The Child Quality-Quantity Tradeoff, England, 1780-1879: A Fundamental Component of the Economic Theory of Growth is Missing

Gregory Clark and Neil Cummins*

January 19, 2019

Abstract

In recent theorizing, modern economic growth was created by substituting child quality for quantity. However evidence for any substantial tradeoff of child quality for quantity is minimal. In England the Industrial Revolution occurred in a period of substantial human capital investment, but no fertility control, huge random variation in family sizes, and uncorrelated family size and parent quality. Yet family size variation had minor effects on educational attainment, occupational status, and child health, for both prosperous and poor families. More children did substantially reduce wealth at death, but only for rich families with inherited wealth. In families with no parental wealth, wealth at death was unaffected by family size. Even for richer families the wealth effect substantially diminishes by the grandchild generation. There is no significant quality-quantity tradeoff in human capital even well into the modern growth era. Growth theory must proceed in other directions.

1 Introduction

Modern high income societies have a combination of low fertility levels and high levels of nurture and education for children. There is much human capital. Poor societies have high fertility levels, lower levels of nurture for children, and less education. Recent economic theory has taken this regularity, and made it central to the theory of economic growth. Growth, it is argued, stems at base from higher levels of human capital (see, for example, Willis (1973); Becker et al. (1990); Galor and Weil (2000); Galor and Moav (2002); Lucas (2002); Galor (2011); O'Rourke et al. (2013)).

Thus the abstract of a recent article published in the *Journal of Economic Growth*, edited by Oded Galor, the main proponent of unified growth theory, states: "The trade-off between child quantity and quality is a crucial ingredient of unified growth models that explain the transition from Malthusian stagnation to modern growth" (Becker et al. (2010), 177). Similarly, Matthias Doepke in a 2015 survey of the contributions of Gary Becker to the concept of the quantity-quality

*Data for this paper was in part collected with NSF grant SES 09-62351 2010-13 (Clark), and PSC-CUNY Award # 64125-00 4 (Cummins). We thank Oded Galor for comments on an earlier version of the paper which stimulated us to make significant improvements in this version. Peter Lindert, Doug Miller, Marianne Page, Ann Stevens, and participants in seminars at UC-Berkeley, Copenhagen University, UC-Davis, UC-Santa Barbara, University of Colorado, Boulder, University of Southern Denmark, LSE, and Warwick all made helpful suggestions.

tradeoff noted: "In unified growth models ... endogenous fertility is a crucial element to account for the phase of Malthusian stagnation, and in most theories Becker's quantity–quality tradeoff is a key element to account for the joint demographic transition and economic takeoff that characterizes successful economies" (Doepke (2015), 63). Doepke further notes that there has been little evidence for a QQ tradeoff in modern developed economies, which have compulsory and free education, and bans on child labor, but "In developing countries, however, there is less provision of high-quality public education, and child labor is still common. These features imply a high opportunity cost of education, and we would expect the quantity–quality tradeoff to be central in shaping fertility and education choices" (Doepke (2015), 63).

Only when circumstances arose in which parents chose to have smaller family sizes was it possible to increase human capital. Parents have limited time and money. The more children parents have, the less input each child can receive, and the less effective will these children be as future economic agents. Economic growth did not come to the world until the last 250 years because before then women gave birth to many children, and these children received little nurture or education to create capable economic agents.

Yet this crucial underlying assumption - that the more children a given set of parents have, the less productive the children will be – rests on the flimsiest empirical evidence. In many modern societies there is a negative correlation between family size and measures of child quality[1]. But modern family sizes are determined by parental choices, choices that correlate with unobservable features of parents which influence child quality. So any observed correlation between quality and quantity will be potentially biased as measures of the causal effects of quantity. Tests of the quality-quantity tradeoff which control for this by using the accidents such as twining of second births, or the gender of first births, mostly fail to detect a significant tradeoff[2]. For modern families, however, the range of variation in family size is mostly 1-3. Pre-industrial families had much greater size variance, so more potential of size eventually having significant effects on quality.

Some theories, such as that of Galor and Moav (2002) have created part of the shift in the modern regime towards higher quality and lower quantity through the selective survival in the Malthusian regime of agents with a greater preference for child quality. In such a setting if families choose the number of children then it is possible to construct models where in cross section there is no negative correlation between child numbers and child quality. However, these models still embody a substantial quantity-quality tradeoff. So even with a heterogenous set of agents with different quality-quantity preferences, if fertility depends mainly on random shocks we will observe a significant quality-quantity tradeoff.

In this paper we utilize a dataset containing the histories of a set of English families which had rare surnames 1780-2016, described below. Using birth, death and marriage records, probate records, censuses, and other sources we reconstructed the histories of 10,000 men marrying 1780-1879 and their 59,000 children. We measure family size with a high accuracy. In England for marriages commencing 1780-1879 there is minimal raw association between fertility and parent "quality". But more importantly, we show that nearly all family size variation lay outside the control of parents, so that the bias caused by correlations between family size and unmeasured "quality" is minimized. From the perspective of the parents, family size was an exogenous shock.

[1]But note that for pre-industrial societies even that raw correlation between quality and quantity is either absent or weak. See Carmichael et al. (2015), Figure 2, p. 21.

[2]Mogstad and Wiswall (2016), argue however that this is because size effects are non-linear, being positive going from 1-2, and 2-3 children, but then becoming negative for larger family sizes. For the sample of families examined here another advantage is that, with an average sibship size of 6 they are concentrated in the area where Mogstad and Wiswall find negative size effects are currently large.

We get thus get largely unbiased estimates of the causal association between size and education, occupation, longevity and wealth. The conclusion is that family size had little effect on education, occupation, longevity, or even on wealth, though in this case it is wealth at death relative to wealth inherited.

The period of study, marriages 1780-1879 is already one where there were considerable investments in education and training. Figure A.2 shows, for example, the percentage of males born from marriages before 1880 described as "scholars", "apprentices," in the censuses of 1851-1911. This number is a lower bound since frequently, even for rich families where we know education would be provided at younger ages, the census enumerators simply left blank the "occupation" column for children without a job or trade, without specifying explicitly "scholar." The families in this sample are divided into those whose rare surnames were on average wealthy, and on average poor[3]. For families from richer lineages 64-80% of sons were still in education or training at age 15. Even for the poor ones it was 22-34%. The age we are looking at here is already the modern one of significant human capital.

The outline of the paper is as follows. We review the existing literature that has attempted to measure the quality-quantity tradeoff, and find little evidence of quantitatively significant effects. We describe the Families of England database, and show that for marriages 1780-1879 there was no targeting of family size, and no correlation of fertility with parent social status. We then show that variations in family size were associated either with no decline in child quality, or with very modest effects, except for the case of wealth at death, which reflects physical and not human capital. We also show that these effects are not being disguised by non-linearities, whereby their are quality gains moving from 1 to 2 or 3 children, and only then loses from greater quantity. We lastly consider how this result was possible in a world where there were already significant quantities of human capital even in poorer households.

1.1 Measuring the Quality-Quantity Tradeoff

The empirical evidence for a quality-quantity tradeoff is based on negative correlations between family size and the measurable 'quality' (educational attainment, health) of offspring. Studies of modern populations show a negative correlation between child numbers and educational and economic achievement[4]. These studies also recently highlighted differing trade-offs for groups at different socioeconomic levels. Grawe (2009) for the US, and Lawson and Mace (2010) for Britain, for example, find a stronger quality-quantity tradeoff for richer families. However, to capture the causal quality-quantity trade-off, researchers must control for the endogeneity of modern family size. Parent influences on child "quality" can follow two potential routes: there is transmission of characteristics through genes or culture, and investments in child quality. Since in the modern world high 'quality' parents often also have smaller numbers of children, the observed negative correlation between N and child quality may stem just from the positive correlation of parent and child quality.

To uncover the true relationship investigators have sought exogenous variation in family size. The most common is that caused by the accident of twin births (e.g. Rosenzweig and Wolpin (1980a); Angrist et al. (2006); Cáceres-Delpiano (2006); Li et al. (2008); Rosenzweig and Zhang

[3]The status of each surname lineage as rich, intermediate, or poor was determined by their average wealth at death 1858-1887.

[4]See Grawe (2003); Lawson and Mace (2010) for Britain, Rosenzweig and Wolpin (1980b); Kaplan et al. (1995) for the US, Rosenzweig and Wolpin (1980a), Jensen (2005) for India, Lee (2008) for Korea, Grawe (2003) for Germany, Desai (1995) for 15 developing countries (using heights as a measure of child quality).

(2009); Ponczek and Souza (2012); Mogstad and Wiswall (2016)). Others, in societies where males are preferred, use first child gender as a source of exogenous variation (Kugler and Kumar (2017)). Liu (2014), uses the relaxation of the one-child policy in China. These strategies, however, typically allow for only modest variation in family size, in the range 2-4. This variation is at levels well below the typical average family size of pre-industrial Europe. The average sibship in our sample below was 6.

These studies in general produce remarkably weak evidence of a significant quality-quantity tradeoff. The typical study either finds no effect, or finds effects that are marginally significant statistically, effects that show up only for some quality measures, and effects that are quantitatively minor. And there is suspicion here, given the centrality of the quantity-quality tradeoff in economics, that there will be a publication bias against studies finding no effect. Angrist et al. (2006), for example, which finds "no evidence of a quality-quantity trade-off" for Israel using census data, is an unpublished working paper, as is Qian (2009), which finds a positive quality-quantity tradeoff in China. Even for the studies that report significant size effects in their abstracts, such as Cáceres-Delpiano (2006), for the USA and Ponczek and Souza (2012), for Brazil, the magnitude of the effects are minor, and the statistical precision weak. Cáceres-Delpiano (2006), finds that an extra child reduces mother's labor force participation, and the chance a child attends private school. But in terms of child academic outcomes there is no effect. In Brazil where the authors announce "negative effects on educational outcomes" an additional child reduces the probability of school attendance at ages 10-15 by 1.3% for boys, and 1.6% for girls. But the standard error for both is around 0.9%. At age 18-20 the probability of school attendance for boys rises by 1.3%, and for girls declines by 1.6%, but now with standard errors of 2.1% and 1.6%. Kugler and Kumar (2017), use as an instrument for family size in India the first born being a girl. In such cases family size is typically 0.3 children larger. With a sample of more than 393,000 families they find that an additional child leads to a -0.08 years reduction per child in schooling, but with a standard error of 0.033. And additional children have no effect on health outcomes.

Black et al. (2005) report a negative family size–child quality relationship for Norway, but find that it completely disappears once they include controls for birth order (quality here is educational attainment) (p. 670). Mogstad and Wiswall (2016), argue that Black et al. (2005), fail to detect size effects because these are non-linear, with increases in size benefiting children going from 1 to 2 children, and hurting them only at larger family sizes. Our sample is well placed in terms, of the size range, to test for such non-linear effects.

In summary, there is often a raw negative correlation in modern populations between child numbers and various measures of child quality. However, once controls to deal with the endogeneity of child quality and quantity are included, studies often find weak or non-existent quality-quantity relationships. The reported effects are often a very low levels of statistical significance. A substantial quality-quantity tradeoff, so vital to most theoretical accounts of modern economic growth is, at best, unproven.

There has been a group of recent economic history papers testing for a quality-quantity trade-off using pre-1914 data (Becker et al. (2010); Lynch (2016); Diebolt et al. (2017); Fernihough (2017); Hansen et al. (2017); Shiue (2017); Klemp and Weisdorf (2018)). Whilst they all show a statistically significant relationship between child quality and quantity, the economic size of the effects is often tiny, and quite consistent with our estimates below. For example, Fernihough (2017) uses the Irish 1911 census to test whether a child recorded as being in school has fewer siblings, as predicted by QQ models. Though the paper claims substantial effects, a variety of techniques have to be used to control for the endogeneity of family size in this period, and the average of the estimates are that a child in school has .12 fewer surviving siblings than one not in school. This translates into a 2% decline in the probability that a child attends school for each additional sibling. Such modest effects are completely consistent with our results below.

In the study closest in method and population to this, Klemp and Weisdorf (2018) use parish records in England 1540-1837 (generated originally by the Cambridge Group in History and Population) to estimate the effects of family size measured by siblings aged 5+ on adult literacy and occupational status. They report substantial effects of total births on literacy and occupational status. But their sample sizes are very small: 1,248 children for literacy and 652-686 for occupational status. This, combined with instrumenting for sibling size using the time to first birth, means the standard errors of their estimates are large. No effect, a coefficient of 0, is just outside the 1% confidence interval for both estimates. Thus for the fraction of men holding a skilled occupation, Klemp and Weisdorf (2018) estimate that this declines by 0.079 with an additional surviving sibling, but with a standard error of 0.028 (table 6, 652 observations). For occupational status we have about 12 times as much data. We can thus reproduce the Klemp and Weisdorf (2018) instrumented estimates, and compare those to our preferred OLS estimates. Our estimated coefficient on births instrumenting on the first birth interval is a decline of proportion in skilled occupations of 0.008 per additional sibling, with a standard error of 0.008 (6,901 observations). If we estimate the effects using OLS the coefficient estimate is a decline of 0.009 with standard error of 0.003. These estimates are are within the Klemp and Weisdorf (2018) 1% confidence intervals. Thus the results in this paper are perfectly consistent with Klemp and Weisdorf (2018), but show much smaller and more precisely estimated effects. And instrumenting in the way they suggest has no effect on the coefficients (see section B in the appendix to this paper).

2 Data

The data used in this study for marriages pre 1880 comes from a genealogical database of 314,363 English and Welsh people who had rare surnames, born 1750-2016. Rare surnames were used since then it is much easier to link people across generations. Since the initial data was collected to study social mobility in England from 1800 to 2012, the surnames used deliberately oversampled from the top and bottom of the wealth distribution for those dying 1858-1887[5]. We thus also estimate separately the effects of family size for the rich versus the poor/average families, in case the quality-quantity effect only appears in part of the educational or income distribution.

All births, deaths and marriages were registered in England from 1837 on. After 1865 the death register includes age at death. So for rare surname individuals we can link their births, deaths and marriages (though less easily for births before 1865). The censuses of 1841-1911, and a 1939 population register, are also available, providing information on parentage (see the list of data sources below). For marriages before 1880 there is considerable information available from parish records of baptisms, which recorded parents' names, and from parish records of marriages, which recorded the names and ages of those marrying as well as their fathers' names. There are many ancillary records which show, particularly for higher status families, family relationships: accounts, for example, of all men matriculating at Oxford and Cambridge universities prior to 1893, their fathers and their marriages, and also probate records.

By focusing on rare surnames, and by employing the whole set of records available for England

[5]See, for example, Clark and Cummins (2015). There are 49,387 individuals from the rich lineages (more than 2.5 times average wealth), 25,728 from the poor (no wealth of any kind), and 239,208 of intermediate wealth.

we achieve much higher matching rates than is typical for linking parents and children in 19th century censuses[6]. But the nature of the sources means we cannot identify parentage for all the people in our sample. Thus for 7,626 recorded rare surname births 1860-1879, we identify a father or mother for 88%[7]. The reasons for failing to find at least one parent in the other 12% of cases are various. In some cases the name likely was misspelled in the birth record, and the person does not belong in the surname lineages used to form the sample. Of those not linked 60% show no further appearance in any record after their birth under the birth name. Likely in most of these cases the name is just misspelled on the birth register. In others the child dies before appearing in a census, or their father dies, or they are living with grandparents in the census, or the family emigrates[8]. Thus one third of those born not linked to a parent died before age 10. However, for children identified as living to at least 21, 4,843 births 1860-79, the match rate is much better, with only 1.8% without at least one parent identified. In part for this reason our preferred measure of family size is the number of children living to age 21[9]. There will be error associated with this measure, but that error will be modest.

Though the numbers of recorded births for men and women is similar, and the match rate to fathers for the births is also similar by gender, the final dataset of family size by father is missing at least 12-14% of girls. This is because children in families can also be identified from the existence of a death record, or from their presence in a census or other record, where the birth was not recorded under the correct family surname. But adult women will only appear in a death or census record if they remain unmarried. Thus more sons are identified from such records, absent the birth record. Table A.1 shows for men and women of the target rare surnames the numbers linked to fathers in total and by gender and type for births 1860-79, for all births and for those attaining age 21. Though an equivalent number of women are matched to fathers in the births sample, many more men are identified from ancillary records. This implies that at least 12% of girls are missing from the sample of births, and 14% from the sample of those attaining age 21.

The evidence, however, is that once we account for omitted daughters, we are capturing most children in these families. Using the dataset we can estimate female fertility rates by age. These fertility rates can then be compared with those calculated by Wrigley and Schofield for England and Wales as a whole from parish records pre 1837, as is done in figure A.1 (reported in the appendix). This comparison suggests that the reconstructed families in this dataset are potentially missing 3% of male births, and 17% of female births. But for children reaching age 21 the percentages of sons and daughters missing will be smaller.

For children reaching age 21 where at least 12% of daughters are missing a factor that limits the error in the data is that a significant number of these missing daughters appear to be in daughter only families, where all the children are missing, so that they not appear at all in our estimations. To see this consider table A.2. This shows by family size the number of sons and daughters recorded. The share of women missing from smaller recorded families is much larger. A part of this will be just a statistical effect (missing women make families on average smaller), but a substantial part seems to be that there are significant numbers of missing all-female families of size 1, 2, or 3. Such omissions will not affect the estimated family size effects in the paper.

We have five measures of quality for children. Table 2.1 (a) and (b) shows the numbers of

[6] Ferrie and Long (2013), for example, link only 20% of adult sons to their fathers in England between 1851 and 1881.

[7] In some cases, where the child is illegitimate, only the mother is listed on birth records.

[8] We could identify the father by getting the birth certificate, but this is prohibitively costly.

[9] For children identified as dying before age 21 the numbers not matched with a father is 41%.

observations for each of these quality measures for fathers, and for children.

Table 2.1: Summary Statistics, Marriages 1780-1879

Statistic	N	Mean	St. Dev.	Min	Max
Fathers					
Age at Death	10,950	63.29	17.44	0.00	102.71
Ln(Real Wealth)	5,651	−1.42	3.70	−9.21	8.96
Marriage Year (1st)	10,383	1,847.29	24.83	1,779.60	1,945.00
Occupational Status	7,580	31.86	24.90	.08	94.12
Higher Education	8,329	.15	.36	0	1
Births (N0)	10,046	5.90	3.18	1	18
Adult Children (N21)	9,100	4.09	2.62	1	18
Number of Wives	10,665	1.13	.38	1	4
Children					
Female	58,974	.48	.50	0	1
Age at Death	36,392	52.28	29.98	0.00	115.53
Died Over 21	45,208	.82	.39	0	1
Ln(Real Wealth)	14,110	−1.28	3.08	−9.21	8.96
DSchool 11-20	13,216	.33	.47	0.00	1.00
Occupational Status	15,982	27.27	21.42	.08	94.12
Higher Education	17,697	.18	10.88	0	1,447

Schooling 11-20 - For a subset of all children, male and female, we have a measure of whether they were explicitly in school or in an apprenticeship aged 11-20 then they appear in a census 1851-1911 at these ages[10]. We also have a measure of whether they were explicitly in employment (exclusive of apprenticeships) ages 11-20 for these cases.

Higher Educational Attainment - For sons only we can construct an indicator variable for higher educational attainment. This is set at 1 under the following: the son enrolled at a university (Oxford, Cambridge, or London)[11]; enrollment at the Army Officer training school at Sandhurst; training as an attorney (1756-1874); enrollment as a registered doctor (1859-1956); was a member of an engineering society (Civil Engineers, 1818-1930, Mechanical Engineers, 1847-1930, Electrical Engineers, 1871-1930); was a trained cleric.

Occupational status - Occupations are given in the censuses of 1841-1911 as well as the population register of 1939. There are also occupation statements in some marriage registers for both grooms and the fathers of the marriage parties, for fathers in birth registers, for the deceased in death registers, and also in some years for the

[10] In the census some children have their occupational status just left blank.
[11] This measure looks only at those probated. But it does provide a ranking of occupations by wealth.

deceased or for executors in probate records. We translated these various occupational statements into 242 occupational categories – carpenter, laborer, solicitor, dealer, stockbroker etc. We gave these occupations a social status score between 0 and 100. That score was created as an equally weighted average of three elements: average normalized *ln* wealth at death by occupation, average fraction of people in each occupation with a university degree or equivalent, and average fraction of males in each occupation who were in school or in training when observed ages 11-20 in the censuses of 1811-1911, and the population register of 1939.

Wealth at death - For all children dying 1858 and later we have whether they were probated or not, and estimated wealth at death for the probated and non-probated. We normalize for changes in wealth over time by dividing wealth by the average wealth at death of the entire population for the decade of death. Again we use in the estimations the natural log of this real wealth measure (to have an outcome variable that is closer to normal in distribution).

Child Survival Rates and Adult Life Span - For all children we have measures of mortality rates (0-21), and adult longevity. In this period social status was strongly associated with infant and child mortality. It was more weakly associated with adult mortality. Table 4.3 shows child survival rates to 21 and life expectancy at age 21 by rare surname groups. Survival rate 0-21 is the fraction of those at born known to live to at least age 21. e21 is expected further years of life at age 21.

Given their educational status, longevity and wealth did parents with more children produce children who were of lower "quality" on the above five dimensions in terms of human capital?

Family size in this period is measurable in at least three different ways. First is the number of children born per father ($N0$). But a child who dies immediately after birth, as would most of the children dying in childhood in this period, makes few claims on parental time and attention. So another measure is children surviving to age 21 ($N21$). This, as noted above, has the advantage of also being measured with the least amount of error as any of the family size measures. Since children die at all ages from 0 to 14, when they can typically begin to support themselves, another measure is the number of child-years per father aged 0-14. For children dying ages 0-14, the child years is the age at death. For those dying 15+ it is 14. We normalize this variable, $N14$, by dividing by 14. It is thus the number of age 14 equivalent children a father has. $N14$ turns out to be typically a weighted average of $N0$ and $N21$, so we report most results just for $N0$ and $N21$.

Because we trace families using rare surnames, which are passed on from father to child, we measure family size as the numbers of children per father. In some cases the fathers will have these children with multiple mothers where the first wife dies early. One factor that will cause smaller family sizes would be the death of the father at an early age. Such variation we do not want to include since it will be associated with adverse childhood circumstances for the children. Thus we limit the sample to fathers who died aged 50 and above, by which time their first wife would normally have ceased to be fertile. This will exclude most variation in family size caused by adverse circumstances. We do not impose the same requirement of survival to age 50 and above on wives, since men whose wives died young often remarried quickly, reducing any adverse effects of death on children.

3 Fertility in England, 1780-1879: A Natural Experiment

Strong evidence of the absense of any fertility control within marriage in England, for marriages 1780-1879, comes from the accident of twin births. Consider a population where there is no control of fertility within marriage. In this case, whenever and however, the marriage terminates, the expected number of births will be increased by 1 by a twin birth, assuming the twin birth has no effect on the length of the subsequent birth interval. Also the increase in the final number of births will be the same whatever is parity at the time of the twin birth. With control, and a target family size for each family, in contrast the increase in births will be less than 1, as families reduce subsequent births to achieve their target family size. Taking account of the higher infant mortality rate of twin as opposed to singleton births, with control a twin in England before 1879 on average would induce only 0.43 additional births (Clark et al. (2019)).

To estimate the effects of a twin birth on total births, we need to control for the parity at which the birth occurs, as well as the age of the mother at the birth, since both higher parity and mother age increase the frequency of twinning. We thus estimate the twin effect on total births for English marriages 1780-1879 in the FOE database using the following equation

$$N_{pk} = \alpha_b DTWIN + \sum \beta_j DPARITY_j + \sum \delta_l DMAGE_l + \varepsilon \tag{1}$$

where N_{pk} is the total number of births in a family where a birth is observed at parity p and mother age k, $DTWIN$ is an indicator for that birth being a twin, the $DPARITY_j$ are indicators which are 1 at parity p, 0 otherwise, and the $MAGE_l$ are indicators which are 1 at the mother's age k, 0 otherwise. The estimated value of the coefficient on $DTWIN$ for marriages 1780-1879, using 54,362 births with complete family size known, 447 of which were twin deliveries, is 0.973, with a standard error of the estimate of 0.109. Had fertility been controlled, with families having a target family size, the effect of twins on births would have been an increase of only 0.44 .[12] There was also a substantial increase of .679 (standard error 0.127) in the total number of children surviving to age 14 when a twin birth occurred. Again the expected rise with families controlling fertility would be 0.07 (Clark et al. (2019)). Thus the evidence is that few, and most likely none, of these families were exercising fertility control. Figure **3.1** also shows that the increase in births with twinning was the same whatever the parity of the twin birth. There was no sign of any behavioral response to twinning across the whole parity range. Thus we can convincingly demonstrate that the recent claim of Cinnerella et al. of substantial parity-dependent fertility control in England before 1850 is mistaken (Cinnirella et al. (2017), Clark and Cummins (2019), Clark et al. (2019)). The conventional literature on fertility in England before 1880, which finds no sign of fertility control within marriage, is correct (see for example, Wilson and Woods (1991), and Woods (2000)).

This raised the possibility that we can use twinning as an instrument for family size in estimating the effect of size on child outcomes. However, since the amount of total size variance induced by twinning is small relative to total size variation, the estimates from twinning while consistent with the estimates below, and while insignificantly different from 0, have a standard error that is about twenty times as large, and so are very imprecise.

Fertility was not chosen. But it still might correlate with measured and unmeasured social characteristics so that it would bias an estimate of the effects of family size on outcomes. The bias in estimating β in the equation

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

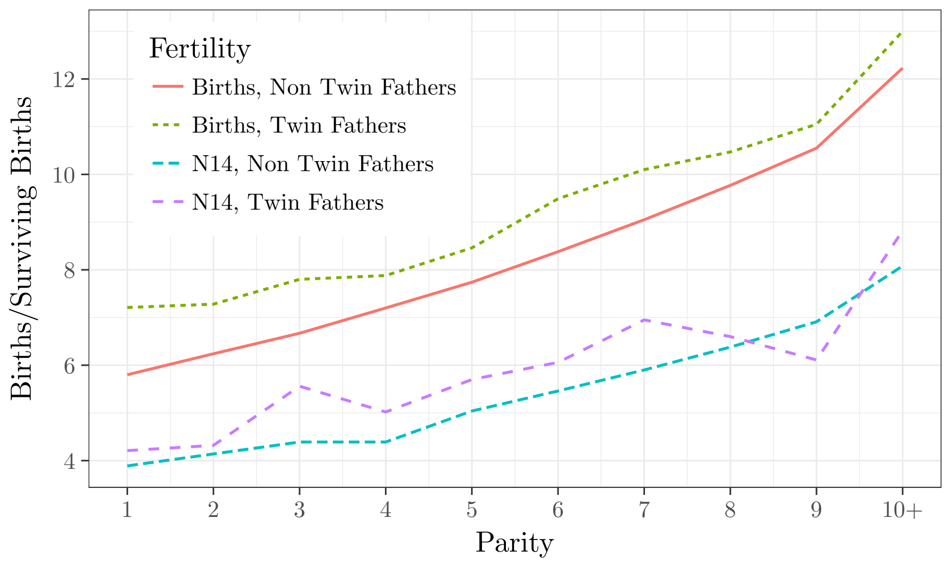


Figure 3.1: Observed Births by Parity at Twin Birth

$$q = \beta N + u \tag{2}$$

by OLS, is the ratio of the covariance of N and u, relative to the variance of N.

$$E(\hat{\beta}) = \beta + \frac{cov(N,u)}{var(N)} \tag{3}$$

However, in terms of measured characteristics, fertility for marriages 1780-1880 was not correlated with measured social status. Before 1880 the one clear element in determining family size that parents chose was age at marriage. If we regress for our sample children per father, measured as either gross fertility ($N0$) or children attaining age 21, net fertility ($N21$), on various quality measures, as well as ages at marriage, as in table 3.1, younger husbands and wives did produce more children. But even the one element of control, age at marriage, was uncorrelated with observed quality. The age of the wife at the husband's first marriage was not significantly correlated with the husband's wealth at death, occupational status, or educational status. Also, the amount of the variance in fertility determined even by age at marriage before 1880 was tiny, less than 1%. Only after 1880 do we see in table 3.1 a strong negative association between gross and net fertility and quality. So in terms of measured social characteristics there will be no bias in estimating the effects of size on social outcomes from observed family size variation. The variation here is orthogonal to social status.

What about correlation between N and unmeasured aspects of quality? These unmeasured aspects, however, would have to be unmeasured aspects of quality that were not correlated with measured aspects such as occupational status or wealth. For we know the measured aspects of status are uncorrelated with fertility. Thus characteristics such as thrift, industry, and low rates

Table 3.1: Determinants of Children per father, marriages 1780-1879, and 1880-1929

	Births, (N0)	N21	Births, (N0)	N21
	Marriages Pre 1880		Marriages 1880-1929	
Ln(Wealth)	−.001	.01	−.06***	−.04***
	(.004)	(.004)	(.01)	(.01)
Occupational Status	−.001	.001	−.004***	−.004***
	(.001)	(.001)	(.001)	(.001)
Higher Education	.003	−.01	.32***	.29***
	(.04)	(.05)	(.05)	(.05)
Number of Wives	.14***	.07**	.17***	.14***
	(.03)	(.04)	(.04)	(.04)
Age at Marriage (husband)	−.01***	−.01***	−.01***	−.01***
	(.002)	(.002)	(.002)	(.002)
Age at Marriage (first wife)	−.03***	−.03***	−.03***	−.03***
	(.002)	(.002)	(.003)	(.002)
Observations	2,538	2,491	4,387	4,325
Note:	*p<0.1; **p<0.05; ***p<0.01 Negative Binomial Regression Fathers die over 50			

of time preference would also correlate with wealth at death, education, and occupational status. But if the unmeasured aspects of quality that help determine fertility are not correlated with the measured aspects of quality, then in equation (2) they are not part of the measurement error u. The absense of any measured quality-quantity correlation at the parent level is what makes marriages in England in the period 1780-1880 a nice natural experiment with which to determine the causal effects of family size on child quality.

Another attractive aspect of this period for measuring the effects of family size on quality is that as figure A.4 illustrates, the range in family sizes for marriages 1780-1879 was enormous. This is further shown in figure A.6, which shows the distribution of children 21+ by family size[13]. The median adult child in this period had 5 adult siblings. Sibship sizes in nineteenth century England at the time of the Industrial Revolution were thus among the largest observed across all societies with well recorded demography. As a reflection of the huge variance in average family size in this period. 10% of children were in families of 2 or less, 13% of children in families of 9 or more.

4 Results

4.1 Family Size and Human Capital

We have two measures of human capital, educational status when observed ages 11-20, and a measure of higher educational attainment in adulthood as discussed above. The first measure indicates educational status well across the whole social spectrum, and for both genders. The last measure is a better proxy for educational success for higher status families, but only for men. Thus among the rich surname lineages 39% of men born before 1850 who lived to age 21 attended Oxford or Cambridge, or attained some other higher educational qualification. But for the poor group they are not such a good measure. Only 1.4% of men reaching age 21 in the average and poor surname lineages born before 1850, for example, attained a higher educational qualification.

The census reports 1841-1911 give occupations at all ages, including "scholar," "pupil" or "apprentice." Thus for each child we can potentially observe what their occupation was at some time from age 11-20, and whether they were acquiring education or training. In the period we are considering there was no compulsory schooling age until 1880. From 1880-93 the school leaving age was 10, 1893-1899 11, and 1899-1918 12. Only in 1891 did primary education become free for all students at state schools. 85% of our sample of children reached age 11 in the period where there were still school fees, and where education after age ten was voluntary. This was a period with costs for education, and where beyond age ten education was entirely at parental discretion. For our sample families, 41% of children 11-20 are explicitly at school or in training. There are other children in these age ranges where no occupation is indicated, where some will be in education. So the schooling variable is actually a lower bound estimate of those in education.

For the indictor variable $Dsch$, 1 if the child is in education or training, we estimate the effect of family size separately for sons and daughters of rich and average/poor family lineages. In the estimation we control for the census year, the age of the child, whether their father was dead, the wealth of their father, the occupational status of their father, and whether the father is had some

[13] Significant numbers of men and women had more than one marital partner in the course of their lifetime because of the early death of a spouse. We take family size throughout as the total number of children per father. We do this because earlier and later husbands of wives in our sample often had common surnames, making their children difficult to identify in the various records.

kind of higher educational attainment. We used the two measures of family size discussed above: $N0$ and $N21$. These two have a correlation of 0.88, so there is meaningful independent variation.

Figure 4.1 (a) present the raw averages of probability a child is observed in education aged 11-20 by numbers of children surviving to age 21, grouped into sizes 1-2, 3-4, 5-6, 7-8, 9-10, and 11+. For the families in the rich lineages family size has no apparent effect on education probabilities for either boys or girls. For average/poorer lineage families there are signs of a quality-quantity tradeoff. Children in average or poor lineages were less likely to be observed in school where family size was larger. The basic regression we estimate for $Dsch$ is

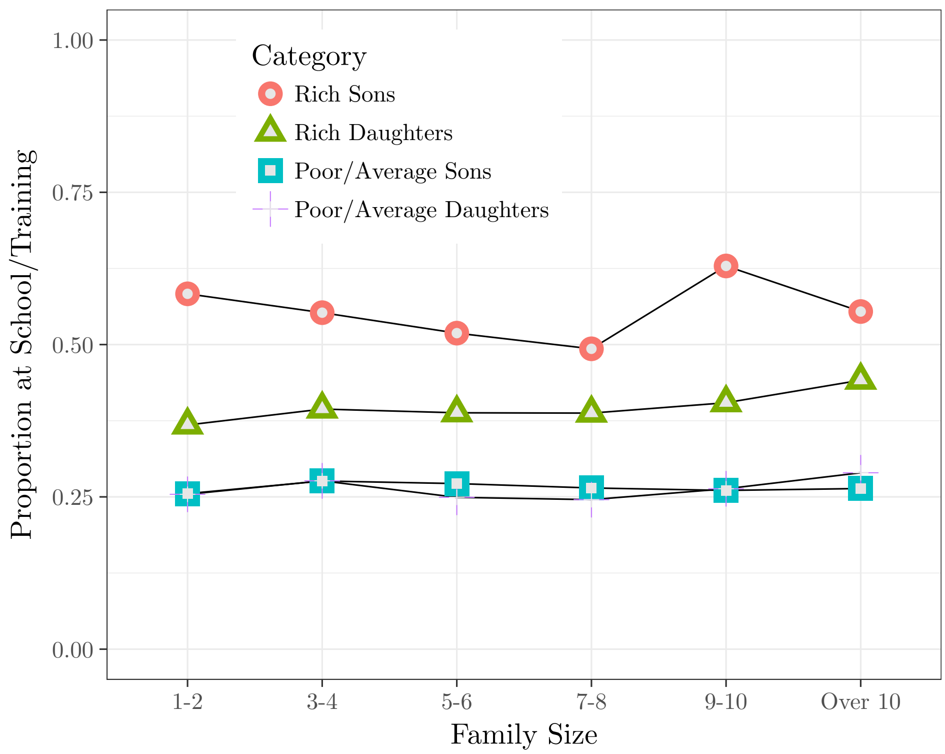
$$Dsch = \alpha + \sum b_i X_{if} + c_1 N + c_2 Brank \tag{4}$$

where N is one of the three child measures, and Brank is a child's birth order divided by the number of births in the family plus 1. Brank thus constructed is thus independent of family size, and any observed effects are orthogonal to those of family size. X is a set of characteristics of fathers: educational attainment, log of wealth at death, occupational status, and whether the father was alive at the time the child was observed. The key parameter of concern here is c_1, but the value of c_2 is also interesting. On a theory where parental inputs matter to success, the oldest child would be expected to receive more such inputs than later children, and to have better outcomes in terms of delaying entry to work and of continuing education.

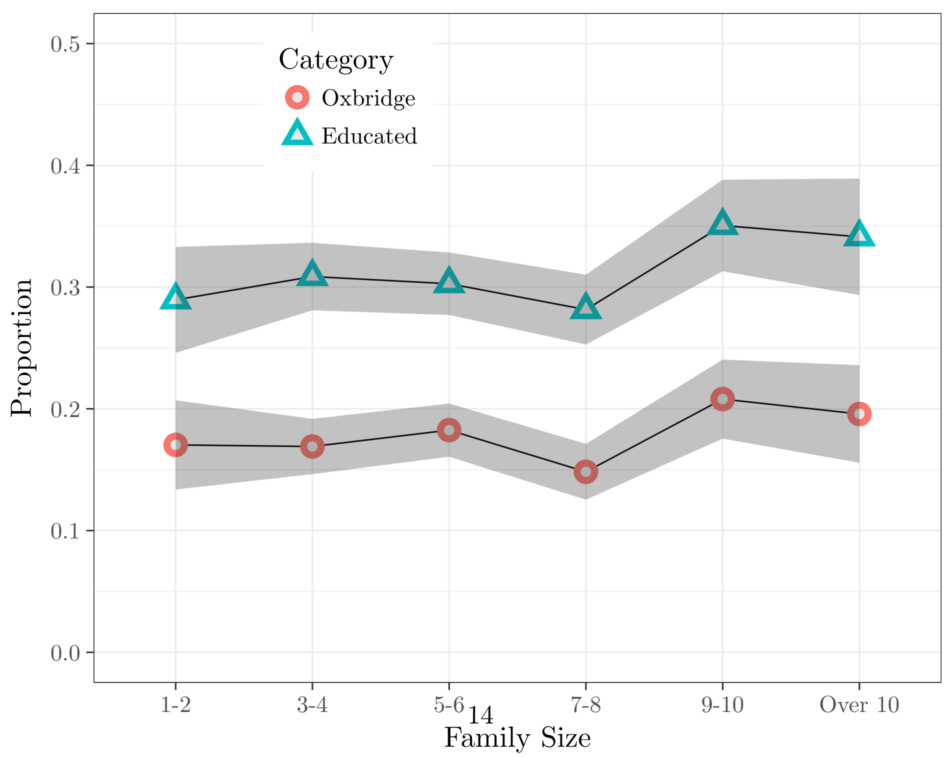
Table 4.1 reports the results of this estimation, using OLS, looking at the sons in the average/poor surname lineages. Since the dependent variable has a value of 0 or 1, statistically the correct estimation would be a logit. But the OLS results are always very similar and much easier to interpret. Table 4.1 also reports the same results for sons from the rich lineages. Note first that there is a strong connection between economic and occupational status and the probability that a son is observed at school or in training for both sets of families. The higher is status the greater the probability that a son is getting formal education. For the average/poorer families there is a significantly lower chance a son is observed at school 11-20 the larger the family size on either measure $N0$ or $N21$. This is the strongest human capital effect we observe across all groups and outcomes. However, the quantitative magnitude of the effect is still modest.

For $N21$ Table 4.1 implies that the elasticity of schooling for sons from average or poor family lineages with respect to family size was -0.13 (see table 5.1 for the summary elasticities). That implies that a family with 10 adult children invested in 6 times as much total education as a family of 1. The total supply of education by parents was very elastic with family size. These results also imply that declining family sizes on their own will explain little of the secular increase in years of education per child since the Industrial Revolution. The decline in size from an average of 5.3 adult children per family to 2 children per family would explain an increase of about 20% in the level of education per child for the average/poor families. But even at a family size of 2 these children would receive significantly less education than the children in the rich families of sizes 5-6 and above. Even with completed family size of 1-2 more two thirds of the children in average/poor families left formal education before were 16 prior to WWI. The class differences in education per child for rich and poor families for marriages 1780-1879 is also completely unexplained by any consideration of family sizes. It was feasible for poorer fathers with smaller family sizes to supply as much education as children received in richer families. After all, the fathers average/poor families in the cases where there were more children did so. The families with smaller family sizes simply chose not to.

Note also that birth rank is never significant, and the estimated coefficient is always positive. Children lower in rank were if anything advantaged in terms of education, despite the older children



(a) Percent at school or apprenticeship aged 11-20, by gender and lineage wealth



(b) Fraction of sons with higher education attainment, rich lineages, marriages 1780-1879

Figure 4.1: Family Size and Human Capital

Notes: Families grouped by sizes 1-2, 3-4, 5-6, 7-8, 9-10, and 11+. 95% Confidence Intervals indicated by shading in panel (b) in their earlier years facing less competition for attention from parents. Where birth order effects are found in other studies, they tend to disfavor higher order children. So the results here are counter to the typical pattern. For daughters in average/poor families the effects of family size are even weaker than for sons. For daughters from rich families there is no cost in terms of education from larger family size.

Figure 4.1 (b) summarizes the data, with no controls, of higher educational attainment by sons (Oxford, Cambridge or London University attendance, Sandhurst attendance, medical qualification, engineering qualification, legal qualification, or clergy training) by family size for families in the richer lineages. The figure shows also Oxford and Cambridge attendance only. The pattern is very clear that higher educational attainment was independent of family size. Table A.4 reports the effects of family size for families from the richer lineages, controlling for co-variates, again by OLS regressions (with logit results in the appendix). Many of the family background controls are significant, but they do not change the effects of size in any significant way. Both with and without the control variables, neither $N21$ nor $N0$ has any significant correlation with higher educational attainment for the sons in the rich lineages. Birth rank is again not significantly predictive. Table A.3 reports the estimated coefficients, by OLS, controlling for co-variates, of higher educational attainment by sons from average/poor families. As noted the probability of such attainment for average/poorer families averaged only 1.4%. But family size on either measure has no discernible impact on this probability, even though father's higher educational attainment and occupational status are significantly predictive. Thus there is no sign for any group that family size had significant effects on higher educational attainment.

4.2 Family Size and Occupational Status

Figure A.7 shows the raw connection between family size and occupational status, for rich and average/poor lineages where family size is measured as $N21$ in bins of 1-2, 3-4, 5-6, 7-8, 9-10, and 11+. The raw data suggests no connection between size and occupational status for either richer or average/poorer families.

The basic regression we estimate for occupational status is, as for education,

$$S_s = \alpha + \sum b_i X_{if} + c_1 N + c_2 Brank \tag{5}$$

where S_s is the occupational status of sons. Again we do the estimates separately for the rich and the average/poor cohorts. As a control we also include the age at which occupational status was measured, which enters strongly positively. Table 4.2 show the estimated effects of $N21$ and $N0$ on occupational status for sons in the rich and the average/poor lineages . For the average/poor lineages with controls there is no statistically significant effect of either family size measure on occupational status. The point estimates on $N0$ and $N21$ are negative, but the size of the effect is very modest with an implied elasticity at the mean of -0.04 to -0.05. Thus while sons from average/poor lineages get less schooling in larger families, that does not translate into any substantial loss in occupational status as adults. For the rich lineages, once we add controls, statistically significant negative family size effects appear. The magnitude of these size effects is still small however, an elasticity on both $N0$ and $N21$ at mean family size of -0.11. For both rich and average/poor lineages there is no significant effect of birth rank.

Table 4.1: Family Size and Probability at School 11-20, Marriages 1780-1879, Sons

	Probability at School 11-20			
	Poor/Average		Rich	
	(1)	(2)	(3)	(4)
N0, Father	$-.006^{***}$		$-.0002$	
	(.002)		(.003)	
N21, Father		$-.005^{*}$		$-.0002$
		(.002)		(.004)
Birth Rank	$-.010$	$-.009$	$-.093^{**}$	$-.093^{**}$
	(.024)	(.024)	(.043)	(.043)
Ln(Wealth), Father	$.018^{***}$	$.018^{***}$	$.015^{***}$	$.015^{***}$
	(.004)	(.004)	(.004)	(.004)
Occupational Status	$.004^{***}$	$.004^{***}$	$.004^{***}$	$.004^{***}$
	(.001)	(.001)	(.001)	(.001)
Higher Education, Father	.087	.084	$-.044$	$-.044$
	(.074)	(.075)	(.040)	(.040)
Father Dead	.017	.013	$-.037$	$-.037$
	(.027)	(.027)	(.039)	(.038)
Census Year	$-.001^{**}$	$-.001^{**}$	$-.002^{**}$	$-.002^{**}$
	(.0004)	(.0004)	(.001)	(.001)
Age Observed	$-.080^{***}$	$-.080^{***}$	$-.046^{***}$	$-.046^{***}$
	(.002)	(.002)	(.014)	(.014)
Female	$-.006$	$-.007$	$-.143^{***}$	$-.143^{***}$
	(.011)	(.011)	(.020)	(.020)
Observations	5,416	5,416	2,702	2,702
R^2	.280	.280	.210	.210
Note:	*p<0.1; **p<0.05; ***p<0.01			
	Ordinary Least Squares Regression			
	Errors Clustered on Fathers			

Table 4.2: Family Size and Occupation Status, sons, marriages 1780-1879

	Occupational Status			
	Poor/Average		Rich	
	(1)	(2)	(3)	(4)
N0	$-.095$		$-.585^{***}$	
	(.134)		(.134)	
N21		$-.059$		$-.579^{***}$
		(.155)		(.155)
Birth Rank	1.402	1.411	-1.044	-1.015
	(1.446)	(1.451)	(1.446)	(1.451)
Higher Education, Father	9.348^{***}	9.330^{***}	.600	.618
	(1.549)	(1.554)	(1.549)	(1.554)
Ln(Wealth), Father	1.106^{***}	1.108^{***}	1.970^{***}	1.987^{***}
	(.173)	(.175)	(.173)	(.175)
Age Observed	$.121^{***}$	$.121^{***}$	$.183^{***}$	$.183^{***}$
	(.030)	(.030)	(.030)	(.030)
Occupational Status, Father	$.418^{***}$	$.419^{***}$	$.425^{***}$	$.427^{***}$
	(.031)	(.031)	(.031)	(.031)
Observations	4,107	4,107	2,551	2,551
R^2	.325	.325	.507	.505

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$
Ordinary Least Squares Regression
Errors Clustered on Fathers

4.3 Family Size and Child Survival Rates

As shown in table 4.3, the fraction of children surviving to age 21 was strongly correlated with social status for marriages 1780-1879. 85% of children survived to age 21 in the richest lineages, compared to 73% in the poorest. In table 4.4 we estimate the average survival rate to 21 of grandchildren as a proxy for the adult economic and social circumstances of children from marriages 1780-1879. If larger family sizes reduced the abilities and human capital of children, it should show in terms of the survival rates of their children. Table 4.4 reports the results of this estimation, using OLS, looking at the sons in all surname lineages. Since the dependent variable has a value of 0 or 1, statistically the correct estimation would be a logit. But as before the OLS results are always very similar and much easier to interpret. Controlling for father social characteristics, and time trends, sons from larger families on average had better child survival rates. There is here no sign that they labored under any disadvantage from larger sibship sizes in terms of the health and welfare of their own children.

Table 4.3: Observed Survival Rates and Lineage Wealth Class, Marriages 1840-79

Group	N	Survival Rate, Children	Survival Rate, Grand-children	Death Age
Richest	5764	0.84	0.88	66.24
Rich	5359	0.80	0.84	65.80
Average	28108	0.62	0.75	65.45
Poorest	4024	0.73	0.81	65.66

4.4 Family Size and Child Wealth

We have estimates of wealth at death for all people dying 1858 and later. This comes from the Principal Probate Registry, and is from 1858-1893 a statement just of the personalty of the deceased (assets aside from real estate), and after 1894 a statement of all assets. For those not probated we have to attribute a probate value. [14]

In the years 1799-1857 there is more selective information on the value of personalty available for wills probated in the highest of the ecclesiastical probate courts, the Prerogative Courts of the Archbishop of Canterbury and of York. However, only about 5% of men were probated in these courts, and quite wealthy men might be probated elsewhere. Thus for this period we only included men as fathers in the wealth regression if they had a probate value in this court. Since this involves selection just on the Xs it should not lead to bias in the results.

[14] In each period after 1858 there was a minimum estate value at which probate was legally required: £10 (1858-1900), £50 (1901-1930), £50-500 (1931-1965), £500 (1965-1974), £1,500 (1975-1983), and £5,000 (1984-2012) (Turner (2010), 628). We thus 1858 and later took as the value of estate for those not probated as typically half the minimum requiring probate: £5 (1858-1900), £10 (1901-9), £15 (1910-9), £20 (1920-30), £25 (1931-9), £50 (1940-9), £100 (1950-9), £250 (1960-1974), £750 (1975-1983), and £2,500 (1984-2012). We did not increase the attributed value in 1901 to £25 because the rise in the probate limit to £50 in that year had little effect on the implied value of the omitted probates in 1901 compared to 1900. Thus whatever the exact cutoff the bulk of the omitted probates were closer to 0 in value than to £50.

Table 4.4: Child Survival and Father's Family Size, marriages 1780-1879

	Child Survival Rate			
N0, Father	−.001		−.001	
	(.001)		(.001)	
N21, Father		.01***		.002
		(.001)		(.001)
Birth Rank	.01	.02	.02*	.02*
	(.01)	(.01)	(.01)	(.01)
Ln(Wealth), Father			.01***	.01***
			(.001)	(.001)
Occupational Rank, Father			.001***	.001***
			(.0002)	(.0002)
Higher Education, Father			.001	.001
			(.01)	(.01)
Female			−.01	−.01
			(.08)	(.08)
Observations	9,118	9,117	4,680	4,680
R^2	.05	.06	.07	.07

Note: *p<0.1; **p<0.05; ***p<0.01
Ordinary Least Squares Regression
Errors Clustered on Fathers
Weighted by Child Births
Quadratic Time Trend included

Since the nominal value of average wealth increased greatly between 1858 and 2012 we normalized by the estimated average wealth at death in each decade. Also since wealth at death has a very skewed distribution, we use the logarithm of normalized estimated wealth to produce a distribution closer to normal. We thus construct for each person i dying in year t a measure of normalized wealth at death which is

$$lnW_{it}^{*} = lnW_{it} - ln\bar{W}_{it}$$

where $ln\bar{W}_{it}$ is the estimated average wealth at death by decade[15]. For each decade lnW_{it}^{*} will thus have an average expected value for the population as a whole of close to 0

We can thus estimate the effect of family size on wealth through

$$lnW_{c}^{*} = b_{0+}b_1 lnW_f^{*} + b_2 D_{falive} + b3ln(Age_D) + b_4 DFem + b_5 Brank + b_6 ln(N21) \tag{6}$$

where lnW_c^{*} is the natural logarithm of normalized wealth for each child of a given father, lnW_f^{*} is the natural logarithm of normalized wealth of the father. D_{falive} is an indicator for when the father is still alive at the time of the child's death, $Ln(Age_D)$ is the natural log of age at death, $Dfem =$ indicator of 1 for daughter and $Brank$ is birth rank as defined above.

$Dfalive$ is a control for the effects of children who die before fathers, and thus receive smaller wealth transfers from fathers. $Ln(Age)$ controls for the observed rise in wealth with age at death. With this formulation, b6 is the elasticity of son's wealth as a function of the number of surviving children the father left. The coefficient b1 shows the direct link between fathers' and sons' wealth, independent of the number of children.

Column (1) of table 4.5 shows the estimated coefficients from equation 6, for all children. The coefficient on $Ln(N21)$, is negative and strongly statistically significant. Family size could have an influence on wealth for two reasons. First inherited wealth has to be divided across more recipients. But secondly the quality of children could decline with family size in terms of their ability to earn and accumulate wealth. We can check which of these forces is at work here by dividing the sample of fathers into two groups. First there are fathers from rich rare surname lineages where significant wealth passes between generations, and the issue is how much of this is dissipated by the receiving generation. But there are also poor rare surname lineages where wealth is absent or inconsequential. Here differences in child wealth were the product of the earnings of children, and their saving behavior. In columns (2)-(4), we estimate the parameters of equation (3) separately for three groups: families from the rich lineages where children mostly inherit some wealth from parents or other relatives, families where the father specifically had some wealth at death, and families where the father had no wealth, as revealed by the father not being probated.

For the children of the wealthy lineages, and for children whose father was probated, the effects of family size are even stronger than those for children overall. However, when we look at children whose fathers were not probated we find that family size now has no significant effect on wealth. Where inheritance of wealth matters, as in the rich lineages, the size effect is important. Where only human capital matters, size no longer matters. Thus the data is consistent with family size mattering in this case not because of effects on human capital, but because of the consequences for the amounts inherited. Notice also that where there was no wealth to inherit the effects of the father still being alive at the time of the child's death also changes from substantial and negative to insignificant.

[15] This was estimated 1895 and later from aggregate probate values reported by Atkinson (2013). 1858-1894 this was estimated from a sample of probate values.

Table 4.5: Child Wealth and Family Size, marriages 1780-1879

	Ln(Wealth)			
	All	Rich	Probated (f)	Non Probated (f)
	(1)	(2)	(3)	(4)
Ln(N21), Father	$-.278^{***}$	$-.464^{***}$	$-.455^{***}$	$-.043$
	(.058)	(.089)	(.086)	(.074)
Birth Rank	$.286^{***}$	$-.034$	$-.193$	$.768^{***}$
	(.087)	(.148)	(.143)	(.095)
Age at Death, Child	$.033^{***}$	$.029^{***}$	$.030^{***}$	$.036^{***}$
	(.002)	(.003)	(.003)	(.002)
Ln(Wealth), Father	$.324^{***}$	$.328^{***}$	$.392^{***}$	
	(.016)	(.022)	(.036)	
Occupational Status, Father	$.031^{***}$	$.028^{***}$	$.028^{***}$	$.050^{***}$
	(.003)	(.004)	(.004)	(.005)
Higher Education, Father	$-.326^{**}$	$-.455^{***}$	$-.295^{*}$	$-.577$
	(.145)	(.157)	(.162)	(.429)
Female	$-.327^{***}$	$-.491^{***}$	$-.553^{***}$	$-.076$
	(.054)	(.082)	(.080)	(.065)
Father Alive	$-.447^{***}$	$-.861^{***}$	$-.934^{***}$	.008
	(.098)	(.171)	(.171)	(.100)
Observations	9,652	4,657	4,951	4,825
R^2	.461	.337	.310	.229
Note:	*p<0.1; **p<0.05; ***p<0.01			
	Ordinary Least Squares Regression			
	Errors Clustered on Fathers			

If we rewrite the relevant parts of equation 6 above as

$$lnW_c^* = b_{0+}b_1 ln(\frac{W_f^*}{N21}) + \ldots + (b_6 - b_1)ln(N21) \tag{7}$$

$\frac{W_f^*}{N21}$ is an estimate of the average amount of bequest a child received from the father, so that then b_1 is also the estimate of the effect of bequest size on child wealth, while (b_6- b_1) is the estimate of the residual effect of family size, through the human capital channel, on wealth retention or accumulation. For all children with probated fathers (b_6- b_1) is -0.063 with a standard error of .086. Thus the data for probated fathers is consistent with the idea that there is no significant effect of family size on the ability of children to retain inherited wealth or accumulate new wealth. Family size again played little role in the human capital or other characteristics of children. Thus family size for either rich and for average/poor families had no significant effect on wealth accumulation, net of what is predicted by the expected bequest.

The multigenerational nature of our data allows us to look at the longer run effects of family size on wealth. In table A.5 the five columns show estimates of the wealth of the grandchildren of men who first married 1780-1879, as a function of grandfather wealth, and family size, $N21$, in the grandfathers' generation, when family size was a random shock. The wealth of grandchildren is still strongly associated with the wealth of the grandfather. But the shock to wealth caused by the accident of family size in the grandfather generation becomes insignificantly different from zero even for families with significant bequests by the grandfather. The wealth shock from family size in the grandfather's generation dissipates quickly across three generations. It does not have any permanent effect on family status.

4.5 Non-Linear Family Size Effects?

In tables 4.1-4.5 we have considered the effects of family size on schooling, higher education, occupational status, child survival rates and wealth in a linear setting. As noted above Mogstad and Wiswall (2016), argue that Black et al. (2005), fail to detect size effects because these are non-linear, with increases in size benefiting children going from 1 to 2 children, and hurting them only at larger family sizes. Could it be that family size here has little overall effect because larger size benefited children at small family sizes, and hurt them more significantly at larger sizes?

We have enough data that we can estimate non-linear size effects by comparing outcomes for completed family sizes of 2, 3, ..., 10+ compared to a family size of 1. Do we see a pattern of gains for sizes 2-3, and then larger costs for sizes 4-10? Figures A.8 (a) and (b) summarize these results for six main outcomes for richer lineages and for average/poorer lineages, controlling for father characteristics. These outcomes are "at school 11-20, boys", "at school 11-20, girls," "higher education", "occupational status", "child survival, children" and "ln(wealth)". The figures show the estimated coefficient on each family size, relative to size 1, and the associated 95% confidence intervals.

As can be seen there is little sign of any substantial non-linearity in family size effects of the type proposed by Mogstad and Wiswall (2016) for either the average/poor lineages or the rich lineages.

5 Why is the Quality-Quantity Tradeoff so Weak?

Table 5.1 summarizes the estimates in tables 11-18 for rich and average/poor families for the six measures "at school 11-20, boys", "at school 11-20, girls," "higher education", "occupational status", “child survival, children”, “skilled occupation” and “ln(wealth)” as the estimated elasticity at mean family size, for both births and $N21$. The general pattern, except for wealth in richer lineages, is of very modest negative effects from family size. Across all seven measures the average elasticity for poor/average families is -0.06 (N0) and -0.005 (N21). Across all but wealth the average elasticity for rich families is -0.05 (N0) and -.04 (N21). At maximum adding another child reduced average child quality by 1%.

Table 5.1: Summary Elasticities, 1780-1879

Variable	Rich	.	Poor/Avg.	.
	N0	N21	N0	N21
At School, 11-20, Boys	−.01	.01	−.25***	−.13***
At School, 11-20, Girls	−.01	−.02	−.06	−.02
Higher Education	−.08***	−.06**	.15	.15
Occupational Status	−.11***	−.11***	−.05***	−.04***
Child Survival, Children	.01	.01	−.01	.02*
Skilled Occupation	−.10***	−.08***	−.04	.002
Wealth	−.52***	−.46***	−.17***	−.04

Note: *p<0.1; **p<0.05; ***p<0.01
Wealth elasticities for Poor/Avg. are for non-probated fathers

For richer families the absence of any substantial quality-quantity tradeoff in human capital is not surprising. In nineteenth century England the children in such families were raised and educated at home with the aid of nursemaids and governesses. The annual salary of a governess at £35-50 was modest relative to the annual income of such families. For married men dying in the wealthier lineages before 1914, for example, assets at death averaged more than £55,000. When older, sons often went to private boarding schools such as Eton, Harrow, Rugby and Winchester. The annual fee for a private boarding school circa 1861 would be £120-£250 per year, and day schools would be considerably cheaper (Turner (2015) p.249). Tuition at Oxford in the 1840s was only £16-£25, though many students paid £40-£50 in addition for private tutoring, again modest relative to the resources of even middle class families (Curthoys (1997) p.50-151).

The absence of significant quality-quantity effects for the average/poor lineage families is more surprising. In larger families parents would have less time to interact with young children with more children to care for. More children also implied less ability of women to contribute through market work to family incomes, and less food consumption, space and clothing for other family members. The costs of formal schooling, even before the era of free provision of compulsory education after 1891, were mainly the foregone wages of children. In the mid nineteenth century the actual weekly costs of school attendance were typically 6d or 12d a week. By age 13 or 14 a son or daughter could expect to earn 36-72d a week at work, depending on the nature of the local industries (Humphries (2010) p.230-33, 316-7). Apprenticeships, typically entered into around age 14, would provide maintenance for children. But these usually required premiums, though in lieu of that payment the length of apprenticeship could be increased, and the earnings above maintenance in later years reduced. Why didn’t these costs of schooling and training force families with larger numbers of children to set their children to unskilled labor at a young age, limiting their future occupational status and earnings?

We do see in table 4.1 that the boys from such families were less likely to be observed at school aged 11-20. But in general quantity has weak detrimental effects on quality even for poorer families – measured as either higher education attainment, occupational status, child survival rates, or even wealth at death. We find, for example, families such as that of Alfred Albert and Eliza Wimbleton who produced 9 children between 1866 and 1888, where Alfred was a general laborer, and Eliza had no occupation. In 1881 when 6 of those children were at home with the parents, aged 15 to 1 the oldest four, respectively 15, 13, 11, and 10 were all still at school. Large families, even at the bottom of the earnings ladder, could afford to keep children in school. In contrast Charles William Wressell, son of an agricultural laborer, in 1861 was at work age 12 also as a laborer, despite his parents having only one other dependent. Thus in nineteenth century England the choice whether to educate children or not was driven more by parental attitudes, and the abilities of children, rather than being dictated by any binding budget constraint.

One reason that there was no binding budget constraint on children's care, nutrition and education, even at family sizes larger than 10, was the typical spacing of surviving children at birth intervals of 2 or more years. The typical age of marriage was 25-27 in nineteenth century England. Before marriage older children would often contribute financially, and with child care, to their families. Thus the younger children in large families, those occurring at birth orders 7 and above, often got transfers of care and resources not just from their parents, but from siblings.

6 Implications

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

The evidence above shows clearly that the costs to families from having more children were, at their strongest, modest in terms of the human capital of the children. Sons in poorer lineages were more likely to be observed at work aged 11-20 if they came from a larger family. But this effect does not result in them having lower occupational status as adults, or lower lifespans, or less wealth at death. Otherwise the effects of family size on all human capital outcomes are marginal if any. Thus for the daughters of the poor, as well as all the children of the rich, family size had no effect on schooling 11-20. For all groups it had no affect on occupational status as an adult, on attaining higher education for sons. Size was significantly associated with less wealth at death. But given their estimated average inheritance, the wealth at death of children in larger families was not any less than in that of smaller families. Thus the children of larger families show no sign of being less capable or less educated. And even the effect of family size on child wealth was transitory. Grandchildren in families with larger size in the first generation are no poorer relative to their grandfather than grandchildren of smaller families in the first generation[16].

All of this calls into question the strong reliance of most theories of the emergence of modern economic growth on the quality-quantity tradeoff with children. The whole Beckerian apparatus finds no counterpart in reality. Modern growth consequently cannot be explained by a switch to

[16] There are studies for children born in the late nineteenth century in England or Ireland that do find a quality-quantity tradeoff. See Bailey et al. (2017), and Fernihough (2017). But this is for the period where fertility was now correlated with parent social status.

smaller family sizes accompanied by more investment in child quality. Modern growth in England began 100 years before there were significant reductions in average family sizes, and there is no sign that larger pre-industrial families involved a sacrifice of child quality.

References

Angrist, Joshua D., Victor Lavy, and Analia Schlosser, "Multiple Experiments for the Causal Link between the Quantity and Quality of Children," September 2006. MIT Department of Economics Working Paper No. 06-26.

Bailey, Martha, Connor Cole, Morgan Henderson, and Catherine Massey, "How Well Do Automated Methods Perform in Historical Samples? Evidence from New Ground Truth," Working Paper 24019, National Bureau of Economic Research November 2017.

Becker, Gary S., Kevin M. Murphy, and Robert Tamura, "Human Capital, Fertility, and Economic Growth," *The Journal of Political Economy*, 1990, *98* (5), S12–S37.

Becker, Sascha O., Francesco Cinnirella, and Ludger Woessmann, "The trade-off between fertility and education: evidence from before the demographic transition," *Journal of Economic Growth*, Sep 2010, *15* (3), 177–204.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education*," *The Quarterly Journal of Economics*, 2005, *120* (2), 669–700.

Cáceres-Delpiano, Julio, "The Impacts of Family Size on Investment in Child Quality," *The Journal of Human Resources*, 2006, *41* (4), 738–754.

Carmichael, Sarah Guilland, Auke Rijpma, and Lotte van der Vleuten, "Quantity versus Quality: Household structure, number of siblings, and educational attainment in the long nineteenth century," Working Papers 0073, Utrecht University, Centre for Global Economic History November 2015.

Cinnirella, Francesco, Marc Klemp, and Jacob Weisdorf, "Malthus in the Bedroom: Birth Spacing as Birth Control in Pre-Transition England," *Demography*, 2017, *54.2*, 413–436.

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

_ and _, "Randomness in the Bedroom: There is No Evidence for Fertility Control in Pre-Industrial England," *Demography (Forthcoming)*, 2019.

_ , _ , and Matthew Curtis, "Twins Reveal Absence of Fertility Control in Pre-Industrial Western European Populations," 2019.

Curthoys, M. C., "The unreformed colleges," in "History of the University of Oxford," Vol. VI, Oxford: Oxford University Press., 1997, pp. 146–173.

Desai, Sonalde, "When Are Children from Large Families Disadvantaged? Evidence from Cross-National Analyses," *Population Studies*, 1995, *49* (2), 195–210.

Diebolt, Claude, Audrey-Rose Menard, and Faustine Perrin, "Behind the fertility–education nexus: what triggered the French development process?," *European Review of Economic History*, 2017, *21* (4), 357–392.

Doepke, Matthias, "Gary Becker On The Quantity And Quality Of Children," *Journal of Demographic Economics*, 2015, *81* (1), 59–66.

Fernihough, Alan, "Human capital and the quantity–quality trade-off during the demographic transition," *Journal of Economic Growth*, Mar 2017, *22* (1), 35–65.

Ferrie, Joseph and Jason Long, "Intergenerational Occupational Mobility in Great Britain and the United States since 1850," *American Economic Review*, June 2013, *103* (4), 1109–37.

Galor, Oded., *Unified Growth Theory*, Princeton: Princeton University Press., 2011.

Galor, Oded and David N. Weil, "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and beyond," *The American Economic Review*, 2000, *90* (4), 806–828.

_ **and Omer Moav**, "Natural Selection and the Origin of Economic Growth," *The Quarterly Journal of Economics*, 2002, *117* (4), 1133–1191.

Grawe, Nathan D., "Testing Alternative Models of the Quality-quantity Trade-off in Fertility," 2003.

_ , "Bequest Receipt And Family Size Effects," *Economic Inquiry*, 2009, *48* (1), 156–162.

Hansen, Casper Worm, Peter Sandholt Jensen, and Lars Lï¿œnstrup, "The Fertility Decline In The United States: Schooling And Income," *Macroeconomic Dynamics*, 2017, pp. 1–29.

Humphries, Jane., *Childhood and Child Labour in the British Industrial Revolution*, Cambridge: Cambridge University Press, 2010.

Jensen, Robert, "Equal Treatment, unequal outcomes? Generating gender inequality through fertility behavior.," 2005.

Kaplan, Hillard S., Jane B. Lancaster, Sara E. Johnson, and John A. Bock, "Does observed fertility maximize fitness among New Mexican men?," *Human Nature*, Dec 1995, *6* (4), 325–360.

Klemp, Marc and Jacob Weisdorf, "Fecundity, Fertility and The Formation of Human Capital," *The Economic Journal*, 2018, *Forthcoming.*

Kugler, Adriana D. and Santosh Kumar, "Preference for Boys, Family Size, and Educational Attainment in India," *Demography*, Jun 2017, *54* (3), 835–859.

Lawson, David W. and Ruth Mace, "Optimizing Modern Family Size," *Human Nature*, Mar 2010, *21* (1), 39–61.

Lee, Jungmin, "Sibling size and investment in children's education: an asian instrument," *Journal of Population Economics*, Oct 2008, *21* (4), 855–875.

Li, Hongbin, Junsen Zhang, and Yi Zhu, "The Quantity-Quality Trade-Off of Children in a Developing Country: Identification Using Chinese Twins," *Demography*, 2008, *45* (1), 223–243.

Liu, Haoming, "The quality–quantity trade-off: evidence from the relaxation of China's one-child policy," *Journal of Population Economics*, Apr 2014, *27* (2), 565–602.

Lucas, Robert E, "The industrial revolution: Past and future," *Lectures on economic growth*, 2002, pp. 109–188.

Lynch, Robert Francis, "Parents Face Quantity–quality Trade–Offs between Reproduction and Investment in Offspring in Iceland," *Royal Society Open Science*, 2016, *3.5* (160087).

Mogstad, Magne and Matthew Wiswall, "Testing the quantity–quality model of fertility: Estimation using unrestricted family size models," *Quantitative Economics*, 2016, *7* (1), 157–192.

O'Rourke, Kevin Hjortshøj, Ahmed S. Rahman, and Alan M. Taylor, "Luddites, the industrial revolution, and the demographic transition," *Journal of Economic Growth*, Dec 2013, *18* (4), 373–409.

Ponczek, Vladimir and Andre Portela Souza, "New Evidence of the Causal Effect of Family Size on Child Quality in a Developing Country," *Journal of Human Resources*, 2012, *47* (1), 64–106.

Qian, Nancy, "Quantity-Quality and the One Child Policy:The Only-Child Disadvantage in School Enrollment in Rural China," Working Paper 14973, National Bureau of Economic Research May 2009.

Rosenzweig, Mark R. and Junsen Zhang, "Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China's "One-Child" Policy," *The Review of Economic Studies*, 2009, *76* (3), 1149–1174.

_ **and Kenneth I. Wolpin**, "Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models," *The Journal of Political Economy*, 1980, *88* (2), 328–348.

_ **and** _ , "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment," *Econometrica*, 1980, *48* (1), 227–240.

Shiue, Carol H., "Human capital and fertility in Chinese clans before modern growth," *Journal of Economic Growth*, Dec 2017, *22* (4), 351–396.

Turner, David., *The Old Boys: The Decline and Rise of the Public School*, New Haven: Yale University Press., 2015.

Turner, John D., "Wealth concentration in the European periphery: Ireland, 1858-2001," *Oxford Economic Papers*, 2010, *62* (4), 625–646.

Willis, Robert J., "A New Approach to the Economic Theory of Fertility Behavior," *The Journal of Political Economy*, 1973, *81* (2), S14–S64.

Wilson, Chris and Robert Woods, "Fertility in England: A Long-Term Perspective," *Population Studies*, 1991, *45* (3), pp. 399–415.

Woods, Robert, *The Demography of Victorian England and Wales* Cambridge Studies in Population, Economy and Society in Past Time, Cambridge University Press, 2000.

List of Tables

List of Figures

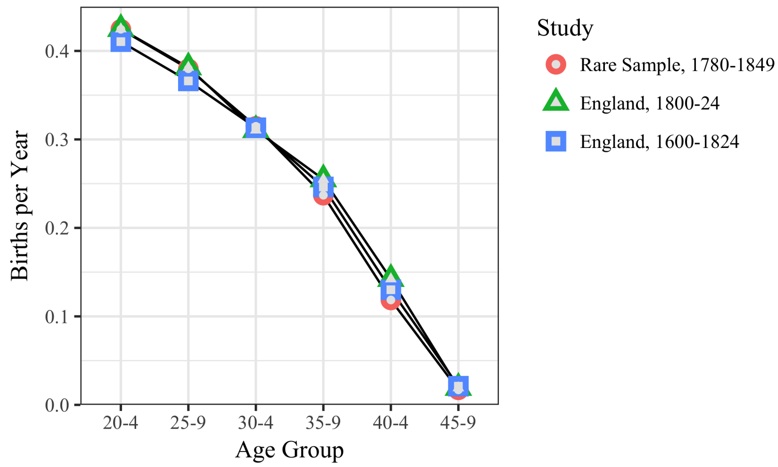


Figure A.1: Fertility by Wives' Age, all marriages 1780-1849

Notes: Because there are more missing female children, fertility rates are estimated from male births for our sample. These are multiplied by 1.953 to estimate total fertility rates, assuming the ratio of male to female births is 1.05 (UK Department of Health, 2013). Sources: England, 1800-24 and 1600-1824, Wrigley et al. 1997 p.355, table 7.1.

A Supplementary Results

Table A.1: Share of Men and Women in Family Size Sample, 1860-79

	All	Men	Women
Births - all	6,205	3,218	2,987
Births - Birth record	5,826	2,877	2,949
Births - no Birth record	379	341	38
21+ - all	4,788	2,529	2,259
21+ - birth record	4,455	2,226	2,229
21+ - no birth record	333	303	30

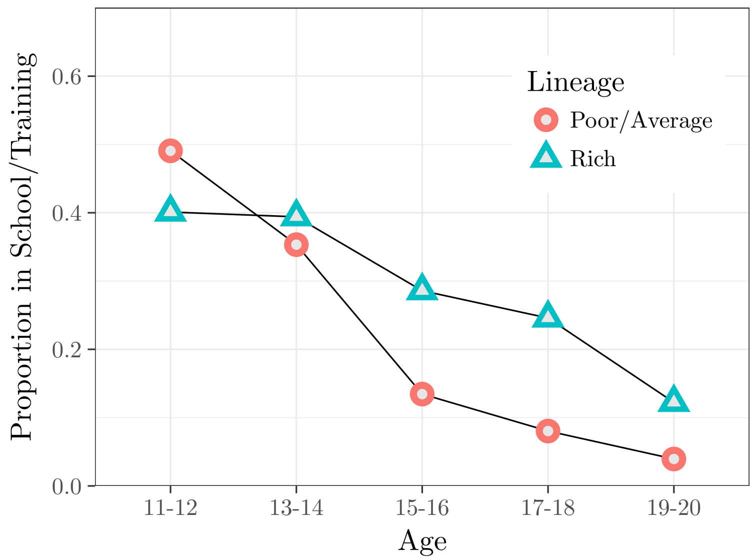


Figure A.2: Fraction of males in education or training, age 11-20, marriages pre 1880

Notes: See Data Description below. Status as reported in the censuses 1851-1911. As noted in the text this underestimates the fraction in schooling or training for both groups.

Table A.2: Missing Women by Family Size, pre-1880 marriages, children 21+

Family Size	All	All Children	Male	Female	% missing females
0	803	0	0	0	0
1	367	367	211	156	26.1
2	452	904	511	393	23.1
3	514	1,542	862	680	21.1
4 or 5	906	4,039	2168	1,871	13.7
6 or 7	554	3,560	1,876	1,684	10.2
8+	433	4,054	2,057	1,997	2.9
All	3,990	14,466	7,695	6,771	12

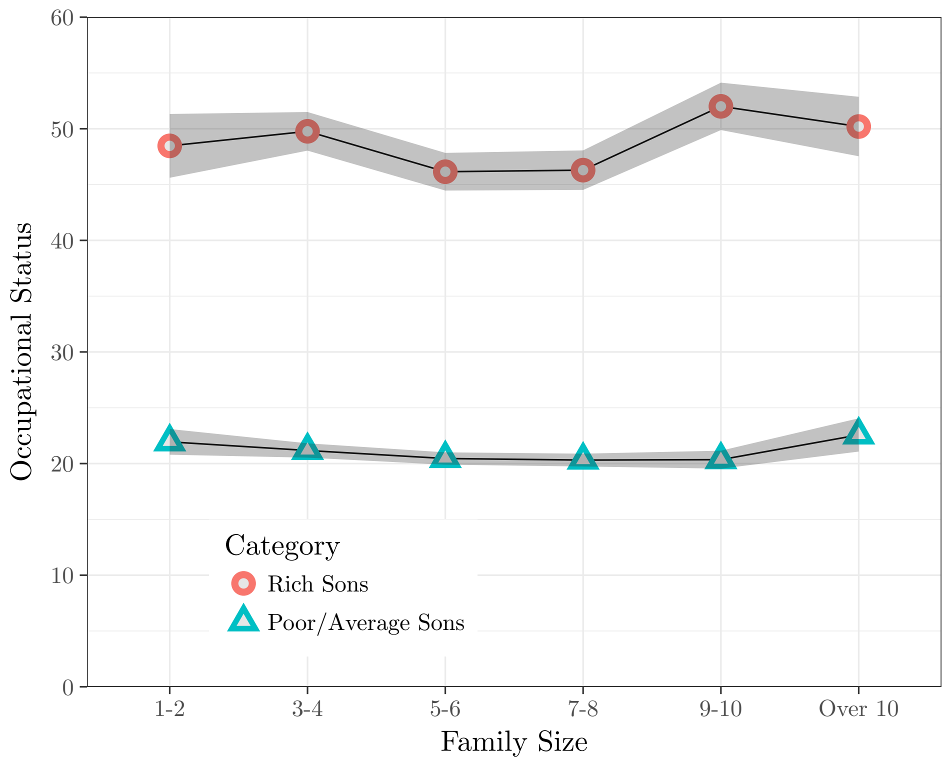


Figure A.3: Son Occupational Status, rich and poor lineages, marriages 1780-1879
Notes: 95% Confidence Intervals indicated by shading.

Table A.3: Size and Higher Education, Poorer/Average Lineages, sons, marriages 1780-1879

	Probability of Higher Education			
	(1)	(2)	(3)	(4)
N0	−.001		.0001	
	(.001)		(.001)	
N21		.0001		.0003
		(.001)		(.001)
Birth Rank	.006	.005	.008	.007
	(.008)	(.008)	(.008)	(.008)
Father Dies Before Child is 21			.005	.005
			(.009)	(.009)
Higher Education, Father			.301***	.301***
			(.087)	(.087)
Ln(Wealth)			.006***	.006***
			(.002)	(.002)
Occupational Status			.002***	.002***
			(.001)	(.001)
Observations	4,270	4,270	4,270	4,270
R^2	.0003	.0001	.202	.202
Note:	*p<0.1; **p<0.05; ***p<0.01			
	OLS Regression			

Table A.4: Size and Higher Education, Rich Lineages, sons, marriages 1780-1879

	Probability of Higher Education			
	(1)	(2)	(3)	(4)
N0	−.002		−.006**	
	(.003)		(.002)	
N21		.006*		−.004
		(.004)		(.003)
Birth Rank	−.041	−.041	−.023	−.022
	(.030)	(.030)	(.028)	(.028)
Father Dies Before Child is 21			−.029	−.031
			(.023)	(.023)
Higher Education, Father			.085***	.085***
			(.031)	(.031)
Ln(Wealth)			.018***	.019***
			(.003)	(.003)
Occupational Status			.005***	.005***
			(.001)	(.001)
Observations	3,302	3,302	3,302	3,302
R^2	.001	.002	.207	.206

Note: *p<0.1; **p<0.05; ***p<0.01
Ordinary Least Squares Regression
Errors Clustered on Fathers

Table A.5: Grandchild Wealth and Grandfather Family Size, marriages 1780-1879

	Ln(Wealth)			
	All	Rich	Probated	Non Probated
	(1)	(2)	(3)	(4)
Ln(N21), Grandfather	−.067	−.172	−.134	.077
	(.068)	(.111)	(.106)	(.085)
Birth Rank	.102	−.312*	−.173	.348***
	(.104)	(.189)	(.186)	(.114)
Age at Death	.022***	.021***	.024***	.022***
	(.002)	(.004)	(.003)	(.003)
Ln(Wealth), Grandfather	.186***	.187***	.260***	
	(.017)	(.023)	(.032)	
Occupational Status	.025***	.021***	.019***	.042***
	(.003)	(.003)	(.003)	(.005)
Higher Education, Grandfather	.125	.018	.166	−.183
	(.141)	(.146)	(.148)	(.480)
Female	−.237***	−.262**	−.363***	−.139*
	(.062)	(.104)	(.103)	(.076)
Father Alive	−.839***	−1.088***	−.988***	−.712***
	(.121)	(.217)	(.199)	(.142)
Observations	6,717	2,790	2,979	3,911
R^2	.334	.235	.243	.188

Note: *p<0.1; **p<0.05; ***p<0.01
Ordinary Least Squares Regression
Errors Clustered on Grandfathers

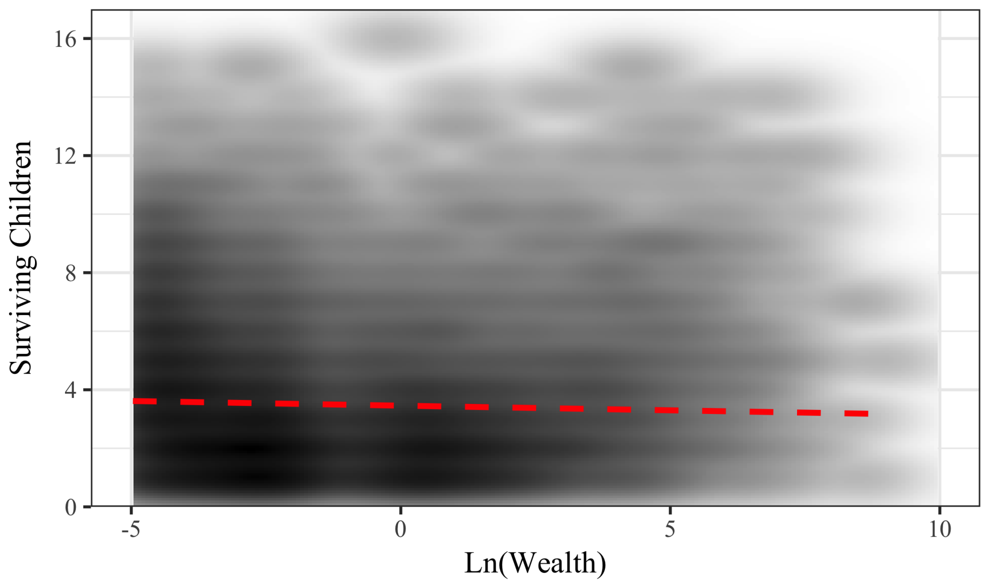


Figure A.4: Family Size versus Father Wealth, England, marriages 1780-1879

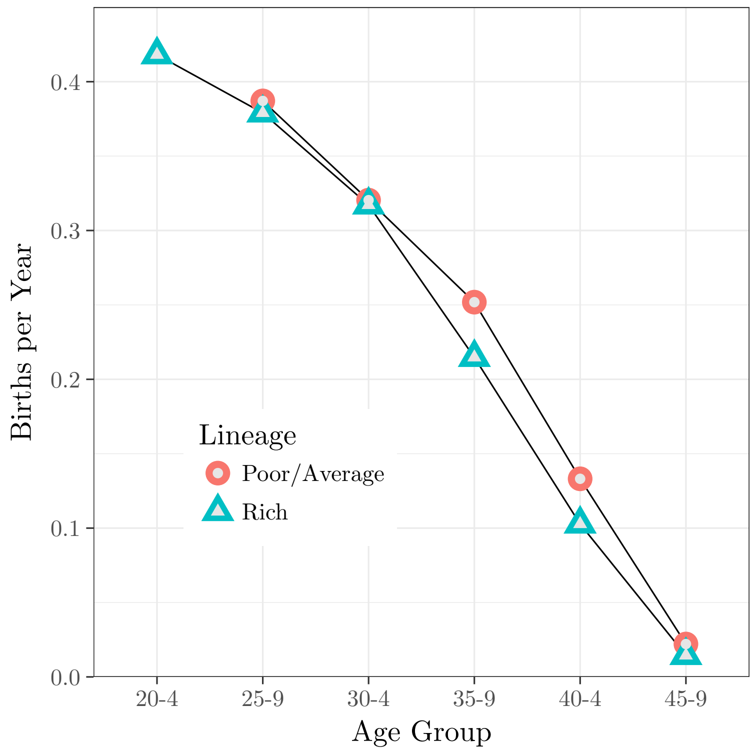


Figure A.5: Fertility by Age, Rich and Poor, Marriages 1780-1879
Notes: Because there are more missing female children, fertility rates are estimated from male births for our sample. These are multiplied by 1.953 to estimate total fertility rates, assuming the ratio of male to female births is 1.05 (UK Department of Health, 2013).

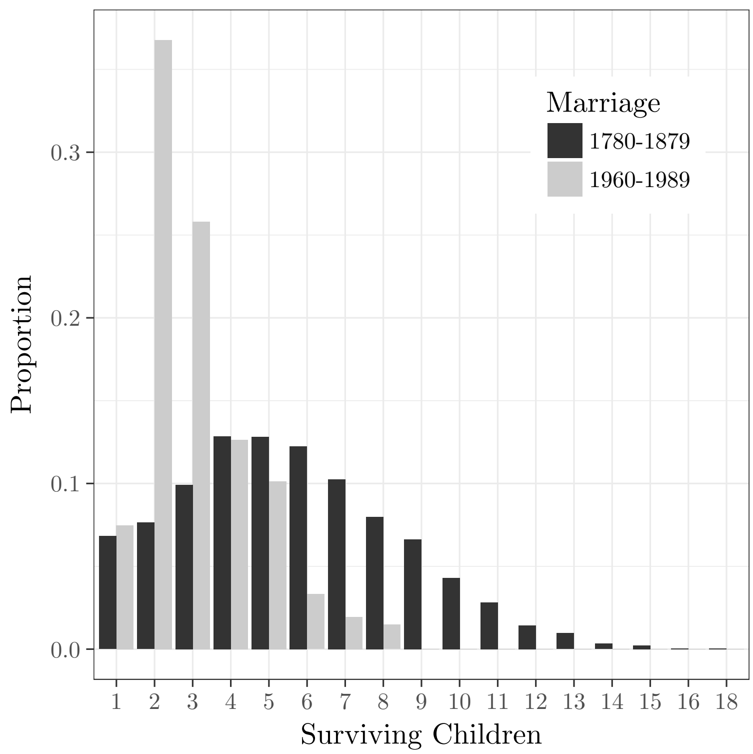


Figure A.6: Share of Children in each Family Size, 1780-1879, 1960-89

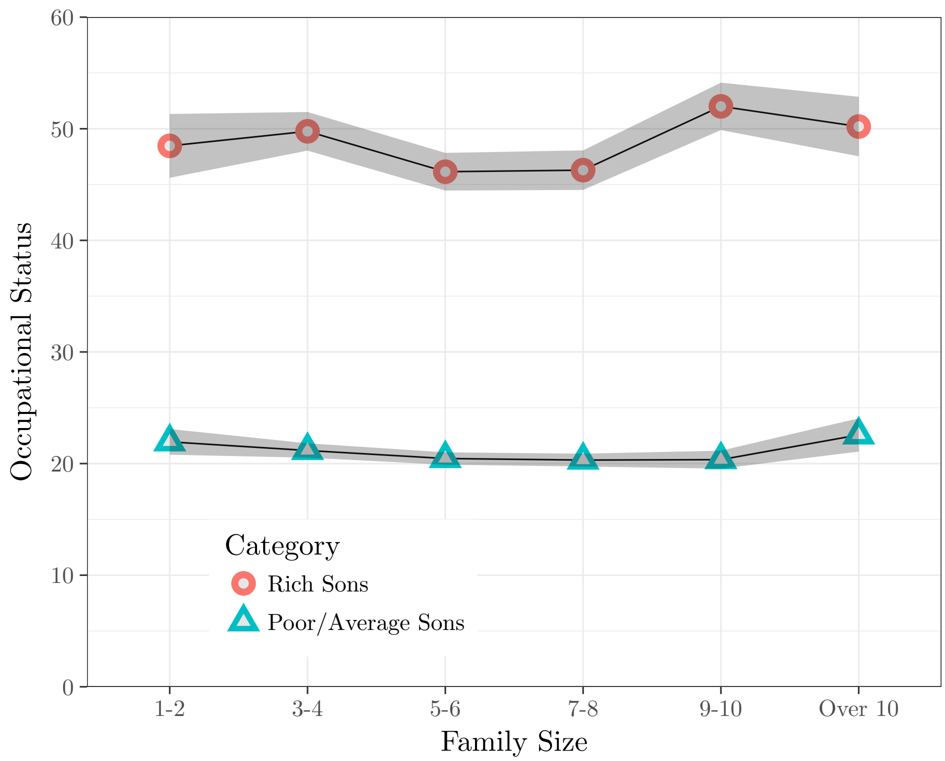
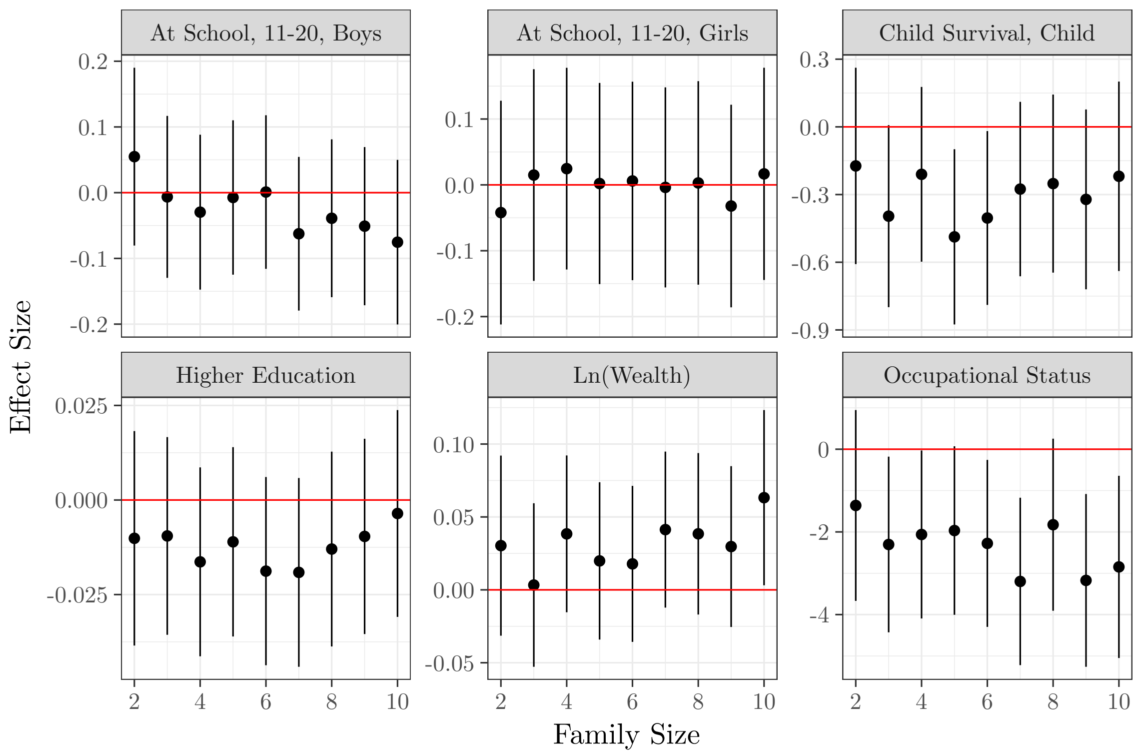
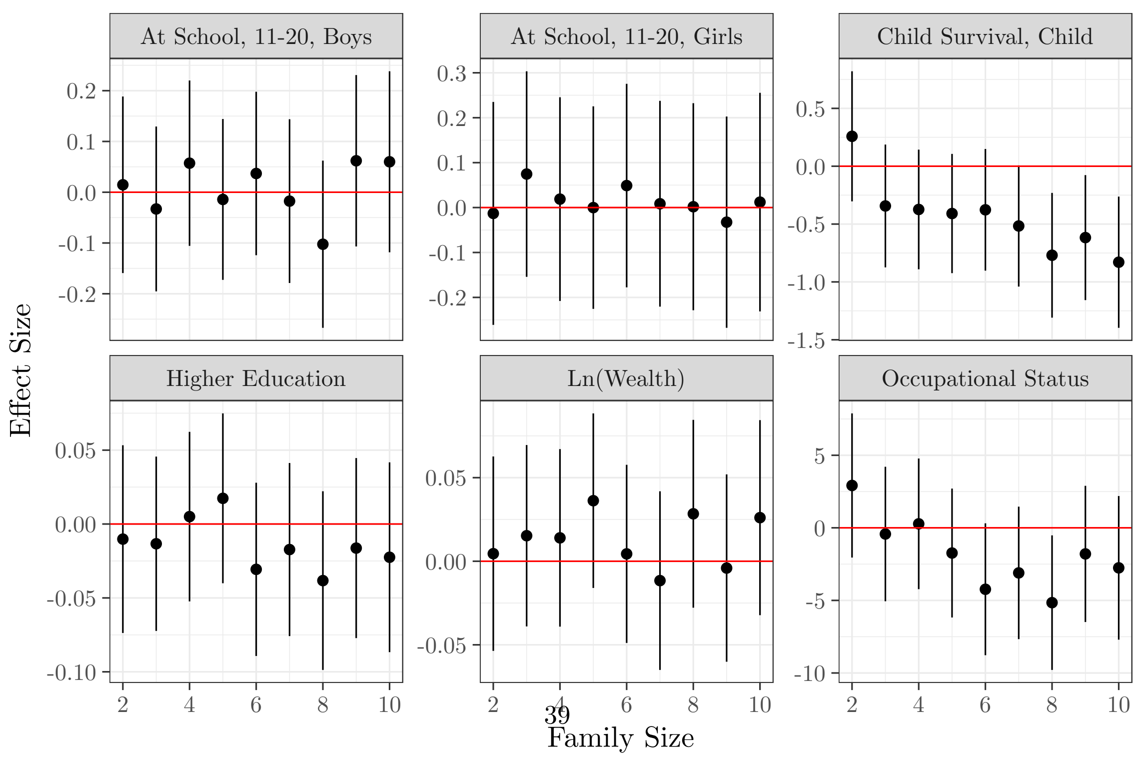


Figure A.7: Son Occupational Status, rich and poor lineages, marriages 1780-1879

Notes: 95% Confidence Intervals indicated by shading.



(a) Poorer/Average Lineages



(b) Rich Lineages

Figure A.8: The QQ Effect, By Lineage

B Klemp-Weisdorf

Table B.1 reports the correlation of family size with the probability of a skilled occupation while table C.7 reports the same regression, instrumenting fertility with the marriage -> first birth interval as is done by Klemp and Weisdorf (2018). Table B.3 reports the correlation of father status and this interval. Table B.4 reports the IV results for the complete sample.

Table B.1: Size and Probability of a Skilled Occupation, All Lineages, Sons of Marriages 1780-1880

	Probability of a Skilled Occupation							
	Rich Lineages				Poor Lineages			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
N0	$-.012^{***}$		$-.009^{***}$		$-.010^{***}$		$-.008^{***}$	
	(.003)		(.002)		(.003)		(.002)	
N21		$-.001$		$-.007^{***}$		.0002		$-.006^{**}$
		(.003)		(.002)		(.003)		(.002)
Birth Rank			.001	.002			$-.007$	$-.006$
			(.021)	(.021)			(.021)	(.021)
Father Dies Young			$-.027$	$-.030$			$-.031^{*}$	$-.035^{*}$
			(.019)	(.019)			(.019)	(.019)
Higher Education, Father			$-.219^{***}$	$-.220^{***}$			$-.224^{***}$	$-.225^{***}$
			(.023)	(.023)			(.023)	(.023)
Ln(Wealth), Father			$.026^{***}$	$.027^{***}$			$.025^{***}$	$.025^{***}$
			(.002)	(.002)			(.002)	(.002)
Occupational Status,Father			$.008^{***}$	$.008^{***}$			$.009^{***}$	$.009^{***}$
			(.0005)	(.0005)			(.0005)	(.0005)
Observations	6,450	6,450	6,450	6,450	6,541	6,541	6,541	6,541
R^2	.006	.00005	.241	.239	.005	0.00000	.231	.229
Note:	*p<0.1; **p<0.05; ***p<0.01 Ordinary Least Squares Regression Errors Clustered on Fathers							

Table B.2: Size and Probability of a Skilled Occupation, All Lineages, Sons of Marriages 1780-1880, Instrumented

	Probability of a Skilled Occupation							
	Rich Lineages				Poor Lineages			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
N0	−.016		−.018		−.015		−.018	
	(.018)		(.014)		(.018)		(.014)	
N21		−.018		−.021		−.017		−.021*
		(.021)		(.016)		(.021)		(.012)
Birth Rank			−.002	.003			−.004	.002
			(.024)	(.024)			(.024)	(.025)
Father Dies Before Child is 21			−.016	−.022			−.015	−.021
			(.024)	(.023)			(.024)	(.020)
Higher Education, Father			−.216***	−.216***			−.216***	−.216***
			(.026)	(.027)			(.026)	(.026)
Ln(Wealth)			.025***	.026***			.025***	.026***
			(.003)	(.003)			(.003)	(.002)
Occupational Status			.008***	.008***			.008***	.008***
			(.001)	(.001)			(.001)	(.0005)
Observations	4,472	4,472	4,472	4,472	4,470	4,470	4,470	4,470
R^2	−.004	−.018	.244	.241	−.003	−.017	.244	.242

Note: *p<0.1; **p<0.05; ***p<0.01
2SLS Regression
IV is Interval between Marriage and First Birth

Table B.3: Marriage to First Birth Interval and Status, All Data

	Marriage to 1st Birth Interval
Rich Families (H)	.094*
	(.049)
Richest Families (HH)	.100**
	(.047)
Poor Families (P)	.141**
	(.056)
Age of Mother	−.001
	(.001)
Birth Year	−.001
	(.001)
Observations	1,448
R^2	.014
Note:	*p<0.1; **p<0.05; ***p<0.01
	Ordinary Least Squares Regression
	Errors Clustered on Fathers

Table B.4: Size and Probability of a Skilled Occupation, All Lineages, Sons of Marriages 1780-1880, Instrumented

	Probability of a Skilled Occupation					
	OLS		*instrumental variable*			
			IV: Marr.-> 1st Birth		IV: Wife's Age at Marriage	
	(1)	(2)	(3)	(4)	(5)	(6)
N0	−.009***		−.008		−.017	
	(.003)		(.008)		(.015)	
N21		−.009***		−.010		−.019
		(.003)		(.011)		(.016)
ln(Wealth)	.024***	.024***	.024***	.025***	.025***	.025***
	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)
Occupational Status	.008***	.008***	.008***	.008***	.008***	.008***
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Higher Education, Father	−.227***	−.226***	−.224***	−.224***	−.228***	−.227***
	(.029)	(.030)	(.030)	(.031)	(.029)	(.030)
Observations	7,347	7,347	6,901	6,901	7,347	7,347
R^2	.260	.260	.257	.257	.257	.256
Note:	*p<0.1; **p<0.05; ***p<0.01					
	Father's with one Wife Only					

C Extra Results

Table C.1 shows the estimated effects of family size on adult life span for children, controlling for the fathers social status and the average lifespan of the parents. Parental lifespan significantly predicts that of children. But in these period adult lifespan is not significantly influenced by social status. Adult lifespan is independent of births, but significantly positively predicted by $N21$. The implication may be that there is just a biological mechanism at work here where families with greater physical robustness produce more surviving children, and those children in turn have longer lifespans.

Table C.2 gives the elasticities at the mean reported for each of the six main characteristics of children in table 5.1 in the body of the paper, but this time restricted to marriages 1780-1849, since for marriages 1850-1879 there are modest signs of fertility control. The results do not change substantially. If anything the reported elasticities are even closer to 0.

Tables C.3-C.6 perform the estimations of tables 4.1-A.3 in the body of the text using the correct logit specification. As can be seen the results are very similar in terms of when a negative family size effect appears, and in terms of the effects of birth rank.

Table C.1: Child Adult Lifespan and Family Size, Marriages 1780-1879

	Adult Age at Death			
	(1)	(2)	(3)	(4)
N0	.029		−.008	
	(.042)		(.062)	
N21		.247***		.199***
		(.049)		(.070)
Average Age at Death of Parents			.181***	.177***
			(.021)	(.021)
Father Dies Before Child is 21			1.148*	1.073*
			(.603)	(.601)
Higher Education, Father			1.001	.980
			(.800)	(.802)
Ln(Wealth)			.124	.116
			(.078)	(.078)
Occupational Status			−.021	−.021
			(.014)	(.014)
Birth Rank			.927	.937
			(.749)	(.748)
Female	3.883***	3.822***	4.610***	4.554***
	(.258)	(.258)	(.382)	(.382)
Observations	23,774	23,773	11,112	11,111
R^2	.011	.012	.025	.026

Note: *p<0.1; **p<0.05; ***p<0.01
Ordinary Least Squares Regression
Errors Clustered on Fathers

Table C.2: Summary Elasticities, 1780-1849

Variable	Rich	.	Poor/Avg.	.
	N0	N21	N0	N21
At School, 11-20, Boys	−.02	−.05	.04	.06
At School, 11-20, Girls	.05	−.02	−.36**	−.15
Higher Education	−.15**	−.11*	.67*	.12
Occupational Status	−.10***	−.08***	−.05*	−.05*
Child Survival, Children	.005	.01	−.03	.01
Wealth	−.48***	−.45***	−.13	−.01

Note: *p<0.1; **p<0.05; ***p<0.01
Wealth elasticities for Poor/Avg. are for non-probated fathers

Table C.3: Family Size and Probability at School 11-20, Marriages 1780-1879, Sons, Poorer/Average Lineages

	Probability at School 11-20			
	(1)	(2)	(3)	(4)
N0, Father	−.007		−.002	
	(.247)		(.146)	
N21, Father		−.006		−.002
		(.263)		(.168)
Birth Rank	−.008	−.008	−.111	−.111
	(2.084)	(2.083)	(1.595)	(1.600)
Ln(Wealth)	.017	.017	.015	.015
			(.191)	(.192)
Father Dead	.003	.003	.004	.004
			(.032)	(.032)
Occupational Status	.091	.088	−.047	−.048
			(1.693)	(1.698)
Census Year	.034	.029	−.005	−.005
Age Observed	−.001	−.001	−.001	−.001
tage	−.082	−.082	−.067	−.067
dfem_c	−.002	−.003	−.141	−.141
Observations	4,461	4,461	2,647	2,647
R^2	.288	.287	.249	.249
Note:	*p<0.1; **p<0.05; ***p<0.01			
	Logistic Regression			
	Errors Clustered on Fathers			

Table C.4: Family Size and Probability at School 11-20, Marriages 1780-1879, Sons, Rich Lineages

	Probability at School 11-20			
	(1)	(2)	(3)	(4)
N0, Father	.0001		$-.019$	
	(.020)		(.022)	
N21, Father		.033		$-.012$
		(.022)		(.025)
Birth Rank	.180	.164	.244	.237
	(.228)	(.229)	(.281)	(.280)
Ln(Wealth)			$.087^{***}$	$.087^{***}$
			(.024)	(.024)
Father Dead			$-.121$	$-.125$
			(.224)	(.223)
Occupational Status			$.025^{***}$	$.025^{***}$
			(.003)	(.003)
Census Year			$-.006$	$-.006$
			(.004)	(.004)
Age Observed			$-.302^{***}$	$-.302^{***}$
			(.027)	(.027)
Observations	1,528	1,528	1,528	1,528

Note: *p<0.1; **p<0.05; ***p<0.01
Logistic Regression
Errors Clustered on Fathers

Table C.5: Size and Higher Education, Rich Lineages, sons, marriages 1780-1879

	Probability of Higher Education			
	(1)	(2)	(3)	(4)
N0	$-.016$		$-.045^{***}$	
	(.017)		(.014)	
N21		.025		$-.034^{**}$
		(.017)		(.017)
Birth Rank	$-.242$	$-.238$	$-.171$	$-.164$
	(.152)	(.151)	(.181)	(.181)
Father Dies Before Child is 21			$-.166$	$-.182$
			(.138)	(.139)
Higher Education, Father			.215	.221
			(.163)	(.164)
Ln(Wealth)			$.145^{***}$	$.145^{***}$
			(.022)	(.022)
Occupational Status			$.028^{***}$	$.028^{***}$
			(.003)	(.004)
Observations	3,245	3,245	3,245	3,245
Note:	*p<0.1; **p<0.05; ***p<0.01 Logistic Regression Errors Clustered on Fathers			

Table C.6: Size and Higher Education, Poorer/Average Lineages, sons, marriages 1780-1879

	Probability of Higher Education			
	(1)	(2)	(3)	(4)
N0	−.013		.098	
	(.073)		(.060)	
N21		.024		.117*
		(.084)		(.065)
Birth Rank	.415	.409	.960	.919
	(.469)	(.464)	(.717)	(.717)
Father Dies Before Child is 21			−.051	.046
			(.658)	(.703)
Higher Education, Father			−.955	−.951
			(.707)	(.697)
Ln(Wealth)			.246***	.248***
			(.068)	(.069)
Occupational Status			.085***	.084***
			(.012)	(.011)
Observations	3,588	3,588	3,588	3,588

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

Table C.7: Size and Probability of a Skilled Occupation, All Lineages, Sons of Marriages 1780-1880, Instrumented

	Probability of a Skilled Occupation							
	Rich Lineages				Poor Lineages			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
N0	−.016		−.018		−.015		−.018	
	(.018)		(.014)		(.018)		(.014)	
N21		−.018		−.021		−.017		$-.021^{*}$
		(.021)		(.016)		(.021)		(.012)
Birth Rank			−.002	.003			−.004	.002
			(.024)	(.024)			(.024)	(.025)
Father Dies Before Child is 21			−.016	−.022			−.015	−.021
			(.024)	(.023)			(.024)	(.020)
Higher Education, Father			$-.216^{***}$	$-.216^{***}$			$-.216^{***}$	$-.216^{***}$
			(.026)	(.027)			(.026)	(.026)
Ln(Wealth)			$.025^{***}$	$.026^{***}$			$.025^{***}$	$.026^{***}$
			(.003)	(.003)			(.003)	(.002)
Occupational Status			$.008^{***}$	$.008^{***}$			$.008^{***}$	$.008^{***}$
			(.001)	(.001)			(.001)	(.0005)
Observations	4,472	4,472	4,472	4,472	4,470	4,470	4,470	4,470
R^2	−.004	−.018	.244	.241	−.003	−.017	.244	.242

Note: *p<0.1; **p<0.05; ***p<0.01
2SLS Regression
IV is Interval between Marriage and First Birth

Families of England: Database Construction

Gregory Clark and Neil Cummins

January 19, 2019

1 Description

The Families of England database aims to constuct a complete genealogy of a representative set of English families from 1750 to 2018, a period of 8 generations, using public data sources. The database currently contains 314,322 individuals. The database is still very much a work under construction. The intergenerational linkages for these individuals are substantially complete for those born before 1930, but for those born later there is more work to be done on establishing these links. Currently there are 194,880 children linked with a father, 132,280 linked with a grandfather, 97,551 linked with a great-grandfather, 70,607 linked with a great-great-grandfather, 49,706 linked with a great-great-great-grandfather, 35,116 linked with a great-great-great-great-grandfather, and 25,092 linked with a great-great-great-great-great-grandfather (7 generations). However there is substantial ongoing work on establishing occupations, educational status, dwelling values, and wealth at death for each individual. Table 1 reports the currents sample sizes for various individual characteristics. We expect to add considerably more data on all the social outcome variables, and also on fertility.

To enable high linkage rate with the sources we have (described below) we adopt the strategy of following families with rare surnames, and follow descent along the male line. The vagaries of English spelling, and the varied ethnic background of the population in different parts of England, ensures that a substantial minority of the English population, even in 1800, held surnames that were shared with modest numbers of other individuals. To ensure that there is no bias in this procedure we will also link many of the daughters to their husbands, and wives to their fathers, to

Outcome	N
People	314,322
Death Location	118,655
Birth and Death Location	94,477
Age at Death,	125,393
Wealth at Death,	47,922
Higher Education	40,625
Occupation,	34,564
School 11-20	22,656
House Value 2017	9,678

Table 1: Current Status of Database

Table 2: Linkage Rates for Men born 1800-1999

Birth Period	Men 21+	Linked	Link Rate
1800-49	6,723	5,446	.81
1850-99	11,482	10,529	.92
1900-49	11,676	10,517	.90
1950-99	8,385	6,702	.80

Calculated from the Families of England Data

check that mobility and other characteristics along the female line have the same character as with the male line. Using such rare surnames we can achieve very high linkage rates between parents and children. Table 2, for example, shows (for a random sample) the fraction of men bearing rare surnames who can be linked to their father.

For men born 1850-1949, and living to reproductive age, the linkage rate is greater than 90 percent. Typical linkage rates for historical intergenerational databases, using all surnames, at least in the US, are only around 20%.[1] These linkages are also of high reliability in the years 1800-1930, since there are multiple sources in many cases identifying parents - censuses, birth records, marriage records, passenger lists - and there are few alternative candidates who can get confused with the target individual. Thus for a sample of 7,626 recorded rare surname births 1860-1879, we identify a father or mother for 88%[2]. The reasons for failing to find at least one parent in the other 12% of cases are various. In some cases the name likely was misspelled in the birth record, and the person does not belong in the surname lineages used to form the sample. Of those not linked 60% show no further appearance in any record after their birth under the birth name. Likely in most of these cases the name is just misspelled on the birth register. In others the child dies before appearing in a census, or their father dies, or they are living with grandparents in the census, or the family emigrates[3]. Thus one third of those born not linked to a parent died before age 10. Again, in contrast, historical intergenerational databases in the US using the general population are claimed to mismatch one third of individuals to their parents (Bailey et al., 2017). A reflection of the likely high success rate in making linkages is the observed intergenerational correlation of occupational status. This is 0.67, which is much higher than that observed in other census based historical linked samples.

Though the numbers of recorded births for men and women is similar, and the match rate to fathers for the births is also similar by gender, the final dataset of family size by father is missing at least 12-14% of girls. This is because children in families can also be identified from the existence of a death record, or from their presence in a census or other record, where the birth was not recorded under the correct family surname. But adult women will only appear in a death or census record if they remain unmarried. Thus more sons are identified from such records, absent the birth record. Table 3 shows for men and women of the target rare surnames the numbers linked to fathers in total and by gender and type for births 1860-79, for all births and for those attaining age 21. Though an equivalent number of women are matched to fathers in the births sample, many more men are identified from ancillary records. This implies that at least 12% of girls are missing from the sample

[1] Long and Ferrie, 2018, for example, link only 20% of adult sons to their fathers in England between 1851 and 1881.

[2] In some cases, where the child is illegitimate, only the mother is listed on birth records.

[3] We could identify the father by getting the birth certificate, but this is prohibitively costly.

Table 3: Share of Men and Women in Family Size Sample, 1860-79

	All	Men	Women
Births - all	6,205	3,218	2,987
Births - Birth record	5,826	2,877	2,949
Births - no Birth record	379	341	38
21+ - all	4,788	2,529	2,259
21+ - birth record	4,455	2,226	2,229
21+ - no birth record	333	303	30

Table 4: Missing Women by Family Size, pre-1880 marriages, children 21+

Family Size	All	All Children	Male	Female	% missing females
0	803	0	0	0	0
1	367	367	211	156	26.1
2	452	904	511	393	23.1
3	514	1,542	862	680	21.1
4 or 5	906	4,039	2168	1,871	13.7
6 or 7	554	3,560	1,876	1,684	10.2
8+	433	4,054	2,057	1,997	2.9
All	3,990	14,466	7,695	6,771	12

of births, and 14% from the sample of those attaining age 21.

To ensure a respresentative sample of people in each generation we have followed the strategy of including in the database all individuals bearing one of the target surnames whenever there is a birth, death or marriage record under that surname. We also try and follow the lineages of those who emigrate from England, typically to Canada, Australia, the USA, and New Zealand.

The genealogical linkages have been established in two ways. For a substantial subset of the data, 67,305 individuals we constructed the genealogical links ourselves. The other 193,690 individuals are from genealogies constructed by members of the Guild of One-Name Studies, a society devoted to studying the history and genealogy of rare surnames. The use of these Guild genealogies raises issues of selectivity, since it is more likely that a rare surname will be included in a Guild study if there is a current bearer of higher social status. But we can do extensive checks on the representativeness of these Guild contributed surnames, and find that at least for the 19th century they have average social status.

In both our reconstructions and those of the Guild genealogies the familial linkages - assigning fathers, mothers, and spouses - are established using a wide range of evidence. For England there are census records 1841, 1851, 1861, 1871, 1881, 1891, 1901, 1911. There is the Population Register of 1939. There is the register of births, deaths and marriages 1837-2005. The birth register 1912-

2005 gives the surname of the mother. There are selective parish registers of births and marriages 1750-1930. There are probate records nationally, 1858-2018, and for the Canterbury and York Ecclesiastical courts 1750-1858. There are passenger lists for those leaving the UK 1890-1960, and for those entering the UK 1878-1960. There are Electoral Registers 1900-2012.

In recalcitrant cases in England we can, at cost, order the actual birth certificate which list the father and mother, or marriage certificate which lists marriage partners, their occupations and those of the fathers. We plan on doing this for a select sample of people marrying around 1990, so that we can get their occupational status, where they would typically be born circa 1960, as well as the occupational status of their fathers born circa 1930.

It is possible in many cases to check proposed familial linkages against genealogies uploaded by ancestry.com members. These genealogies are not always reliable. But the better ones cite source documents which can be inspected to see if the link is sound.

Ancestry.com records the age at death of many migrants from the England to Canada, Australia, NZ and USA. For Australia the voting rolls 1903-1983 give occupations. For the US the censuses 1850-1960 record occupations. Canada and New Zealand also have some occupational information from voting rolls. However, wealth at death is generally not available for migrants to these countries.

2 Status Measures

The social status indicators we have are age at death, wealth at death, schooling, occupation, location, and first names of children.

Wealth at Death: For England and Wales the Principle Probate Registry records whether someone was probated, and the value of their estate for all deaths in England 1858-2018. This information is the most comprehensive and unusual outcome result that we have for this database. The probate information is searchable at https://probatesearch.service.gov.uk/#wills. However, the estate values 1996-2018 are now obtainable only at cost of 10 pounds per person.

Schooling and Training: The censuses of 1851-1911, and population register of 1939, record whether anyone aged 10-19 is still attending a school, which gives us a measure of education for the earlier years. From the previous NSF project we have a database of all students who attended Oxford or Cambridge, 1750-2015. But this constitutes only 1-2% of each cohort. Complete records are available for attendees at the Royal Military Academy Woolwich (1790-1839) and Royal Military College Sandhurst (1800-1946). Complete records are available for Masters and Mates Certificates, 1850-1927, UK Medical Registers, 1859-2015, UK, Civil Engineer Lists, 1818-1930, UK, Electrical Engineer Lists, 1871-1930, UK, Mechanical Engineer Records, 1847-1930, UK, Articles of Clerkship, 1756-1874. From all these measures we can construct indices of educational attainment for people in the database born before 1900.

Occupation Status: The censuses of 1851-1911, and the Population Register of 1939 record occupations, so we can estimate adult occupations for the cohorts born 1920 and before. Passenger lists give occupations for international travellers up to 1960. Birth certificates record the occupation of father's, and from 1995 on that of mothers also. Marriage certificates record the occupations of husband and wife, and of fathers. So for a select sample we can estimate occupations for people born up to around 1980.

Dwelling Value: From the electoral census of 1999-2012 we have the addres where adults were living in 1999-2012, from which we can infer using the Land Registry the property value. This gives an indirect measure of family income.

Children's First Names: Children's first names are a good proxy for family social status in modern generations. Using records of Oxbridge attendance and property values we can assign status measures to parents based on their child name choices.

After completing the genealogical links, and the status information, we will have potentially the following information for each person in the database:

Date of birth, longevity, wealth at death, educational attainment, occupation, birth location, fertility, child mortality, death location, birth order, number of siblings, age at marriage

3 Representativeness

Figure 1 compares the singulate mean age at marriage, by sex, for the FOE data with the population (Office for National Statistics (2015)). As there is a strong status gradient on marriage age we only compare our "PP" (poorest group) with the population. There is a close correspondence in all 4 series. However after 1950, our estimates of female singulate mean age at marriage are higher than that of the rest of the population. This is due to the difficulty of calculating the marriage age of currently living women (for many women marrying into the sample, we typically observed year of birth through their death record).

Figure 2 compares the median age at death, by sex, for the FOE data with the population (Office for National Statistics (2010)). Here our the sample consistently reflects the population.

Figure 3 compares the probate rate for our 4 surname type lineages with the general population (calculated from Cummins (2019)). Here, our poorest group, "PP", track the level and trend of the general population, 1892-1992. Our elite groups are consistently above the population average and maintain their relative position over time.

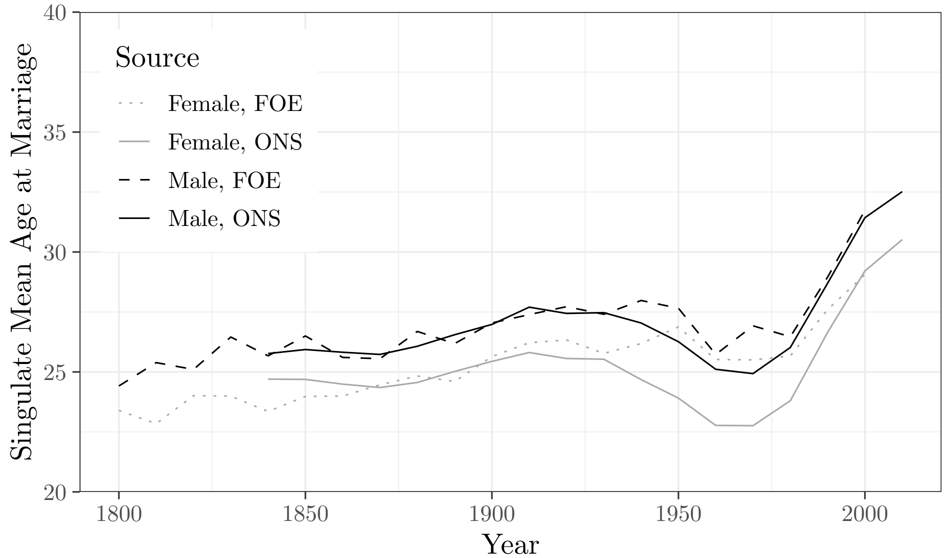


Figure 1: Age at Marriage: FOE data v Population, Male and Female

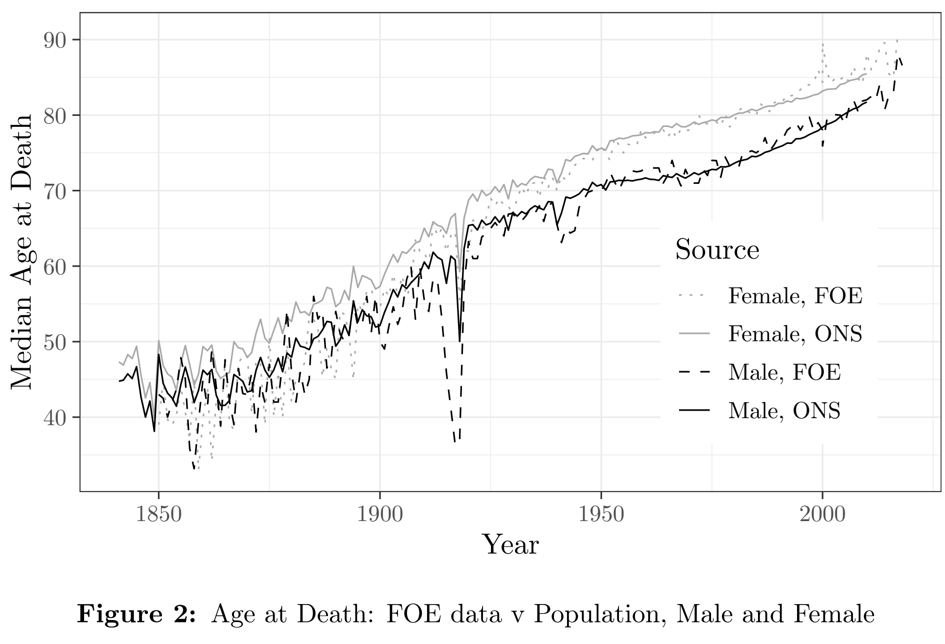


Figure 2: Age at Death: FOE data v Population, Male and Female

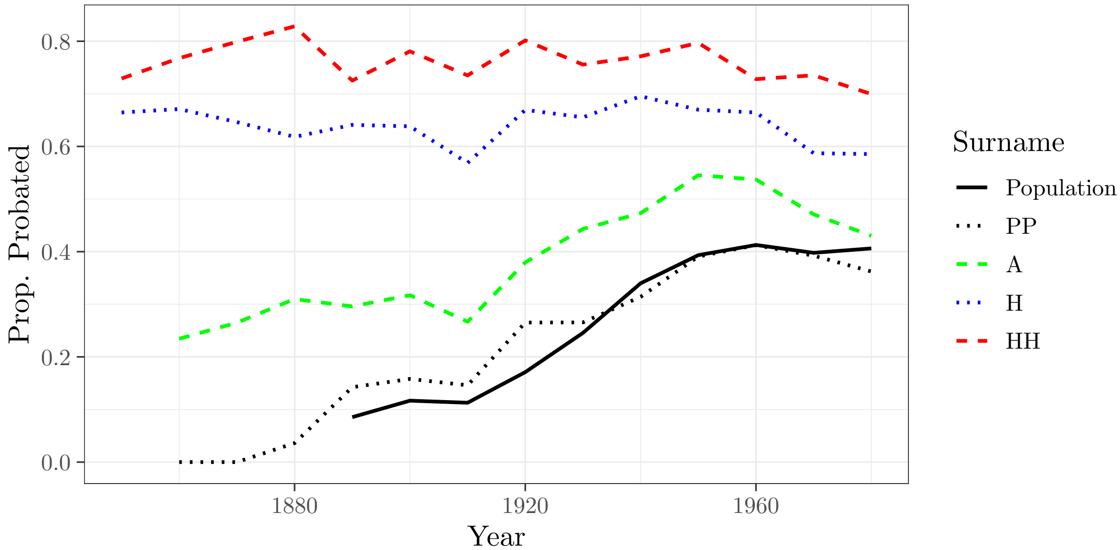


Figure 3: Proportion Probated: FOE data v Population, Male and Female

References

Bailey, Martha, Connor Cole, Morgan Henderson, and Catherine Massey, "How Well Do Automated Methods Perform in Historical Samples? Evidence from New Ground Truth," Working Paper 24019, National Bureau of Economic Research November 2017.

Cummins, Neil, "Where is the Middle Class? Inequality, Gender and the Shape of the Upper Tail from 60 million English Death and Probate Records, 1892-2016," *Centre for Economic Policy Research Discussion Paper 13436*, 2019.

Long, Jason and Joseph Ferrie, "Grandfathers Matter(ed): Occupational Mobility Across Three Generations in the US and Britain, 1850â??1911," *The Economic Journal*, 2018, *128* (612), F422–F445.

Office for National Statistics, "Cohort Fertility, England and Wales, 2009 (Excel sheet 430Kb).," 2010.

_ , "Dataset: Marriages in England and Wales 2015 Updated," 2015.

A Data Sources

A.1 Births, Deaths, Marriages

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

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

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

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

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

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

A.2 Social Outcomes

Censuses, 1841-911 and Population Register 1939

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

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

A.3 Wealth at Death

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

Gov.UK. Wills and Probate 1858-1996 [database on-line]. Original data: Principal Probate

Registry. Calendar of the Grants of Probate and Letters of Administration made in the Probate Registries of the High Court of Justice in England. London, England © Crown copyright. Available at `https://probatesearch.service.gov.uk/Calendar#calendar` (last accessed: 01 Apr 2016).

General Register Office. 1861. Annual Report of the Registrar General 1859. (pp. 173-181) Available online at `http://www.histpop.org` (last accessed: 01 Apr 2016).

Prerogative Court of Canterbury and Related Probate Jurisdictions: Probate Act Books. Volumes: 1850-57. Held at the National Archives, Kew. (Catalogue Reference: PROB 8/243-250.)

A.4 Educational Status

Ancestry.com. UK, Articles of Clerkship, 1756-1874 [database on-line]. Provo, UT, USA: www.ancestry.com Operations, Inc., 2012. Original data: Court of King's Bench: Plea Side: Affidavits of Due Execution of Articles of Clerkship, Series I, II, III (KB 105-107). The National Archives, Kew, Richmond, Surrey. Registers of Articles of Clerkship and Affidavits of Due Execution (CP 71). The National Archives, Kew, Richmond, Surrey. Available at www.ancestry.com(last accessed: 01 Apr 2016).

Ancestry.com. Cambridge University Alumni, 1261-1900 [database on-line]. Provo, UT, USA: Ancestry.com Operations Inc, 1999. Original data: Venn, J. A., comp.. Alumni Cantabrigienses. London, England: Cambridge University Press, 1922-1954. Available at www.ancestry.com (last accessed: 01 Apr 2016).

Ancestry.com. Oxford University Alumni, 1500-1886 [database on-line]. Provo, UT, USA: Ancestry.com Operations Inc, 2007. Original data: Foster, Joseph. Alumni Oxonienses: The Members of the University of Oxford, 1715-1886 and Alumni Oxonienses: The Members of the University of Oxford, 1500-1714. Oxford: Parker and Co., 1888-1892. Available at www.ancestry.com (last accessed: 01 Apr 2016).

Ancestry.com. UK, Crockford's Clerical Directories [database on-line]. Provo, UT, USA: Ancestry.com Operations Inc, 2009. Original data: Crockford Clerical Directory. England: 1868-1932. See title page image of each directory for original source information. Available at www.ancestry.com (last accessed: 01 Apr 2016).

Ancestry.com. UK, Electrical Engineer Lists, 1871-1930 [database on-line]. Provo, UT, USA: Ancestry.com Operations, Inc., 2014. Original data: Society of Telegraph Engineers (later, Institution of Electrical Engineers) Membership Lists, 1887–1930. Institution of Engineering and Technology, Savoy Place, London, England. UK, Civil Engineer Lists, 1818-1930. Available at www.ancestry.com (last accessed: 01 Apr 2016).

Ancestry.com. UK, Mechanical Engineer Records, 1847-1930 [database on-line]. Provo, UT, USA: Ancestry.com Operations, Inc., 2013. Original data: Mechanical Engineering Records, 1847-1930. London, UK: Institution of Mechanical Engineers. Available at v (last accessed: 01 Apr 2016).

Ancestry.com. UK Medical Registers, 1859-1959 [database on-line]. Provo, UT, USA: Ancestry.com Operations Inc, 2008. Original data: General Medical Council, comp. UK Medical Registers, 1859-1959. London: General Medical Council, 1859-1959. This data is provided in partnership with the General Medical Council. Available at www.ancestry.com (last accessed: 01 Apr 2016).

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

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