales. The first generation of migrants, however, may only have experienced temporary loss of occupational status from migration.

This loss of status might stem from differences in employment structure across countries. With employment structures in the receiving countries for most of this period more focused on the primary sector, average male occupation scores using the English occupational scores might have been lower than in England. The modest drop in occupational status for immigrants might also have come from differences in job qualifications for higher level jobs in the new country. Their children, however, might still be the ultimate beneficiaries of moving country.

To check this I consider the outcomes for their children, born and educated in the country of migration. How did they fare compared to their compatriots whose parents and they themselves stayed in England? To measure this we classify grandchildren within England as falling into two categories. The first is “stayers”. This consists of grandchildren born in the same region as their father was born and died in, and as the grandfather died in. This is shown in table 16.4, Panel A. The second group are internal movers, table 16.4, Panel B. Here the son dies in a different region within England than the grandfather, but the grandson is born in England. The third group are external movers, table 16.4, panel C. These are those born abroad whose grandfather died in England, but whose father died in the country/region they were born in.[95]

Because of more limited numbers of observations, we combine the external movers into two just groups: those born in the major countries of English settlement - Canada, USA, and Australia/New Zealand - “New World” - and those born in other countries such as Ireland, Scotland, or British India.

[95] Since there was significant movement between Australia and New Zealand, and between Canada and the USA, I counted , for example, a grandson born in Canada to a father dying in the USA in this category, or a grandson born in New Zealand whose father died in Australia.

Table 16.4: Second Generation – Three Types of Grandchildren

Panel A - Stayers

Generation	Born	Died
0	-	England – region x
1	England - region x	England - region x
2	England - region x	

Panel B - Internal Movers

Generation	Born	Died
0	-	England – region x
1	England – region x	England – region y
2	England	

Panel C – External Movers

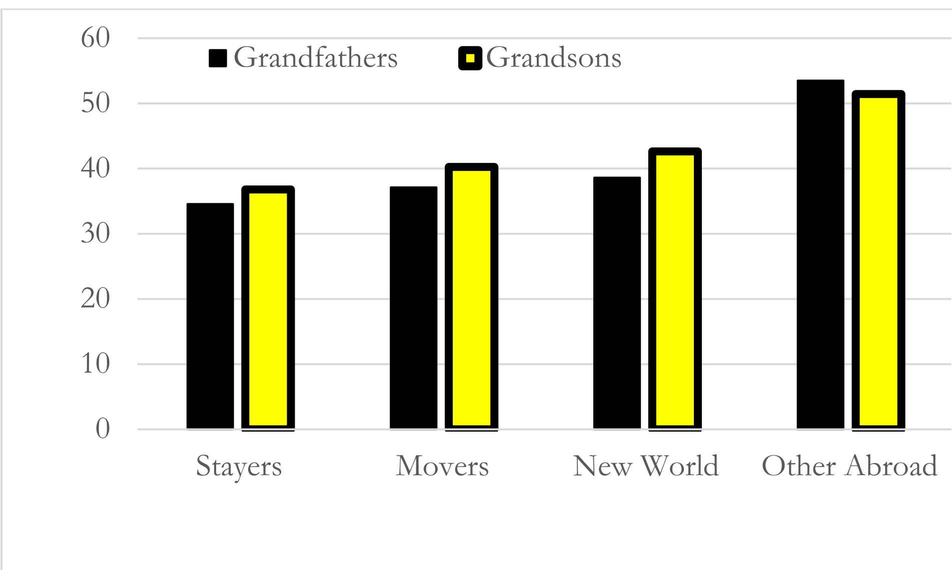
Generation	Born	Died
0	-	England
1	England	Country x
2	Country x	

Table 16.5: Migration, Occupational Status, and Higher Education, Second Generation

Death Location	N	Occupational Status	s. e.	N	Higher Education	s. e.
Internal Movers	5,533	43.77	-	6,331	0.096	-
Stayers	17,029	41.48**	0.398	17,729	0.085	0.008
New World	914	42.85	0.859	846	0.063*	0.016
Abroad-other	150	47.19	1.879	149	0.055	0.032

Notes: Estimated with grandfather fixed effects. Includes both average and elite lineages. ** significantly different from Internal Movers at 1% level, * significantly different from Internal Movers at 5% level.
Source: FOE database.

Figure 16.6: Occupational Status of Grandchildren versus Grandfather



Source: FOE database
Notes: 752 grandsons in New World, 61 in other abroad.

Table 16.5 compares the occupational and higher educational status of grandchildren, who are second generation migrants, with their cousins who were born in England. For comparability, the comparison for those migrating abroad is done with those who migrated within England. Because of the smaller amounts of data in this comparison, I amalgamate migrants to the major receiving countries of Canada, Australia and New Zealand, and the USA, into one "New World" category.

For occupational status, cousins who were the children of migrants to Australia/New Zealand, Canada and the USA attained an occupational status similar to that of their cousins who were the children of migrants within England and Wales. This result is seen both in the broader sample and in the "average" lineages. The table also finds that cousins whose parents migrated to other foreign countries attained higher occupational status, though this result is not statistically significant in either sample. Table 16.5 also shows New World cousins had no gain in educational status relative to their "mover" cousins in England and Wales. Also in this case they are not doing any better than the stayers in England.

In line with the above figure 16.6 shows occupational status of grandsons relative to their grandfathers, for those in the average lineages. Those located in the New World countries do not show any substantial gains relative to their grandfathers, compared to grandsons whose fathers moved within England and Wales. Countries such as the USA did not offer any significant upwards mobility bonus to English migrants.

Though the US thinks of itself as a land of boundless opportunity, in the first generation migrants generally had lower than expected occupational and educational status compared to their fathers in England. For the second generation of migrants, those born in the US, occupational status rebounded. But it was no higher than for their cousins who stayed in England. And educational status was still lower than would be predicted by the status of their cousins in England.

Conclusions

The fact that social status in a society like England is largely genetically determined has implications for the role of migration in generating economic prosperity. If social abilities are largely socially determined, then there can be a world of plasticity, where immigrants in the long run play no substantial role in social and economic outcomes. As they become acculturated to the receiving society, they adopt the mores and social abilities of that society. Similarly who leaves a society has no long run implications also. Whatever roles they played will be occupied by a fresh group from within the society.

If, however, each person in a society, and each emigrant and immigrant, carries in their DNA the material that determines social outcomes, then migration can play a substantial role in determining long run outcomes. The arrival of groups with substantial social abilities like the Huguenots in 1685, or the Ashkenazi Jews from Russia and Poland 1880-1914, then represented a permanent improvement in the social and economic abilities of the society. This gain is also preserved even when the group, as with Huguenot descendants, is completely culturally absorbed into English society.

This has the implication that societies have a significant material incentive to attract high status immigrants. As illustrated in the case of the Netherlands, different immigrants carry with them substantially different net fiscal costs or benefits, especially in modern social welfare societies.

England historically was a net gainer from migrant flows all the way from 1685 to the present. The emigrants it exported to the New World settlement societies were only modestly positively selected. But at the same time the immigrants it has attracted have generally been of much higher social ability.

17. Social Class and Reproductive Success: 1690-2026

With status outcomes largely determined by additive genetically inheritance in England, a second force shaping social abilities for the society, along with selective migration, is the relative reproductive success of different social classes. This will have potentially significant effects on both average social abilities, as well as the distribution of those abilities within England.

In this chapter I estimate the reproductive success of three elite subgroups in England: those who enrolled at Oxford or Cambridge, those with high occupational status, and the wealthy, compared with the population as a whole.

I focus on the reproductive success of men in earlier years, since the structure of the lineage data makes male descendants easier to identify. Since for men reproduction can occur late in the life-course I restrict these estimates in the modern period to men and women born before 1964.[96]

The three measures we have of high status occur at different periods in the life cycle. Enrollment at Oxford or Cambridge is a measure for ages 16-21 typically, early in the life course. Occupational status is typically measured at age 40, so this is a status measure from mid-life. Wealth here is measured at death.

To measure reproductive success we need a baseline for England as a whole. This we get from the estimated Net Reproduction Rate (NRR) by birth cohort. The standard NRR is the average number of daughters that would be born to a woman if she passed through her lifetime conforming to the age-specific fertility and mortality rates of a given year. The NRR used here was calculated for England by the Cambridge Group for 1670-1837 from parish records, then using the national birth and death registration data 1837-2020.

---

[96]We only have complete fertility records up until 2007, so for the last cohort of men born 1960-4 we get estimates of their fertility only up to ages 43-47. I extrapolate to cover their expected lifetime fertility.

To estimate a cohort NRR we adjust the annual estimates by taking the average NRR for women by birth cohort as the annual NRR in their most reproductive ages of 20-39 for the birth years 1670-1919. For the years 1920-1979 we have detailed estimates of births, by birth cohort and year, that we can convert into an estimated NRR by birth cohort.

This cohort NRR is shown in figure 17.1. As can be seen English population was growing by nearly 50 percent for each cohort for the cohorts born 1750-1849, but overall fertility in England was at less than replacement rates for the cohorts born 1880-1929, and 1950-1979.

With genealogical data we need to ensure an equal probability for inclusion in fertility measures of men who leave no descendants, as for those with descendants. In the genealogies used here following everyone born with a given surname, this is not an issue except for the earliest generations. In those childless men with the surname may be less often recorded. To avoid this I include men in the fertility calculation only if their brothers' fertility is also recorded.[97]

Within the FOE genealogy it is difficult to directly measure the NRR. For many women we observe them till marriage, but then do not have records of their subsequent fertility, or age of death. Also, for many children, we observe they do not reach adulthood, but do not know the exact age of death. Thus we employ a simpler measure of net fertility. This is N21, which is the number of children who lived to age 21 born by each man who lived to age 21.

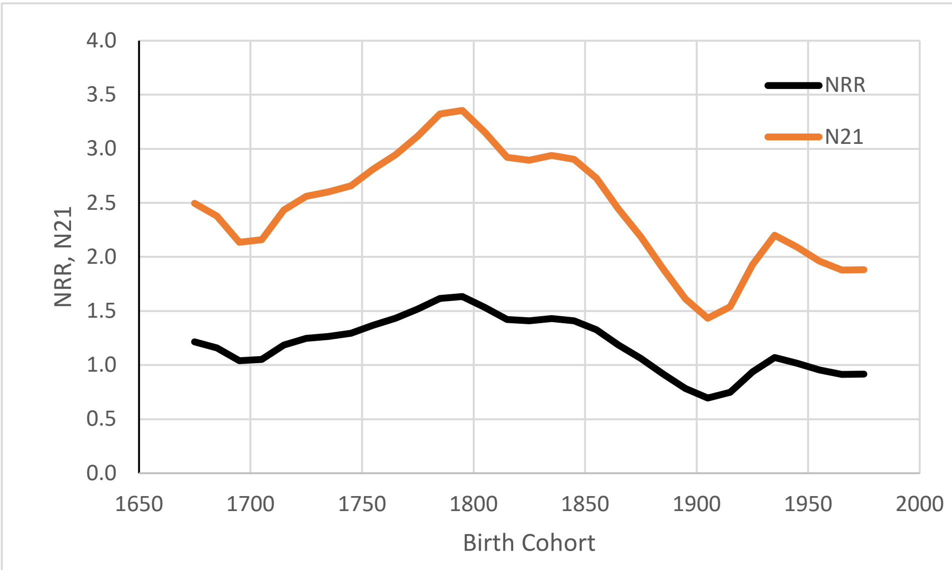
Assuming all fertility for men occurs at age 21 and above (close to reality for England in this period), this measure should be very close to the NRR multiplied by 2. In figure 17.1 we also convert the cohort NRR measure into an estimated number of children who reach ages 21+ per mother who reaches age 21+. This is 2.055 times the NRR, to allow for the higher fraction of males at birth.

Thus for the birth cohort of the 1890s where the national estimated NRR was 0.78, N21 estimated from the FOE database for men in families of average lineages was 1.66 which implies a NRR of 0.81. This is not significantly different from the national estimated NRR.

---

[97] This still leaves the possibility of exclusion of men where all brothers in a generation were childless. But such exclusions will be rare.

Figure 17.1: NRR and N21, birth cohorts 1670-1979, England and Wales



Sources: NRR, 1670-1837 – Wrigley et al., 1997, 614.

In earlier years there was more chance of a daughter being undetected than for a son. In elite family genealogies constructed at the time there was more attention to sons than daughters. Thus for men born prior to 1800 their fertility is calculated as double the number of adult sons they produce.

In later years, the 1950s-1960s we need to adjust fertility upwards to account for the fact that we observe all births for men only until early 2007. This is done by observing for men in this group born in the previous decade what fraction of their net fertility occurs after the cutoff date for their fertility to be observed.

Oxford and Cambridge Matriculants, 1690-1964

The first well defined elite group we examine are those who were born 1690-1964 and attended Oxford or Cambridge universities as students at some point in their career. This is mostly men in these years, though with some women in the later decades. Throughout these years Oxford and Cambridge were highly selective institutions admitting typically just 1-2 percent of each cohort of males. Thus we are observing the fertility of an educational elite.

This sample included, first, all persons in the FOE database who attended Oxford or Cambridge, and for whom we have a complete record of their siblings. This is to avoid bias in favor of higher measured fertility in earlier years where men only get recorded because they have offspring who themselves also produce offspring. Second the sample used here includes a sample of Oxford and Cambridge matriculants born 1680-1799 and 1920-64 with known fertility to extend the data in these years where the FOE data is thinner.

We are able to trace the numbers of adult children of many of these attendees, even in the years before 1841 and census records, because of a number of factors. First many of these higher status men left wills proved in the Probate Court of the Archbishop of Canterbury, whose texts have been preserved. These wills often reveal how many surviving children the testator had at the time of writing. Second, because many of these men played significant roles in the Church, in politics, in law, and in the military, biographical details are available for many of them. Third many of them were from elite families where detailed family trees were already being constructed, even in the eighteenth and nineteenth centuries.

Figure 17.2 summarizes the results for the sample of 3,884 attendees I have been able to assemble. Half this sample were all holders of a set of rare surnames that was used to form the FOE database (2,129 persons). But this sample was substantially supplemented for the birth years 1680-1799 and 1920-1964 by a further sample of 1,829 Oxbridge attendees with rare surnames whose fertility could be traced.

In earlier years there was much less attention to documenting the careers of daughters than those of sons. Daughters were more likely to be omitted from family trees and biographies than were sons. Or daughter's births were recorded, but with no detail on whether they achieved adulthood. For those in our FOE sample born 1680-1799, for example we find 671 adult sons recorded, but only 555 adult daughters. To correct for this I infer the number of children per attendee for those born before 1800 as the number of sons recorded times two. This will, however, inflate the standard errors of the estimates in these years.

For the last cohorts there is truncation of observed fertility. For England we only have complete public birth records through the end of 2006. In my sample we have people born up until 1964 where we are relying on complete birth records, so we only observe fertility for people before ages 42-46 in last decade. Similarly in the previous 1950s decade we observe fertility completely only up until ages 47-56. I adjust the reported fertilities for 1950-64 using information on net fertility rates for those born in the previous decades at ages 42 and above. For 1960-4 this raises net fertility by 6.3 percent, and for 1950-9 by 1.1 percent.

Figure 17.2 shows the pattern of these estimates for the educational elite, compared to estimates of comparable measures of fertility for the general population. There is an initial period, births 1680-1779, where the net fertility of the Oxbridge elite averaged 2.96 adult children, compared with 2.51 for England as a whole. In pre-industrial England the educated elite had a reproductive advantage compared to the general population.

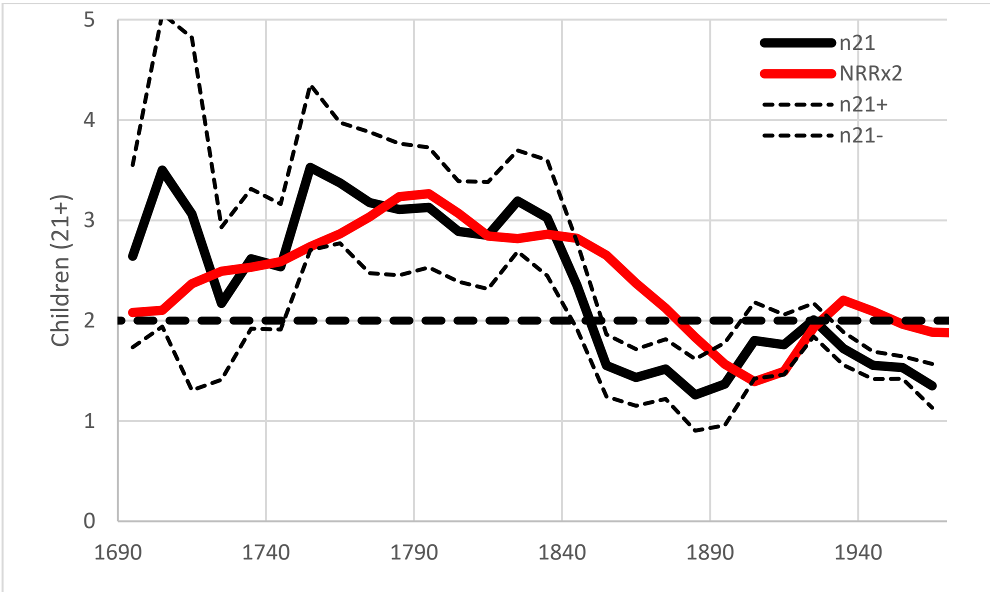
This reproductive advantage is surprising in light of figure 17.3, which shows the percentage of men in each period who never married for the three elite groups. Those attending Oxford and Cambridge had a notably low marriage rate. The never married fraction was nearly 25 percent in 1680-1899.

One constraint on the marriage rate of Oxbridge matriculants before 1882, was that Oxford and Cambridge faculty had to be unmarried. These Oxbridge faculty were mostly recruited from Oxbridge graduates. But even for those Oxbridge matriculants born 1900-69 the share never married remained very high, between 17 and 20 percent.[98] These rates greatly exceeded the never married share for those of high occupational status.

This initial period of high Oxbridge fertility 1680-1779 was followed by a second interval in the early Industrial Revolution, for births 1780-1849, where the net fertility of the educational elite was now modestly below that of the general population. Net fertility among the elite changed little in this period, at 2.93 adult children per Oxbridge attendee, still well above replacement fertility. But in the general population net fertility had risen to 3.07.

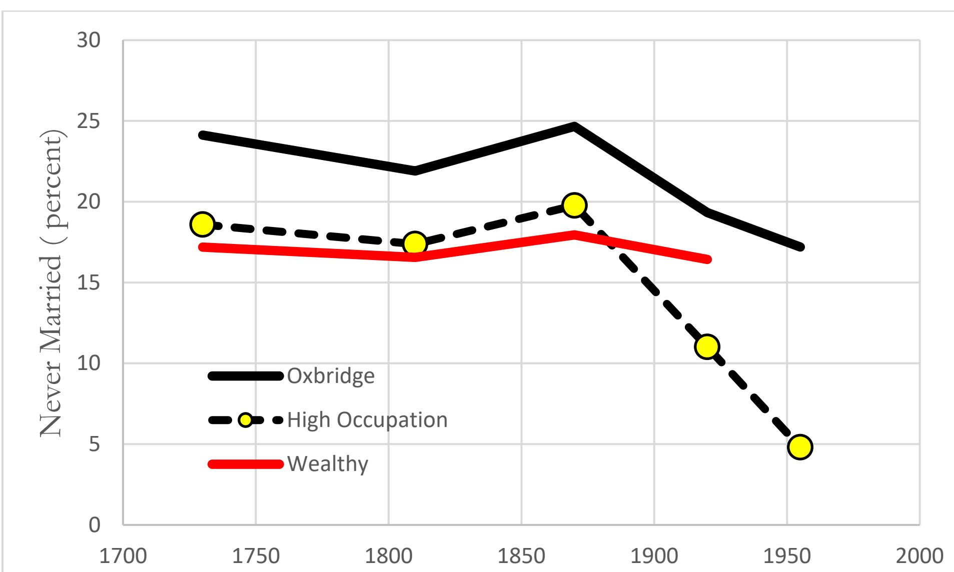
[98] By the birth cohort 1950-69 unmarried cohabitation was becoming more common. So I counted all persons living with a person of the opposite gender on the electoral register at any time 2002-2024 as having been married.

Figure 17.2: Net Fertility Oxbridge Attendees versus General Population



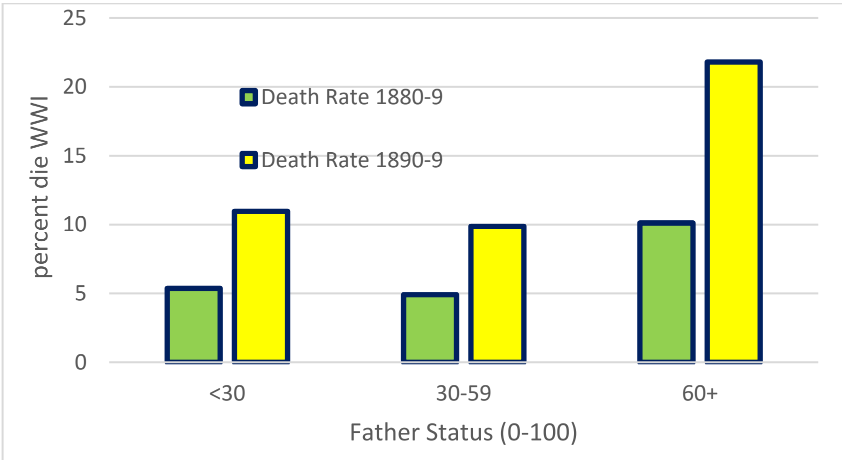
Note: The solid black line shows the decadal average number of children living to age 21+ for enrollees at Oxford and Cambridge who themselves lived to age 21+. Dotted lines show the 95 percent confidence interval of this estimate. The solid red line shows the period NRR for women times 2.

Figure 17.3: Percent of Men Never Married by Period



Note: The periods are 1680-1779, 1780-1839, 1840-1899, 1900-39, and 1940-69.

Figure 17.4: Share of Males Dying in WWI by birth decade and social class



Note: Table shows the death rate in WWI by birth decade and father occupational status on a scale of 0-100.

However, very quickly, for cohorts born 1850 and later, who would be first marrying typically around 1880, fertility dropped sharply. While net fertility was 2.86 for the 1830s birth cohort, it was 1.60 for the 1850s cohort. In only one birth decade from the 1850s to the 1960s, the 1920s, does net fertility for the Oxford and Cambridge attendees attain 2.00 or above. Net fertility was at its lowest for the cohort born in the 1880s, where it was 1.25. It is notable that Francis Galton coined the term *Eugenics* in his book *Inquiries into Human Faculty and its Development*, published in 1883. From the 1930s on net fertility of Oxbridge enrollees was well below 2, and well below that of the general population.

The very low net fertility of the cohorts born in the 1880s and 1890s, and low marriage rates, may also be partly explained by deaths in WWI. As figure 17.4 shows the share of men with fathers in professional occupations dying in WWI was 10 percent for those born in the 1880s and 22 percent for those born in the 1890s.

For the modern period we can estimate fertility separately for male and female enrollees at Oxford and Cambridge. For those born 1920-1964 the average net fertility of men was 1.69, and for women 1.57.

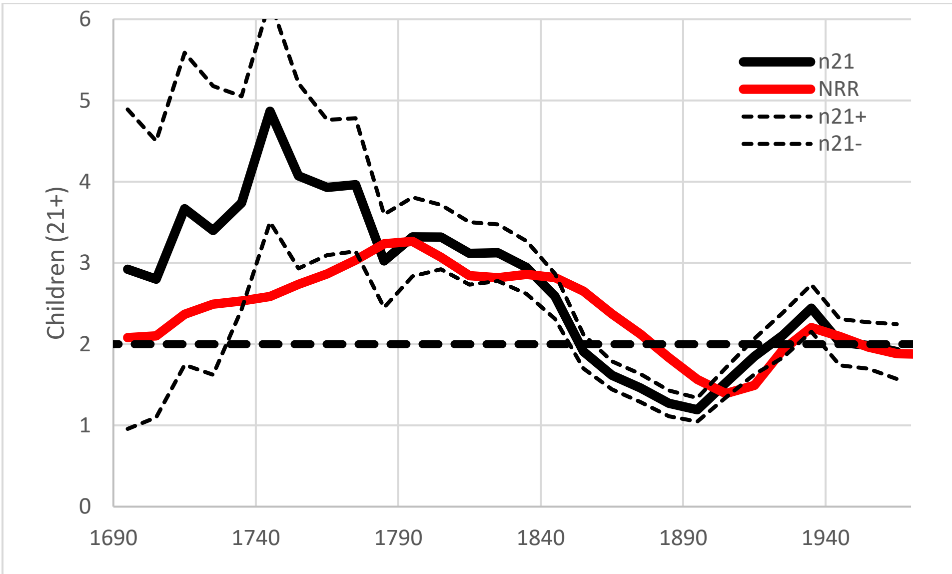
High Occupational Status Men, born 1690-1964

We include here men whose occupational status was 60 and above on a 0-100 scale of occupational status.[99] These occupations include Academic, Accountant, Architect, Army Officer, Auctioneer, Bank Official, Bookseller, Chemist, Church of England Clergyman, Higher Civil Servants, Civil Engineer, Company Director, Company Secretary, Dentist, Diplomat, Esquire Or Gentleman, Estate Agent, Grazier, Large Farmer, Head Teacher, Insurance Official, Ironmaster, Journalist, Landed Proprietor, Lawyer Or Jurist, Manufacturer, Medical Doctor, MP, Merchant, Navy Officer, Outfitter, Publisher, Scientist, Secretary, Stationer, Stenographer, Stockbroker, Tea Planter, Titled, University Student, and Veterinarian.

Figure 17.5 shows the net fertility of this group by decade. For births in the period 1700-1779 high occupational status men had net fertility which greatly exceeded that of the general population: 3.71 versus 2.53 adult children per man. But for those born in the 1850s to 1890s net fertility, as with Oxbridge enrollees, was far below the general average. However, for this group net fertility for men born 1900-1964 tended to be at or above the social average. This primarily stemmed from the much greater marriage rates of these men born 1900-1964 compared to Oxbridge graduates.

[99] See Clark, Cummins and Curtis, 2024.

Figure 17.5: Net Fertility High Occupational Status Men versus General Population

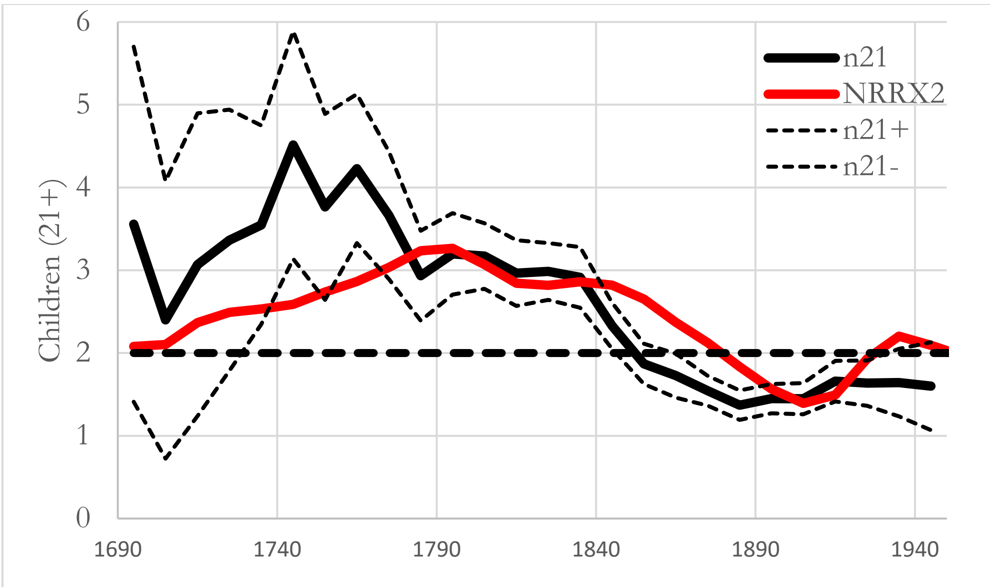


Note: The solid black line shows the number of children living to age 21+ for high occupational status men. Dashed lines show the 95 percent confidence interval of this estimate. The solid red line shows the period NRR for women times 2.

Wealthy Testators, born 1690-1949

For those dying 1858 and later we have well observed average wealth at death from probate records. We classify testators as wealthy if their wealth above 5 times average wealth for their decade of death. For deaths in the years 1796-1857 the Prerogative Court of Canterbury (PCC), as well as the Prerogative Court at York (PCY), where the wealthiest testators were probated, gives a measure of the estate value at death. For those dying in the years 1720-1857 we project average wealth at death backwards from the £291 of the 1860s in a relatively ad hoc manner to be 200 in the 1720s. Since for men the great majority of those probated in the PCC had wealth greater than 5 times average wealth in the period 1796-1857, we assume any man probated in the PCC met the wealth standard of 5 times average wealth where the estate values were not listed. We only included men in the PCY court if their wealth at death 1797-1857 exceeded 5 times the assumed average wealth.

Figure 17.6: Net Fertility Wealthy Men versus General Population



Note: The solid black line shows the number of children living to age 21+ for wealthy men. Dashed lines show the 95 percent confidence interval of this estimate. The solid red line shows the period NRR for women times 2.

Figure 17.6 shows the number of births in each decade, the average number of children reaching age 21, the 95 percent confidence interval of this estimate, and the Net Reproduction Rate (NRR) times two for women born in each decade for England as a whole.

For wealthy men born 1690-1779 net fertility is very similar as for men with high occupational status, and substantially exceeds the national average. For births 1780-1849 net fertility is close to the national average. As with both Oxbridge attendees and high occupational status men, net fertility for births 1850-1899 is well below the social average. Finally in the period of births 1920-49, net fertility of the wealthy again falls well below the social average, as in the Oxbridge case.

Overall Fertility of the Elite, born 1690-1964

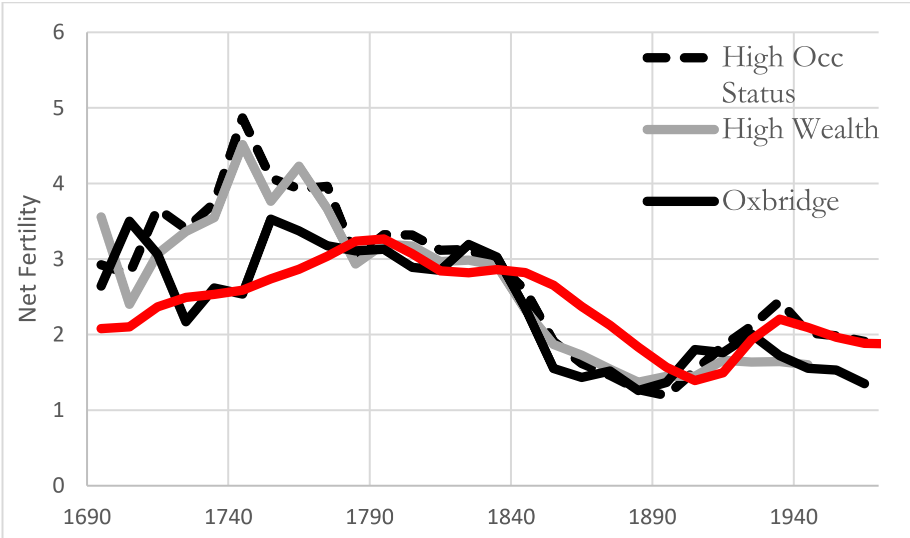
Figure 17.7 shows for all three elites net fertility for men born 1690-1964. It is clear that there was a period, births 1690-1780, where elite fertility substantially exceed average fertility. This was followed by the period of birth 1780-1850 where elite equaled average fertility. Then we see a two generation period, births 1850-1910 where elite fertility was substantially below average fertility. In the modern period the picture is less clear. For men born 1910-64 with high occupational status, their net fertility equals that of the general population. But for men and women attending Oxford or Cambridge, and for the wealthy, their fertility 1930 and later is substantially below average fertility.

An alternative measure of fertility levels for those born 1920-1979 can be derived using people with rare surnames. Since for England we have comprehensive birth registers available 1920-2005, we can approximate the NRR per surname as just the number of births in a 30 year period divided by the number of births in the previous 30 year period. This measure has the advantage of allowing also for "illegitimate" fertility, where only a mother is recorded on the birth certificate.

We can assign status to surnames in a number of ways. One is to link the surnames to addresses listed in the 1999-2026 electoral register. From the electoral registers we get an estimate of average house value by surname, normed to 2017. We divide the 448 surnames in the FOE database with evidence on house values 1999-2017 into those with house value less than £200,000, £200,000-299,999, £300,000-499,999, and £500,000 and above. The £500,000 and above group represents about 8 percent of all houses in England and Wales, so represents roughly the highest status 8 percent of the population. The other three groups are respectively 54, 22, and 15 percent of the population.

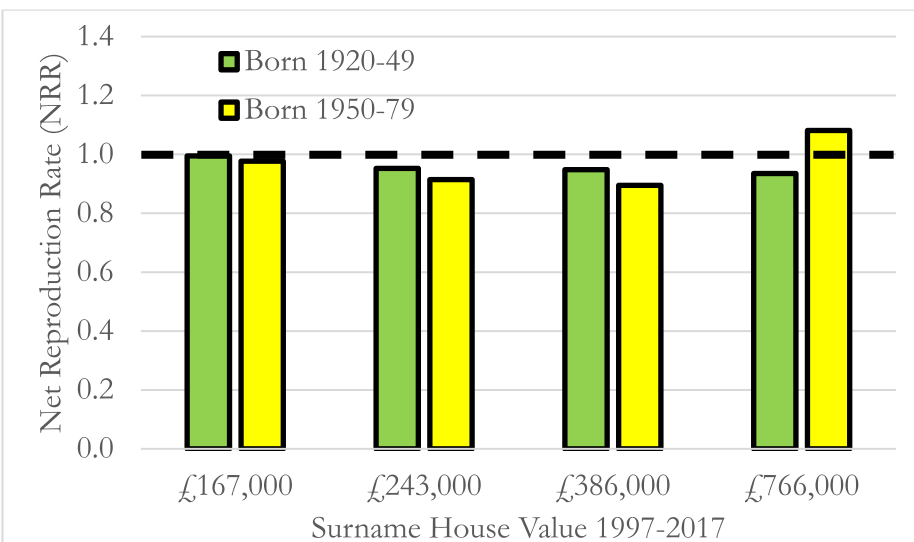
Figure 17.8 shows the approximate NRR for each category of surnames for those born 1920-49, and 1950-79. Overall for the non-elite surnames in the FOE database, the NRR was 0.96 for those born 1920-49, and 0.94 for those born 1950-79. Thus in the figure those born 1920-49 show a slight negative fertility gradient with status. But for those born 1950-79 there is a U-shaped fertility pattern, with the highest status surnames showing 15 percent higher than average fertility, but the second highest status group showing 5 percent below average fertility. But overall for the most recent generations with completed fertility the evidence is of very modest status gradients in fertility.

Figure 17.7: Net Fertility of Three Elites, 1690-1969



Source: Figures 17.2, 17.5, 17.6.

Figure 17.8: Surname NRR by Surname House Value



Note: Births 1980-2006 adjusted to a 30 year basis.

Hugh-Jones and Abdellaoui, 2022, using UK Biobank data on the fertility of those born 1940-69, find a modest gradient of those with lower polygenic indexes for educational attainment having higher fertility. They also observe a decline in average polygenic scores for educational attainment for those born 1940-69 of 0.055 of a standard deviation.

This study also finds that the fertility differences driving the decline in polygenic scores were concentrated in the parents of lowest socioeconomic status. These differences were thus leading to a widening dispersion of polygenic scores within the population.

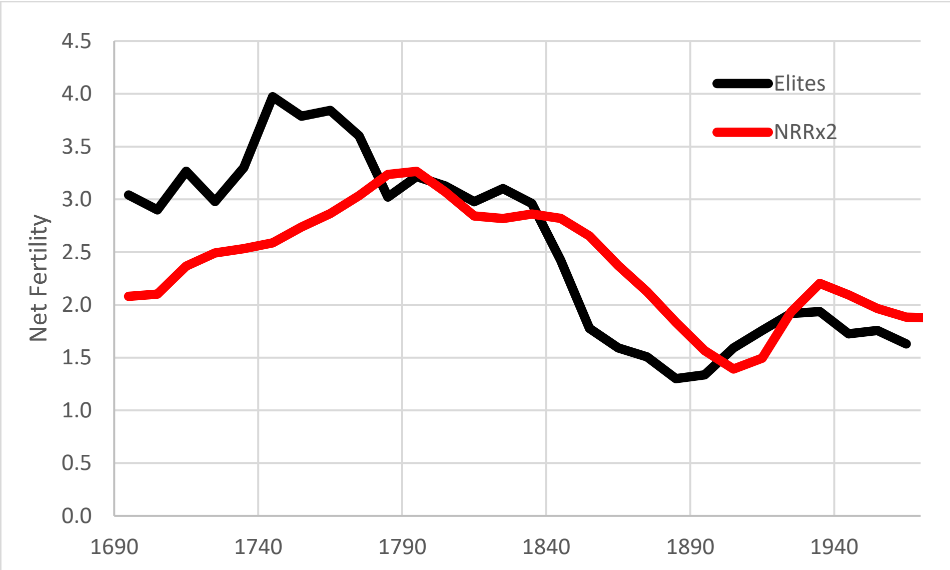
Elite Fertility and the Distribution of Abilities in England, 1500-2026

Figure 17.9 shows the average net fertility of the three elite groups, versus the population as a whole. This suggests the overall narrative is one where elite net fertility substantially exceeded general net fertility for birth cohorts of the 1690s to 1770s. This conforms with a pattern of "survival of the richest" found in pre-industrial England all the way from the 1250s to the 1770s. The wealthiest testators in England born from the 1490s to the 1760s showed strong net fertility advantages over the poorest testators (Clark and Cummins, 2015). That advantage was stronger the further back we go in time. Other measures of net fertility, which include the chances of a man having a surviving son at death, imply that elites in medieval England 1250-1450 also showed substantially higher than average fertility (Clark, 2007, chapter 6).

Grouping people by 30 year birth generations, figure 17.10 shows the size of a series of synthetic cohorts of children of the different social cohorts, with a base in each case of 100 in 1680-1709. These are synthetic cohorts in the sense that they do not follow the descendants of the original elites 1680-1709, who would over time regress to the mean. Instead they look at the number of descendants of each elite group had all their descendants remained in an elite position. We do see in chapter 4 that persistence was actually very strong, so this counterfactual will not violate reality too strongly.

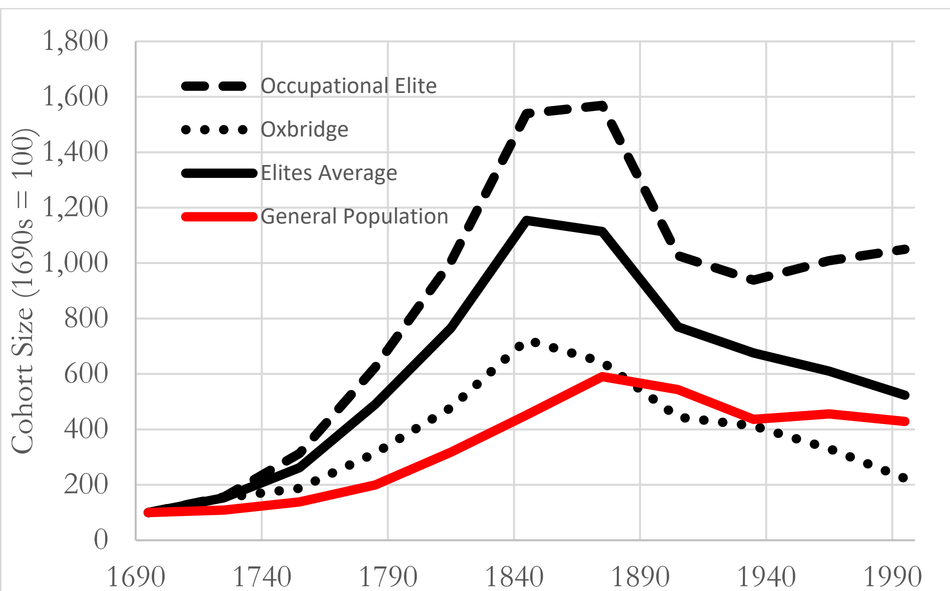
As the figure shows, there was a very substantial increase in the descendants of the elite compared to the population in general from those born 1680-1709 to 1830-59. The elite group overall grew relative to the average population by 2.7 times. But from births 1830-59 the reproductive success of the elite was mostly below the general population. By births 1980-2009 the size of the elite cohort was now only 1.25 times the size of the general population compared to 1680-1709.

Figure 17.9: Elite Average Net Fertility, 1690-1969



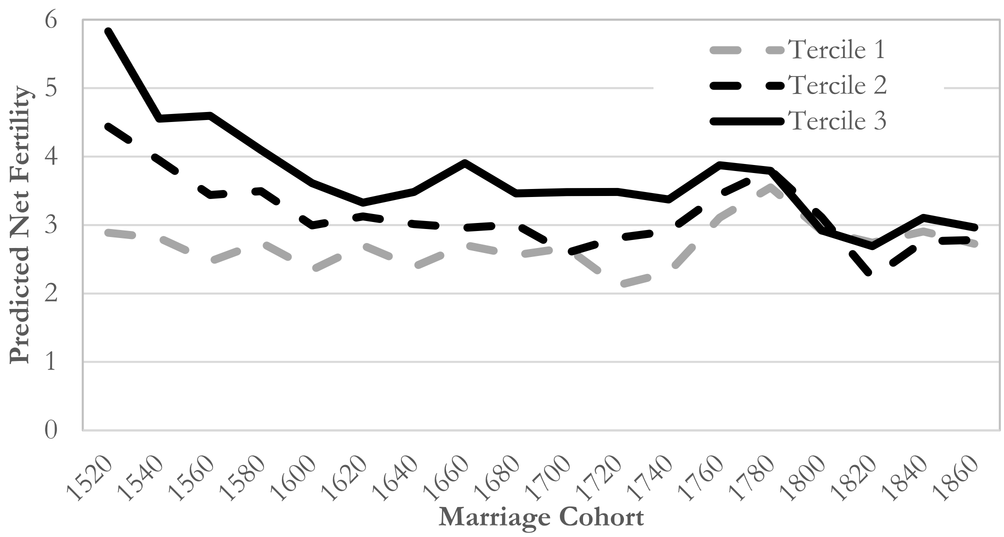
Source: Figures 17.2, 17.5, 17.6.

Figure 17.10: Cohort Size and Social Class by Generation, England 1690-1994



Source: Figure 17.7, 17.9

Figure 17.11: Net Fertility by Wealth Terciles, marriage cohorts, 1520-1879



Source: Clark and Cummins, 2015, figure 7.

In earlier work with Neil Cummins (Clark and Cummins, 2015) we estimated the net fertility of wealthier men in England first marrying in the period 1520-1879 from the numbers of children named in their wills. Will makers themselves were a modest share of all men 1540-1858 at the upper end of the status distribution. But we further divided up the wills by wealth terciles, so that the top tercile would be typically the top 4-8% of the male wealth distribution. Figure 17.11 shows the estimated net fertility of that top wealth tercile, by twenty-year marriage periods 1520-1879. This shows that the period 1680-1709 was just the continuation of a long pre-industrial era in England, back at least to 1250, where social elites had much higher net fertility than the general population.

It is clear from the discussions in the chapters above that the sustained period of superfecundity by elites in England, for those born 1250-1780, must have led to significant changes in the genetics underlying social abilities. The share of the population with the highest level social abilities is calculated here to have maximized for the cohorts born 1780-1849. This is also, of course, the period where England had the most impact on the world stage, though that may be no more than coincidence.

Conclusions

The suggestion that there were very significant changes in social abilities for England as whole driven by differences in net fertility with social status will be uncomfortable reading for many. It should not be taken as implying that society should promote fertility by the advantaged groups, and discourage it for the disadvantaged. Such policies would not command approval from the vast majority of people in modern societies. And anyway, recent policies to promote fertility, such as extending parental leave, have had little effect (see, for example, Thomas *et al.,* 2022).

18. Marital Assortment

Chapter 4 gave evidence of very strong marital assortment by social status in England 1600-2022. That assortment showed a phenotype correlation in underlying status of 0.8, and a genotype correlation of close to 0.6.

Marital assortment is important both for rates of intergenerational social mobility, and for the distribution of abilities within society. Further, with genetic transmission of social status marital assortment will have maximal effects on both intergenerational persistence, and on the social distribution of abilities. For with genetic transmission mothers and fathers have equal effects on child outcomes, and that is the condition under which assortment has the biggest effect.

Here I argue that the high degree of assortment evident in English marriages 1754-2021 was unusual relative to most pre-industrial societies. It would have produced an unusually large variation in social abilities in English society, compared to most pre-industrial societies. In recent years in the eternal search for the sources of the Industrial Revolution some scholars have claimed that the focus should be on the attitudes and behaviors of an "upper-tail" of 1-2 percent of the population who drove technological and institutional innovation.[100]

If the Industrial Revolution was an upper tail event in European society, driven by the attitudes and capabilities of the top 1-2 percent of the population, such marriage patterns could explain the abundant supply of upper-tail talent in England, in particular, and northern Europe in general, circa 1800.

Marital Assortment and the Dispersion of Ability

If x_f and x_w are the genotype value of the father and mother, then the genotype value of the child will be

[100] The term was introduced by Mokyr, 2005. See also Mokyr, 2017 and Squicciarini and Voigtländer, 2015.

$$x_c = \frac{1}{2}(x_f + x_w) + e \tag{1}$$

where e is a random shock (needed also to maintain genotype value variance). These assumptions comport with the evidence of chapter 4 on status transmission in England 1600-2026.

With a genotype assortment between parents of m, then

$$x_w = mx_f + v \tag{2}$$

with v again a random shock. This in turn implies

$$x_c = \frac{(1+m)}{2}x_f + \frac{1}{2}v + e \tag{3}$$

Thus the greater is genetic assortment, the lower will be social mobility rates.

From (1) the variance of child genetic value x_c will be

$$\sigma_{x_c}^2 = \frac{1}{4}var(x_f) + \frac{1}{4}var(x_w) + \frac{1}{2}cov(x_f x_w) + \sigma_e^2 \tag{4}$$

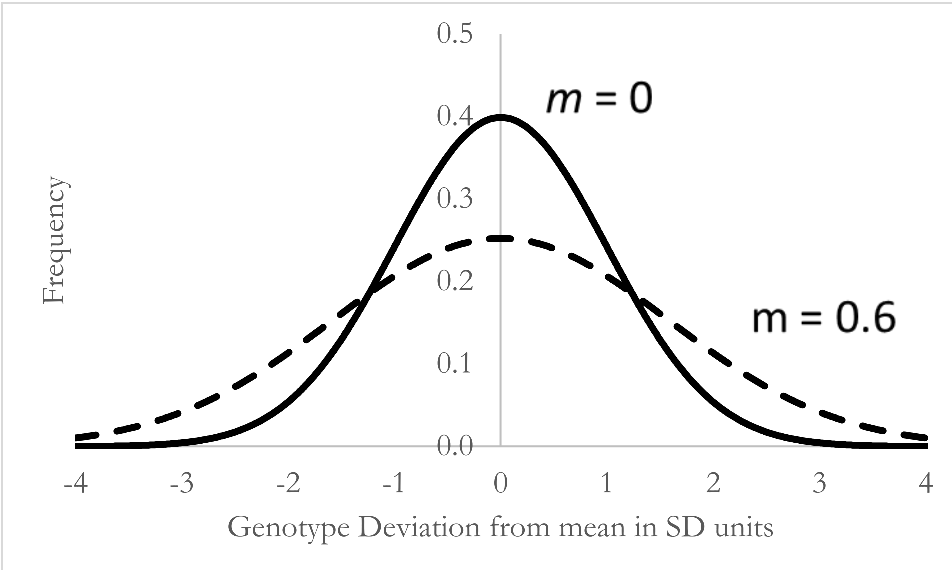
In long run social equilibrium, σ_x^2 will be the same across generations, and across men and women. Since $var(x_f) = var(x_w) = \sigma_x^2$, and $cov(x_f x_m) = m\sigma_x^2$, where m is the correlation between x_f and x_w, then (1) implies that in equilibrium

$$\sigma_x^2 = \frac{1}{2}\sigma_x^2 + \frac{1}{2}m\sigma_x^2 + \sigma_e^2 \tag{5}$$

$$\Rightarrow \quad \sigma_x^2 = \frac{2}{(1-m)}\sigma_e^2 \tag{6}$$

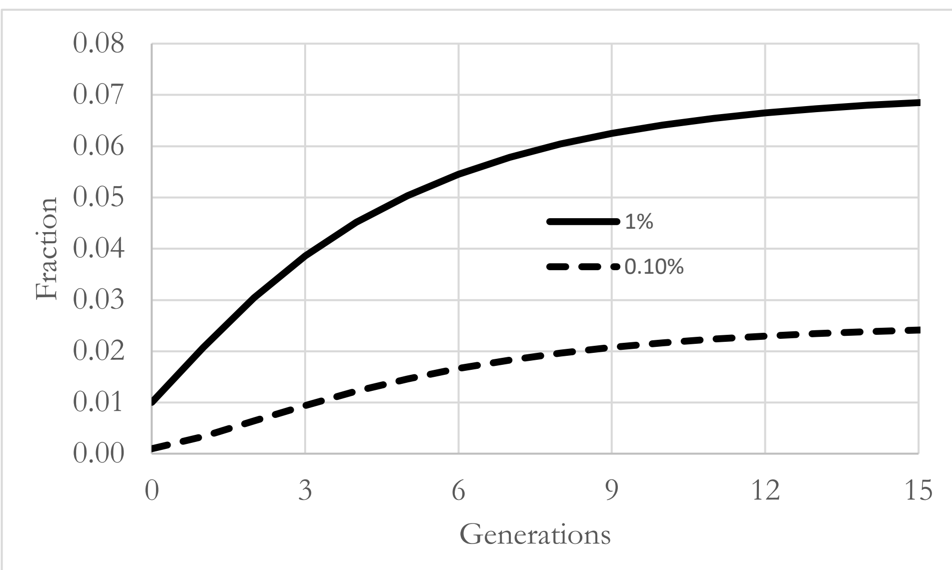
This implies that the variance of underlying social abilities will depend strongly on the degree of assortative mating. If m increases from 0 to 0.6, then the variance of underlying abilities under (6) would increase by a factor of 5. Assortment would, of course, have no effect on the average level of abilities. Figure 18.1 shows the implied equilibrium distribution of social abilities with no marital assortment, versus a 0.6 genetic correlation between parents.

Figure 1: Marital Assortment and the Social Distribution of Abilities



Note: Deviation from mean shown in standard deviation units where there is no parental assortment.

Figure 2: Dynamics of an Increase in Marital Assortment



Notes: The figure shows the share of the population with abilities in the top 1 percent, and 0.1 percent of the distribution where there is no marital assortment.

However, a change in marital assortment from no assortment to a correlation between parents of 0.6 would expand the distribution of social abilities only over the course of many generations. It would take some time for the distribution to widen to the new equilibrium. If we rewrite (5) in terms of the dynamics of σ_x^2 over generations, indexed by t, then

$$\sigma_{x_{t+1}}^2 = \frac{1+m}{2}\sigma_{x_t}^2 + \sigma_e^2 \tag{7}$$

Figure 18.2 shows the spread of the distribution across generations with the arrival of marital assortment. It shows by generation the share of the distribution that would be in the top 1 percent, and top 0.1 percent of the old distribution of abilities where there was no marital assortment. The share in the old top 1 percent would rise eventually to 7 percent of the population. But it would be less than four fifths of the way to that higher share 6 generations, 180 years, after the increase in assortment. The change in the top 0.1 percent share would be more dramatic: a twenty-five-fold increase. But after six generations only two thirds of that increase would be achieved.

Thus changes in marital patterns in the late middle ages in Europe would potentially affect the population distribution of abilities fully only 300 years later at the time of the Industrial Revolution. We show also that earlier marriage systems where families arranged the match would typically display much less assortment, and thus less variance within the population in social abilities.

Such a change is of interest because of our difficulties in explaining the long delay in the Industrial Revolution (Clark, 2007). In the previous chapter I emphasize one element in this, which was the reproductive success of those more economically successful in pre-industrial England from 1250 onwards. This would increase average ability level in the society. But a second factor that would also potentially influence the rate of technological advance would be a widening of the ability distribution caused by more assortative marriage.

Why was the degree of social status assortment in marriage so high in England by 1754? Men and women could instead choose to value physical appearance most highly, or age, or height, or family lineage social status. There is no clear reason why marital assortment should have focused so heavily on social status.

But there is the intriguing possibility of a link to the pre-industrial period in England. I noted in the previous chapter that for at least 500 years before 1800 there was a strong positive gradient between socio-economic status and fertility (Clark, 2007, chapter 4, Clark and Cummins, 2015). We also saw in the previous chapter that from 1680-1780 the occupational and wealth elites in England had much greater net fertility than the average family. Since both parents mattered equally for child outcomes, those who married a partner of the highest possible social status would have descendants who were more successful, and consequently had more reproductive success. So it is even possible that highly assortative mating is a hard wired consequence of a pre-industrial world of "survival of the richest."

The History of Marriage

We saw in chapter 4 that English marriage was highly assortative from 1754 onward. We also see in chapter 7 the equal influence of mothers and fathers on child outcomes, for outcomes associated with human capital rather than physical capital. This implies that the distribution of social abilities would be strongly influenced by marital assortment.

This raises the question of whether one source of greater abilities in the upper tail in England in the Industrial Revolution period was the earlier cultural adoption of highly assortative marriage, associated with the European Marriage Pattern.

The European Marriage Pattern, which present in England from at least 1538 had one important feature for marital assortment. This was that most marriages were between women and men who were legally, and also economically, independent of their parents. Men and women aged 21 or less had to have their father's consent to marry.[101] But in the years 1538-1750 the average age of marriage for women was close to 24, and for men 26. So the great majority of marriages were ones where parental consent was not required. In the Marriages of England (MOE) database, 77 percent of marriages 1837-1879 involved a groom and bride both aged 21 or older.

[101]In 1987 the age below which parental consent was needed was reduced to 18.

Another sign of the independence of the parties getting married was that throughout these years a significant fraction of men and women never married, in contrast to other pre-industrial societies where marriage was near universal, in particular, for women.

By the time of marriage many lower-class women had been living independently from their parents as servants in the homes of others. Upper class women were more dependent on their parents, as they typically did not have occupations, and would often receive a financial gift upon marriage from their parents. But these upper-class marriages were typically based on an arrangement between a couple, rather than as arrangements between families. Later marriages also allow men and women more time to reveal their realized adult social status than when people marry just after, or even before, puberty.

In contrast in many other pre-industrial societies marriage for men and women was at much earlier ages, and was much more determined by social conventions and family wishes. The Tsimane, for example, are a shifting-cultivation subsistence agriculture society in Bolivia, has been extensively studied by anthropologists (Godoy et al., 2008, Gurven et al. 2009). The average marriage age was sixteen for both men and women, and men and women typically correlate closely in age at marriage. As with other native Amazonians, the strongly preferred marriage partner is a specific cousin, with three quarter of couples in the study following this rule. For a man the preferred partner is the daughter of his mothers' brother, or his fathers' sister (Godoy, 2008, 204-5). This would imply a limited set of potential marriage partners for each party in a marriage.

Male suitors ask the parents of a woman for permission to marry the daughter. In a survey asking reasons for marriage among married Tsimane men and women, 38 percent of women reported the marriage was against their will, 7 percent reported they had no other option, and 9 percent reported it was arranged (Gurven, 2009, 161).[102]

If we assume that in cousin marriage, because of the constraints imposed, a person typically marries a random cousin, what will be the underlying correlation in traits such as education or occupational abilities? As we saw in chapter 4, where m is the marital correlation, the genetic correlation of cousins is

$$\left(\frac{1+m}{2}\right)^3$$

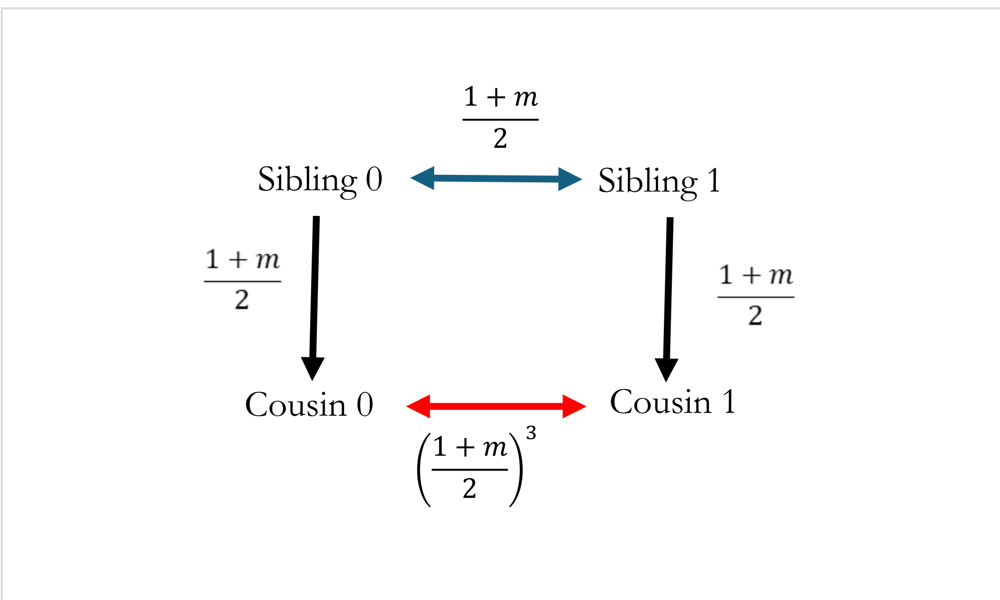
[102] Among men, 11 percent reported the marriage was against their will, 14 percent reported it was arranged

As illustrated in figure 18.3, because the link between cousins involves three steps, each with a correlation of (1+m)/2, we get the above correlation between cousins. This implies that in equilibrium in a society where everyone marries a random cousin we must have

$$m = \left(\frac{1+m}{2}\right)^3$$

This implies m = 0.24, which is a surprisingly low genetic correlation. But it is the only correlation between marital partners that is consistent under this social arrangement. Note that if cousins in England married, then the typical genetic correlation would be much higher, at 0.51, because m is much higher under English marriage patterns, so even randomly matched cousins would be much more closely related in terms of social outcomes.

Figure 18.3: Genetic correlation between cousins



Thus a system of cousin marriage, surprisingly, is much less assortative for social traits than marriage as practiced in England 1754-2021. In line with this, among the Tsimane the correlation in years of education at marriage was only 0.14.[103]

Cousin marriage was a surprisingly common marital arrangement in the pre-industrial world. All Muslim communities, for example, seem to have had high shares of marriages taking this form, a traditional that continues to the present day.[104] Korotayev argues this stems from the Islamic Law tradition that daughters are entitled to inherit from their fathers half of what their brothers inherit (Korotayev, 2000, 400). Cousin marriage keeps such inheritance within the extended family. However, at least in the modern era, cousin marriage is as common among the property-less, as among the propertied, in Islamic communities. Its association with Islam thus seems more cultural than economic.

There is, however, good evidence against the claim of Jack Goody (Goody, 1983) that cousin marriage was common in the Roman Empire in Western Europe. Shaw and Saller show from a study of noble genealogies and tombstone inscriptions that cousin marriage was rare (Shaw and Saller, 1984). But we do see in the Roman world another common feature of pre-industrial marriage that would reduce marital assortment. This was the practice of parents of contracting marriages for their daughters at young ages. Keith Hopkins estimates that the average age of marriage in pagan Roman society was well below 18, with even girls as young as 10-11 marrying (Hopkins, 1965).

In marital systems where the match was arranged by the families, suppose there is a correlation, m_f, in the genetics of occupational status between the fathers. In that case the correlation between the spouses will be $m_f\left(\frac{1+m}{2}\right)^2$, since the genetic correlation between each father and their child is $\left(\frac{1+m}{2}\right)$. This implies the spousal correlation, m_c, in social equilibrium would be

$$m_c = m_f\left(\frac{1+m}{2}\right)^2$$

[103] This was the correlation adjusting for age, since there is a strong age trend. Average years of education for women were 1.24, and for men 2.56 (Godoy, 2008, table 1). The correlation in knowledge of plant names was high, at 0.63, however.
[104]See Bittles, 2025. Cousin marriage is also common in parts of Hindu Southern India.

Assuming that the fathers match with the same strong genetic correlation of 0.6 we observe in English marriages 1837-2021, then in equilibrium

$$m = 0.6\left(\frac{1+m}{2}\right)^2 \rightarrow m = 0.23$$

Thus a marital system where the matching is between families, as represented by the fathers, will again produce a much lower correlation in spouses than observed in England 1837-2021, even if the status matching between fathers is as close as is observed for spouses in England. This equilibrium correlation for family matching is also very similar to that implied by cousin marriage.

There is evidence that in places such as medieval and early modern Italy, 1300-1650, girls continued to marry relatively young, and among higher status families, parents played a substantial role in their choice of husbands (Dean and Lowe, 1998, 7-11). Matching can only be as strong as is observed for England if the qualities of the spouses themselves that determine the match. But in marital systems where families do the matching, and the brides are very young, these qualities will play little role.

Thus, for example, in 1244 Henry III of England arranged with Alexander II of Scotland that his daughter Margaret would marry his son, the future Alexander III. At this time Margaret was 4 years, and Alexander 3 years old. The marriage itself took place in 1251. Henry himself had married Margaret's mother, Eleanor of Provence, when she was 12, and he 28. She and Henry met for the first time at their wedding. As late as 1641 in Protestant England Mary, daughter of Charles I, married at age 9 to 14 year old William, future Prince of Orange. The marriage was not consummated, however, until 1644.

So potentially the development of the European Marriage Pattern in England, sometime before 1538, where women were mostly living independently of their families by the time of marriage, and were marrying spouses who were contemporaries in age, led to greater inequalities in social abilities and a fatter tail of children with higher talents. It will also be the case that a switch from a regime of little marital assortment across marriage partners to one of high assortment would take many generations to have full effect in terms of the distribution of abilities.

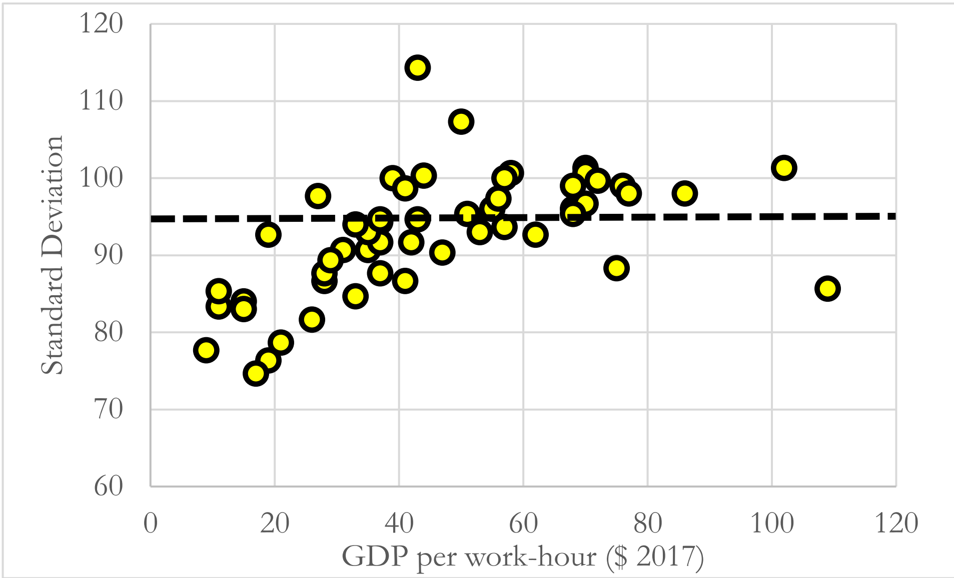
Henrich (2020), and Schulz et. al. (2019), argue that the medieval church's prohibitions on cousin marriage weakened kinship structures in Western Europe. This eventually resulted in psychologically different populations when compared to other global regions where cousin marriage preserved intensive kinship ties. But here we are focused just on its effects of marriage systems on the degree of assortment, and hence the distribution of abilities. There is no implication that the European Marriage Pattern changed the general level of abilities.

We do not know when the European Marriage Pattern was adopted in England, or elsewhere in Northern Europe. We know it was in place in England by 1538 when a mandate of Henry VIII required each parish to keep a record of weddings, christenings, and burials. But knowledge of medieval demographic conditions is too limited to allow an estimate of whether this pattern was an innovation of the fourteenth century, the thirteenth century or earlier. But we see above in figure 18.2 that a switch to greater assortment even in 1300 would not have full effect in terms of population distributions of abilities until more than 300 years.

Matthew Curtis, who first applied the measures of marital assortment used here to Quebec, 1700-1970, finds a similar degree of marital assortment in the French derived population of Quebec from 1700 on. He also is able to show that women were as important as men for children's educational outcomes, measured as ability to sign the marriage register (Curtis, 2026). Thus it seems likely that a high degree of marital assortment was a general feature of countries with the European Marriage Pattern across northern Europe.

So there is an intriguing possibility that there is an underlying connection between the European Marriage Pattern, greater assortment in marriage, and the timing and location of the Industrial Revolution in Europe.

Figure 18.4: Standard Deviation of PISA scores versus GDP per work-hour



Sources: OECD, 2019. Penn World Tables, Version 10.01 (Feenstra et al., 2015).

Notes: Average Standard Deviation across Mathematics, Language, and Science tests. Dashed line shows average of OECD members. The upper outlier in standard deviation is Israel.

The PISA data on educational performance of 15 year olds gives some support to the idea that high income countries now tend to have a higher variance of academic ability than lower income countries. Figure 18.4 shows the standard deviation of PISA scores in 2018 versus GDP per work-hour in 2017. For the high income Western and Asian societies the variance is similar. But it is clear that, on average, the lower income countries have a lower variance of PISA scores. The countries with standard deviation below 80 are Costa Rica, Dominican Republic, and the Philippines. Other countries standard deviation below 80 with relatively low incomes, but not included in the figure because of data limitations in the Penn World Tables are Albania, Bosnia, Kosova, Indonesia and Morocco.

Conclusions

The degree of assortment in marriage, though it will not change mean social abilities, should be an important determinant of the variance. This will be true particularly when as we saw in chapter 7, mothers and fathers contribute equally to social outcomes for children. This connection between the strength of assortment and variance of social outcomes will exist whether or not there is social causation of child outcomes. So if economic outcomes depend in strong part on the extent of upper level ability in a society, marriage patterns can potentially play an important role.

As chapter 4 showed, by 1754 when measures of marital assortment become possible in England, marriage was as highly assortative as it is today. But what we know of earlier marriage patterns, such as for the ancient Romans, or medieval Italy, suggests that likely northern European marriage had become much more assortative sometime between the middle ages and 1754. Marriage patterns would then provide another potential impetus for such developments as the Scientific Revolution of 1543 and later, the Enlightenment of 1685-1815, and eventually the Industrial Revolution.

19. Conclusions

The message of this book, based on English experience 1600-2026, is concise and direct.

First the persistence of social status across generations is much greater than conventionally estimated. The true intergenerational correlation of status is around 0.8. This implies that elite and underclass families only slowly approach mean status across generations. Even after 10 generations there will still be a significant correlation in social status between a person and their descendants. Effectively English society is segmented into persistent social strata. Mobility between these strata is consistently observed, but this process is slow within each single generation. However, there are no poverty or wealth traps. Social mobility is at the same rate across the entire social spectrum.

Second, underlying rates of status persistence have stayed the same all the way from 1600 to 2026 in England. Despite England 1600 to 2026 experiencing profound social changes, the physics of social position have remained unchanged.

The average person in 1600 lived in a small town or a village. Immigration and emigration were both limited, so the population was mainly composed of the native English. Education was available, but was mostly privately supplied. Nearly half of all men, and a majority of women, as witnessed by the ability to sign their name were illiterate. The modal person had no formal education. Differences in nutrition between elite and underclass families were such that poorer men were shorter than those from the elite.

The technologies employed in the home, farm , and workshop had changed little over the previous 400 years. Production skills for more skilled workers were transmitted across generations using apprenticeships. There was a sharp gender division of labor, with married women mainly engaged in rearing children, and men engaged in producing goods and services. The political franchise was limited to propertied men. The typical person in this society came from a sibship of six, so that parental attention to each child was necessarily limited.

In contrast in 2026 England is a society of large urban concentrations. The technologies of home, office and factory have changed substantially within each persons' lifetime. Migration into England has been extensive. The English population now is composed of a mix of those who derive their ancestry generations earlier from England, and immigrants from a whole variety of other societies: Ireland, Poland, Romania, India, Pakistan, Bangladesh, and Nigeria. Some of these emigrant groups have fared extremely well socially, some have fared poorly. Formal education occupies on average 16-18 years per person. The typical sibship size is two, so that children command more of their parents' attention.

Further there have been persistent attempts to improve social mobility. Most education and health care is now publicly provided. But further, local school funding is increased where there are more children from disadvantaged families, or children in social care. Poorer children also get free meals at school. There are also, admittedly modest, income subsidies to poorer families with children.

Yet this menu of social changes – such as moving to a society with technological dynamism, dramatically smaller family sizes, the emancipation of women, universal voting, and the growth of the modern welfare state - have not increased social mobility rates in any degree. The underlying structures of inheritance have been highly stable in England. The vast resources poured into increasing rates of social mobility have achieved nothing.

The third feature of social mobility across these years is that it is consistent with additive genetic transmission of social abilities. With such a mechanism of transmission, then to get the underlying intergenerational correlation of status of 0.8 there would have to be a genetic correlation in social abilities of parents of 0.6. Assuming spouses match on their social phenotypes, to produce a 0.6 matching of social genotypes requires an even closer match in terms of spousal phenotypes. But this close matching we do indeed observe. The underlying correlation in social status in England in marriage is 0.8. And in line with the stability of rates of social mobility the correlation in underlying status in marriage has been 0.8 all the way from marriages 1754 to those 2021.

A simple model of additive genetic transmission with strong assortment correctly predicts the relative correlations in status of relatives as distant as fourth cousins. It also predicts outcome correlations in the eighteenth century as well as it predicts correlations measured 1999-2026.

Parental genetics, however, can determine child social outcomes in two different ways. The first is the direct path where the genome dictates child social aptitudes, independent of any interaction with the parents. The second way, *genetic nurture,* is where the parental genome leads parents to create a childhood environments for their offspring which are more or less conducive to fostering social aptitudes. Parents with more favorable genetics create a home environment that enhances social outcomes for their children. Some studies have claimed that this indirect genetic channel is nearly as important as the direct channel.

However, the supposed empirical demonstrations of *genetic nurture* are not strong. They mainly consist of showing that within families, where we would observe only the first direct effect, genetics are a less strong predictor of outcomes such as years of education than they are in the population as a whole, where we would observe both mechanisms. But as has been discussed when we measure noisy outcomes such as years of education, restricting the estimate to genetic differences across siblings increases the ratio of noise to signal, and that will itself reduce the observed strength of relationships.

Furthermore, if *genetic nurture* accounts for half of the effects of parental genetics, then adopted children would be more correlated in outcomes with adoptive parents, and also adopted siblings than we observe in practice. For adult IQ, for example, there is almost no correlation between adopted child and adopting parent, where adoption is random with respect to family status.

Also the second section of the book lays out an extensive set of results showing an absence of the kind of social causation that genetic nurture would imply, but consistent with direct genetic transmission of outcomes.

If direct genetic transmission dominates, for example, then mothers and fathers should play equal roles in predicting child outcomes independent of the social status of women. People inherit near 50% of their genetic material from their mothers and from their fathers. But with important social causation of outcomes then we should observe differential impacts of mothers and fathers. Given that mothers even today spend much more time in childcare than fathers, the general expectation would be for a more important maternal role. But what is found on average for a whole set of measures of social status spanning 1754-2026 – literacy, education, schooling, occupational status, house value, index of multiple deprivation, and adult longevity – is that the status of mothers and fathers are equally predictive of child outcomes.

In marriages formed before 1880 in England there was no attempt at fertility control. Family size was determined only by accidents of biology, and by the age at which women married. Even if we account for parental age at marriage, 97% of the variance in numbers of children is unexplained. It was largely a random variable. Given a fixed parental time budget larger families reduce the amount of parental attention children receive. They also reduce the material living standards of children, particularly in poorer families. Yet family size had no significant effect on child outcomes, other than wealth at death for children in families with significant wealth where the bequest per child falls. Again what is remarkable is the absence of sign of social causation of child outcomes.

Birth order also makes no difference to outcomes. Since your genetic inheritance is a random draw independent of birth order, being the first child should confer no advantage. But it is well documented that older children receive more parental attention. So with social causation there should be a decline in child outcomes with higher birth order. Oldest sons in rich families in the nineteenth century did gain an advantage over younger sons. Most notably they were more likely to attend Oxford or Cambridge as a preparation for life as a country squire. But for the other 99.9 percent of families there was a complete absence of birth order effects both in the nineteenth century, and now.

More surprisingly, the misfortune of losing a parent, mother or father, at a young child while associated with reduced child adult longevity, caused no significant decline in other child social outcomes. Even more surprisingly, losing both parents at a young age had no significant effect on attained social status. Either nurture was not of great impact, or social mechanisms existed to replace lost parents. This might be because many young children with a dead parent saw that parent replaced by a second wife or husband. However, since that replacement parent would be imperfectly correlated in characteristics with the original parent, with social causation of child outcomes, we would expect the correlation in status between child and dead parent to be lower when the parent dies young. But that does not happen. The correlation of a son's adult status with that of their father is, for example, independent of the son ever meeting the father.

The *grandmother hypothesis* from anthropology explains the human menopause and long post-reproductive lifespan in women, by proposing that older women contribute to their family's reproductive success by providing care for grandchildren. Grandfathers can be presumed to play a similar role. However, again there is no sign of lower social outcomes for children born after the death of their grandmother or grandfather. And while grandfather status is a strong predictor of grandchild outcomes, the prediction is just as strong for grandfathers who never met their grandchild.

Wealth, as noted, is inherited differently from other measures of social status. Of all outcomes it has the largest claim to be determined by social causation – how much money did you parents and other relatives bequeath you? Wealth at death is determined by these bequests, but also by your abilities and attitudes towards accumulation. Thus in wealthier families, family size influences the wealth of children. However, for marriages formed before 1800 family size was almost entirely a random draw. Illustrating this, the correlation in family size across generations was minimal: 0.05 or less. Family size shocks on wealth were thus a one-time event.

We can trace the effects of family size shocks on wealth across as many as 5 generations. This shows that accidental wealth gains from an unusually small family get dissipated within three generations. Similarly accidental wealth losses from having an abundance of siblings get compensated by more wealth accumulation, again within three generations. In the long run even wealth is largely a function of the genetic endowment of social abilities.

Another outcome correlated with socio-economic status is adult longevity. The rich and educated live long, the poor and uneducated die young. It is again widely believed that there is social causation of longevity, that work and home environments determine health. But we have seen that once we control for attained occupational status, childhood environment has no effect on longevity. Years of education also has no effect, again controlling for attained status. Attained status correlates with longevity, but that seems to be because of an inherent correlation within individuals of social status with physical robustness.

Another symptom of social causation would be if status persistence was unusually pronounced for children with parents at the top or the bottom of the status hierarchy. Wealth traps or poverty traps would indicate social causation of status. With additive genetic inheritance there should be a consistent, linear connection between parent and child status where status is measured on a continuum as with occupational status, house value, neighborhood quality, or educational attainment. The rate of regression to the mean should be constant across status distributions. This is what we indeed find for outcomes such as occupational status, adult longevity, house value, and neighborhood quality.

Another accidental social outcome was where in England a child was born. For recent years, social outcomes on average are much better for those born in London and the South

East than elsewhere in England. The chances of an 18 year old in London attending Oxford or Cambridge, for example, are three times those of an 18 year old in the North. Did accidents of birth have significant consequences on child outcomes, once we control for the social background of parents? Here we can show in a variety of ways that there were no disadvantaged locations in England in terms of social possibilities. The advantage of the South over the North in social outcomes came only from sorting of the English population where for the past 200 years those of greater ability moved to the South, and those of lesser ability moved North.

All these findings reinforce the view that it is direct genetic factors that shape social outcomes, challenging the prevailing views that emphasize social causation and interventions.

The last section of the book looks at the implications of these findings of the importance of direct genetic causation for social and economic policy.

One surprising implication is that English society since even the eighteenth century has been largely a meritocracy. If we were to create a society in which there was fundamental equality of opportunity, a true meritocracy, then we would find that genetic endowments, along with random accidents, would be the sole forces determining social outcomes. Parents and family would not matter, except as a signal of the genetic inheritance of individuals.

The finding both in modern and historical England, and also in modern Denmark and Sweden, that genetic transmission can explain observed social outcomes, is indeed evidence of the meritocratic nature of these societies. If significant social barriers to advancement of children of less advantaged families are removed, then the only force left to determine who has high social status will be people's genetic inheritance. England was a meritocracy even before the Industrial Revolution.

A second implication of the patterns of inheritance found here is that there is clear indication that modern societies spend too much on education. One of the pieties of the modern age is that more education is always good for any recipient, and that education promotes greater social mobility. But we see no evidence that successive expansions of education generated greater social mobility in England. Social mobility rates are no greater in the England of today, which now has compulsory state-provided education to age 18, than it was in the England of 1800 when there was very little public education, and no requirements of any kind for schooling.

Studies of the effect of increases in compulsory years of education in England for lower ability students in 1918-22, 1947, and 1972 find no effect on future employment, income or health for these students. There is every indication that corralling students in schools against their will imposes significant social costs on these students, who show by revealed preference that they do not value this education, and no actual benefits either.

The second implication of direct genetic transmission is that social outcomes will change relative to other societies – educational and other social abilities, and even health – only when the genetic endowment of the society changes. Change is only possible with differential fertility by social class, or differential emigration and immigration by social class. This statement does not endorse in any way such changes as the ones societies should pursue, but merely outlines the constraints of social possibilities.

This reality gives societies, for example, an incentive to select high status immigrants, and reject low status ones. Societies that pursue this strategy – Australia, Canada, New Zealand – will gain permanent improvements in social outcomes relative to those such as the USA that generously accept many low status migrants. The *melting pot* imagery beloved in the USA, where the social background of migrants becomes irrelevant to their status within the USA with one or two generations, is incorrect. As witnessed by the experience of English migrants to the New World, migrants of greater and lesser social abilities mostly retain their social rank in the new society. This assumes the migrants came also from a relatively meritocratic society.[105]

This is not advocacy that societies should adopt restrictive immigration policies. There are plenty of reasons for societies to favor low status immigrants fleeing conflict or oppressive social regimes. But it does emphasize that if social causation is limited, the relative fiscal costs of different types of immigrants will endure across all future generations.

---

105 Exceptions to this generalization, for example, would be the Jewish emigrees to England and the USA from Tsarist Russia, whose economic and social prospects were severely limited by social hostility and legal restrictions in their homeland.

Another thing that will affect future social outcomes is the relative fecundity of people at the top and the bottom of social abilities. Here most people would rightly decry any attempt to restrict fertility by social class. But we now live in a world where there is fear that all societies are doomed to extinction from our wealth of material goods and our poverty of children. The popular obsession with social causation of outcomes makes parenting seem an exhausting and endlessly demanding task. A reminder that good parenting does not involve very much effort might be a suitable tonic to encourage fecundity.

Postscript

20. Eugenics

To my surprise, two readers of the draft manuscript of this book expressed disappointment that it did not contain a discussion of the history of eugenics. One of them indeed stated:

> *There is a reason that the kinds of arguments made in this book are so taboo – because there is such a long track record of flawed work. One wouldn't know that this history existed from this book. (The word "eugenics" appears only once in the entire manuscript.)*

That surprise was because, in terms of evidence for or against the empirical claims above, a history of the eugenics movement adds nothing. Whatever the violations perpetuated in the name of eugenics, those violations speak in no way to the question of what determines social outcomes. Suppose Francis Galton never launched the Eugenics movement. The data presented above in that case would have exactly the same evidential value as it has in a world where eugenic ideas were indeed used and abused.

Another idea implicit in the quote above is that the eugenics movement was overthrown because the belief of Galton and his followers in the importance of inheritance in determining social outcomes was demonstrated to be pseudoscience. This demonstration was so thorough that any claims for strong genetic inheritance of characteristics is inherently suspect. Further such claims must meet a much higher standard of evidence than other empirical claims relating to the social world, because of the "long track record of flawed work."

However, in practice, the decline of the eugenic movement in England had nothing to do with correction of flawed work, or demonstration that Galton's belief in the primacy of nature over nurture was indeed pseudoscience. Enthusiasm for eugenics worldwide went into sharp decline after 1945, as the horrors of the Nazi regime were publicized. But the decline in interest in eugenics did not owe to any significant developments in the study of the nature of inheritance pre- or post- 1945.

Dalton and his followers were confident that nature dominated nurture in the determination of intellectual abilities. When Dalton introduced the term "eugenics" in his 1883 book *Inquiries into the Human Faculty and its Development* he noted that "there is no escape from the conclusion that nature prevails enormously over nurture" (Galton, 1883, 241). But the evidential basis for this claim was minimal before the 1920s. The primary ways of investigating the nature/nurture division until recently has been first the study of monozygotic twins versus dizygotic twins, and second the study of adoptions.

Although Francis Dalton reported a first twin study in 1875, his lack of understanding of the mechanisms of genetic inheritance allowed for only partial and inconclusive use of the twin data to point towards nature. Canonical modern twin studies were only introduced in 1924, independently, by Curtis Merriman in the USA, and Hermann Siemens in Germany. But thereafter a slew of such studies followed (Rende, Plomin, and Vandenberg, 1990). These studies consistently supported the claim that genetic inheritance was the main determinant of IQ. They certainly provided enough evidence of genetic inheritance to support the eugenics claim that by selective reproduction within the population, average intelligence could be raised.

Adoption studies as a method of estimating the relative contribution of nature versus nature were first proposed by Richardson in 1912 (Richardson, 1912-3). Such studies are limited in their ability to detect the effects of home environment because in most societies adoption placements are to some degree selective. But an extensive survey of such studies by Harry Munsinger located a handful that could be used to assess the relative contribution of genetics and environment to IQ in children (Munsinger, 1975). Three out of five such canonical studies date to 1928, 1935 and 1936-49 (Munsinger, 1975, table 6, p. 657). Munsinger's conclusion from these is that

> *the available data strongly suggest that under existing circumstances, heredity is much more important than environment in producing individual differences in IQ.*

Thus again in the 1920s, 1930s and 1940s the new evidence from adoptions was strongly supportive of the eugenics movement's claim that selective reproduction would change population characteristics.

The high point of the eugenics movement in England was perhaps the first International Eugenics Congress, held in London in 1912, with 800 participants, and presided over by Charles Darwin's son, Leonard Darwin. Attendees included the former Prime Minister, Arthur Balfour, the future Prime Minister Winston Churchill, and John Maynard Keynes.

Other prominent eugenicists included William Beveridge, director of the London School of Economics 1919-37, and a founder of the British National Health Service in 1948, George Bernard Shaw, Marie Stopes, Sidney and Beatrice Webb, Bertrand Russell, H. G. Wells, and Julian Huxley.

But eugenics became a toxic subject post WWII, not because of the correction of erroneous beliefs about the importance of genetic inheritance, but because of a general moral switch that strongly rejected earlier racist beliefs associated with eugenics. Thus the history of the Eugenics movement offers no insight into the claims made in this book about the strength of inheritance of social status, or the mechanisms of that inheritance.

Technical Appendix

Additive Genetic Inheritance

Table A1 shows the expected correlation between relatives where we have noisy measures of a phenotype, and parents match based on the actual phenotype. *m* is the correlation between the genetic values of the parents, *r* the correlation between the phenotypic values. h^2 is the heritability of the trait (i.e., the fraction of its variance attributable to genetic causes).

Table A1: Correlations between relatives with assortative mating on a phenotype

Relative to Child	Parents Match on Phenotype	Relative to Child	Parents Match on Phenotype
Average of parents	θh^2	Ave. Grandparents	$\theta h^2 \left(\frac{1+m}{2}\right)$
Full Sibling	$\theta h^2 \left(\frac{1+m}{2}\right)$	Single parent	$\theta h^2 \left(\frac{1+r}{2}\right)$
Uncle/Aunt	$\theta h^2 \left(\frac{1+m}{2}\right)^2$	Single grandparent	$\theta h^2 \left(\frac{1+m}{2}\right)\left(\frac{1+r}{2}\right)$
Cousin	$\theta h^2 \left(\frac{1+m}{2}\right)^3$	Cousin removed	$\theta h^2 \left(\frac{1+m}{2}\right)^4$
Second Cousin	$\theta h^2 \left(\frac{1+m}{2}\right)^5$	Second Cousin removed	$\theta h^2 \left(\frac{1+m}{2}\right)^6$
Third Cousin	$\theta h^2 \left(\frac{1+m}{2}\right)^7$	Third Cousin removed	$\theta h^2 \left(\frac{1+m}{2}\right)^8$
Fourth Cousin	$\theta h^2 \left(\frac{1+m}{2}\right)^9$		

Sources: Clark, 2023, table 1, revised to incorporate measurement errors.

Using observed correlations in status across relatives shown in table 4.5, *m* and h^2 can be estimated. For all but linear descendants of one parent the expected correlation on the Fisher formulae, ρ_n has the form (from table A1)

$$\ln(\rho_n) = \ln(\theta) + \ln(h^2) + \ nln\left(\frac{1+m}{2}\right) \qquad \text{(A1)}$$

In the other cases the expected correlation is

$$\ln(\rho_n) = \ \ln(\theta) + \ln(h^2) + \ nln\left(\frac{1+m}{2}\right) + \ ln\left(\frac{1+r}{1+m}\right) \qquad \text{(A2)}$$

where r is the phenotype correlation between spouses.

This means that we can estimate *m* and h^2 from the parameters of a linear regression

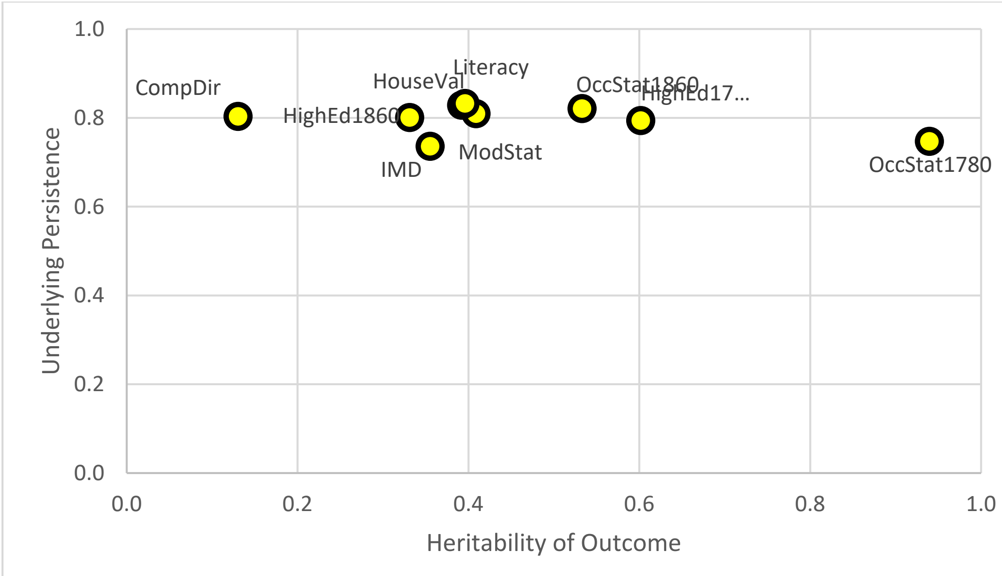
$$\ln(\rho_n) = a \ + \ \ln(b)\,n \ + cd_{lin} \qquad \text{(A3)}$$

where d_{lin} is an indicator which is 1 for the cases where the phenotype marital correlation appears. $b = \frac{1+m}{2}$ is the persistence rate of the correlation as we move one step down the family tree, or one step across between full siblings. To allow for the much greater standard errors of the estimated $\ln(\rho_n)$ for more distant relatives expression (A3) was fitted using weighted least squares, weighted by the inverse squared standard error of each correlation.

Once *m* is estimated we can graph the link between the implied fraction of shared genotype and the correlation of social outcomes, as is shown in figures 4.10-4.12.

Figure A1 shows the estimated values of $b = \left(\frac{1+m}{2}\right)$ for each of the outcome measures from equation (A3). It also shows the 95% confidence interval of each estimate. This is plotted against the estimated heritability of each trait, also from (3). SI table S2 gives the estimated values of *b, m* and h^2 from these estimates, as well as the R^2 of the fit, which averages 0.88.

Figure A1: Estimates of persistence versus heritability, births 1780-1997



Notes: Company Director (2002-22), "CompDir", Ln House Value, 1999-2022, "HouseVal", Index of Multiple Deprivation, 1999-2022, "IMD", Modern Social Status, 1999-2022, "ModStat", Occupational Status births 1780-1859, "OccStat1780", Occupational Status births 1860-1919, "OccStat1860", Higher Education births 1780-1859, "HighEd1780", Higher Education births 1860-1919, "HighEd1860", Literacy Marriages 1754-1879, "Literacy." Lines indicate 95% confidence intervals for the estimates.

As figure A1 shows, the estimates of *b*, social status persistence, cluster around 0.785 for all nine measures, even though the measured heritability of the measured traits varies substantially. The high R^2 of the fit implies the Fisher formulas predict well the correlations. Figure A1, which shows the same underlying rate of social mobility from the eighteenth century to the present, suggests that possibly all social status shows the same persistence parameter of 0.79. If this is through additive genetic transmission, then also throughout this period marital partners had to be correlated 0.57 on the relevant genetics.

Additive Genetic Inheritance and Extended Family Lineages

Table A2 shows the expected correlation between relatives where we have noisy measures of social status, and also affine relatives.

Table A2: Correlations between relatives with Affine Relationships

Relationship	Definition	Implied Correlation
Spouses	MF	θr
Siblings	M-M, etc.	$\theta h^2 \left(\frac{1+m}{2}\right)$
Parent-Child	M←F, etc.	$\theta h^2 \left(\frac{1+r}{2}\right)$
Sibling in-law	MF-M, etc.	$\theta h^2 r \left(\frac{1+m}{2}\right)$
Aunt/Uncle	F←F-M, etc.	$\theta h^2 \left(\frac{1+m}{2}\right)^2$
Spouse of Sib-in-law	MF-FM, etc.	$\theta h^2 r^2 \left(\frac{1+m}{2}\right)$
Aunt/Uncle-in-law	MF←M-M, etc.	$\theta h^2 r \left(\frac{1+m}{2}\right)^2$
Cousins	M←F-F→M, etc.	$\theta h^2 \left(\frac{1+m}{2}\right)^3$
Sibling of sib-in-law	F-MF-M, etc.	$\theta h^2 m \left(\frac{1+m}{2}\right)^2$
Sibling-in-law (2nd degree)	MF-MF-M, etc.	$\theta h^2 rm \left(\frac{1+m}{2}\right)^2$
Child of sib-in-law (2nd degree)	M←MF-MF-M	$\theta h^2 m^2 \left(\frac{1+m}{2}\right)^3$
Spouse of sib-in-law (2nd degree)	MF-MF-MF	$\theta h^2 r^2 m \left(\frac{1+m}{2}\right)^2$
Sibling of sib-in-law (2nd degree)	F-MF-MF-M	$\theta h^2 m^2 \left(\frac{1+m}{2}\right)^3$
Child of spouse of sib-in-law (2nd degree)	M←MF-MF-MF	$\theta h^2 rm^2 \left(\frac{1+m}{2}\right)^3$
Child of sib-in-law (2nd degree) twice rem.	M←MF-MF-MF→M	$\theta h^2 m^2 \left(\frac{1+m}{2}\right)^5$
Sibling-in-law (3rd degree)	MF-MF-MF-M	$\theta h^2 rm^2 \left(\frac{1+m}{2}\right)^3$
Spouse of sibling-in-law (3rd degree)	MF-MF-MF-MF	$\theta h^2 r^2 m^2 \left(\frac{1+m}{2}\right)^3$

Sibling-in-law (4th degree)	MF-MF-MF-MF-M	$\theta h^2 r m^3 \left(\frac{1+m}{2}\right)^4$
Spouse of sibling-in-law (4th degree)	MF-MF-MF-MF-MF	$\theta h^2 r^2 m^3 \left(\frac{1+m}{2}\right)^4$
Sibling-in-law (5th degree)	MF-MF-MF-MF-MF-M	$\theta h^2 r m^4 \left(\frac{1+m}{2}\right)^5$
Spouse of sibling-in-law (5th degree)	MF-MF-MF-MF-MF-MF	$\theta h^2 r^2 m^4 \left(\frac{1+m}{2}\right)^5$

Notes: M/F indicates gender, MF or FM denotes spouses, - indicates a sibling link, and → indicates a parent–child link. Thus M-MF-MF-M indicates the pair brother-brother-wife-

References

192.com People Finder [internet]. Directory of UK addresses from a variety of official sources 2002-2023. [cited 2023 April 4]. Available from: https://www.192.com/people/search/.

Abdellaoui, Abdel, Oana Borcan, Pierre Chiappori, and David Hugh-Jones. 2022. "Trading Social Status for Genetics in Marriage Markets: Evidence from UK Biobank." HCEO Working Paper.

Abdellaoui, Abdel, Hugh-Jones, D, *et al.* 2019. Genetic correlates of social stratification in Great Britain. *Nat Hum Behav* **3**, 1332–1342. https://doi.org/10.1038/s41562-019-0757-5.

Abramitzky, R., Platt Boustan, L. and Eriksson, K. 2012. Europe's tired, poor, huddled masses: self-selection and economic outcomes in the age of mass migration. *American Economic Review, 102(5)*: 1832–1856.

Abramitzky, R., Platt Boustan, L. and Eriksson, K. 2013. Have the poor always been less likely to migrate? Evidence from inheritance practices during the age of mass migration. *Journal of Development Economics, 102*, 2–14.

Althoff, Lukas, Harriet Brookes Gray, and Hugo Reichardt. 2023. "The Missing Link Women and Intergenerational Mobility" Working Paper.

Angrist, Joshua, Victor Lavy, and Analia Schlosser. 2010. "Multiple Experiments for the Causal Link between the Quantity and Quality of Children," *Journal of Labor Economics*, *28* (4), 773–824.

Arendt, J.N. 2005. 'Does education cause better health? A panel data analysis using school reform for identification'. *Economics of Education Review* 24:149-160.

Atkinson, Anthony B. 2018. "Wealth and Inheritance in Britain from 1896 to the Present," *Journal of Economic Inequality*, 16(2): 137–169.

Bagger, Jesper, Javier A Birchenall, Hani Mansour, Sergio Urzúa. 2021. Education, Birth Order and Family Size, *Economic Journal*, Volume 131(633)33–69, https://doi.org/10.1093/ej/ueaa089.

Bailey, Martha, Connor Cole, Morgan Henderson, and Catherine Massey. 2017. "How Well Do Automated Methods Perform in Historical Samples? Evidence from New Ground Truth," Working Paper 24019, National Bureau of Economic Research November.

Balaj, Mirza et al. 2024. "Effects of education on adult mortality: a global systematic review and meta-analysis." *The Lancet Public Health, 9*(3), e155 - e165.

Barclay, Kieron J. 2015. A within-family analysis of birth order and intelligence using population conscription data on Swedish men. *Intelligence, 49*: 134-143. doi.org/10.1016/j.intell.2014.12.007.

Barclay, Kieron J. 2015. Birth order and educational attainment: evidence from fully adopted sibling groups, *Intelligence, 48:* 109-122. doi.org/10.1016/j.intell.2014.10.009.

Bates T C et al. 2018. The Nature of Nurture: Using a Virtual-Parent Design to Test Parenting Effects on Children's Educational Attainment in Genotyped Families. Twin Res Hum Genet. *21*(2): 73-83. doi: 10.1017/thg.2018.11.

Becker, G.S. 1960. "An Economic Analysis of Fertility." Pp. 135–87 in *Demographic and Economic Change in Developed Countries*, edited by G.S. Becker. Princeton, NJ: Princeton University Press.

Becker, G.S. and H.G. Lewis. 1973. "On the Interaction Between the Quantity and Quality of Children." *Journal of Political Economy 81*(2): S279–S288.

Becker, Gary S., Kevin M. Murphy, and Robert Tamura. 1990. "Human Capital, Fertility, and Economic Growth," *The Journal of Political Economy*, *98* (5), S12–S37.

Becker, Sascha O., Francesco Cinnirella, and Ludger Woessmann. 2010. "The trade-off between fertility and education: evidence from before the demographic transition," *Journal of Economic Growth*, *15* (3): 177–204.

Bell, B., Blundell, J. and Machin, S. 2023. 'Where is the Land of Hope and Glory? The Geography of Intergenerational Mobility in England and Wales', Scandinavian Journal of Economics 125(1), 73–106.

Bell, Jordana T and Tim D Spector. 2012. "DNA methylation studies using twins: what are they telling us?" *Genome Biology* 13:172- http://genomebiology.com/2012/13/10/172.

Belmont L, and Marolla FA 1973. Birth order, family size, and intelligence: A study of a total population of 19-year-old men born in the Netherlands is presented. Science, 182(4117): 1096–1101.

Benning, John W, Carlson, Jedidiah, Smith, Olivia S, Shaw, Ruth G, and Harpak, Arbel. 2025 Confounding fuels misinterpretation in human genetics. Proc. R. Soc. B 292: 20251615. https://doi.org/10.1098/rspb.2025.1615

Bhalotra, Sonia, Clarke, Damian. 2019. Twin Birth and Maternal Condition. *Review of Economics and Statistics*, 101: 853-865.

Bhalotra Sonia R, Clarke Damian 2020. The Twin Instrument: Fertility and Human Capital Investment. *Journal of the European Economic Association 18*: 3090–3139. https://doi.org/10.1093/jeea/jvz058.

Biggar, Robert J., Jan Wohlfahrt, Tine Westergaard, and Mads Melbye. 1999. "Sex Ratios, Family Size, and Birth Order." American Journal of Epidemiology, 150(9): 957-962.

Bittles A. 2025. The global prevalence of consanguinity. http://www.consang.net accessed January, 2025.

Bjerkedal, T., Kristensen, P., Skjeret, G. A., and Brevik, J. I. 2007. Intelligence test scores and birth order among young Norwegian men (conscripts) analyzed within and between families. *Intelligence*, *35*, 503–514.

Björklund, Anders, Mikael Lindahl, and Erik Plug. 2004. Intergenerational effects in Sweden: What can we learn from adoption data?", IZA discussion paper 1194.

Björklund, Anders, Mikael Lindahl, and Erik Plug. 2006. "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data." *Quarterly Journal of Economics* 121: 999-1028.

Bjorklund, Anders, Jantti M. 2009. Intergenerational income mobility and the role of family background. In: Salverda W, Nolan B, Smeeding T, editors. *Oxford Handbook of Economic Inequality*. Oxford University Press; Oxford: 491–521.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education*," *The Quarterly Journal of Economics*, *120* (2): 669–700.

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

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2016. Healthy (?), wealthy, and wise: birth order and adult health. Econ Hum Biol 23:27–45.

Blake, Judith. 1981. "Family Size and the Quality of Children," *Demography*, *18* (4): 421–442.

Blake, Judith. 1985. "Number of Siblings and Educational Mobility," *American Sociological Review*, *50* (1): 84–94.

Bolton, Paul. 2012. Education: Historical statistics *UK Parliament, Standard Note*: SN/SG/4252, 27 November 2012.

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

Botzet, L. J., Rohrer, J. M., and Arslan, R. C. 2021. Analysing effects of birth order on intelligence, educational attainment, big five and risk aversion in an Indonesian sample. *European Journal of Personality*, *35*(2), 234-248. https://doi.org/10.1002/per.2285.

Bowles, Samuel and Herbert Gintis, 2001. "The Inheritance of Economic Status: Education, Class, and Genetics," Working Papers 01-01-005, Santa Fe Institute.

Branigan, Amelia R., Kenneth J. McCallum and Jeremy Freese. 2013. "Variation in the Heritability of Educational Attainment: An International Meta-Analysis." *Social Forces*, 92(1): 109-140.

Brea-Martinez, Gabriel. 2023. *Materfamilias*: the association of mother's work on children's absolute income mobility, Southern Sweden (1947–2015), *European Review of Economic History*, Volume 27(1): 1-23, https://doi.org/10.1093/ereh/heac010.

Bruze, Gustaf. 2015. "Male and Female Marriage Returns to Schooling." *International Economic Review*. 56.10.1111/iere.12100.

Buckles K, Kolka S. 2014. Prenatal investments, breastfeeding, and birth order. Soc Sci Med 118: 66-70. doi: 10.1016/j.socscimed.2014.07.055.

Cáceres-Delpiano, Julio. 2006. "The Impacts of Family Size on Investment in Child Quality," *The Journal of Human Resources*, *41* (4): 738–754.

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., and Pan, Y. 2014. 'Early Childhood Investments Substantially Boost Adult Health.' *Science, 343*: 1478 – 1485. DOI: 10.1126/science.1248429.

Card, David, 1999. The causal effect of education on earnings. In: Ashenfelter, Orley and David Card (Eds.), *Handbook of labor economics, vol. 3A*. Elsevier, Amsterdam, NL.

Card, David, 2008. Returns to schooling. In: Durlauf, S.N., Blume, L.E. (Eds.), *The New Palgrave Dictionary of Economics*. Palgrave Macmillan UK, London, pp. 1–11.

Carmichael, Sarah Guilland, Auke Rijpma, and Lotte van der Vleuten. 2015. "Quantity versus Quality: Household structure, number of siblings, and educational attainment in the long nineteenth century," Working Papers 0073, Utrecht University, Centre for Global Economic History November.

Carrier, N H. 1955. "An Examination of Generation Fertility in England and Wales." *Population Studies*, 9(1): 3-23.

Christiansen L, Lenart A, Tan Q, et al. 2016. DNA methylation age is associated with mortality in a longitudinal Danish twin study. *Aging Cell*, *15*(1):149–54.

Cinnirella, Francesco, Marc Klemp, and Jacob Weisdorf. 2017. "Malthus in the Bedroom: Birth Spacing as Birth Control in Pre-Transition England," *Demography*, *54.2*: 413–436.

Clark, Damon. 2018. "Children and Their Parents: A Review of Fertility and Causality." *Journal of Economic Surveys*, 32, 518–540.

Clark, Damon. 2023. "School quality and the return to schooling in Britain: New evidence from a large-scale compulsory schooling reform." *Journal of Public Economics* 223: 1-10.

Clark, Damon and Royer, Heather 2013. The effect of education on adult mortality and health: Evidence from Britain. *American Economic Review, 103*(6):2,087–2,120.

Clark, Gregory. 2007. *A Farewell to Alms: A Brief Economic History of the World.* Princeton University Press.

Clark, Gregory. 2023. "The Inheritance of Social Status; England, 1600-2022." *Proceedings of the National Academy of Sciences.* 120 (27) e2300926120.

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

Clark, Gregory and Neil Cummins. 2015. "Malthus to modernity: wealth, status, and fertility in England, 1500–1879." *Journal of Population Economics, 28*(1): 3-29.

Clark, Gregory and Neil Cummins. 2019. "Randomness in the Bedroom: There is No Evidence for Fertility Control in Pre-Industrial England," *Demography*, 56(4): 1541-1555.

Clark, Gregory, Neil Cummins Matthew Curtis. 2019."Twins Support the Absence of Parity-dependent Fertility Control in Pretransition Populations." *Demography*, 57(4): 1571-1595.

Clark, Gregory, Neil Cummins Matthew Curtis. 2024. "Three new occupational status indices for England and Wales, 1800–1939." *Historical Methods*, 57(1): 41-66.

Clark, Gregory and N. Cummins, 2022. Assortative Mating and the Industrial Revolution: England, 1754-2021. https://hub.cepr.org/discussion-paper/138817.

Clark, Gregory and Neil Cummins. 2024. Does Education Matter? Tests from Extensions of Compulsory Schooling in England and Wales 1919-22, 1947 and 1972. Working Paper.

Clark, Gregory and Neil Cummins. 2025. Hypergamy reconsidered: marriage in England, 1837-2021. *Plos One,* 20(2), e0316769.

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

Clark, Gregory and Gillian Hamilton. 2006. "Survival of the Richest. The Malthusian Mechanism in Pre-Industrial England." *Journal of Economic History, 66*(3): 707-36.

Clark, Gregory and Martin Kristensen. 2026. "The Myth of Nordic Mobility: Social Mobility Rates in Modern Denmark and Sweden." *Comparative Economic Studies, 68(1): 6*.

Clark, Gregory and Christian Nielsen. 2026. The Returns to Education: A Meta-Study. *Kyklos*. **https://doi.org/10.1111/kykl.70041**.

Cohn, R.L. 1992. The occupations of English immigrants to the United States, 1836–1853. *Journal of Economic History, 52(2)*: 377–387.

Collado, M. D., I. Ortuño-Ortín, J. Stuhler. 2022. Estimating Intergenerational and Assortative Processes in Extended Family Data, *Review of Economic Studies.* https://doi.org/10.1093/restud/rdac060.

Conti, G. and Heckman, J. J. 2010. 'Understanding the early origins of the education– health gradient: A framework that can also be applied to analyze gene–environment interactions.' *Perspectives on Psychological Science*, *5*(5): 585–605.

Creamer, Ella. 2024. "Booker prize urged to consider name change over slavery link" *The Guardian*, 25/4/2024.

Crow, J. F. and J. 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.

Cummins NJ, Ó Gráda C. The Irish in England. 2024. *The Journal of Economic History*. Published online: 1-35. doi:10.1017/S0022050724000536.

Curtis, Matthew. 2026. "The her in inheritance: How Marriage Matching has Always Mattered, Quebec, 1800-1970." *Journal of Economic History*, 86(1): 38-66.

Curthoys, M. C., 1997. "The unreformed colleges," in "History of the University of Oxford," Vol. VI, Oxford: Oxford University Press: 146–173.

DeLong, Brad. 2026. Big Trouble in Little Human Genetic Difference-Worshipper Land. Substack Essay, 24/6/2026. https://braddelong.substack.com/p/big-trouble-in-little-human-genetic.

Dean, Trevor and Kate J.P. Lowe. 1998. *Marriage in Italy, 1300-1650*. Cambridge: Cambridge University Press.

Debiasi, E., Rosenbaum-Feldbrügge, M., and Eriksson, B. 2021. "The long-term consequences of parental death in childhood on mortality and the role of socioeconomic status: Evidence from Sweden at the turn of the 20th century." *History of the Family, 26,* 657–681.

Delaney, Judith M., and Paul J. Devereux. 2019. "More Education, Less Volatility? The Effect of Education on Earnings Volatility over the Life Cycle." *Journal of Labor Economics* 37 (1): 101-137.

Desai, Sonalde, 1995. "When Are Children from Large Families Disadvantaged? Evidence from CrossNational Analyses," *Population Studies*, *49* (2): 195–210.

Devereux, Paul J., and Robert A. Hart. 2011. "Forced to be Rich. Returns to Compulsory Schooling in Britain." *The Economic Journal* 120: 1345-1364.

Dickson, Matt. 2013. "The Causal Effect of Education on Wages Revisited." *Oxford Bulletin of Economics and Statistics* 75 (4): 477-498.

Diebolt, Claude, Audrey-Rose Menard, and Faustine Perrin, 2017. "Behind the fertility–education nexus: what triggered the French development process?," *European Review of Economic History*, *21* (4): 357–392.

Dolton, Peter, and Sandi, Matteo. 2017. "Returning to returns: Revisiting the British education evidence." *Labour Economics* 48: 87-104.

Domingue, B.W., Fletcher, J. 2020. Separating Measured Genetic and Environmental Effects: Evidence Linking Parental Genotype and Adopted Child Outcomes. *Behav Genet* **50**, 301–309. https://doi.org/10.1007/s10519-020-10000-4.

Dotti Sani, Giulia M and Judith Treas. 2016. Educational Gradients in Parents' Child-Care Time Across Countries, 1965–2012. *Journal of Marriage and Family*, 78(4): 1083-96.

Dowd, Jennifer B., Amanda M Simanek and Allison E Aiello. 2009. "Socio-economic status, cortisol and allostatic load: a review of the literature." *International Journal of Epidemiology 38(5)*: 1297–1309.

Downey, Douglas B., 2001. "Number of siblings and intellectual development: The resource dilution explanation.," *American Psychologist*, *56* (6–7): 497–504.Draper, Nicholas. 2010. *The Price of Emancipation: Slave-Ownership, Compensation and British Society at the End of Slavery*. Cambridge: Cambridge University Press.

Dribe, Martin, Enrico Debiasi, and Björn Eriksson. 2022. The Effect of Parental Loss on Social Mobility in Early Twentieth-Century Sweden. *Demography* 59(3): 1093–1115. https://doi.org/10.1215/00703370-9962514.

Duckenfield, Bridget. 2013. *College Cloisters – Married Bachelors*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Duncan, R. 1963. Case studies in emigration: Cornwall, Gloucestershire and New South Wales, 1877–1886. *Economic History Review, 16(2)*, 272–289.

Erickson, Charlotte J., 1972. "Who Were the English and Scots Emigrants to the United States in the Late Nineteenth Century?," in D. V. Glass and R. Revelle, eds., *Population and Social Change* (New York): 347-81.

Erickson, Charlotte J., 1981. "Emigration from the British Isles to the U.S.A. in 1831," *Population Studies*, *25* (July): 175-97.

Erickson, Charlotte J., 1986. "The Uses of Passenger Lists for the Study of British and Irish Emigration," in Ira A. Glazier and Luigi De Rosa, eds., *Migration across Time and Nations* (New York): 318-35.

Erickson, Charlotte J., 1989. "Emigration from the British Isles to the U.S.A. in 1841: Part I. Emigration from the British Isles," *Population Studies, 43* (Nov.): 347-67.

Erickson, Charlotte J., 1990. "Emigration from the British Isles to the U.S.A. in 1841: Part II. Who Were the English Emigrants?" *Population Studies, 44* (Mar.): 21-40.

Espín-Sánchez, José-Antonio, Joseph P. Ferrie, and Christopher Vickers. 2023. "Women and the Econometrics of Family Trees" NBER Working Paper No. 31598.

Espín-Sánchez J-A, Gil-Guirado S, Vickers C. 2022. "La "Doña" è Mobile: The Role of Women in Social Mobility in a Pre-Modern Economy." *Journal of Economic History*, 82(1): 1-41. doi:10.1017/S0022050722000018.

Esping-Andersen G. 2004. Untying the Gordian Knot of social inheritance. *Research in Social Stratification and Mobility.*, 21:115–38.

Esping-Andersen G, Garfinkel I, Han WJ, Magnuson K, Wagner S, Waldfogel J. 2012. Child Care and School Performance in Denmark and the United States. *Child Youth Serv Rev.* 34(3):576-589. doi: 10.1016/j.childyouth.2011.10.010.

Espinoza, Héctor et al. 2020. "Post-18 Education: Who is Taking Different Routes and How Much do they Earn?" *Centre for Vocational Education Research*, Briefing Note 013.

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

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table" *American Economic Review, 105(10)*: 3150-3182

Feldman Marcus W, Ramachandran Sohini. 2018 Missing compared to what? Revisiting heritability, genes and culture. *Phil. Trans. R. Soc. B 373*: 20170064. http://dx.doi.org/10.1098/rstb.2017.0064.

Fernihough, Alan, 2017. "Human capital and the quantity–quality trade-off during the demographic transition," *Journal of Economic Growth*, *22* (1): 35–65.

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

Fishman SH, Min S. 2018. Maternal Age and Offspring's Educational Attainment. J Marriage Fam. 80(4): 853-870. doi: 10.1111/jomf.12490.

Free UK Genealogy "FreeREG" [internet] [cited 2023 April 4] Available from: https://www.freereg.org.uk.

Galor, Oded., 2011. *Unified Growth Theory*, Princeton: Princeton University Press.

Galor, Oded and David N. Weil, 2000. "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and beyond," *The American Economic Review*, *90* (4), 806–828.

Galor, Oded and Omer Moav, 2002. "Natural Selection and the Origin of Economic Growth," *The Quarterly Journal of Economics*, *117* (4), 1133–1191.

Galton, Francis. 1874. *English Men of Science: their Nature and Nurture.* London: Macmillan and Co.

Galton, Francis. 1883. *Inquiries into Human Faculty and its Development.* London: Macmillan.

Galton, Francis. 1904. Distribution of Successes and of Natural Ability among the Kinsfolk of Fellows of the Royal Society. *Nature* **70**, 354–356 (1904). https://doi.org/10.1038/070354a0.

Gershuny, J., and O. Sullivan. 2017. United Kingdom Time Use Survey, 2014-2015. Centre for Time Use Research, IOE, University College London. [data collection]. UK Data Service. SN: 8128, http://doi.org/10.5255/UKDA-SN-8128-1.

Gervin, K., Hammerø, M., Akselsen, H. E., et al. 2011. Extensive variation and low heritability of DNA methylation identified in a twin study. *Genome research*, *21*(11), 1813–1821. doi:10.1101/gr.119685.110.

Gihleb, Rania and Kevin Lang. 2020. "Educational Homogamy and Assortative Mating Have Not Increased." NBER Working Paper W229.

Gimelfarb, A. 1981. A general linear model for the genotypic covariance between relatives under assortative mating. *J. Math. Biology 13*, 209–226. https://doi.org/10.1007/BF00275215.

Godoy, Ricardo, et al. 2008. "Assortative mating and offspring well-being: theory and empirical findings from a native Amazonian society in Bolivia." *Evolution and Human Behavior 29*: 201–210.

Goldberger, Arthur S., 1979. Heritability. *Economica*, *46*(184): 327–347. https://doi.org/10.2307/2553675.

Goldin, C. 1992. "The Meaning of College in the Lives of American Women: The Past Hundred Years," Working Paper no. 4099, National Bureau of Economic Research, Cambridge, MA.

Goldthorpe, John H. 1983. Women and class analysis: In defence of the conventional view. *Sociology* 17: 465-488.

Goldthorpe, John H. 1984. Women and class analysis: A reply to the replies. *Sociology* 18: 491–499.

Goody, Jack. 1983. *The development of the family and marriage in Europe*. Cambridge: Cambridge University Press.

Gormley WT, Gayer T, Phillips D, Dawson B. 2005. The effects of universal pre-K on cognitive development. Developmental Psychology. 41(6):872–884. doi: 10.1037/0012-1649.41.6.872.

Grawe, Nathan D., 2003. "Testing Alternative Models of the Quality-quantity Trade-off in Fertility."

Grawe, Nathan D. 2009. "Bequest Receipt And Family Size Effects," *Economic Inquiry*, *48* (1), 156–162.

Greenwood, J, N Guner, G Kocharkov, and C Santos. 2014. "Marry your like: Assortative mating and income inequality." *American Economic Review 104* (5), 348-53.

Grenet, Julien. 2013. "Is Extending Compulsory Schooling Alone Enough to Raise Earnings? Evidence from French and British Compulsory Schooling Laws." *The Scandinavian Journal of Economics* 115 (1): 176-210.

Guild of One-Name Studies [internet] [cited 2023 April 4]. Available from: https://one-name.org/

Guo, G., and VanWey, L. K. 1999. Sibship size and intellectual development: Is the relationship causal? *American Sociological Review, 64*, 169–187.

Gurven, Michael, et. al. 2009. "A Bioeconomic Approach to Marriage and the Sexual Division of Labor." *Human Nature*, 20:151–183. DOI 10.1007/s12110-009-9062-8.

Guryan, Jonathan, Erik Hurst, and Melissa Kearney. 2008. "Parental Education and Parental Time with Children." Journal of Economic Perspectives, 22 (3): 23-46. DOI: 10.1257/jep.22.3.23

Hall, Catherine, Nicholas Draper, et al. 2014. *Legacies of British Slave-Ownership: Colonial Slavery and the Formation of Victorian Britain*. Cambridge: Cambridge University Press, 2014.

Hansen, Casper Worm, Peter Sandholt Jensen, and Lars L?nstrup, "The Fertility Decline In The United States: Schooling And Income," *Macroeconomic Dynamics*, 2017, pp. 1–29.

Harmon, Colm, and Ian Walker. 1995. "Estimates of the Economic Return to Schooling for the United Kingdom." *American Economic Review* 85 (5): 1278-1286.

Haskey, John. 2015. "Marriage rites: trends in marriages by manner of solemnisation and denomination in England and Wales, 1841-2012." In Rebecca Probert, Joanna Miles and Perveez Mody (eds.), **Marriage rites and rights.** London, Hart Publishing.

Hattie, J. A. C. 2009. *Visible learning: A synthesis of 800+ meta-analyses on achievement*. Oxford, UK: Routledge.

Hatton, Timothy J 2004. Emigration from the UK, 1870–1913 and 1950–1998. *European Review of Economic History, 8(2)*, 149–169.

Hatton, Timothy J. 2021. Emigration from the United Kingdom to the United States, Canada and Australia/New Zealand, 1870–1913: Quantity and quality. *Australian Economic History Review*, 61(2), 136–158. doi.org/10.1111/aehr.12218.

Hayward, M.D., and B.K. Gorman. 2004. The long arm of childhood: The influence of early-life social conditions on men's mortality. *Demography, 41(1)*: 87-107.

Healey, Edna. 1986. *Wives of fame: Mary Livingstone, Jenny Marx, Emma Darwin* London : Sidgwick and Jackson.

Healey, Edna. 2001. *Emma Darwin: The inspirational wife of a genius* London: Headline.

Heckman, James J. and Hotz, V. Joseph. "The Sources of Inequality for Males in Panama's Labor Market." *Journal of Human Resources*, Fall 1986, 21(41): 507–42.

Heckman, James, and Rasmus Landersø. 2022. Lessons for Americans from Denmark about inequality and social mobility. *Labour Economics*, 77.

Heckman, James, Lochner L. 2000. Rethinking education and training policy. In: Danziger S, Waldfogel J, editors. Securing the Future. Russell Sage Foundation; New York.

Henrich, Joseph. 2020. *The WEIRDest People in the World: How the West Became Pyschologically Peculiar and Particularly Prosperous.* New York: Farrar, Straus and Giroux.

HM Land Registry [internet]. Sales prices for all buildings England and Wales 1996-2023. [cited 2023 April 4]. Available from: https://landregistry.data.gov.uk/app/ppd.

Holmlund, Helena, Mikael Lindahl, and Erik Plug. 2008. "The Causal Effect of Parent's Schooling on Children's Schooling: A Comparison of Estimation Methods." IZA DP No. 3630.

Holt, Sarah B. 1961. "Quantitative Genetics of Fingerprint Patterns." *British Medical Bulletin*, 17 (3), 247–250.

Hopkins, M. Keith. 1965. "The Age of Roman Girls at Marriage." *Population Studies*, 18(3): 309-327.

Howe, L. J. et al. 2022. Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nat. Genet. 54*: 581–592.

Hugh-Jones, David, and Abdellaoui, Abdel. 2022. Human Capital Mediates Natural Selection in Contemporary Humans. *Behav Genet* **52**, 205–234. https://doi.org/10.1007/s10519-022-10107-w.

Humphries, Jane., 2010. *Childhood and Child Labour in the British Industrial Revolution*, Cambridge: Cambridge University Press.

Iacovou, Maria 2008. Family Size, Birth Order, and Educational Attainment, *Marriage and Family Review*, 42:3, 35-57. DOI: 10.1300/J002v42n03_03.

Jensen, Robert, 2005. "Equal Treatment, unequal outcomes? Generating gender inequality through fertility behavior."

Jepsen, Lisa K, Christopher A Jepsen 2004. "An empirical analysis of the matching patterns of same-sex and opposite-sex couples." *Demography*, 39: 435−453.

Kantarevic, Jasmin and Stéphane Mechoulan. 2006. Birth Order, Educational Attainment and Earnings: An Investigation Using the PSID. *Journal of Human Resources, 41*(4): 755-777.

Kaplanis, J., A. Gordon, T. Shor, et al. 2018. Quantitative analysis of population-scale family trees with millions of relatives. *Science*. DOI: 11.1126/science.aam9309.

Kelly, Morgan, Joel Mokyr, and Cormac O'Grada. 2023. "The Mechanics of the Industrial Revolution," *Journal of Political Economy*, University of Chicago Press, vol. 131(1): 59-94.

Kemptner, D., H Jürges, S Reinhold. 2011. Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany. *Journal of Health Economics 30(2)*: 340-354.

Kessler, Daniel. 1991. "Birth Order, Family Size, and Achievement: Family Structure and Wage Determination." *Journal of Labor Economics,* 9(4): 413–26. http://www.jstor.org/stable/2535077.

Klemp, Marc, Chris Minns, Patrick Wallis, Jacob Weisdorf. 2013. Picking winners? The effect of birth order and migration on parental human capital investments in pre-modern England, *European Review of Economic History*, 17(2): 210-232, https://doi.org/10.1093/ereh/het004.

Klemp, Marc and Jacob Weisdorf. 2018. "Fecundity, Fertility and The Formation of Human Capital," *Economic Journal.*

Kong, A. et al. 2018. The nature of nurture: Effects of parental genotypes. *Science*, 359, Issue 6374: 424-428. DOI: 10.1126/science.aan6877.

Korotayev , Andrey. 2000. "Parallel-Cousin (FBD) Marriage, Islamization, and Arabization." *Ethnology*, 39(4): 395-407.

Kotlikoff, L. and Summers, L. 1981. 'The role of intergenerational transfers in aggregate capital accumulation', *Journal of Political Economy*, *89(4)*: 706–32.

Kristensen, Petter and Tor Bjerkedal. 2007. "Explaining the Relation Between Birth Order and Intelligence" Science, 22 Jun 2007, Vol 316, Issue 5832, p. 1717. DOI: 10.1126/science.1141493.

Kristensen, Petter and Tor Bjerkedal. 2010. Educational attainment of 25year old Norwegians according to birth order and gender. Intelligence, 38(1): 123-136. doi.org/10.1016/j.intell.2009.08.003.

Kugler, Adriana D. and Santosh Kumar. 2017. "Preference for Boys, Family Size, and Educational Attainment in India," *Demography*, *54*(3): 835–859.

Lala Kevin N, Brown G, Twyman K, Feldman Marcus W. 2025. Impediments to countering racist pseudoscience. *Evolutionary Human Sciences,* 7:e24. doi:10.1017/ehs.2025.10012.

Lala, Kevin N. and Marcus W. Feldman. 2024. Genes, culture, and scientific racism, *Proc. Natl. Acad. Sci.* 121 (48) e2322874121. https://doi.org/10.1073/pnas.2322874121.

Lambert, Paul S, et al. 2013. "The Construction of HISCAM: A Stratification Scale Based on Social Interactions for Historical Comparative Research," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 46:2, 77-89. DOI: 10.1080/01615440.2012.715569.

Lawson, David W. and Ruth Mace, 2010. "Optimizing Modern Family Size," *Human Nature*, *21*(1): 39–61.

Lee, James J. et al. 2018. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet. 50*: 1112–1121.

Lee, Jungmin, 2008. "Sibling size and investment in children's education: an asian instrument," *Journal of Population Economics*, *21*(4): 855–875.

Lefgren, L, and F McIntyre. 2006. "Examining the Relationship between Women's Education and Marriage Outcomes," *Journal of Labor Economics* 24: 787–830.

Leibowitz, Arleen. 1974. "Marriage, Family Human Capital, and Fertility." *Journal of Political Economy*, *82*(2), Part 2: S111-S131.

Li, Hongbin, Junsen Zhang, and Yi Zhu. 2008. "The Quantity-Quality Trade-Off of Children in a Developing Country: Identification Using Chinese Twins," *Demography*, *45*(1): 223–243.

Lindert, Peter. 2004. *Growing Public: Social Spending and Economic Growth since the Eighteenth Century*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511510717.

Litchfield, H. E. ed. 1915. *Emma Darwin, A century of family letters, 1792-1896*. London: John Murray.

Liu, Haoming, 2014. "The quality–quantity trade-off: evidence from the relaxation of China's one-child policy," *Journal of Population Economics*, *27*(2): 565–602.

Liu H. 2018. Social and genetic pathways in multigenerational transmission of educational attainment. Am Sociol Rev. 83: 278–304. https://doi.org/10.1177/0003122418759651.

Lleras-Muney, A. 2005. 'The relationship between education and adult mortality in the United States'. *Review of Economic Studies* 72: 189-221.

Long, Jason, and Joe Ferrie. 2013. "Intergenerational occupational mobility in Great Britain and the United States since 1850." *American Economic Review*, *103*(4) 1109-37.

Lucas, Robert E, 2002. "The industrial revolution: Past and future," *Lectures on economic growth*: 109–188.

Lundborg, Petter , Anton Nilsson and Dan-Olof Rooth. 2014. "Parental Education and Offspring Outcomes: Evidence from the Swedish Compulsory School Reform." *American Economic Journal: Applied Economics, 6(1)*: 253-78.

Lynch J W, Kaplan George A, Cohen Richard D, et al. 1994. Childhood and Adult Socioeconomic Status as Predictors of Mortality in Finland. *The Lancet, 343*: 524–527.

Lynch, Robert Francis, 2016. "Parents Face Quantity–quality Trade–Offs between Reproduction and Investment in Offspring in Iceland," *Royal Society Open Science*, *3.5* (160087).

Mare, R. D. 1991. "Five Decades of Educational Assortative Mating," *American Sociological Review*, 56(1), 15–32.

Marks, G. N. 2008. "Are Father's or Mother's Socioeconomic Characteristics More Important Influences on Student Performance? Recent International Evidence." *Social Indicators Research*, *85*(2): 293–30.

Marmot, Michael. 2004. *The Status Syndrome: How Social Standing Affects Our Health and Longevity*. *New York: Times Books. ISBN 9780805073706*.

Marmot, Michael; Siegrist, Johannes. 2006. *Social Inequalities in Health: new evidence and policy implications. Oxford: Oxford University Press. ISBN 9780198568162*.

Marmot, M., Rose, G., Shipley, M. and Hamilton, P. J. 1978. "Employment grade and coronary heart disease in British civil servants." *Journal of Epidemiology and Community Health* 32(4): 244-9.

Marmot, M.G. and M.J. Shipley. 1996. "Do Socioeconomic Differences in Mortality Persist After Retirement? 25 Year Follow Up of Civil Servants from the First Whitehall Study." *British Medical Journal 313*: 1177-1180.

Marmot, M.G., M.J. Shipley and G. Rose. 1984. "Inequalities in Death -- Specific Explanations of a General Pattern?" *Lancet* 323: 1003-1006.

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

Mazumder, B. 2008. 'Does education improve health? A reexamination of the evidence from compulsory schooling laws'. *Federal Reserve Bank of Chicago Economic Perspectives* Q2: 2-16.

Meyer, Michelle N et al. 2023. Wrestling with Social and Behavioral Genomics: Risks, Potential Benefits, and Ethical Responsibility. *Hastings Center Report 53(S1).*

Miller, Paul, Charles Mulvey and Nick Martin. 1995. "What Do Twins Studies Reveal About the Economic Returns to Education? A Comparison of Australian and U.S. Findings." *American Economic Review*, 85(3): 586-599.

Millimet, Daniel L. and Le Wang, 2011. "Is the Quantity-Quality Trade-Off a Trade-Off for All, None, or Some?," *Economic Development and Cultural Change*, *60*(1): 155–195.

Mogstad, Magne and Matthew Wiswall, 2016. "Testing the quantity–quality model of fertility: Estimation using unrestricted family size models," *Quantitative Economics*, *7*(1): 157–192.

Mokyr, Joel. 2005. "Long-term Economic Growth and the History of Technology." In Aghion, P. and Durlauf, S. (eds), *Handbook of Economic Growth*. Amsterdam: Elsevier, pp. 1113–80.

Mokyr, Joel. 2018. "Bottom-up or top-down? The origins of the Industrial Revolution." *Journal of Institutional Economics*, 14: 1003-24.

Montenegro, Claudio E., and Harry Anthony Patrinos. 2014. "Comparable Estimates of Returns to Schooling around the World." *World Bank Policy Research Working Paper No. 7020*.

Munsinger, H. 1975. The adopted child's IQ: A critical review. *Psychological Bulletin, 82*(5), 623–659. https://doi.org/10.1037/0033-2909.82.5.623.

Nagylaki, T. 1982. Assortative mating for a quantitative character. *J. Math. Biology*, 16(1):57-74. doi: 10.1007/BF00275161.

Nivard, M.G., Belsky, D.W., Harden, K.P. *et al.* 2024. More than nature and nurture, indirect genetic effects on children's academic achievement are consequences of dynastic social processes. *Nat Hum Behav* **8**, 771–778. https://doi.org/10.1038/s41562-023-01796-2

OECD 2019. *PISA 2018 Results (Volume I): What Students Know and Can Do*, PISA, OECD Publishing, Paris, *https://doi.org/10.1787/5f07c754-en*.

Office for National Statistics (ONS), 13 February 2024, ONS website, statistical bulletin, Labour market in the regions of the UK: February 2024.

Ohlsson, H., Roine, J. and Waldenström, D. 2014. ' Inherited wealth over the path of development: Sweden, 1810–2010', Working Paper, Uppsala University.

Okbay, A., Y. Wu, N. Wang, *et al.* 2022. Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nat Genet* **54,** 437–449. https://doi.org/10.1038/s41588-022-01016-z.

Olneck, Michael R., and Bills, David B. 1979. "Family Configuration and Achievement: Effect of Birth Order and Family Size in a Sample of Brothers." *Social Psychology Quarterly* 42: 135-48.

Oreopoulos, P., 2006. Estimating average and local average treatment effects of education when compulsory school laws really matter. *Am. Econ. Rev. 96* (1), 152–175.

Oreopoulos, P., 2008. Estimating average and local average treatment effects of education when compulsory schooling laws really matter: Corrigendum. *American Economic Review.* https://assets.aeaweb.org/asset-server/articles-attachments/ aer/contents/corrigenda/corr_aer.96.1.152.pdf.

O'Rourke, Kevin Hjortshøj, Ahmed S. Rahman, and Alan M. Taylor. 2013. "Luddites, the industrial revolution, and the demographic transition," *Journal of Economic Growth*, *18*(4): 373–409.

Outhit, C. 1933. A study of the resemblance of parents and children in general intelligence. *Archives of Psychology*, *149*, 1–60.

Papageorge NW, and Thom K. 2020. Genes, Education, and Labor Market Outcomes: Evidence from the Health and Retirement Study. J Eur Econ Assoc.,18(3): 1351-13910. doi: 10.1093/jeea/jvz072.

Paserman, M. Daniele, and Claudia Olivetti. 2015. "In the Name of the Son (and the Daughter): Marriage and Intergenerational Mobility in the United States, 1850-1940," *American Economic Review*, 105(8): 2695-2724.

Paserman, M. Daniele, Claudia Olivetti and Laura Salisbury. 2018. "Three-generation Mobility in the United States, 1850-1940: The Role of Maternal and Paternal Grandparents." *Explorations in Economic History*, October: 73-90.

Perez, Santiago. 2019. "Intergenerational Occupational Mobility Across Three Continents" The Journal of Economic History*, 79(2): 383-416*.

Pfouts, J. H. 1980. Birth order, age-spacing, IQ differences, and family relations. *Journal of Marriage and the Family*, 42, 517–531.

Piketty, T., Postel-Vinay, G. and Rosenthal, J.-L. 2014. 'Inherited vs self-made wealth: theory and evidence from a rentier society (Paris 1872–1927)', *Explorations in Economic History*, *51*: 21–40.

Plomin, Robert. 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, Robert. 2018. *Blueprint: How DNA makes us who we are.* London: Allen Lane.

Plomin, Robert. 2024. "Nonshared environment: Real but random." *JCPP Advances, 4(1).* https://doi.org/10.1002/jcv2.12229.

Plug Erik. 2004. "Estimating the effect of mother's schooling on children's education using a sample of adoptees" *American Economic Review 94(l)*:358-68.

Ponczek, Vladimir and Andre Portela Souza. 2012. "New Evidence of the Causal Effect of Family Size on Child Quality in a Developing Country," *Journal of Human Resources*, *47*(1), 64–106.

Prandy, K, and Paul S Lambert. 2003. "Marriage, Social Distance and the Social Space: An alternative derivation and validation of the Cambridge Scale." *Sociology*, 37(3): 397-411.

Price, Jacob. 2008. Parent-child quality time: does birth order matter? *Journal of Human Resources, 43 (1):* 240-265.

Qian, Nancy. 2017. "Quantity-Quality and the One Child Policy: The Positive Effect of Family Size on School Enrollment in China," in "Gender and Development," UNU Wider Oxford University Press.

Qin, Zhaoxiong. 2001. 'Rethinking Cousin Marriage in Rural China." *Ethnology, 40(4)*: 347-360.

Rablen, Matthew D. and Andrew J. Oswald. 2008. Mortality and Immortality: The Nobel Prize as an experiment into the effect of status upon longevity. *Journal of Health Economics, 27(6)*: 1462-71.

Record, R. G., McKeown, T., and Edwards, J. H. 1969. The relation of measured intelligence to birth order and maternal age. *Annals of Human Genetics, 33*: 61−69.

Redelmeier, D.A. and S.M. Singh 2001a. "Survival in Academy Award-Winning Actors and Actresses." *Annals of Internal Medicine 134*, 955-962.

Redelmeier, D.A. and S.M. Singh 2001b. "Longevity of Screenwriters Who Win an Academy Award: Longitudinal Study." *British Medical Journal 323*, 1491-1496.

Redmayne, R.A.S. 1904. The Coal-Mining Industry of the United Kingdom. *The Engineering Magazine, 26*.

Rende RD, Plomin R, Vandenberg SG. 1990. Who discovered the twin method? *Behavioral Genetics 20*: 277–85.

Retherford, R. B., and Sewell, W. H. 1991. Birth order and intelligence: Further tests of the confluence model. *American Sociological Review, 56*, 141–158.

Richardson, L. F. The measurement of mental "nature" and the study of adopted children. *Eugenics Review, 1912-13*(4): 391-394.

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

Rodgers, J. L. 1984. Confluence effects: Not here, not now! *Developmental Psychology, 20*, 321–331.

Rodgers, Joseph Lee. 2014. "Are birth order effects on intelligence really Flynn Effects? Reinterpreting Belmont and Marolla 40 years later." *Intelligence 42*: 128–133.

Rodgers, J. L., Cleveland, H. H., van den Oord, E., and Rowe, D. C. 2000. Resolving the debate over birth order, family size, and intelligence. *American Psychologist*, *55*, 599–612.

Rohrer JM, Egloff B, Schmukle SC. 2015. "Examining the effects of birth order on personality." Proc National Academy of Science. 112(46):14224-9. doi: 10.1073/pnas.1506451112.

Rosenzweig, Mark R. and Junsen Zhang. 2009. "Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China's "One-Child" Policy," *The Review of Economic Studies*, *76*(3): 1149–1174.

Rosenzweig, Mark R and Kenneth I. Wolpin. 1980. "Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models," *The Journal of Political Economy*, *88*(2): 328–348.

Rosenzweig, Mark R and Kenneth I. Wolpin, 1980. Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment," *Econometrica*, *48*(1): 227–240.

Schulz, Jonathan F., Duman Bahrami-Rad, Jonathan P. Beauchamp, and Joseph Henrich. 2019. "The Church, intensive kinship, and global psychological variation." *Science* 366, no. 6466: eaau5141.

Schwartz, Christine R. 2013. Trends and Variation in Assortative Mating: Causes and Consequences. *Annual Review of Sociology, 39(1)*: 451-470.

Schwartz, Christine and R. Mare. 2005. Trends in educational assortative marriage from 1940 to 2003. *Demography 42*(4): 621-646.

Scottish Council for Research in Education 1949. *The Trend of Scottish Intelligence.* London: London University Press.

Selzam, S. et al. 2019. Comparing within- and between-family polygenic score prediction. *Am. J. Hum. Genet. 105*, 351–363.

Shaw, Brent D. and Richard P. Saller. 1984. "Close-Kin Marriage in Roman Society?" *Man*, 19(3): 432-444.

Shen, Hao and Marcus W. Feldman. 2020. Genetic nurturing, missing heritability, and causal analysis in genetic statistics. *Proc. Natl. Acad. Sci. USA, 117*: 25,646-25,654.

Shiue, Carol H. 2017. "Human capital and fertility in Chinese clans before modern growth," *Journal of Economic Growth*, *22*(4): 351–396.

Silventoinen, K., Sammalisto, S., Perola, M., et al. 2003. Heritability of Adult Body Height: A Comparative Study of Twin Cohorts in Eight Countries. *Twin Research, 6*(5), 399-408. doi:10.1375/twin.6.5.399.

Smith, K. R., Hanson, H. A., Norton, M. C., Hollingshaus, M. S., and Mineau, G. P. 2014. Survival of offspring who experience early parental death: Early life conditions and later-life mortality. *Social Science and Medicine*, 119, 180–190.

Squicciarini, M P, and N Voigtländer. 2015. "Human Capital and Industrialization: Evidence from the Age of Enlightenment", *Quarterly Journal of Economics* 130(4), 1825–83.

Sulloway, F. J. 1996. *Born to rebel: Birth order, family dynamics, and creative lives.* Pantheon Books.

Tambs, K., Sundet, J.M., Magnus, P. *et al.* 1989. Genetic and environmental contributions to the covariance between occupational status, educational attainment, and IQ: A study of twins. *Behav Genet* **19**, 209–222. https://doi.org/10.1007/BF01065905.

Tan, Hui Ren, 2018. "More is Less? The Impact of Family Size on Education Outcomes in the United States, 1850–1940," *Journal of Human Resources*.

Thomas, J., Rowe, F., Williamson, P. *et al.* 2022. The effect of leave policies on increasing fertility: a systematic review. *Humanit Soc Sci Commun* **9**, 262.

Tiikkaja S, Hemström O. 2008. Does intergenerational social mobility among men affect cardiovascular mortality? A population-based register study from Sweden. *Scand J Public Health. 36(6)*: 619-28. doi: 11.1177/1403494808090635.

Tiikkaja S, Hemström O, Vågerö D. 2009. Intergenerational class mobility and cardiovascular mortality among Swedish women: a population-based register study. *Soc Sci Med.* 2009. *68(4)*: 733-9. doi: 11.1016/j.socscimed.2008.11.017.

Tiikkaja S, Olsson M, Malki N, Modin B, Sparén P. 2011. Familial risk of premature cardiovascular mortality and the impact of intergenerational occupational class mobility. *Soc Sci Med. 75(10)*: 1883-90. doi: 11.1016/j.socscimed.2011.06.031.

Torvik, F. A. *et al.* 2022. Modeling assortative mating and genetic similarities between partners, siblings, and in-laws. *Nat Commun* **13**, 1108. https://doi.org/10.1038/s41467-022-28774-y.

Trejo, Sam, Daniel W. Belsky, et al. 2018. "Schools as Moderators of Genetic Associations with Life Course Attainments: Evidence from the WLS and Add Health." *Sociological Science 5*: 513-40.

Turkheimer, Eric and Mary Waldron. 2000. Nonshared Environment: A Theoretical, Methodological, and Quantitative Review. *Psychological Bulletin*, 126(1): 78-108.

Turner, David., *The Old Boys: The Decline and Rise of the Public School*, New Haven: Yale University Press., 2015.

Turner, John D., "Wealth concentration in the European periphery: Ireland, 1858-2001," *Oxford Economic Papers*, 2010, *62* (4), 625–646.

UK Companies House Register of UK Company Directors [internet] [cited 2023 April 4] https://find-and-update.company-information.service.gov.uk/.

UK, Department of Education, 2024. Key Stage 4 Performance.

UK, Department of Education. 2024. Participation in education, training and employment age 16 to 18. 26/9/2024.

UK Ministry of Housing, Communities and Local Government. English Indices of Deprivation 2019. 9/26/2019 [internet] [cited 2023 April 4]. Available from: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019.

UK, Office of National Statistics. 2021. *Childbearing for women born in different years, England and Wales: 2020*.

van de Beek, Jan H., et al.. 2023. Borderless Welfare State: The Consequences of Immigration for Public Finances. University of Amsterdam, Amsterdam School of Economics.

von Stumm, Sophie, and Robert Plomin. 2018. Monozygotic twin differences in school performance are stable and systematic. *Developmental Science*, *21(6)*, DOI: 10.1111/desc.12694.

Waldfogel J. 2006. *What Children Need*. Harvard University Press, Cambridge, Mass.

Waller, John C. 2012. "Commentary: The birth of the twin study—a commentary on Francis Galton's 'The History of Twins'," *International Journal of Epidemiology*, 41(4): 913-917. https://doi.org/10.1093/ije/dys100.

Wallis, Patrick and Cliff Webb. 2011. "The education and training of gentry sons in early modern England." *Social History, 36*(1): 36-53.

Wang, Biyao et al. 2021. "Robust genetic nurture effects on education: A systematic review and meta-analysis based on 38,654 families across 8 cohorts." *American Journal of Human Genetics, 108*(9): 1780-1791. doi.org/10.1016/j.ajhg.2021.07.010.

Wertz, J. et al. 2019. Genetics of nurture: a test of the hypothesis that parents' genetics predict their observed caregiving. Dev. Psychol. 55, 1461–1472.

Wichman, A. L., Rodgers, J. L., and MacCallum, R. C. 2006. A multilevel approach to the relationship between birth order and intelligence. *Personality and Social Psychology Bulletin, 32*, 117–127.

Willis, Robert J. 1973. "A New Approach to the Economic Theory of Fertility Behavior," *The Journal of Political Economy*, *81*(2): S14–S64.

Willoughby, E.A., McGue, M., Iacono, W.G. *et al.* The role of parental genotype in predicting offspring years of education: evidence for genetic nurture. *Mol Psychiatry* **26**, 3896–3904 (2021). https://doi.org/10.1038/s41380-019-0494-1.

Wilson, Chris and Robert Woods, 1991. "Fertility in England: A Long-Term Perspective," *Population Studies*, *45*(3): 399–415.

Winchester, Nicole. 2022. Educational attainment gap: Regional Disparities in Focus. UK Parliament, House of Lords Library.

Woods, Robert, 2000. *The Demography of Victorian England and Wales* Cambridge Studies in Population, Economy and Society in Past Time, Cambridge University Press.

Wrigley et al., 1997. *English Population History from Family Reconstitution, 1580-1837.* Cambridge: Cambridge University Press.

Wu, Y. et al. 2021. Estimating genetic nurture with summary statistics of multigenerational genome-wide association studies. *Proc. Natl Acad. Sci. USA 118*, e2023184118.

Xi, Song, et al. 2020. "Long-term decline in intergenerational mobility in the United States since the 1850s." *Proceedings of the National Academy of Sciences*, 117(1): 251-258; DOI: 10.1073/pnas.1905094116.

Young, A. I. et al. 2018. Relatedness disequilibrium regression estimates heritability without environmental bias. *Nat. Genet., 50*: 1304-10.

Young, A I, Stefania Benonsdottir, Molly Przeworski, and Augustine Kong. 2019. "Deconstructing the sources of genotype-phenotype associations in humans." *Science*, 365(6460): 1396-1400.

Young, A. I. et al. 2022. Mendelian imputation of parental genotypes improves estimates of direct genetic effects. *Nat. Genet. 54*: 897–905.

Young, Michael. 1958. *The Rise of the Meritocracy 1870-2033: An Essay on Education and Equality*. London: Thames and Hudson.

Zietsch, Brendan P. 2011. Variation in Human Mate Choice: Simultaneously Investigating Heritability, Parental Influence, Sexual Imprinting, and Assortative Mating. *The American Naturalist*, *177(5)*.