European
Historical
Economics
Society

EHES Working Paper | No. 285 | October 2025

Hereditarian Fallacies? Inheritance of Social Status in England, 1600-2022

Gregory Clark[1]
Southern Denmark University

https://www.ehes.org/

EHES Working Paper | No. 285 | October 2025

Hereditarian Fallacies? Inheritance of Social Status in England, 1600-2022

Gregory Clark[1]
Southern Denmark University

Abstract

A recent article in PNAS "The Inheritance of Social Status in England, 1600-1822" (Clark, 2023) argued that social status in England throughout these years showed a strong and stable pattern of inheritance. This pattern was consistent with the additive genetic inheritance of social abilities, in the presence of strong genetic assortment in these characteristics in parenting. This article has induced criticism that it committed "Hereditarian Fallacies" (Benning et al., 2023, 2024, 2025). Similar criticism also appeared in blog postings by the geneticist Sasha Gusev, and in Lala and Feldman, 2024, and Lala et al., 2025. This paper offers additional evidence for why patterns of status inheritance in England 1600-2022 are consistent with direct genetic inheritance, but not with cultural transmission of status.

JEL Codes: J62
Keywords: Intergenerational mobility, status inheritance, cultural transmission

[1] Corresponding Author: Gregory Clark (gclark@sam.sdu.dk)

Notice
The material presented in the EHES Working Paper Series is property of the author(s) and should be quoted as such. The views expressed in this Paper are those of the author(s) and do not necessarily represent the views of the EHES or its members.

Hereditarian Fallacies? Inheritance of Social Status in England, 1600-2022

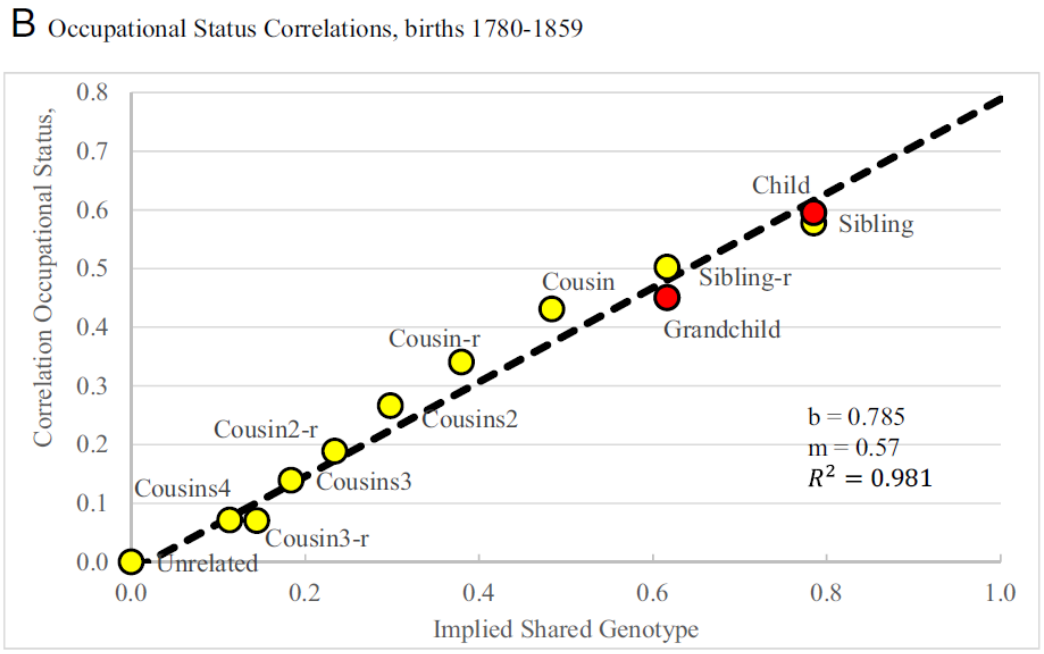
Gregory Clark
University of Southern Denmark and LSE
12/10/2025

Cultural Transmission

The alleged hereditarian fallacy in Clark, 2023, is the supposed confusion of genetic transmission of characteristics across generations with cultural transmission (Benning et al., 2023, 2024, 2025).[1] Consider, for example, the patterns observed in figure 2b of Clark, 2023 (here figure 1). This shows a correlation between relatives that is proportional to their estimated shared genotype in the genetics relevant to occupational status, for men born 1780-1859.

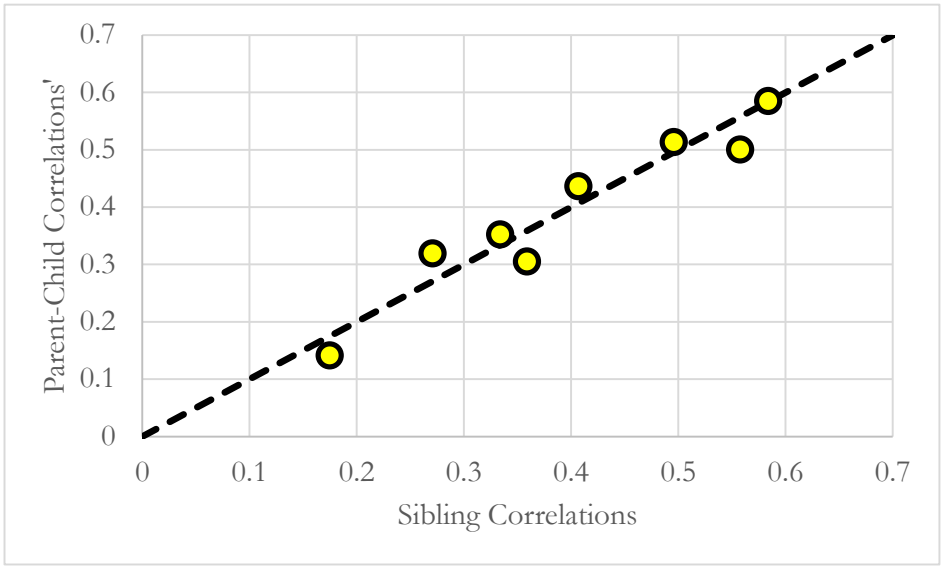
[1] Lala et al., 2025, argues similarly that Clark, 2023, is fallacious because "this interpretation entails a model of causation that the evolutionary human sciences have long rejected. For instance, a major conclusion of contemporary gene–culture coevolution research is that the interactions between genes and culture are bidirectional, and that the causes of cultural variation cannot be reduced to genetics" (Lala et al., 2025, section 3.2).

Figure 1: Correlations of Occupational Status, Males, England



Source: Clark, 2023, figure 2b.

Figure 2: Sibling versus Parent-Child Correlations, England, 1754-2022



Notes: The sibling versus parent-child correlations for the nine different outcomes in Clark, 2023.

Specifically a general characteristic we observe is that the parental correlation in social outcomes with their children is the same as the correlation in outcomes across their children. Yet the children were reared in a common home environment, while the parent-child pairs would experience different home environments. These child environments of parent and child would be correlated, but would differ. They have to differ if we are to observe regression to the mean across generations in social outcomes, converging on average towards the mean, and if social outcomes are the product of childhood environments. So if what matters to child outcomes is the family environment they were reared in, the parent-child correlation must be less than the sibling correlation.

To put this formally, suppose that the observed social status of child j in family i, y_{ij}, is derived from an underlying family status x, culturally transmitted, such that

$$y_{ijt} = x_{it} + u_{ijt}$$

However, to ensure regression to the mean in cultural status across generations, the underlying family cultural status must regress to the mean across generations. This also implies that there is a random component in the transmitted family status, e, to ensure that the variance of family cultures across families remains constant across generations. Thus

$$x_{it} = bx_{it-1} + e_{it}$$

In this case the sibling correlation in observed social status will be

$$\rho_{sib} = \frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2},$$

But the parent-child correlation will be lower by a factor b,

$$\rho_{parent-child} = b\frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2}.$$

In the case of father's occupational status, for example, with cultural transmission the correlation father's status between child and parent would be 0.8 that of the sibling correlation. In contrast, with additive genetic transmission the correlation in genetics between parent and child is the same as between siblings.

Figure 2 shows for the nine outcomes reporter in Clark, 2023, over the years 1780-2022, the correlations between siblings versus that between parent and child. As can be seen there is a basic equality for each status outcome, even though the child-parent correlation ranges from 0.15 to 0.60. This seems an important obstacle to cultural explanations of status transmission.

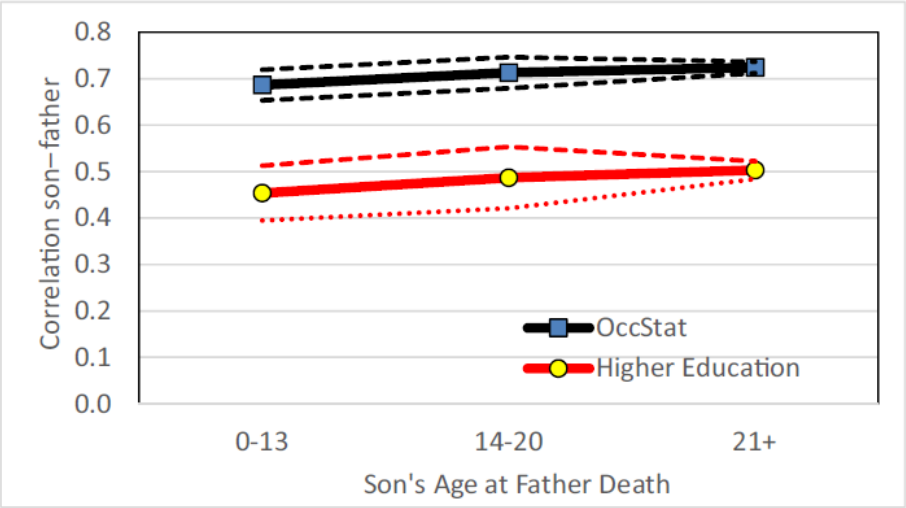
A second obstacle cultural theories face comes when we look at the correlation in status between children and their parents and grandparents in cases where the children and parents or grandparents interacted, versus cases where early death of parents or grandparents removed possible social interaction. With cultural transmission contact with relatives matters. With direct genetic transmission, contact can play a lesser role, or even inconsequential, role.

Clark, 2023, shows the surprisingly unimportant role of contact with fathers on son outcomes. Figure 4 of Clark, 2023, reproduced as figure 3 below, shows that the correlation of sons in occupational status and higher education attainment with their fathers. This correlation is almost as high for those whose fathers die after they reach age 21 compared to before they reach age 14. An extra 10-15 years of personal exposure to fathers as a child has no statistically significant effect on how much sons resemble fathers.

Sons who lost their fathers early fell into three categories. Those orphaned and raised by relatives or the state. Those raised by their widowed mother alone. And those raised by their mother and stepfather. If we assume that both parents had equal effect in forming a child, then the degree to which children losing their father would show a declining correlation with that father depends on how correlated fathers are with their wives, siblings, parents, and in-laws.

Under the conventional assuming that these correlations are less than 0.5, then the correlation of a son with an absent father should decline very substantially. If they are raised by the widowed mother alone, the correlation with the father should decline by a third. A stepfather would have a correlation in social characteristics with the biological father of only 0.25 or less, so that if the son is raised by their mother plus stepfather, the correlation with the

Figure 3: Correlations of Occupational Status, and Higher Education, Sons to Fathers



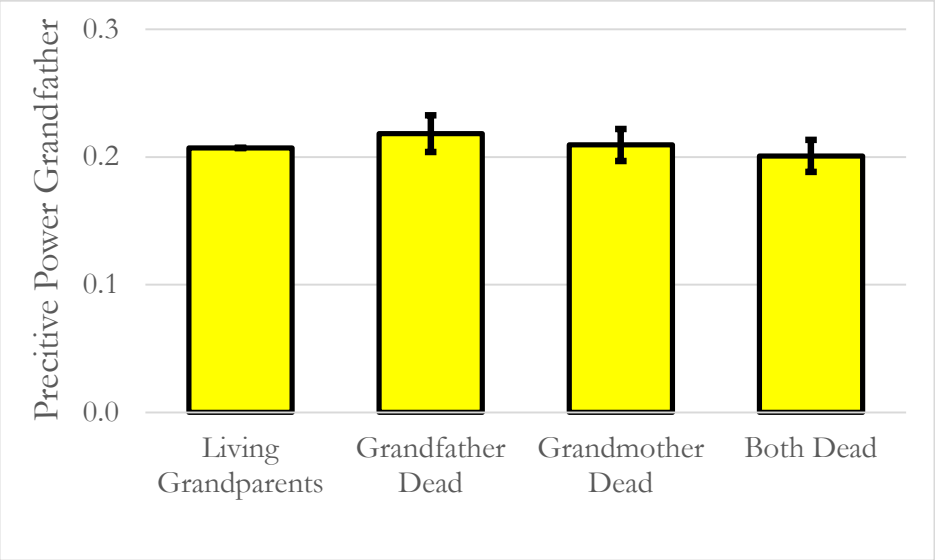
Source: Clark, 2023, Figure 4.

father would fall a full half. If the child is an orphan raised by grandparents or siblings from the patriline, the decline in correlation would again be one third. Rearing by grandparents or siblings from the patriline, the decline in correlation would again be now two-thirds. Clearly no such declines appear in the correlations in figure 3.

Clark, 2023, however argues that marriage in England was much more assortative in terms of social abilities than generally assumed, with a correlation of around 0.8 throughout centuries (Clark, 2023). The correlation in social characteristics between each parent and their parents or siblings would also be around 0.8.

In this case of tight assortment if orphaned children were raised by a grandparent or sibling, then the correlation between them and their father's characteristics would decline by only 10-30% depending on which side of the family fostered the son. The decline in correlation would be even greater for children reared by state institutions. With only the mother raising the son, the correlation between son and father should decline by only 10%. In the case of a stepfather, the correlation of the stepfather with the dead father in terms of social characteristics will be only 0.6. The correlation between son and biological father should then decline by 20%. But overall there should be at least a 15% decline in the correlation between son and father with the early death of a father.

Figure 4: 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 3 shows clearly that, even for the case of tight marital assortment for occupational status, there is clearly no 15% decline in father-son correlation.

We can also look at cultural transmission through grandparents. In this case we estimate the predictive power of grandfather occupational status on grandson occupational status, controlling for father occupational status. We divide grandsons into four types: those with both grandparents living at their birth, those with the grandfather dead at birth, the grandmother dead at birth, and both grandparents dead at birth. This should index the amounts of contact sons had with grandparents.

As figure 4 illustrates, the information about grandson outcomes from grandfathers is unaffected by whether or not the grandson ever met either the grandfather or grandmother. Again this is at odds with the idea of important cultural transmission

Wealth as a Potential Non-Genetic Transmitter of Status

Benning et al., 2023 identify wealth as one omitted non-genetic transmission mechanism. Clark, 2023, they state "conflates genetic and non-genetic transmission (e.g. of wealth) within families" (Benning et al., 2023, p. 1).

In aid of this claim, they detail the strong correlation between father's wealth and son occupational status. This result, figure 2 of their paper, is reproduced for sons born 1780-1859 in figure 5 (Benning et al., 2023, figure 2, p. 6). We observe a strong correlation between father's log wealth and son's occupational status of 0.73. There is also a strong correlation between father's occupational status and son's status of 0.79, and between father's occupational status and father's log wealth of 0.76. Benning et al. conclude from this that non-genetic wealth transmission explains the strong persistence of educational and occupational status, house value, and neighborhood characteristics.

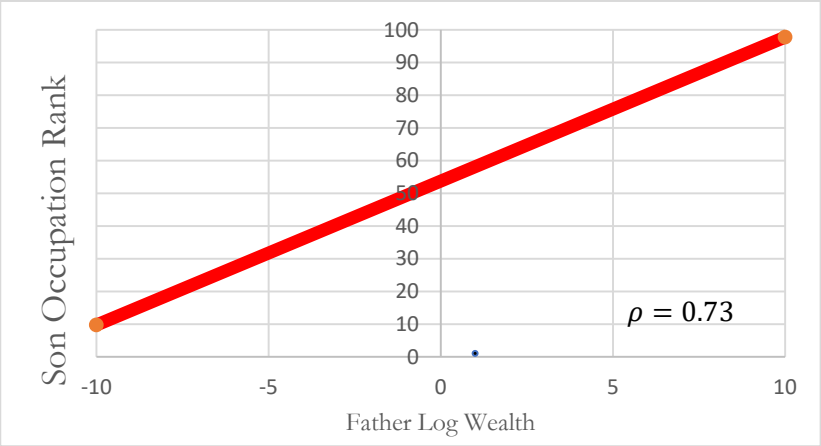
In the Families of England (FOE) lineage database, however, wealth inheritance shows characteristics that demonstrate that it cannot explain the patterns of inheritance of other elements of social status.[2] These patterns of wealth inheritance show its non-genetic nature. But they also show that wealth transmission cannot explain the other elements of status transmission across generations.

For fathers who left wealth at death, the wealth of their children at death was strongly influenced both by their wealth, but also by the number of adult children in the family. The more children, the less each child received. Also daughters in earlier years typically inherited substantially less wealth than sons.

If wealth inheritance was the non-genetic driver of status persistence, even for outcomes such as occupational status, or higher education, then the patterns of persistence for these outcomes would show also deviations from what would be predicted by an additive genetic model of transmission.

[2] Appendix 2 outlines the contents of this database.

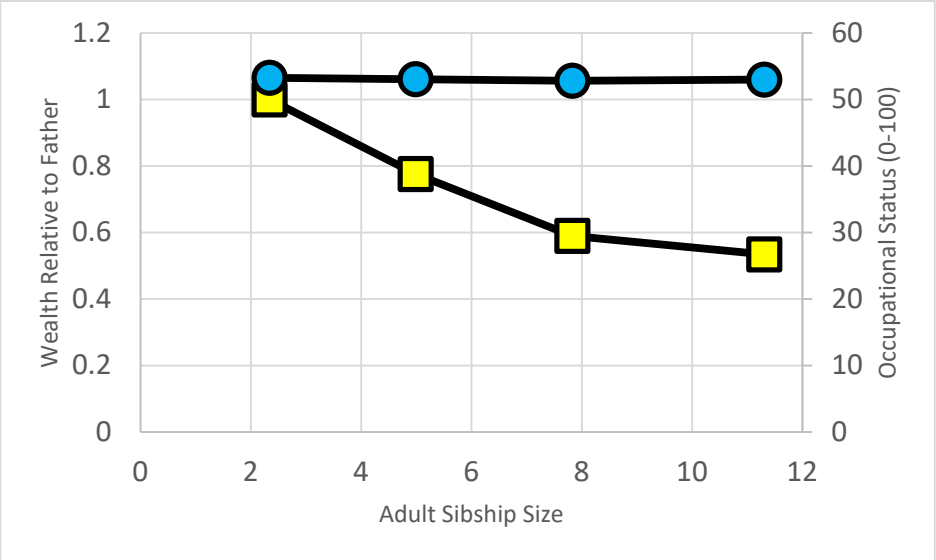
Figure 5: Link Between Son Occupational Status and log Father Wealth, England, 1780-1859



Notes: This reproduces the first panel of Benning et al., 2023, figure 2.

Source: FOE database.

Figure 6: Wealth and Occupational Status of Sons, English marriages pre 1880, where father had wealth.



Notes: Marriages pre-1880. Son wealth at death is measured relative to that of father. Families were included where father wealth as death was above average. Son occupational status is measured on a scale 0-100. Sibship sizes are divided into 4 groups: 1-3, 4-6, 7-9, and 10+.

Source: Families of England (FOE) database.

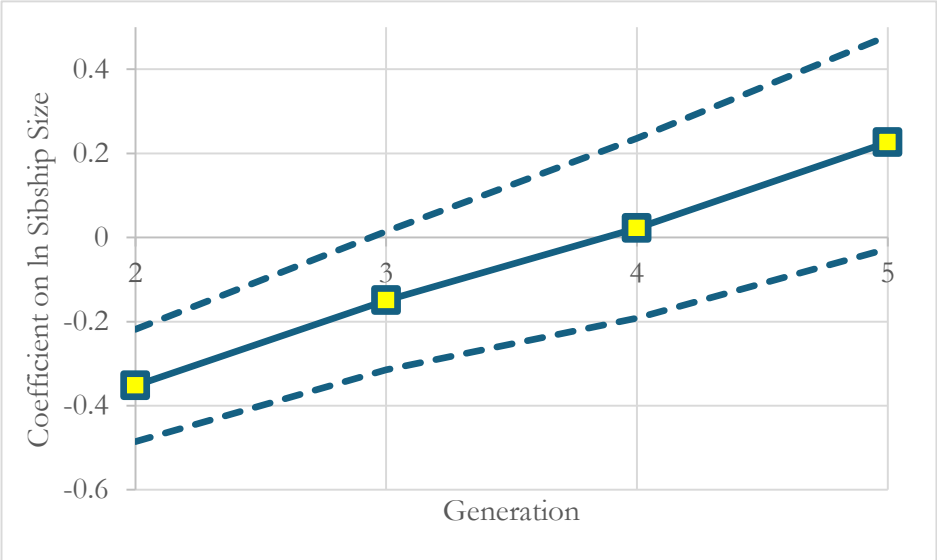
For marriages in England formed before 1880, there is no evidence of any attempt to control fertility. Family size was largely random (Clark, Cummins, and Curtis, 2019), with the only systematic determinant of fertility the age of the wife at marriage. A very small amount, however, of the variation in family size is explained by the age of the wife at marriage. The vast majority of size variation is unpredictable. Remarkably, however, for most social outcomes, family size had little or no effect. This was true both for high status families with wealth to bequeath, and low status families with no wealth. It was true for male occupational status and higher educational status. It is true also for both genders for literacy, adult longevity, and for the chance of being in education aged 11-19.

The only case in which family size mattered was for the wealth of children, in cases where the father bequeathed wealth. Figure 6 shows that for fathers with above average wealth at death and for marriages before 1880, child wealth was substantially influenced by family size. In these same families, however, as the figure shows, son occupational status was completely unaffected by family size.

If wealth is the driver of other outcomes such as occupational status, occupational status would decline with family size for the sons in these families. Figure 6, however, shows that the inheritance of occupational status was not affected by family size. Sons in larger families showed no decline in occupational status compared to sons in small families. Thus the non-genetic inheritance of wealth cannot be the pathway to the inheritance of occupational status.

Note also that in Clark (2023) it was shown that wealth is inherited asymmetrically between the patriline and matriline, with the patriline three times as important in terms of the strength of inheritance. Again if wealth was the driver of social status, we would expect inheritance of status along the patriline to be stronger than along the matriline. Instead what we find is a complete symmetry of the patriline and matriline for social outcomes other than wealth. Paternal grandfathers have the same predictive ability for grandchild outcomes as do maternal grandfathers. This is what would be predicted by genetic transmission not mediated by wealth. Again there is a clear distinction between the pattern of inheritance of wealth (non-genetic) and the pattern for other social outcomes. For these other outcomes the pattern is consistent with genetic transmission. This interesting gender symmetry in inheritance of social outcomes, even when England was a highly patriarchical society, is explored in detail in Clark and Cummins, 2024.

Figure 7: Effect of Initial Sibship Size on Wealth, 2-5 Generations (wealth ≥ 1), all children



Notes: The figure shows the coefficient of ln(adult sibship size) in the second generation on ln(wealth) in the second to fifth generations. The dotted lines show the 95% confidence intervals.

Source: Clark and Cummins, 2025, figure 6.

The random shocks to family size for marriages before 1880 were very weakly correlated across generations. There was very minor correlation across one generation (less than 0.1), and no correlation across two generations. So they acted largely as one time, random, shocks to wealth (Clark and Cummins, 2025). These shocks as we noted affected child wealth. But with the FOE lineage data we can also trace out the effect of these wealth shocks in the third, fourth and fifth generations.

Specifically we estimate the log of child wealth at death on the log of father's wealth in the first generation, as well as the log of completed sibship size in the second generation. Figure 7 shows the results of that estimation for sibship size, for all father's marrying before 1880, and with wealth at death above average.

As can be seen the random wealth shocks from family size in the first generation have an estimated effect of zero by the time of great-grandchildren. The great-grandchildren of a man of given wealth who had 1 adult child have the same predicted wealth as the descendants of

an equivalently wealthy man who had 12 adult children. This suggests that while wealth persists strongly across generations in English society 1750-2025, it is not mainly wealth itself that explains that long run persistence, but instead the personal characteristics associated with wealth. Random shocks to wealth uncorrelated with those personal characteristics themselves do not persist.

Thus there is clear indication in the FOE database that wealth itself is not an important non-genetic driver of the other elements of social status. Instead wealth adapts within 2-3 generations to correlate with the overall social status of families. That underlying social status shows patterns of persistence that are consistent with additive genetic transmission.

Changing Intergenerational Correlations in Status versus Stable Underlying Status Persistence

Clark (2023) finds a decline of various status correlations with genetic distance that is constant across outcomes and time periods. Yet the underlying parent-child correlation on measures such as occupational status, higher education, and literacy, was declining over time for people born 1760-1919. Benning et al., 2023, asset this implies that social institutions were changing the rates of social mobility over time, and there was no constancy in the pattern of social mobility.

To understand why there is no contradiction between changing intergenerational correlations in social outcomes, and a stable persistence of status between generations, we need to appreciate the role of measurement errors. Measures such as occupational status, higher education attainment, and literacy all indicate social status imperfectly, with a degree of error. Their intergenerational correlations are thus not direct measures of the true underlying correlation in social status across generations. The degree to which they convey information about underlying status can easily change over time. In general social status seems transmitted across generations in the following way. Let y be any measure of status (occupational status, education, literacy, house value, neighborhood quality), and x the true underlying status. Then

$$y_t = x_t + u_t$$
$$x_t = bx_{t-1} + e_t$$

If we try and estimate b from the expression

$$y_t = \beta y_{t-1} + v_t$$

Then

$$E(\hat{\beta}) = b\frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2} < b$$

If the random factors connecting underlying status *x* with some measured status *y* change over time, then without any change in *b*, underlying persistence in status, measured intergenerational status correlations will change.

The correlation between average parent literacy and child literacy for the average lineages in the FOE database declines from 0.460 for children born 1748-1829 to 0.320 for children born 1830-1872. This is a statistically significant decline.[3] However, as figure 8 below shows average literacy rates rose from 60% to 90% over these years. When average literacy rates were 60%, literacy conveys much more information about social status than when average literacy rates were 90%. Thus the estimated intergenerational literacy correlations will fall.

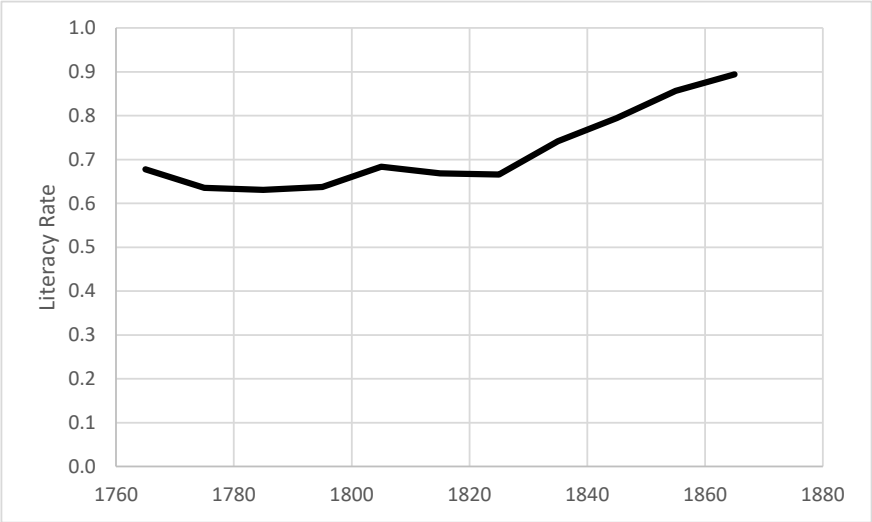
Over the birth period 1760-2021 the structure of occupations in England changed substantially, as is shown in table 1. Occupations such as laborer, whose social status was very clear and well defined steadily declined as a share of employment. At the same time a new group of occupations whose social status was much more diverse for a given occupational status steadily rose as a share of employment. Such occupations included descriptors such as "clerk", "engineer", "manager", "officer", and "director". Managers can manage large complexes, or a modest food stall. An engineer can be university trained, or just a mechanic. A police officer can be a chief constable, or just a constable.

For grooms in weddings 1837-1849 the share of such ill-defined occupations was only 1.6%. But by marriages 1910-9 11.4% of grooms had such occupations, and by 1980-2021, 25.9% of grooms had such diffuse occupational labels.

This change in occupational structures towards occupational labels with less precise social statuses is accompanied by a decline in the measured intergenerational correlation in occupational status between fathers and sons. Figure 9 shows this decline using occupational status indices. This decline is what Benning et al., 2023, focus on as showing there is no constancy is status inheritance 1600-2021.

[3] There are 1,913 observations in the first period, and 2,508 in the second.

Figure 8: Literacy Rates of men and women, births 1760-1869, FOE general lineages



Notes: Average male and female literacy rates at marriage.

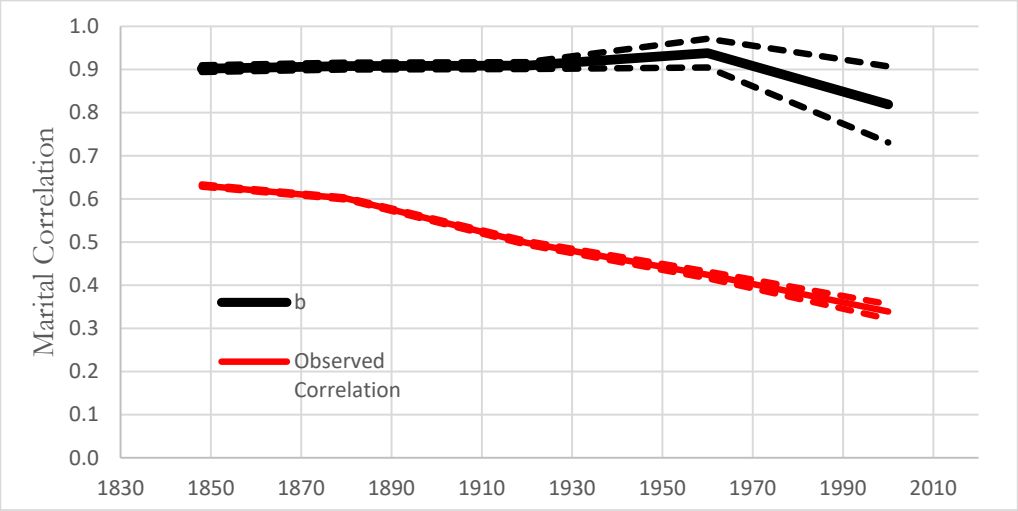
Source: FOE database.

Table 1: Occupational Structures, England, 1837-2021

	1837-49	1837-49 (%)	1910-9	1910-9 (%)	1980-2021	1980-2021 (%)
Laborer	73,259	28.9	10,868	18.1	180	1.1
Clerk	1,838	0.7	3,595	6.0	409	2.6
Engineer	1,963	0.8	2,127	3.5	1,428	9.1
Manager	70	0.0	572	1.0	1,381	8.8
Officer	276	0.1	524	0.9	522	3.3
Director	0	0.0	11	0.0	317	2.0
All	253,152		59,952		15,688	

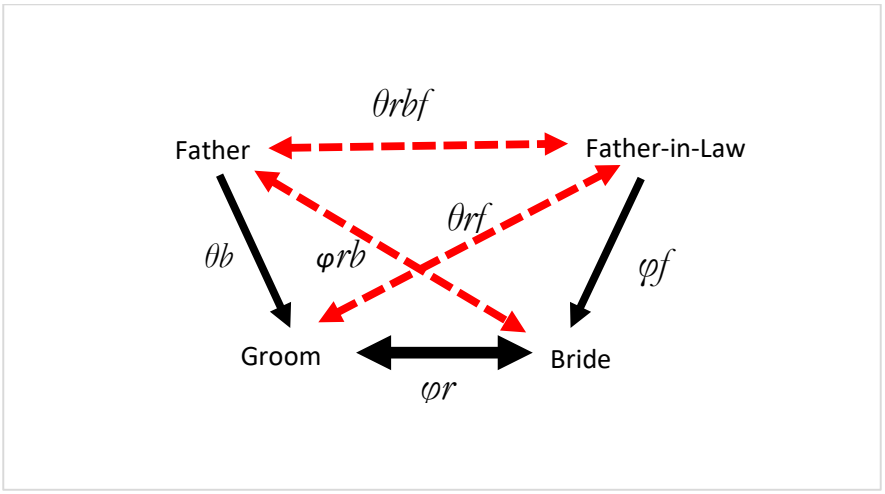
Source: Marriages of England (MOE) Database.

Figure 9: Underlying and Observed Occupational Correlations, England 1837-2021



Note: the dashed lines show the 5 percent confidence intervals. The indices used here to measure occupational status are: 1837-1939, CCC, 1940-2021 CAMSIS. The black line shows the error corrected intergenerational correlation in occupational status father to son. Source: MOE database.

Figure 10: Observed Correlations in Status in Marriage



But how much of the decline is a measurement issue, and how much a change in the strength of real underlying persistence? To estimate that we can use the same data on marriages 1837-2021. From English marriage registers we get the occupation of grooms, their fathers, and their father in laws. Figure 10 shows what the observed correlations between the various parties to the marriage will be if the union is formed only between husband and wife, and they match with correlation *r* on their underlying social status. Assuming an underlying correlation of *b* between father and son, and *f* between father and daughter, and an attenuation factor of θ when we observe a measured occupational correlation between men, then we get the following observed correlations:

Father-son	θb
Son-Father-in-Law	θrf
Father-Father-in-Law	θrbf

But this in turn implies that

$$\frac{corr(father,father-in-law)}{corr(groom,father-in-law)} = \frac{\theta rbf}{\theta rf} = b$$

Thus by taking the ratio of the father-in-law to father, and father-in-law to groom correlations we can get an estimate of the underlying intergenerational father-son correlation independent of measurement errors, even when these errors are changing over time periods as in England 1837-2021.

Table A1 in appendix 1 and figure 9 shows the resulting estimates. For 1837-59, 1860-99, 1900-39 there are separate estimates using the HISCAM and CCC occupational indices, and the line shows the average of these. All through the period 1837-1979 the estimated underlying value of b is close to 0.90, with no sign of any downward trend. The estimated persistence rate b is 0.77 for the last period, marriages 1980-2021. But there is much less data so that the 95% confidence interval for this estimate extends to 0.9, close to the earlier average of 0.90.

Overcounting of distant relationships in familial correlations

Benning et al., 2023, also argues that the consistent pattern of correlations found for all nine social outcomes in Clark, 2023, across close and distant relatives is an artifact of an incorrect way of calculating these correlations.

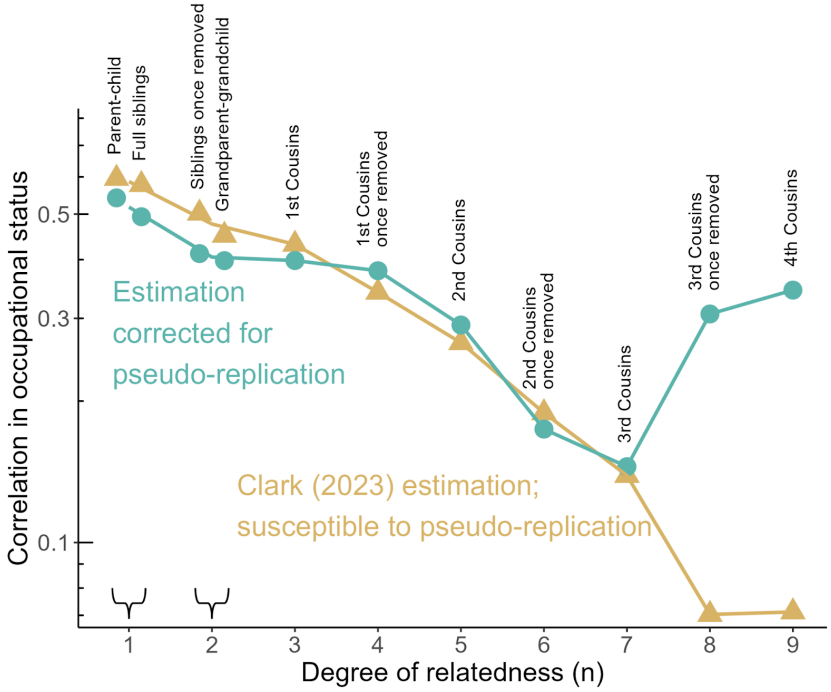
As we go down the family tree from close relatives, like siblings, to more distant ones like fourth cousins, the number of times some individuals feature in a correlation will rise sharply. Thus the nine correlations of occupational status by relationship in 1780-1859 was calculated from 17,382 pairs of individuals, but only 1,878 unique individuals. The randomness of reproduction ensures that while a person will typically appear in only 2-3 sibling correlations, they can appear 100 times or more in fourth cousin correlations. This means that the correlation estimates for more distant relatives are more frequently estimated from interconnected observations.

The interdependence of many of the occupational status pairings in calculating the occupational correlation would certainly increase the sampling error in using this data to calculate the average correlations between relatives of different degrees of affinity. But Clark, 2023 does not make any use of estimated sampling errors. Importantly, if the underlying correlations between relatives are the same, independently of how many pairings they appear in, then the simple average correlation of each pairing, as used in the paper, will be an unbiased estimator of the true correlations.

If the unweighted correlation coefficients have a higher than optimal sampling error, then this would be expected to show up as noise in the 9×11 separate estimated correlation coefficients. That unsystematic noise should not produce a clear and equivalent correlation pattern for all 9 status measures observed. Sampling errors will obscure patterns in the data, not result in them showing in high relief.

Benning et al., 2023, propose an alternative estimator of correlations that involves randomly sampling one relative pair of each type, for each surname lineage, then bootstrapping the estimate 1,000 times. When they do this for one of the 9 outcomes reported in Clark (2023), occupational status 1780-1859, they reproduce the correlations for 9 of 11 familial relationships. But they have substantially different correlation estimates for the most distant relatives, 3rd cousins once removed and fourth cousins. This is shown in figure 11.

Figure 11: Benning et al., 2023, figure 3



The reweighting of pairs in calculating the correlations, as proposed by Benning et al., 2003, is so extreme as to ensure that the sampling error will be substantially larger than with the unweighted average used in Clark, 2003. To see this note that for the case of occupational status of 4th cousins born 1780-1859, only 32 out of 264 surname lineages contribute 4th cousins. And the surname lineages contribute such parings very unequally. One lineage contributed only one pair of 4th cousins, but three each contributed more than 3,000 pairs.

With the Benning et al., 2023, new weighting, the 9 sparsest lineages that previously contributed 1% of the 4th cousin pairs and have a handful of observations now constitute 29% of the sample, and the 4 lineages that contributed 72% of fourth cousin pairs now contribute 13%. The proposed reweighting thus gives very substantial weight to a handful of 4th cousin pairings, the same weight as with lineages with potentially many independent pairings. This will substantially increase the sampling error above that for the simple average used in the paper. It is no surprise, then, that the estimated correlations for more distant relatives change substantially, as we see in figure 3 of Benning et al., 2023.

Circular reasoning in the construction of occupational status measures

The occupational status variable used in Clark, 2023, was created by an algorithm which sought to maximize the father-son and father-in-law-son correlations in occupational status pre-1939. Given that "unusually high correlations of relatives are at the heart of Clark (2023a)'s arguments the strong persistence of correlations across distant relatives, this appears to be, at least in part, circular reasoning." (Benning et al., 2023, p. 8).

Clark, Cummins and Curtis, 2024 describes how the occupational status index used in Clark, 2023 was constructed. The method used to construct the occupational status measures 1800-1939 is the same as that used by sociologists to construct the standard CAMSIS status measures for recent years, and again the standard HISCAM measures for the years before 1939.

All these status measures seek to assign status in such a way as to maximize familial correlations, such as father-son correlations for the years before 1939 (Lambert et al., 2013). As appendix table A1 shows the measure used for Clark (2023), labelled the CCC index, shows significantly higher correlations between relatives than the HISCAM measure in the period 1837-1939. But this is because the CCC index is based on a much larger database of father-son and father-in-law-son occupational pairings than was available to the devisers of the HISCAM measure. That allows for estimation of a more accurate occupational scale, and the better the accuracy the higher become observed correlations between all sets of relatives.

Suppose occupational status is biased upwards for the father-son correlation in particular, because of the way the indices are constructed. This bias would in fact reduce the estimated intergenerational persistence found in Clark, 2023. For that persistence is not estimated using directly the father-son correlation, but is instead estimated from the decline is correlations as we move to more distant relatives. With an upward bias in the father-son correlation, that decline will be greater, and estimated persistence smaller.

It is also notable that the decline in persistence of status correlations with genetic distance is just the same for occupational status as it is for other measures of status such as higher educational attainment, or house value, where there is no possible index issue.

Conclusions

There is no simple and obvious "hereditarian fallacy" in Clark, 2023. Indeed the pattern of correlations between relatives found in the paper, while completely consistent with additive genetic inheritance of status, instead turn out to be very hard to reconcile with any plausible model of cultural transmission.

Specifically also, none of the four major issues that Benning et al., 2023, find with the Clark, 2023, have purchase.

Wealth was not a confounding factor that explains the pattern of correlations within a lineage of such status markers as occupational status, or education, without genetic transmission. Wealth itself when it comes as a wealth shock independent of people's social abilities does not affect other measures of social status. Wealth shocks also dissipated rapidly, within 2-3 generations, while underlying social status shows much greater persistence across ten or more generations.

The decline in father-son correlations across time for various partial status measures does not contradict the argument of Clark, 2023 for a constant strength of status persistence. These partial parent-child status correlations will underestimate true status persistence by varying degrees across time if the measurement errors in partial status measures change. And we can observe exactly such changes in the degree of measurement errors for measures such as literacy or occupational status.

Benning et al., 2023, suggest a reweighting for the relative pairings in calculating correlations in occupational status 1780-1859 which involves giving a weight of nearly 30% to a handful of such pairings, out of thousands of correlations in the database for this measure, in the case of 4th cousins. Not surprisingly rather than produce an estimate with less sampling error, this adds considerable noise to their correlation measures. They also offer no explanation of why the potential noise introduced by the suboptimal sampling error from the current unweighted correlation estimates should produce a consistent pattern of correlations across nine separate social outcome measures.

Finally the supposed circularity of the occupational status measures, had it existed, would again lead to lower estimates of persistence, not the higher estimates they assume.

Appendices

Appendix 1:

Table A1: Underlying Correlations in Intergenerational Occupational Status, marriages 1837–2021

Period	Index	N	Father-in-law to groom	Father-in-law to father	Father to son	b_{Ratio}
1837–1859	CCC	391,244	0.591	0.533	0.706	0.903
			(0.001)	(0.001)	(0.001)	(0.002)
1860–1899	CCC	535,670	0.563	0.510	0.677	0.907
			(0.001)	(0.001)	(0.001)	(0.002)
1900–1939	CCC	250,827	0.491	0.444	0.602	0.905
			(0.002)	(0.002)	(0.002)	(0.003)
1837–1859	HISCAM	379,020	0.408	0.375	0.586	0.919
			(0.002)	(0.001)	(0.001)	(0.004)
1860–1899	HISCAM	516,795	0.386	0.343	0.537	0.887
			(0.002)	(0.002)	(0.002)	(0.006)
1900–1939	HISCAM	241,369	0.329	0.286	0.439	0.867
			(0.002)	(0.002)	(0.002)	(0.006)
1900–1939	CAMSIS	237,824	0.374	0.324	0.492	0.865
			(0.002)	(0.002)	(0.002)	(0.005)
1860–1899	CAMSIS	44,430	0.343	0.324	0.419	0.942
			(0.004)	(0.004)	(0.004)	(0.014)
1980–2021	CAMSIS	10,872	0.280	0.217	0.349	0.773
			(0.009)	(0.009)	(0.009)	(0.063)

Notes: Standard errors in parentheses. For the estimate of b from the ratio of correlations, standard errors from 10,000 bootstrap replications.

Appendix 2: Families of England Database

The database is a genealogy of a set of English families, 1600-2024, who carried one of 499 rarer surnames, where we track every holder of the surname, no matter what their location. This database has 436,191 individuals.

These surnames fall in two groups. The first is a set of 232 surnames, 61,196 people, which had significantly higher wealth than average in the mid nineteenth century that was used to measure the persistence of wealth in England 1858-2024. 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, Mitford, Pigou, and Rusbridger.

The second, larger, group is a set of 267 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 1796-2024, voting rolls 1999-2024, 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.

In constructing the FOE database we directly compiled 425 of the 499 surname lineages (with 118,741 persons). The other 74 lineages (324,161 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. Thus the Mitchelmore lineage, for example, incorporates the surnames Michelmore, Mitchelmore, Mitchamore, Mitchmore, Mouchemore, Muchamore, and Muchmore.[20] Similary 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 any living holders of the surname. In these cases 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 individuals born with a rare surname and living to adulthood we can identify the father and mother. For those born 1800-1939 the percentage linked to their parents is even higher at 92%.

Figure A1 shows an illustrative fragment of the FOE database, showing linkages across 7 generations. 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. Table A1 shows the outline of the source of the data, and its distribution across time, and between general and elite lineages.

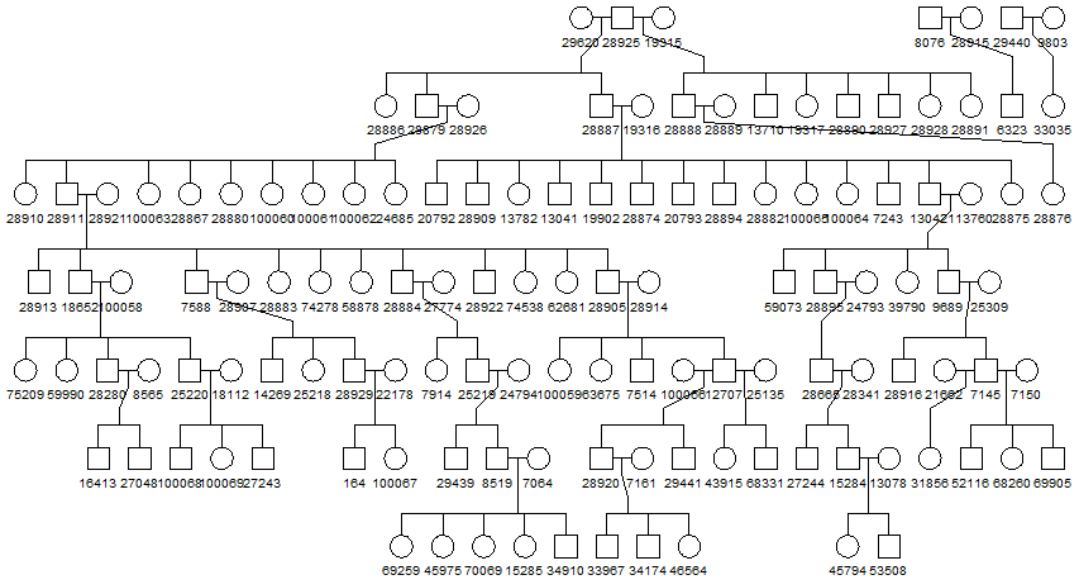
Since the focus of the genealogy is descent on the patriline, family size is measured as the numbers of births associated with each father, as well as the numbers of children attaining ages 14 and 21.

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.

In the years 1796-1858 a value was also attributed to each estate at probate. For wills proved in the highest probate courts, the Prerogative Court of Canterbury and the Prerogative Court of York, we get a record of the probate value also in these years. These wills were typically those of wealthier individuals.

For 84,691 individuals we have an estimate of wealth at death. 21,971 of these had wealth at death above average for their decade of death. In this study where we are seeking to measure the effects of wealth shocks we use individuals both from the general and the elite lineages.

Figure A1: 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 A2: Families of England 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	54,503	10,542
1850-99	115,988	99,540	15,475
1900-49	79,246	69,314	9,190
1950-2025	48,170	41,371	6,405
All	436,190	374,995	61,195

Source: FOE database. Note that significant numbers of people have no date of birth observed.

References

Benning, John W, Jedidiah Carlson, Ruth G. Shaw, and Arbel Harpak. 2023. "Clark (2023) and the Persistence of Hereditarian Fallacies". bioRxiv Preprint, 3 Nov, https://www.biorxiv.org/content/10.1101/2023.11.01.565061v1

Benning, John W, Jedidiah Carlson, Olivia S. Smith, Ruth G. Shaw, and Arbel Harpak. 2024. "Confounding Fuels Misinterpretation in Human Genetics". bioRxiv Preprint, 22 Aug, https://doi.org/10.1101/2023.11.01.565061

Benning, John W, Jedidiah Carlson, Olivia S. Smith, Ruth G. Shaw, and Arbel Harpak. 2025. "Confounding fuels hereditarian fallacies." bioRxiv Preprint July 11 https://doi.org/10.1101/2023.11.01.565061.

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

Clark, Gregory and Neil Cummins. 2024. "Matriline versus Patriline: Determinants of Social Outcomes in England, 1754-2024." LSE Economic History Working Papers, # 366.

Clark, Gregory and Neil Cummins. 2025. "How Long do Wealth Shocks Persist? Less than three generations in England, 1700-2025."

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.

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

European
Historical
Economics
Society

EHES Working Paper Series

Recent EHES Working Papers

EHES 284 How Long do Wealth Shocks Persist? Less than three generations in England, 1700-2025
Gregory Clark, Neil Cummins

EHES 283 Has oil richness been a force for income equality in Venzuela over the long term?
Pablo Astorga

EHES 282 Female labor force participation in historical census microdata
Jørgen Modalsli

EHES 281 Colonial Governance and Resource Allocation in the British West Indies (1838–1938)
Luisito Bertinelli, Fabio Gatti, Eric Strobl

EHES 280 Why railways fail: Colonial railways and economic development in Habsburg Bosnia-Herzegovina
Magnus Neubert, Stefan Nikolić

EHES 279 Text Analysis Methods for Historical Letters, The case of Michelangelo Buonarrotti
Fabio Gatti, Joel Huesler

EHES 278 The Rákóczi family's estate wine production, wine trade, and wine prices at the end of the 17th century
Attila Ulrich

EHES 277 Transport and Biological Living Standards: Evidence from Late 19th-Century Catalonia
Ramon Ramon-Muñoz, Guillermo Esteban-Oliver, Josep-Maria Ramon-Muñoz

EHES 276 Climate adaptation in early Australian wheat farming (1860–1960)
Costanza Maria Fileccia, Eric Strobl

All papers may be downloaded free of charge from: https://www.ehes.org/

The European Historical Economics Society is concerned with advancing education in European economic history through study of European economies and economic history. The society is registered with the Charity Commissioners of England and Wales number: 1052680.