

# DISCUSSION PAPER SERIES

DP17788

**MEASURING MOBILITY:  
INTERGENERATIONAL STATUS  
MOBILITY ACROSS TIME AND PLACE**

Gregory Clark, Neil Cummins and Matthew Curtis

**ECONOMIC HISTORY AND LABOUR  
ECONOMICS**

**CEPR**

# MEASURING MOBILITY: INTERGENERATIONAL STATUS MOBILITY ACROSS TIME AND PLACE

*Gregory Clark, Neil Cummins and Matthew Curtis*

Discussion Paper DP17788  
Published 06 January 2023  
Submitted 20 December 2022

Centre for Economic Policy Research  
33 Great Sutton Street, London EC1V 0DX, UK  
Tel: +44 (0)20 7183 8801  
[www.cepr.org](http://www.cepr.org)

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Economic History
- Labour Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Gregory Clark, Neil Cummins and Matthew Curtis

# MEASURING MOBILITY: INTERGENERATIONAL STATUS MOBILITY ACROSS TIME AND PLACE

## Abstract

There is considerable interest in comparing intergenerational social status mobility across time and place. But such attempts are vitiated by unknown measurement errors in status indices, errors that also vary over time and place. Typically the more error, the more apparent social mobility. Using a new database of 1.7 million marriages in England 1837-2021, we show how improving the quality of an occupational status index generates lower implied social mobility. Without control of the errors embodied in all social status indices, attempts to compare social mobility across time and place are impossible. This paper develops a solution to this problem using not absolute correlations, but the relative correlations between different family members. The method deployed here suggests that occupational status persistence in England 1837-2021 was always much greater than conventionally measured, and was largely unchanging 1837-2021.

JEL Classification: N/A

Keywords: N/A

Gregory Clark - gclark@ucdavis.edu  
*University of California – Davis, LSE and CEPR and CEPR*

Neil Cummins - n.j.cummins@lse.ac.uk  
*London School Of Economics And Political Science and CEPR*

Matthew Curtis - mjdcurtis@gmail.com  
*ECARES - Université Libre de Bruxelles*

# Measuring Mobility: Intergenerational status mobility across time and place

Gregory Clark, Neil Cummins and Matthew Curtis<sup>1</sup>

Dec 20, 2022

## Abstract

There is considerable interest in comparing intergenerational social status mobility across time and place. But such attempts are vitiated by unknown measurement errors in status indices, errors that also vary over time and place. Typically the more error, the more apparent social mobility. Using a new database of 1.7 million marriages in England 1837-2021, we show how improving the quality of an occupational status index generates lower implied social mobility. Without control of the errors embodied in all social status indices, attempts to compare social mobility across time and place are impossible. This paper develops a solution to this problem using not absolute correlations, but the relative correlations between different family members. The method deployed here suggests that occupational status persistence in England 1837-2021 was always much greater than conventionally measured, and was largely unchanging 1837-2021.

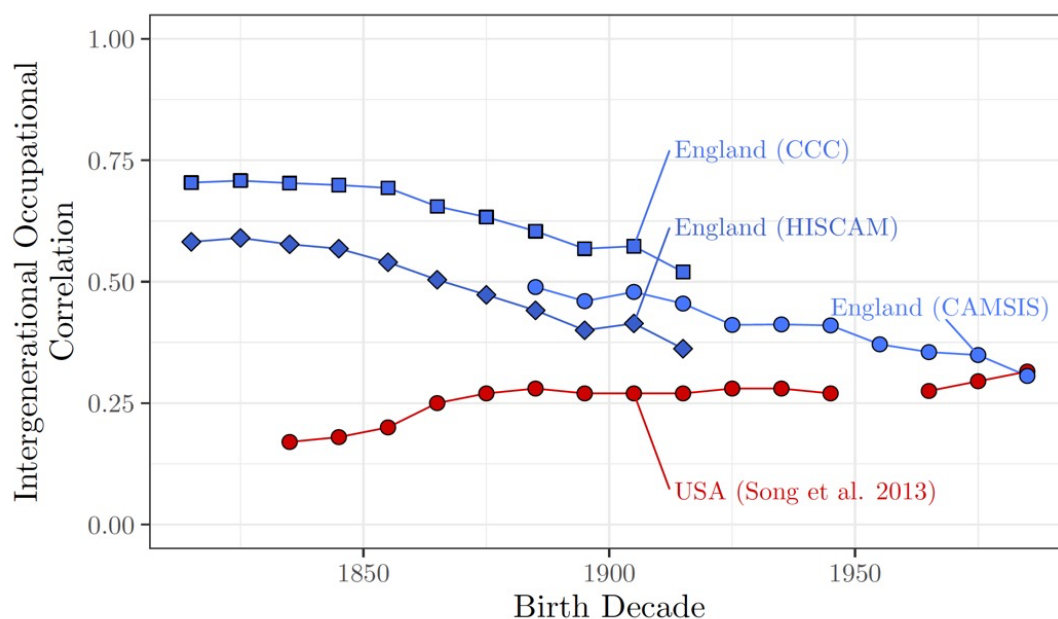
## 1 Introduction

Sociologists and economists have conducted many studies of intergenerational social mobility rates, and have compared these rates across time, and across societies. A recent study, for example, of intergenerational occupational status correlations for the USA for men born 1830-1980 found the results shown in figure 1 (Song et al. (2020)). This suggests rapid social mobility in the US throughout the birth years 1830-1980, with intergenerational

---

<sup>1</sup> University of California – Davis, LSE and CEPR, LSE and CEPR, ECARES - Université libre de Bruxelles. The substantial data collection in this paper was made possible by the generous financial support for Economic History at UC Davis of Michael Dearing.

correlations always less than 0.33.<sup>2</sup> But also we see a clear decline in social mobility between mid-nineteenth century and the modern period.<sup>3</sup>



**Figure 1: Intergenerational Occupational Status Correlations USA Compared with England, 1810s-1980s**

*Notes:* The figure contrasts estimates of intergenerational occupational persistence for the US from Song et al. (2020) with a set of alternative estimates for England. England (CCC) are the intergenerational occupational correlations 1810-1919, calculated by the new occupational status index presented in this paper using the 1.7m marriage records, England (HISCAM) are the same set of correlation, 1810-1919 but instead using the HISCAM index (Lambert et al. (2013)). England (CAMSIS) uses the CAMSIS index to calculate the correlations 1920-1980 (<https://www.camsis.stir.ac.uk/>).

For England social mobility rates can be measured for the modern period using the CAMSIS occupational status scales, and for 1800-1938 using the HISCAM scales.<sup>4</sup>

<sup>2</sup> Similarly Ferrie and Long directly estimate occupational mobility for England 1851-1911, using census data, and imputing average earnings for each occupation as a measure of status. They find for 1851-1911 that the intergenerational correlation of status in England was then 0.27, compared to 0.23 in the USA (Long and Ferrie (2018), table 5).

<sup>3</sup> However, an unpublished study by Zachary Ward (Ward (2021)) reports a different trend of intergenerational correlations for the USA 1830-1980, and much higher intergenerational correlations once measurement error is controlled for.

<sup>4</sup> <https://www.camsis.stir.ac.uk/Data/Britain91.html>, <https://www.camsis.stir.ac.uk/hiscam/>. See Prandy and Lambert (2003); Lambert et al. (2013).

Employing a large new database, which records for 1.7 million marriages in England 1837-2021 the occupation of the groom, and that of his father and his father-in-law, figure 1 shows by decade of birth, 1810 to 1980, for men in England the correlation in occupational status of fathers and sons, using the CAMSIS index from 1990 for the birth decades 1880-1980, and the HISCAM-GB index for the birth decades 1810-1920.<sup>5</sup> The Camsis index used is derived for 371 occupational categories using data on the occupations of partners in marriage in England from a sample from the 1991 census, for 92,021 marriages. The HISCAM-GB index is derived for around 400 occupations using 51,419 occupational pairings between father and son.

These new series suggest much slower rates of social mobility for England across most of the birth decades 1830-1980 than in the USA, but also a consistent trend towards increasing social mobility in England.

However, with the large set of data we have on nearly 3 million father-son and father-in-law son occupational pairs we are also able to construct a new index for 1800-1939, on the same association principles as HISCAM, but with a different set of 442 occupational categories. We call this the CCC index.<sup>6</sup> As figure 1 shows the CCC index produces even higher intergenerational status correlations for the English birth cohorts 1810-1920. By implication there is even more of an increase in social mobility rates 1810-1980 than with the HISCAM index.

Figure 1 illustrates that a problem with all these measures, and with comparisons of social mobility across time and place, is that the occupational status indices used embody substantial errors. These errors are of two types. First the ascribed status categories will have errors in their status ranking. That error will be smaller the more data is employed to estimate the index, as we see in figure 1. If we had more data for marriages 1940-2021 we could produce a new version of the CAMSIS index which would also show higher father-son status correlations. But second, within each ascribed status category, status can vary by differing amounts with different employment structures. This error, however, will not be reduced through more data.

As occupations change over time, the importance of these errors can also change, giving potentially spurious impressions of changes in social mobility rates. Thus for the English marriages 1837-79, the most common occupational descriptor, laborer,

---

<sup>5</sup> Appendix 1 details the source of these marriage records.

<sup>6</sup> Details of its construction are given in Clark et al. (2022).

constituted 27% of grooms, and 30% of their fathers. Laborer is a homogenous work description, with little difference within the category in social status for different types of laborers. For marriages 1980-2021 in contrast the most common occupation for grooms, 10% of occupations, was “manager” or “supervisor”. Manager is a highly heterogenous category, encompassing all kinds of true social status. Thus differences in the structure of occupations over time or across countries can lead to occupational status indices which correspond more or less closely to true social status.

In this paper we show how to measure intergenerational occupational status mobility rates independent of these two types of measurement error. This method employs the ratio of the correlation of status between fathers and fathers-in-law relative to that between son and father-in-law as the measure. Whatever the errors in status attribution to occupations, this ratio will measure the underlying intergenerational correlation in occupational status. This new measure suggests occupational status mobility rates are substantially lower than conventionally measured. These measures also suggest that in England intergenerational occupational status correlations were close to constant across the birth interval 1810-1989.

## **2 Measuring Social Mobility with Imperfect Status Indices**

Researchers on social mobility have long recognized the potential measurement issues we raise above (see, for example, Solon (1992); Braun and Stuhler (2018)). One proposed solution was using abstract measures of movement between social status categories, where social status by category is not parameterized, such as the Altham Index (Altham (1970); Altham and Ferrie (2007); Long and Ferrie (2013); Perez (2019)). But the Altham Index is not easy to calibrate as a measure of social mobility. As has been noted, “one possible reason for the lack of success of the Altham index may be difficulty in interpreting it” (Bouchet-Valat (2022, p.205)).

Further, any exercise in measuring social mobility that has to assign different occupations or educations to a limited set of social categories, as with the Altham Index, still faces exactly the same problem of differing measurement errors across time and space in making these categorizations. A host of detailed occupational descriptions has to be allocated to a small set of categories. The errors in that assignment will be again subject to change over time and across countries, depending on the nature of production technologies and the distribution of employments.

Another proposed solution is using additional observations of an individual over time as an instrumental variable (Altonji and Dunn (1991); Modalsli and Vosters (2019); Ward

(2021)). But this IV strategy will not deal with the second measurement problem we outlined above, which is the varying tightness and looseness of occupational labels across time and place. Also only some individuals will have multiple occupational reports if we follow individuals across censuses. Thus the instrument will only be a partial correction for measurement errors in occupational status.

Here we show how we can use the marriage records in England to estimate intergenerational occupational correlations that will be independent of both measurement errors discussed above. As is detailed in appendix 1, since 1837 the marriage certificate in England has recorded the “rank or profession” of grooms, brides and both fathers. Since all these marriage certificates have survived in the archives of the Registrar General there is thus a trove of 110 million marriage records in England 1837-2022 with which potentially to measure intergenerational social mobility. As noted above we have access to a subset of 1.7 million copies of these records, transcribed from church registers deposited in County Record Offices.

The marriage certificates also recorded much ancillary information about the marriage. It recorded whether the bride or groom were minors, and so needed parental consent for the marriage. Often it gave the age of both bride and groom. It also noted whether this was the first marriage of bride and groom, and if the second or later marriage whether the party was widowed or divorced. It frequently noted whether the bride or groom’s father was dead.

There is also good evidence for England from 1837 onwards that matching in marriage was between grooms and brides. Since the average age of first marriage was high throughout this period, being 24 for women and 26 for men, most parties to marriage were legally and financially independent of their parents, with many women also living separately from their birth families as domestic servants. Arranged marriages were very uncommon. Marriage to close relatives was also uncommon, with cousin marriages well below 1% of all marriages.<sup>7</sup> But below we systematically test whether a significant fraction of marriages involved matches between the fathers or the groom and the father in law, and find no evidence of this.

While brides mostly did not have any listed occupation in the earlier English marriage records, we can think of them having a latent occupational status, that the groom was

---

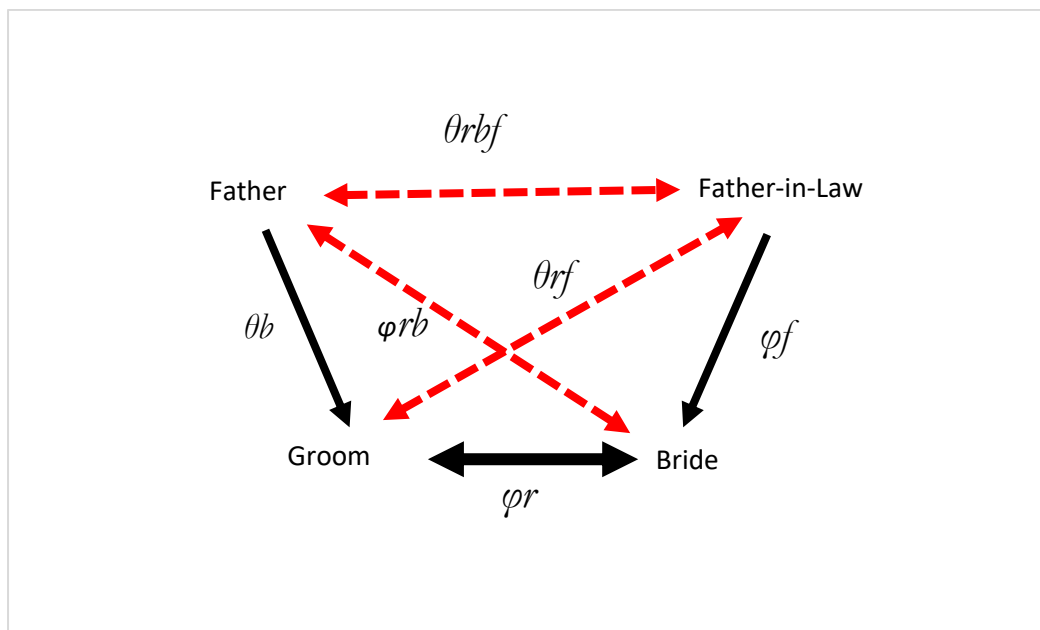
<sup>7</sup> As estimated from the fraction of marriages in the marriage database which involve partners with the same surname.



matching with. This latent status of brides was as important as the manifest status of grooms in predicting child outcomes.<sup>8</sup>

Consider figure 2, which shows the pattern of correlations in occupational status between a groom, his father and his father-in-law, assuming that the matching in marriage is between groom and bride. The true correlations in occupational status between father and son, bride and groom, and bride and her father, are assumed to be  $b$ ,  $r$ , and  $f$ . The correlation in occupational status between groom and father-in-law and father and father in law will be, with matching between bride and groom,  $rbf$  and  $rf$ .

But these correlations, where observed, will be attenuated by measurement errors, measurement errors that vary with time and place. The attenuation will potentially be different where the pair observed is male,  $\theta$ , as opposed to male and female,  $\varphi$ . Figure 2 shows the observed correlations between father and son, father and father-in-law and son and father-in-law.



**Figure 2: Observed Correlations in Status in Marriage**

Notes: The black lines in the figure show observed causal correlations, the red lines resultant correlations, on the assumption of only groom-bride matching.

<sup>8</sup> See Clark and Cummins (2022)

The observed correlation in occupational status between groom and father, and groom and father-in-law will be  $\theta b$  and  $\theta r f$ . The observed correlation in occupational status between father and father-in-law will be  $\theta r b f$ . But this in turn implies that

$$b = \frac{\text{corr}(\text{father}, \text{father-in-law})}{\text{corr}(\text{groom}, \text{father-in-law})} = \frac{\theta r b f}{\theta r f} \quad (1)$$

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.<sup>9</sup>

Table 2 shows these intergenerational correlation estimates for father-in-law to groom and father-in-law to father using the CCC and HISCAM indices for marriages 1837-1979, and the CAMSIS index for marriages 1900-2021. Despite the measured father-in-law to groom correlation dropping substantially on all the indices between 1837-79 and 1980-2021, the implied intergenerational father-son correlation shown in the table is close to 0.9 all the way from 1837 to 1979, independent of what index is used. Thereafter the point estimate drops to 0.77, though with a standard error now because of fewer observations of 0.035, so that the actual value could be in the range 0.72 to 0.84 in the 5% confidence interval. Importantly the substantial decline in the father to son correlations shown by the three different status indices in figure 1 where the decline 1837-2021 is not echoed in the associated estimates of the underlying father-son correlation which goes just from 0.90 down to 0.77.

Note that the underlying marital correlation in underlying social status will be given by

$$r = \frac{\text{corr}(\text{groom}, \text{father-in-law})}{\text{corr}(\text{groom}, \text{father})} \times \frac{b}{f} = \frac{\theta r f}{\theta b} \times \frac{b}{f} \quad (2)$$

If daughters inherit underlying status as strongly as sons, so that  $b = f$ , then the underlying marital status correlation  $r$  will be just<sup>10</sup>

$$r = \frac{\text{corr}(\text{groom}, \text{father-in-law})}{\text{corr}(\text{groom}, \text{father})} \quad (3)$$

---

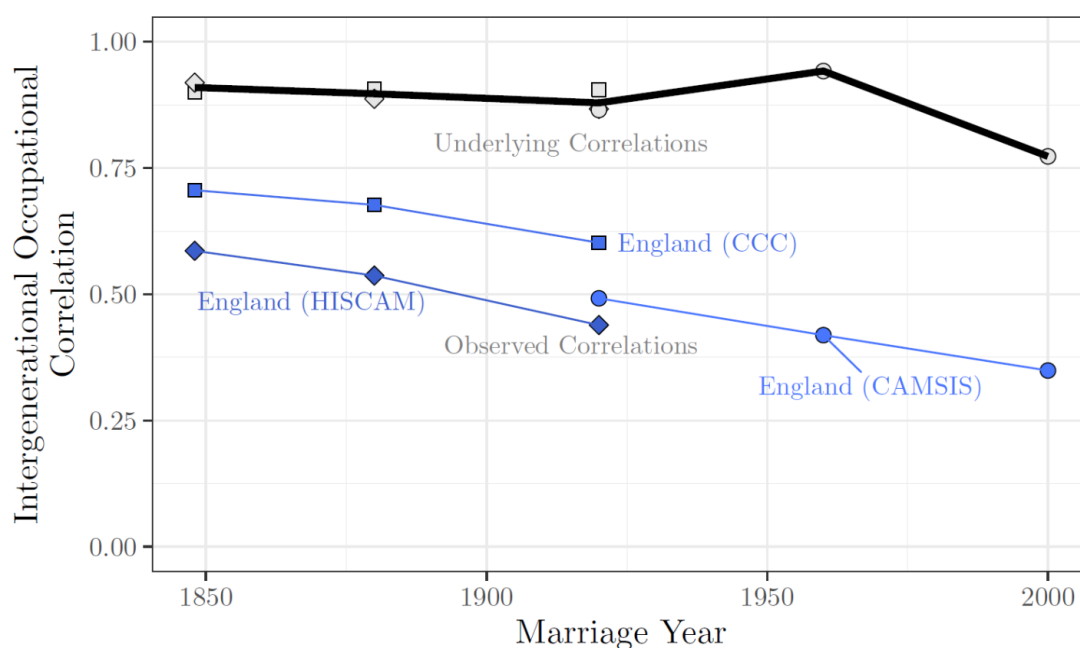
<sup>9</sup> The method used here to correct for errors in status observations echoes that underlying Stuhler et al. (2022).

<sup>10</sup> Clark and Cummins (2022) show other evidence suggesting that  $b = f$ , so that there is gender symmetry in the inheritance of (underlying) occupational status.

**Table 2: 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 (0.001)	0.533 (0.001)	0.706 (0.001)	0.903 (0.002)
1860–1899	CCC	535,670	0.563 (0.001)	0.510 (0.001)	0.677 (0.001)	0.907 (0.002)
1900–1939	CCC	250,827	0.491 (0.002)	0.444 (0.002)	0.602 (0.002)	0.905 (0.003)
1837–1859	HISCAM	379,020	0.408 (0.002)	0.375 (0.001)	0.586 (0.001)	0.919 (0.004)
1860–1899	HISCAM	516,795	0.386 (0.002)	0.343 (0.002)	0.537 (0.002)	0.887 (0.006)
1900–1939	HISCAM	241,369	0.329 (0.002)	0.286 (0.002)	0.439 (0.002)	0.867 (0.006)
1900–1939	CAM SIS	237,824	0.374 (0.002)	0.324 (0.002)	0.492 (0.002)	0.865 (0.005)
1860–1899	CAM SIS	44,430	0.343 (0.004)	0.324 (0.004)	0.419 (0.004)	0.942 (0.014)
1980–2021	CAM SIS	10,872	0.280 (0.009)	0.217 (0.009)	0.349 (0.009)	0.773 (0.035)

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



**Figure 3: Underlying and Observed Intergenerational Correlations, England 1837-2021**

*Note:* The observed correlation are calculated from the 1.7m marriage records 1837-2021. The underlying correlations based on the ratio of the correlation of father-father in law to father in law-groom (as equation 1).

Despite again the declining measured father-son correlations, these implied marital correlations are high, and in this case stable throughout the years 1837-2021, lying always in the range 0.80 to 0.83 whatever status index is used. This very high implied assortment in marriage throughout is consistent with the high intergenerational correlation estimated across these same years.<sup>11</sup>

Thus, despite the observed substantial decline over time in father-son, father to father-in-law, and father-in-law to son correlations, the correlation patterns observed are largely consistent with both very high levels of assortment by status in marriage, and a subsequent very strong intergenerational correlation in status. The observed correlations also suggest no change over time in the strength of marital assortment, despite the rise in female education and employment across these years. Intergenerational mobility in status is very limited throughout, with just a modest increase observed in the last period.

<sup>11</sup> See Clark and Cummins (2022).

### 3. Was Marital Matching only between Spouses?

The results on underlying intergenerational mobility rates are premised on matching in marriage in England 1837-2021 being only between the groom and bride. Figures 4 and 5 show what the pattern of observed correlations between groom, bride, father and father-in-law would be if matching in marriage was between groom and father-in-law, or between father and father-in-law. In each figure we allow for a different underlying correlation in status between son and father ( $b$ ) and daughter and father ( $f$ ). We also allow for different error attenuation when the observation is between men versus between a man and a woman.

As figure 4 shows, the estimate of the underlying intergenerational correlation through

$$b = \frac{\text{corr}(\text{father}, \text{father-in-law})}{\text{corr}(\text{groom}, \text{father-in-law})} = \frac{\theta r b}{\theta r} \quad (1)$$

would not be affected if some marriages were formed by matching in status between groom and father-in-law. The figure also suggests, however, that if this type of matching is important then it produces an asymmetry in the correlations. Now the underlying matching of groom to father-in-law,  $r$ , is much stronger than that from the bride to her father-in-law,  $r b f$ . We show below empirically that there is no evidence for such an asymmetry. Hence an insignificant share of marriages in any period involving matching in status between groom and father-in-law.

If there were significant numbers of marriages coming from matches between father and father-in-law then, as figure 5 shows, expression (1) would produce an estimate of  $1/b$  rather than  $b$  for these marriages. For now (1) becomes

$$\frac{\text{corr}(\text{father}, \text{father-in-law})}{\text{corr}(\text{groom}, \text{father-in-law})} = \frac{\theta r}{\theta r b} = \frac{1}{b}$$

Thus it is important that we show that matching between fathers/families, as represented by father occupational status, was not an important element for English marriages 1837-2021.

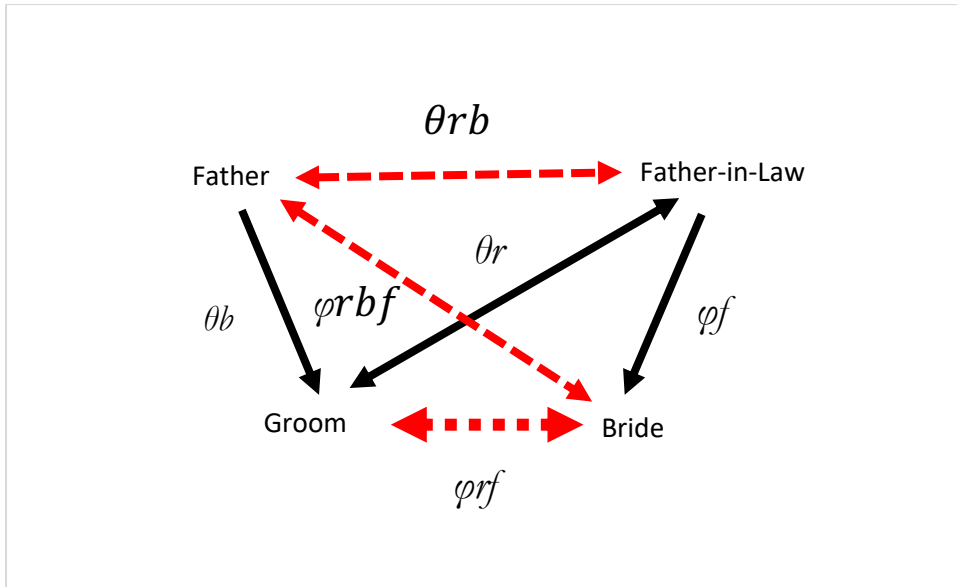


Figure 4: Groom to Father-in-law Matching

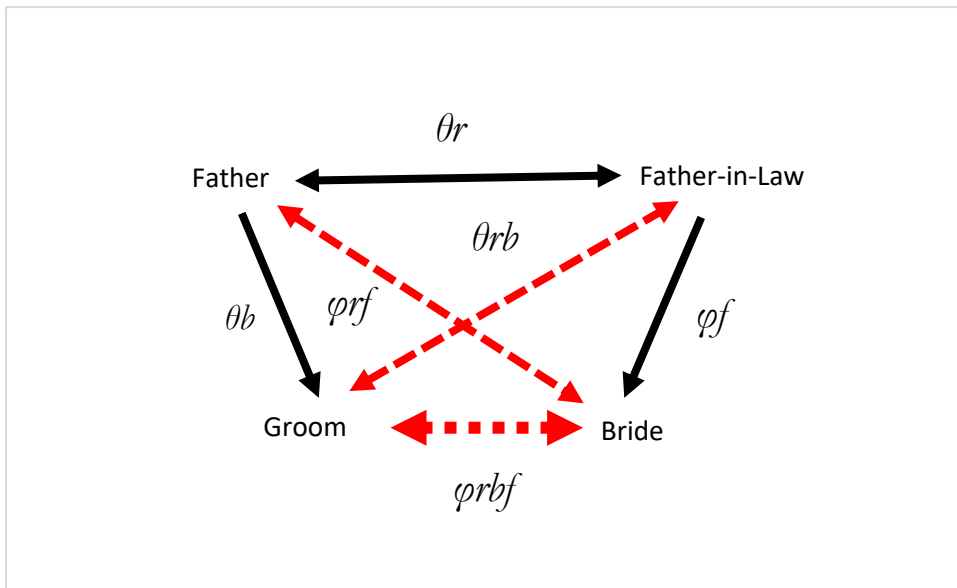


Figure 5: Father to Father-in-law Matching

Empirically, matching across families, if it occurred, would be less prevalent where one or both of the fathers were dead at the time of the marriage. If the fathers, in some subset of marriages, were the ones looking for a match with another suitable family, then the absence of one or more of the fathers should make such matching less important. The characteristics of the son or daughter would matter more than the former status of the dead parent. The matching would be more likely in this case to be from the groom to the father-in-law (which as noted is a type of matching which is consistent with the results above). So the observed status correlation between fathers should be lower where one or more of the fathers are dead, and the correlation between groom and father-in-law higher. Thus the estimated  $b$  in this case should be lower.

Another thing that should reduce father-father matching would be marriages where the bride or groom were entering their second marriage. At least one of the parties to the marriage in this case will have been living much more independently of their parents, and in an independent social circle, making parental matching as the determinant of the pairing less likely.

Conversely parental matching would be more likely if the bride, groom, or both, were underage. Although until 1929 the minimum age for marriage was 14 for men and 12 for women, men and women under age 21 needed the consent of a parent or guardian to marry. This implied that if one of the parties was underage there was much more opportunity for one or both fathers to exert an influence on what match was formed, and again for parental matching being more prevalent.

Another feature of the marriage register is that it gives the parish in which the marriage occurred, as well as the parishes the bride and groom were living in at time of marriage: “residence at the time of marriage.” As table 3 shows a significant number of grooms and brides were not listed as resident in the parish they were marrying in. Marriages seem to have followed a traditional pattern of being in the bride’s home parish, with few brides listed as being resident in another parish.<sup>12</sup> We can thus divide marriages into those which occurred between parties both resident in the bride’s home parish, and marriages where at least one party was resident elsewhere. There would be the greater opportunity of matching between fathers or families where both families came from the same parish.

We thus estimate the father to father-in-law correlation, as well as the groom to father-in-law correlation for each of the four cases

---

<sup>12</sup> There were four times as many grooms residing outside the marriage parish than there were brides.

- (a) At least one father dead versus both fathers alive,
- (b) Both parties of full age, versus at least one child a minor,
- (c) At least one of groom and bride married before, versus both in first marriage,
- (d) At least one party to the marriage resident outside the marriage parish, versus both resident in that parish.

In each case, if matching across fathers in marriage was common, we would expect a lower correlation in status between fathers in the first case than in the second.

We divide the data into the first 50% (1837-1874) versus second 50% (1875-2021) of marriages to examine for any potential trends over time. Table 3 Panels A and B shows the quantity of data available on each measure. In the earlier period it was not consistently recorded when fathers were dead, which explains the smaller numbers of deaths recorded in this interval. The records do consistently report when either a bride or groom is a minor needing parental consent, giving good sample sizes for both periods on this measure. The records are again complete on whether bride or groom has been married before. Where the records are harder to decipher is on whether the bride and groom reside in the parish where the marriage took place. Often the residence of the bride or groom is given as some subunit of the parish, or as a street address. We took all street addresses not followed by an indication of a township or parish as showing the person resided in the marriage parish. Where the residence was listed in another county (or even country), or another parish in the same county, we took that as indicating a residence outside the marriage parish. We also counted all individuals describes as “sojourner in this parish” as resident outside the parish.

Table 3 also shows the characteristics of grooms and brides in each of the subgroups in terms of age and groom’s occupational status. Occupational status is measured 1837-1874 using the CCC scale, and 1875-2021 using the CAMSIS 1990 scale.

There is very little difference in average groom occupational rankings for marriages involving minors versus those where both parties were of full age, and for those marrying for the first time versus those in a second marriage. Grooms in marriages involving minors, however, have lower occupational status than those of full age. But the biggest occupational gap occurs between those where both parties are resident in the parish of the marriage versus those where at least one party resides in another parish. For marriages with both parties resident in the parish, occupational status 1837-75, for



**Table 3: Distribution of Marriage Data**  
**Panel A**

<b>Period</b>	<b>Variable</b>	<b>Fathers Alive</b>	<b>At least one father dead</b>	<b>Both Parties Full Age</b>	<b>Someone a minor</b>
1837-1875	Groom Age	26.4	27.5	28.5	22.1
	Bride Age	24.4	24.9	26.8	19.5
	Groom Occ Stat (CCC)	31.1	28.8	32.0	27.1
	Count	849,008	51,547	581,283	172,951
1875-2020	Groom Age	26.2	29.1	28.4	22.5
	Bride Age	24.2	26.6	26.5	19.5
	Groom Occ (CAMSIS)	38.9	39.5	40.1	35.5
	Count	467,214	193,181	555,997	150,725

**Panel B**

<b>Period</b>	<b>Variable</b>	<b>Both first marriage</b>	<b>Someone married before</b>	<b>Both resident in parish of marriage</b>	<b>Someone resident elsewhere</b>
1837-1874	Groom Age	24.4	38.6	26.3	27.2
	Bride Age	22.6	34.4	24.2	24.7
	Groom Occ Status	32.0	32.1	28.9	38.3
	Count	623,591	150,725	514,052	125,979
1875-2020	Groom Age	25.5	40.8	26.9	27.6
	Bride Age	23.6	36.2	24.8	25.3
	Groom Occ (CAMSIS)	39.1	39.2	36.3	45.0
	Count	564,870	70,351	314,302	72,127

example, averaged 29, versus 38 for those where one party came from outside the parish. In comparing the four groups, thus, we adjust for occupational status for this residence comparison by matching each groom where the marriage involved a party outside the parish with a groom of equivalent occupational status from the group where all parties were within the parish.

In almost all these cases we find the correlations are unaffected in subsamples where father-father-in-law matching would be potentially more prevalent. Figures 6 and 7 show the estimated father-father-in-law correlation, groom-father-in-law correlation, and implied intergenerational correlation  $b$ , as well as the 95% confidence interval for each of the 8 sub-samples of the data, in each of the two time periods.<sup>13</sup> Appendix tables A3 and A4 detail these estimates and associated standard errors.

Dead fathers, wedding parties of full age, or who are entering their second marriage, where father-father-in-law matching would be less prevalent, if it occurred at all, all show just as strong a correlation between fathers and fathers-in-law in occupational status. The resulting estimates of  $b$  are consequently just as high as in the general population of weddings, being still around 0.9.

The only case where the father-father-in-law correlation changes significantly is where we compare marriages where both parties are resident in the marriage parish to those where at least one party is resident elsewhere. Here, because for marriages with a party resident elsewhere the occupational status of the groom is much higher, we control for occupational status in making the comparison by matching grooms in each category by occupational status. But even with this control, as figures 6 and 7 show, the father-father-in-law correlation is greater for the case where a party to the marriage is from outside the parish the marriage is occurring in. However, this difference is the opposite of what we would expect if father-father-in-law matching is driving up the  $b$  estimate. For with an out-of-parish party to the marriage we would expect family matching to be less likely than when both families live in the same parish. But this case is the one that produces the highest estimates of  $b$  in the range 0.95.

Overall figures 6 and 7 suggest that matching in marriage in England, all the way from 1837 to 2020 was completely dominated by bride-groom matches on characteristics. Matching between families seems to have been the determinant in an insignificant fraction of marriages.

---

<sup>13</sup> For the  $b$  estimates the 95% confidence interval was bootstrapped with 10,000 repetitions.

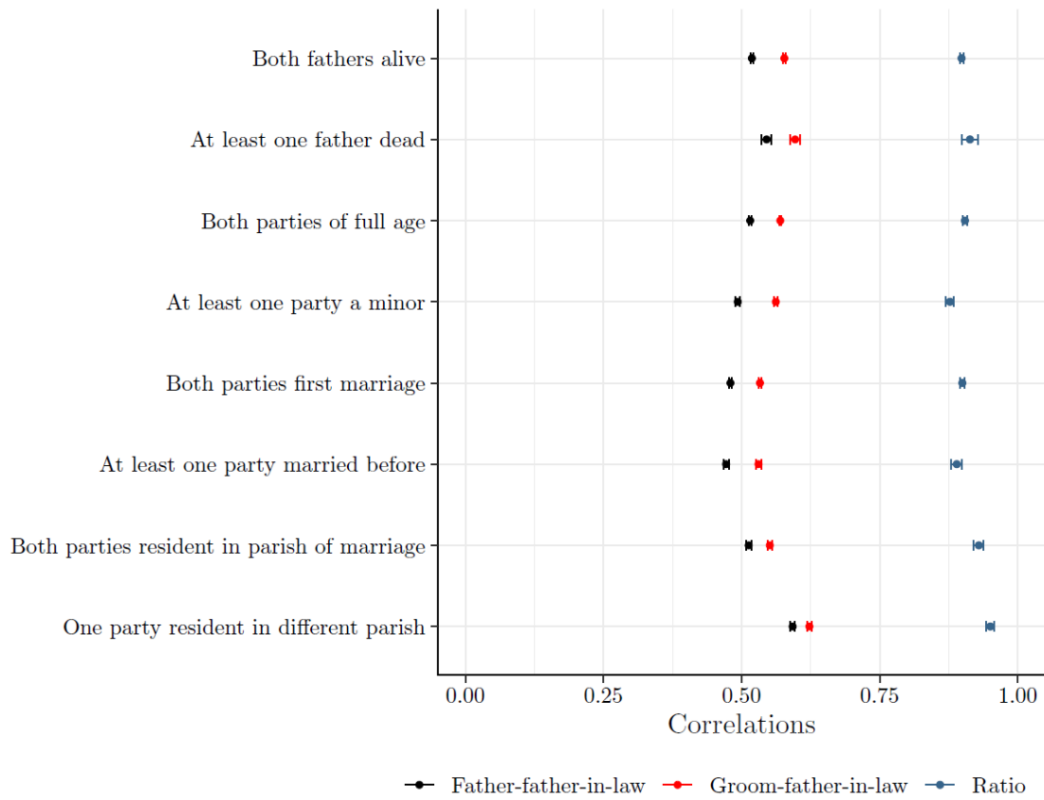


Figure 6: Testing for father-father-in-law matching, 1837-1874

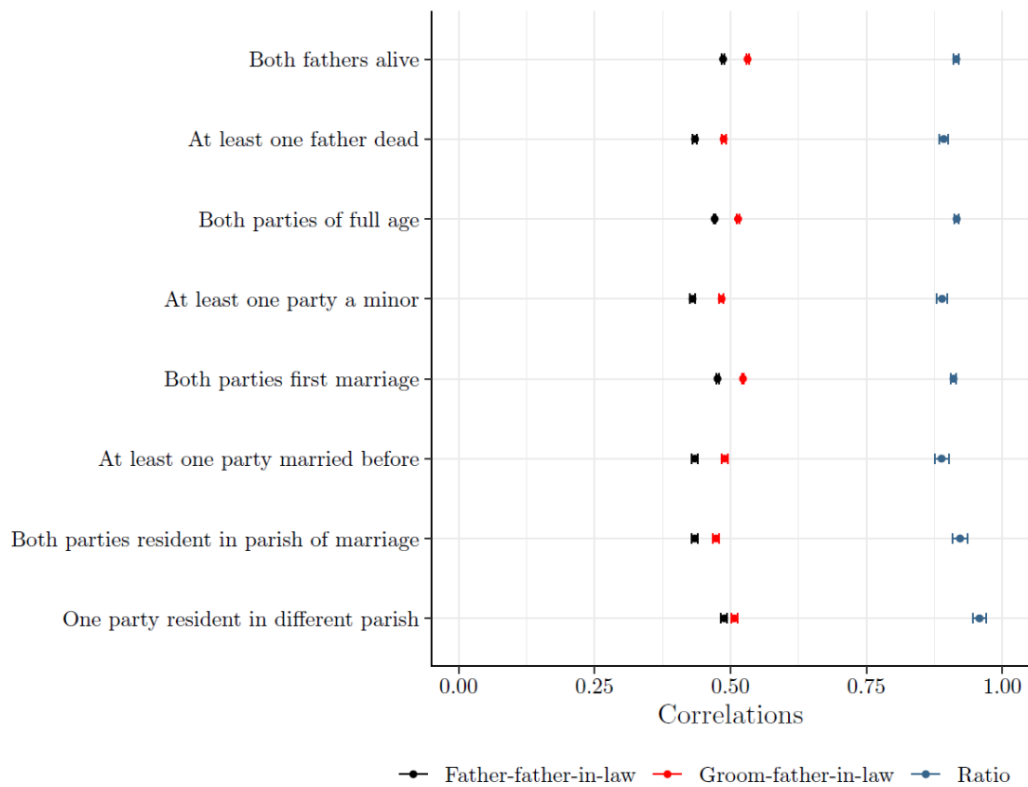


Figure 7: Testing for father-father-in-law matching, 1875-2021

The above evidence strongly supports the idea that father to father-in-law matching was insignificant in marital matches in England 1837-2021. There is also compelling evidence that matching from groom to father-in-law was unimportant. As noted, in that case the underlying matching of groom to father-in-law,  $r$ , is much stronger than that from the bride to her father-in-law,  $rbf$ . For the period 1837-79 we can test for this type of matching because we have a measure of status for both bride and groom, which is their literacy indexed by their ability to sign the marriage register. To estimate the underlying groom-father-in-law correlation we can take the ratio of the correlation of groom literacy to their father-in-law compared to the correlation with their own father. With groom to father-in-law matching this ratio will be, from figure 4,

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

In contrast, again from figure 4, for brides the expected ratio of the correlation of their literacy with their father-in-law versus father is

$$\frac{\text{corr}(\text{bride}, \text{father} - \text{in} - \text{law})}{\text{corr}(\text{bride}, \text{father})} = \frac{\varphi r b f}{\varphi f} = r b$$

This implies that the second ratio, for the bride, should be lower by a factor of  $b^2$  if matching in marriage was groom to father-in-law. In intuitive terms groom to father-in-law matching implies that the relative correlation of a groom to his father-in-law compared to his father should be much stronger than the relative correlation to father-in-law compared to father for brides.

Table 4 shows these estimated correlations and relative correlations using literacy of brides and grooms compared to occupational status of their fathers. For the period 1837-1879 we have 365,371 cases where we observe both bride and groom literacy at marriage, and fathers' occupational status.<sup>14</sup> From this data we can calculate the correlation of the literacy of bride and groom with their father and father-in-law. When we take the ratio of these correlations for bride and groom we find a very similar number. There is no evidence that the groom-father-in-law connection is much stronger than the bride-father-in-law connection, as would be observed with matching between groom and father-in-law. Instead the matching seems close to symmetrical: as we would get with bride-groom or father-father-in-law matching.

---

<sup>14</sup> After 1879 literacy becomes close to universal and so less useful as a status measure.

**Table 4: Testing for Asymmetrical Marital Matching, 1837-1879**

<b>Relative</b>	<b>Bride Correlation</b>	<b>S.E. Correlation</b>	<b>Groom Correlation</b>	<b>S.E. Correlation</b>
Own Father	0.155	0.00007	0.278	0.00006
Father-in-law	0.128	0.00007	0.209	0.00006
Ratio	0.828	0.0003	0.753	0.0003

*Note:* The standard error of the ratio was bootstrapped with 10,000 repetitions.

#### **4. Is an Intergenerational Status Correlation of 0.88 too High?**

The average underlying intergenerational correlation in occupational status of 0.88 calculated above is much higher than most existing estimates. There is reason that it may be potentially biased upwards by about 0.05. This correlation is calculated as the ratio of the correlation of status of the fathers to the son-father-in-law correlation. The father-father correlation is calculated a generation earlier than the father-son correlation. If these correlations are declining over time because of changes in the structure of employments, and in the terms used to describe occupations, then potentially they will be biased upwards by this downward drift in measured correlations. Based on the observed data, however, the maximum upward bias for England would be around 0.05. This would imply the true underlying intergenerational correlation would still be 0.83 or greater.

Is such a high intergenerational correlation plausible? We can show for marriages in England in the period 1837-1879 that the true correlation must be indeed in this range.

The CCC index above shows a measured intergeneration occupational status correlation of 0.70 for marriages 1837-1859. But this measured correlation will still be below the true correlation as a result of two forms of measurement error. The first is the mismeasurement of the exact average status of each of the 442 occupation categories. The second is that people whose occupation is assigned to the same of the 442 categories will

often actually differ in occupational status. The category “clerk,” for example, covers occupations that differ widely in earnings, and in other measures of occupational status.

Suppose a persons true occupational status is  $z$ . Suppose also their assigned status on an occupational index is  $Z$ . Then there will be two independent errors linking their assigned status to their true status.  $Z = z + u + e$ , where  $e$  is the error in measuring the true average occupational status of the assigned occupation  $Z$ .  $u$  is the error caused by the range of occupations that fall under the label  $Z$ , each with a different underlying status.

When we measure intergenerational mobility with such a social status index the estimate is biased downwards by a factor

$$\frac{\sigma_z^2}{\sigma_z^2 + \sigma_u^2 + \sigma_e^2} \quad (4)$$

Using an entirely different set of data which links occupations in the period 1800-1939 in England with measures of education and wealth at death we are able to construct another occupational index for this period for England, which we dub here the CCC2 index.<sup>15</sup> This index performs almost as well as the CCC index in terms of the observed intergenerational correlation of status on the marriage database. Yet it has a correlation of just 0.86 in terms of the occupational status assigned to occupation categories.

For the CCC and CCC2 indices, because of their entirely independent construction, the error component  $e$  attached to errors in the average occupational status by category will be independent, but not the within-category component  $u$ . Assuming the error term  $e$  variance is the same for each of these indices, the correlation between these indices 0.86 will be

$$\rho = \frac{\sigma_z^2 + \sigma_u^2}{\sigma_z^2 + \sigma_u^2 + \sigma_e^2} = 0.86 \quad (5)$$

This implies that the error component in these indices we have derived has to be at least 14% of the variance in measured status. It also implies that if we multiply our father-son correlations by 1.16 we will get an estimate closer, but still not as large as, the true underlying persistence of occupational status across generations. Since that correlation for the CCC index for 1837-1859 marriages is 0.70, the true intergenerational correlation in occupational status in this period has to be at least 0.81. When we add the attenuation caused by the variance within occupational categories, the true underlying correlation of

---

<sup>15</sup> For details of the construction of these indices see Clark et al. (2022).

occupational status in England 1837-1859 could easily equal the 0.83-0.88 estimated by the new method above. Thus for at least the 1837-1859 period we can verify that there is nothing unreasonable about an estimated true intergenerational status correlation of 0.83-0.88.

## 5. Conclusion

In this paper we first illustrate the dangers of trying to compare rates of intergenerational mobility across time and place. Measured mobility rates will be strongly influenced by both the quality of the occupational status index, and the characteristics of employment by time period and location. Any conclusions from such raw intergenerational correlations on mobility differences across time and place are fraught with error, and are unreliable.

Next we show how to derive measures of the true underlying father-son occupational status correlation, independent of measurement errors, using the occupational status of fathers, sons, and fathers-in-law. These new measures of intergenerational correlations for England are much higher than those measured by conventional estimates. They are in the region of 0.9 for all periods but the most recent, marriages 1980-2021, where it is around 0.78. They suggest that the social world shows far less intergenerational social mobility than has been traditionally believed. They also suggest little change in social mobility rates in England over the years 1837-2021.

These new measures will be valid if the overwhelming majority of marriages were formed through matching between grooms and brides, independent of their parents' social status. We run a variety of tests of this assumption, and find that circumstances which would allow more parental input in marital unions, such as grooms or brides who were minors, are not associated with any stronger correlation in status between the fathers. Marital matching in England all the way from 1837 to the present, as it is now, was largely dependent on the characteristics of the bride and groom.

There will be resistance to the idea that true social mobility rates could be so low. But we see from the CCC index that for the period 1837-1859 the implied correlation in underlying occupational status father-son has to be substantially greater than 0.81. So at least in this period we can independently validate these new estimates.

## Appendix 1: The Freereg Marriage Data

Most of these marriage records were transcribed by volunteers to the Freereg organization, and posted on their web page (<https://www.freereg.org.uk/>), but we collected 32,000 additional marriage records from Essex church records 1837-2021.

These marriage records, where the information comes from church registers deposited in local record offices, exclude civil marriages. Though Civil marriage was introduced in England in 1837, such marriages remained a small minority of all weddings before 1914. In 1841 Civil marriages were 1.7% of all marriages. In 1914, they were still only 24%, and in 1952 31% (Haskey, 2015).

The details in the marriage record have remained unchanged from 1837 to the present. Figure A1 shows a marriage from 2020 and one from 1838. The marriage record records for groom, bride and both fathers “Rank or profession.” It also gives the groom and bride age (or just “minor” and “full age”), as well as their ability to sign the register, a measure of literacy for the earlier years.

Page 10

2020 Marriage solemnized at <u>ST ANDREW'S CHURCH</u> in the Parish of <u>ASHINGDON</u> in the <u>COUNTY OF ESSEX</u>								
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
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 [Signature] after BANNS by me, [Signature] CURATE.

This marriage was solemnized between us, [Signature] in the presence of us, [Signature] [Signature] [Signature] [Signature]

---

Page 1.

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 Hephzibah Chopping	19 20	Bachelor Spinster	Labourer -	Bardfield Salting Bardfield Salting	John White Michael Chopping	Labourer Miller

Married in the Church after banns according to the Rites and Ceremonies of the Church of England by me, Richard Victor's Bryon Minister

This Marriage was solemnized between us, Henry White his mark in the Presence of us, Elizabeth White + her mark Hephzibah Chopping + her mark Benjamin Chopping his mark

Figure A1: The Marriage Register in England, 1837-2022



Because transcribing these marriage records is a volunteer effort based on local interests, the numbers of marriages recorded by county varies considerably. Four counties contain about 50% of the Freereg marriages transcribed for England: Kent, Lancashire, Lincolnshire, and Staffordshire. But these counties were very different in terms of occupations and urbanization, so that the overall sample generated is representative of England as a whole. Figure A2 shows the overall coverage of parishes by county.



**Figure A2: Distribution of observations by parish, Freereg marriage database**

The marriage records for the years 1980-2021 came largely from deposited church marriage registers in Essex Record Office that we collected ourselves. In these years only a minority of all weddings were performed in churches. 49% of weddings by 1982 and 68% by 2012 were civil (Haskey, 2015). But there is no reason to expect that the father-son correlations for church weddings would be any different than for the population as a whole.

## Appendix 2: Robustness Checks on Marital Matching

Appendix tables A3 and A4 give the detailed estimated underlying figures 6 and 7 in the main text. The tables show a set of cases, indicated by a \*, where if father to father-in-law occupational status matching was occurring in a significant subset of marriages, then such matches should be more prevalent. Fathers, for example, are more likely to match in marriage where both are alive, as opposed to where one or both are dead. If such matching was occurring then it would drive up the father-father-in-law correlation. This in turn would increase the estimated value of  $b$ , which is derived as the ratio of the father-father-in-law to groom- father-in-law correlations.

**Table A3: Robustness Checks, 1827-74**

Subgroup	N	Father-finl correlation	Groom-finl correlation	$b_{Ratio}$
Both fathers alive*	579,342	0.519 (0.001)	0.577 (0.001)	0.898 (0.002)
At least one father dead	21,794	0.545 (0.005)	0.597 (0.005)	0.914 (0.008)
At least one party a minor*	131,289	0.493 (0.002)	0.562 (0.002)	0.877 (0.004)
Both parties of full age	436,812	0.516 (0.001)	0.570 (0.001)	0.904 (0.002)
Both parties first marriage*	478,014	0.480 (0.001)	0.533 (0.001)	0.900 (0.002)
At least one party married before	92,370	0.472 (0.003)	0.531 (0.002)	0.890 (0.005)
Bride and groom resident in marriage parish*	87,779	0.513 (0.003)	0.551 (0.002)	0.930 (0.005)
One party resident elsewhere	87,779	0.593 (0.002)	0.623 (0.002)	0.950 (0.004)

Notes: \* indicates marriage groups where father-father-in-law matching would be more prevalent, if it was occurring.

**Table A4: Robustness Checks, 1875-2020**

Subgroup	N	Father-finl correlation	groom-finl correlation	$b_{Ratio}$
Both fathers alive*	450,527	0.487 (0.001)	0.532 (0.001)	0.915 (0.002)
At least one father dead	148,396	0.435 (0.002)	0.487 (0.002)	0.892 (0.004)
At least one party a minor*	135,293	0.430 (0.002)	0.484 (0.002)	0.889 (0.005)
Both parties of full age	476,023	0.471 (0.001)	0.514 (0.001)	0.916 (0.002)
Both parties first marriage*	511,272	0.476 (0.001)	0.523 (0.001)	0.910 (0.002)
At least one party married before	65,063	0.434 (0.003)	0.489 (0.003)	0.888 (0.006)
Bride and groom resident in marriage parish*	67,225	0.434 (0.003)	0.474 (0.003)	0.922 (0.007)
One party resident elsewhere	67,225	0.488 (0.003)	0.507 (0.003)	0.958 (0.006)

Notes: \* indicates marriage groups where father-father-in-law matching would be more prevalent, if it was occurring.

## References

- Altham, P. M. E. and J. P. Ferrie, “Comparing Contingency Tables Tools for Analyzing Data from Two Groups Cross-Classified by Two Characteristics,” *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 2007, 40 (1), 3–16.
- Altham, Patricia M. E., “The Measurement of Association of Rows and Columns for an  $r \times s$  Contingency Table,” *Journal of the Royal Statistical Society. Series B (Methodological)*, 1970, 32 (1), 63–73.
- Altonji, Joseph G. and Thomas A. Dunn, “Relationships Among the Family Incomes and Labor Market Outcomes of Relatives,” June 1991.
- Bouchet-Valat, Milan, “General Marginal-free Association Indices for Contingency Tables: From the Altham Index to the Intrinsic Association Coefficient,” *Sociological Methods & Research*, 2022, 51 (1), 203–236.
- Braun, Sebastian Till and Jan Stuhler, “The Transmission of Inequality Across Multiple Generations: Testing Recent Theories With Evidence From Germany,” *The Economic Journal*, March 2018, 128 (609), 576–611.
- Clark, Gregory and Neil Cummins, “Assortative Mating and the Industrial Revolution: England, 1754–2021,” CEPR Discussion Paper DP17074 2022.
- \_\_\_\_\_, \_\_\_\_\_ and Matthew Curtis, “The Mismeasure of Man: Why Intergenerational Occupational Mobility Is Much Lower than Conventionally Measured, England, 1800–2021,” Technical Report DP17346, Centre for Economic Policy Research, London 2022.
- Lambert, Paul S., Richard L. Zijdemans, Marco H. D. Van Leeuwen, Ineke Maas, and Kenneth Prandy, “The Construction of HISCAM: A Stratification Scale Based on Social Interactions for Historical Comparative Research,” *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 2013, 46 (2), 77–89.
- Long, Jason and Joseph Ferrie, “Intergenerational Occupational Mobility in Great Britain and the United States since 1850,” *American Economic Review*, June 2013, 103 (4), 1109–37.
- Long, Jason and Joseph Ferrie, “Grandfathers Matter(ed): Occupational Mobility Across Three Generations in the US and Britain, 1850–1911,” *The Economic Journal*, 2018, 128 (612), F422–F445.
- Modalsli, Jørgen and Kelly Vosters, “Spillover Bias in Multigenerational Income Regressions,” *Journal of Human Resources*, 2019, pp. 0520–10919R2.
- Perez, Santiago, “Intergenerational Occupational Mobility across Three Continents,” *The Journal of Economic History*, 2019, 79 (2), 383–416.
- Prandy, Kenneth and Paul Lambert, “Marriage, Social Distance and the Social Space: An Alternative Derivation and Validation of the Cambridge Scale,” *Sociology*, 2003, 37 (3), 397–411.

- \_\_\_\_\_ and Wendy Bottero, “Social Reproduction and Mobility in Britain and Ireland in the Nineteenth and Early Twentieth Centuries,” *Sociology*, 2000, *34* (2), 265–281.
- Solon, Gary, “Intergenerational Income Mobility in the United States,” *The American Economic Review*, 1992, pp. 393–408.
- Song, Xi, Catherine G. Massey, Karen A. Rolf, Joseph P. Ferrie, Jonathan L. Rothbaum, and Yu Xie, “Long-term decline in intergenerational mobility in the United States since the 1850s,” *Proceedings of the National Academy of Sciences*, 2020, *117* (1), 251–258.
- Stuhler, Jan, Collado, M. Dolores and Ortuno-Ortin, Ignacio, “Estimating Intergenerational and Assortative Processes in Extended Family Data.” Working Papers, Human Capital and Economic Opportunity Working Group, , No 2022-022, <https://EconPapers.repec.org/RePEc:hka:wpaper:2022-022>.
- Ward, Zachary, “Intergenerational Mobility in American History: Accounting for Race and Measurement Error,” NBER Working Paper 29256 September 2021.