Childlessness, Celibacy and Net Fertility in Pre-Industrial England: The Middle-class Evolutionary Advantage

David de la Croix[1]   Eric B. Schneider[2]   Jacob Weisdorf[3]

June 2, 2019

Abstract

This paper reconsiders the fertility of historical social groups by accounting for singleness and childlessness. We find that the middle class had the highest reproductive success during England's early industrial development. In light of the greater propensity of the middle class to invest in human capital, the rise in the prevalence of these traits in the population could have been instrumental to England's economic success. Unlike earlier results about the survival of the richest, the paper shows that the reproductive success of the rich (and also the poor) were lower than that of the middle class, once accounting for singleness and childlessness. Hence, the prosperity of England over this period can be attributed to the increase in the prevalence of middle-class traits rather than those of the upper (or lower) class.

Keywords: Fertility, Marriage, Childlessness, European Marriage Pattern, Industrial Revolution, Evolutionary Advantage, Social Class
JEL Classification numbers: J12, J13, N33

1 Introduction

In attempting to explain the historical economic success of England and Western Europe more generally, scholars have long turned to differences in demographic regimes across countries as a potentially important factor. Malthus (1798) first raised the point arguing that high fertility from low ages at marriage and low celibacy in China prevented any material improvements in living standards whereas the preventive check, which limited fertility, allowed for incomes slightly above subsistence in England. Malthus's preventive check has been refined over time into what Hajnal (1965) defined as the European Marriage Pattern (EMP). Characterised by late age at marriage, high rates of celibacy and the prevalence of nuclear households, Hajnal argued that the EMP was common west of a line from St. Petersburg to Trieste and less common to the east. Thus, the EMP allowed Europeans to control their fertility and experience living standards above subsistence (Voigtländer and Voth 2013). In addition, De Moor and Van Zanden (2010) and Foreman-Peck (2011) have argued that the European Marriage Pattern influenced economic development not only by limiting population growth but also by promoting higher investment in human capital, especially since women were given high levels of agency in the system.

Sitting alongside these arguments, several theoretical contributions (Galor and Moav 2002; de la Croix and Doepke 2003; Clark 2007; Galor and Moav 2001) have argued that the internal demographic dynamics of pre-Industrial societies (or of least developed societies) mattered for human capital formation. In particular, Galor and Moav (2002) has suggested that the struggle for survival during most of human existence had generated an evolutionary advantage to human traits that were complementary to the growth process, and that this eventually triggered the take-off of modern economic growth. Their theory considers two types of families: *quantity* families, prone to having many, but less educated offspring, and *quality* families, prone to having fewer, but more educated children (in a Beckerian sense). Because quality families invested in their children's education, their children earned comparatively higher wages than their low-quality counterparts, which afforded educated individuals more offspring than uneducated ones. Quality families therefore had an evolutionary advantage over quantity families, causing their share in the population to rise and eventually leading to birth-restriction behaviour. On the evidence side, several studies have explored the historical link between various measures of income on the one hand and fertility or population growth on the other (e.g. Galor and Klemp 2014; Vogl 2016; Ashraf and Galor 2011; Clark and Hamilton 2006; Boberg-Fazlic, Sharp, and Weisdorf 2011). Three studies in particular (Clark and Hamilton 2006; Boberg-Fazlic, Sharp, and Weisdorf 2011; Clark and Cummins 2015) have provided pre-Industrial evidence that better educated and thus more wealthy individuals had higher net reproduction rates.

Both of these strands in the literature have contributed to our understanding of demographic change and economic development, but each has some problematic assumptions that need to be addressed more carefully. For the EMP, Dennison and Ogilvie (2014) have found that the differences in demographic regimes were not so different in Eastern and Western Europe, and they also show that demographic regimes were not homogenous within a country. At the same time, the survival of the richest literature has not properly accounted for differential rates of celibacy and childlessness across classes when computing the reproductive success of each group (Boberg-Fazlic, Sharp, and Weisdorf 2011).

This paper addresses these earlier shortcomings by analysing the internal fertility dynamics of pre-Industrial England, leading both to a substantial revision of previous conclusions about what caused the wealth of nations and to a strong empirical basis for the idea that quality families had an evolutionary advantage during the critical stages in the transition from economic stagnation to modern economic growth. We use the family reconstitution data collected from English parish registers and assembled by the Cambridge Group (Wrigley et al. 1997) to estimate the reproductive success of different wealth groups. Compared to previous attempts to measure reproductive success in historical England, which calculate reproductive rates conditional on giving birth, our study considers both the *intensive* margin, i.e. an individual's number of surviving offspring, and the *extensive* margin, i.e. whether or not an individual married or had children at all (Aaronson, Lange, and Mazumder 2014 and Baudin, de la Croix, and Gobbi 2015). Thus, we explore reproductive success across four margins: two extensive margins - the choice to marry and the choice to have children - and two intensive margins - the number of children born and the child death rate.

Our empirical exercise also proposes a novel methodological approach, which enables us to estimate the reproductive contribution of women who died before completing their reproductive period (something that happened often in the past). Whereas Weir (1984), Schofield (1985) and other historical demographers usually evaluate women's reproductive success at the end of their reproductive period (usually after the age of 40 or 50), we use Cox proportional hazard models instead. This model allows us to estimate the risk of demographic events occurring by means of survival curves which start at the age at which a woman becomes at risk of the event (i.e. age 16 for marriage). This way we are able to explore the reproductive potential of a much larger share of a historical population than is normally the case in studies of women's reproductive success.

Our study replicates earlier, class-specific, historical-demography findings by Clark and others that the upper class had the highest numbers of surviving offspring on the intensive margin. However, by accounting for all four margins of reproduction, our findings dispute the previous view that the upper classes had more surviving offspring than the lower classes. We show that

the middle classes of farmers and traders were the most successful socio-economic group in terms of reproduction once taking the extensive margin of fertility into account. Merchants, professionals and gentry, the upper classes, had high rates of celibacy and childlessness driving down their overall reproductive success, whereas labourers, servants, husbandmen and craftsmen had much lower rates of celibacy and childlessness but limited their fertility within marriage through birth spacing instead (Cinnirella, Klemp, and Weisdorf 2017). Thus, the middle classes combined high fertility and low celibacy and childlessness and therefore had the highest net reproduction rates.

We argue that this demographic regime provides a better explanation for historical England's economic success than those put forth in the previous work. Middle class families were more likely to invest in apprenticeships for their children (Leunig, Minns, and Wallis 2011), and according to Crouzet (1985) more than 85 per cent of England's early industrialists came from a middle-class background. This pattern of human capital investment emphasises how Hajnal's European Marriage Pattern might have explained England's economic success: not because all English families universally strove to limit births, but because those classes that invested comparatively little in growth-promoting education were more inclined to engage in birth restricting practises. This conclusion fits with the mechanism described in the quantity-quality model proposed in Galor and Moav (2002). There, differences in fertility between different social-groups' preferences for child quality have an impact on the distribution of traits in the population. Their theory highlights how the right balance between child quality and quantity (and not just child quantity per se) is important in explaining increased investment in human capital in the transition from stagnation to economic growth, consistent with behaviour of the English middle class observed here.

The paper unfolds as follows. Section 2 illustrates the different margins through which social status affects reproduction and reviews the literature on celibacy, childlessness, survival of the richest and the European Marriage Pattern. Section 3 describes the data used for the analysis and how we define social classes. In Section 4 we calculate the class-specific rates of celibacy, childlessness, marital fertility and child mortality and translate these rates into class-specific net rates of reproduction. Section 5 provides robustness checks and argues in particular that our results are not driven by selective migration. Finally, Section 6 discusses the implications of the results and Section 7 concludes.

2 Theoretical and Historiographical Underpinnings

2.1 A Theoretical Decomposition of the Data

As mentioned above, our paper considers two neglected aspects of overall fertility: celibacy and childlessness. We present differences in the four margins of fertility by social group using the following framework (extending the approach in Baudin, de la Croix, and Gobbi (2020)). Households are heterogenous by social class c. We consider two types of marital status: married and singles. The marriage rate depends on c: $m(c)$. We assume singles do not have children. Net reproduction $n(\cdot)$ as a function of social class c is:

$$n(c) = m(c) \, (1 - z(c)) \, b(c) \, (1 - d(c)) \tag{1}$$

where $z(c)$ is the fraction of childless married women, $b(c)$ is the number of births conditional on being married and having children, and $d(c)$ is the mortality rate of children aged 0 to 15. Equation (1) can be seen from the point of view of women or men, and we conduct the analysis for both. However, as in the vast majority of studies of fertility, we take women's fertility as our baseline and report the findings for male fertility in Appendix K. Although it is theoretically possible to estimate $z(c)$ and $b(c)$ for all marriages not just bachelor/spinster marriages, we limit ourselves to bachelor/spinster marriages. This assumption does not influence our final results and is driven by weaknesses in the data explained more fully in Section 4.2. Finally, in order to interpret differences in net reproduction as differences in survival among occupational groups, we must assume high levels of marital homogamy. This assumption will be discussed further in Section 3.2.

Although social class is normally captured by a discrete variable, in the following it is more convenient for the exposition to think about social class as a continuous variable. Usually, the literature finds that in pre-industrial societies the social gradient of births is positive (Skirbekk et al. 2008), $b'(c) \geq 0$, and the social gradient in mortality is negative or nil, $d'(c) \leq 0$, i.e. wealthier households have larger families and experience lower rates of child mortality. In most of the literature, one considers marital fertility, implying that the marriage margin is constant (Wrigley et al. 1997, p. 428), and in some cases the extensive margin of fertility is constant too (Boberg-Fazlic, Sharp, and Weisdorf 2011), $m(c) = \bar{m}$ and $z(c) = \bar{z}$, implying that the effect of social class on fertility is positive,

$$\mathbf{n}'(c) = \bar{m}(1 - \bar{z}) \, ((1 - d(c))\mathbf{b}'(c) - b(c)\mathbf{d}'(c)) > 0,$$

as $\mathbf{b}'(c) \geq 0$ and $\mathbf{d}'(c) < 0$.

In general though, the effect of social class on reproduction n is the sum of four effects:

$$\mathbf{n}'(c) = \underbrace{\underbrace{(1 - z(c))b(c)(1 - d(c))\,\mathbf{m}'(c)}_{\text{marriage margin}} - \underbrace{m(c)b(c)(1 - d(c))\,\mathbf{z}'(c)}_{\text{childlessness margin}}}_{\text{Extensive margins of fertility}}$$

$$+ \underbrace{m(c)(1 - z(c))\left((1 - d(c))\,\mathbf{b}'(c) - b(c)\,\mathbf{d}'(c)\right)}_{\text{intensive margin of fertility}}.$$

If marriage rates decrease with social class and/or childlessness rates increase with social class, it is possible that the rich would not be so reproductively successful after all. Formally, this happens when

$$b(c)(1 - d(c))(m(c)\mathbf{z}'(c) - (1 - z(c))\mathbf{m}'(c)) \quad > \quad m(c)(1 - z(c))((1 - d(c))\mathbf{b}'(c) - b(c)\mathbf{d}'(c)). \quad (2)$$

If marriage rates and/or childlessness rates are non-monotonic functions of social class, the pattern of reproductive success can become more complicated. To know whether condition (2) holds in reality, one needs to estimate $z(c)$, $m(c)$, $b(c)$ and $d(c)$ from data.

The relationship between our decomposition (1) and the so-called European Marriage Pattern (EMP) is explained in Figure 1. The bigger rectangle represents the maximum reproduction obtained when 100 per cent of women (horizontal axis) marry at age 15 (vertical axis). The usual representation of the European Marriage Pattern argues that by delaying marriage and having some dose of celibacy, the span of reproduction is reduced. In this paper, we consider the vertical line $1 - m(c)$ as dependent on social class. We add a second vertical line, representing childless couples. Taking into account child mortality, represented by the inner rectangle, we are left with the grey area representing reproduction. Thus, our paper calculates the size of the shaded box for different social groups to understand the social gradient in reproductive success.

2.2 Previous Literature on the Extensive Margin: Celibacy

We are not the first to address the marriage margin (or celibacy) and its influence on historical demographic change and economic development. Scholars commonly agree that the proportion of unmarried women in early modern England was nontrivial (Dennison and Ogilvie 2014). But there is disagreement about how important celibacy was in slowing population growth and whether celibacy differed across social groups.

The importance of celibacy for the early modern demographic regime has been contested. Both Hajnal (1965) and Wrigley and Schofield (1981) emphasised age at marriage as the most important factor maintaining the preventive check and reducing population pressure. However,
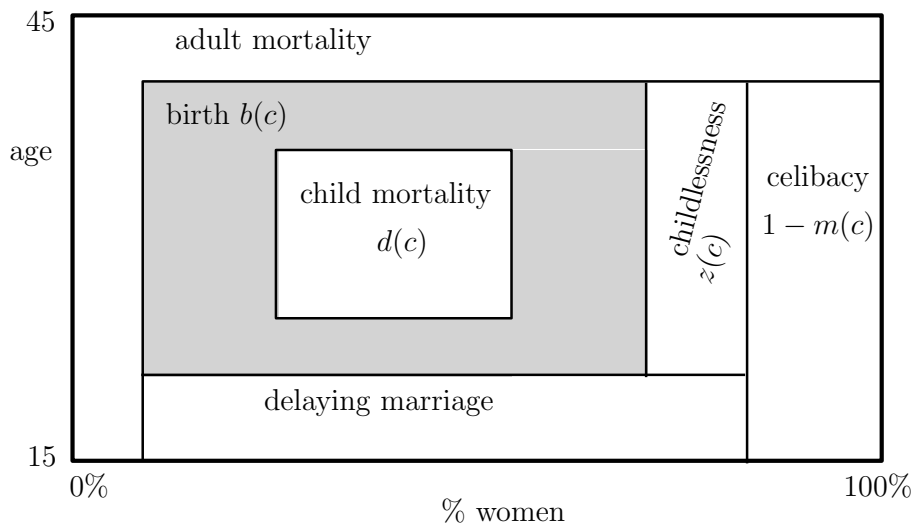
Figure 1: Schematic View of How the EMP Relates to the Four Margins of Reproductive Success

Weir (1984) reworked Wrigley and Schofield's figures arguing that most of the change in fertility in the seventeenth century was driven by changes in the rate of celibacy with a shift toward age at marriage in the eighteenth century. This was quite a provocative idea since it challenged the then dominant servant model of the preventive check in England first proposed by Malthus (1798) and refined by Hajnal (1965) whereby in poor economic conditions, servants delayed marriage until they could accumulate the resources to set up a new household. If fertility was being driven by the decision of whether or not to marry at all, then other mechanisms had to be at play.

Relative to this earlier literature, this paper makes three main contributions. First, it provides the first analysis of celibacy from micro-data rather than from aggregate statistics. Second, it suggests that there are substantial differences in celibacy across social groups and highlights the heterogeneity of experience of marriage patterns within a society. Third, it considers the influence of celibacy on fertility alongside the other three margins influencing fertility, providing a greater understanding of the extent to which celibacy matters to overall fertility.

2.3   Previous Literature on the Extensive Margin: Childlessness

On the second aspect of the extensive margin of fertility, childlessness, very little has been said for pre-industrial societies. It is often assumed that being married implied a desire for children, and the observed childlessness rate is a measure of natural sterility in a population (Leridon 2008). For England, Wrigley et al. (1997) discuss sterility as a part of their discussion of marital fertility, mainly attributing sterility to biological factors. They argued that the main factor influencing entry sterility, i.e. married couples who did not have children, was the

mother's age at marriage since only 2.6 and 3.8 per cent of women who married at ages 15-19 or 20-24 respectively never bore children whereas 69.3 per cent of women who were first married at the age of 40-44 never bore children. Across all ages entry sterility varied between 8.3 and 11.5 per cent during the early modern period generally following overall trends in fertility (p. 403). Hollingsworth (1965), who studied the demographics of British nobles based on genealogical records, finds an average rate of childlessness at 24 per cent, based on all persons ever married, a surprisingly high rate compared with other populations and the average entry sterility rates for English provincial parishes (Wrigley et al. 1997, pp. 395-397).

Beyond biological factors, there might also be economic determinants of childlessness. Three recent papers stress those determinants for US data in a historical perspective. Gobbi (2013) shows how childlessness rates and fertility rates co-move over time as a function of shocks to the gender wage gap and to the cost of having children. Aaronson, Lange, and Mazumder (2014) focus on a quantity-quality trade-off faced by parents and look at how the Rosenwald Rural Schools Initiative in the early twentieth century affected fertility along both the extensive and intensive margins. They show that the expansion of schooling opportunities decreased the price of child quality. This reduction in the price of child quality decreased the proportion of women with the highest fertility rates as expected, but it also led to a decrease in childlessness rates as more women entered on the extensive margin. In addition, Baudin, de la Croix, and Gobbi (2015) provide a framework to understand the deep causes of childlessness and how their importance has changed over time. Using US census data, they find that the main causes of childlessness across the past 100 years have changed from poverty-induced necessity to a choice driven by higher levels of income and education among women.

Our study contributes to this literature by estimating variations in childlessness across social classes in early modern England and seeking to understand what voluntary and involuntary mechanisms may drive social differences in childlessness.

2.4   Previous Literature on the Survival of the Richest

The question of reproduction and its link to the dynamics of development and inequality goes back to work on differential fertility by Althaus (1980) in a neo-classical growth model, de la Croix and Doepke (2003) in an endogenous growth model and Galor and Moav (2002) in a unified growth framework. Galor and Moav (2002) model the industrial revolution as resulting from a change in the composition of the population. They suppose preferences are of two types, quantity (Q) and quality (q). Couples endowed with Q prefer, *ceteris paribus*, having more children with low investment while couples endowed with q like to invest in the human capital of their children. If during the Malthusian period of stagnation *quality* couples benefit from an

evolutionary advantage, then their share in the population rises slowly over time and eventually leads to a take-off in income. These couples would have an evolutionary advantage having more grandchildren than their more fertile counterparts because human capital accumulation would increase earning possibilities for their children and higher earnings would allow their children to have more offspring. This view has been supported by empirical work by Galor and Klemp (2014) using Canadian genealogy data. Galor and Moav (2002) model *quality* people as those who value the education of their children, but these people could also be those with an entrepreneurial spirit (Clark 2007, Galor and Michalopoulos 2012). However, for this mechanism to work, there must be strong persistence in the inherited traits across generations whether the traits are inherited through genetic or cultural transmission (Bowles 2007).

While many studies have explored the determinants of fertility in the past, three studies have focused specifically on the relationship between income and reproduction in England. In their seminal article, Clark and Hamilton (2006) tested the hypothesis that the rich had higher net reproduction rates than the poor. They motivated their study using a basic Malthusian framework involving two components: a positive link between income and fertility and diminishing returns to labour in production. The combination of these components produces a long-run equilibrium where income is maintained at the level of subsistence and population growth is absent. However, because income in this equilibrium was not distributed equally across the population, the rich would leave more surviving offspring than the poor. Clark and Hamilton provided evidence for their theory by analysing wills in South-East England between 1585 and 1638 (see Figure B.1), using wealth at death as a proxy for income and children mentioned in the will as a proxy for surviving offspring. They found that the richest testators left twice as many offspring as the poorest. Clark and Cummins (2015) extended Clark and Hamilton's initial analysis by collecting wills from the same counties from 1500-1914. They found that the rich had higher net reproduction throughout the early modern period until 1780 when fertility began to fall among the wealthiest tercile of testators. Interestingly, they also regressed net reproduction on occupational status and wealth and found that their occupational status dummies became insignificant when controlling for wealth decile.

The probate records that Clark and Hamilton (2006) and Clark and Cummins (2015) rely upon have some strengths: they capture all probated men dying in a given geographical area so the only censoring from migration is for migrants abroad or deaths at sea; fertility is measured across all marriages; and some form of wealth is ascribed to all probated, so there are no issues related to the recording of occupations. However, the will evidence is not perfect. It only exists for men who left a will or were probated. Most rich men were probated, but fewer poor people were. Thus, if the poor people who were probated were non-randomly selected from the population of poor people, there could be potential biases in the probate sample. Clark and

Cummins (2015) show that there were not substantial differences in marriage rates and fertility between probated and non-probated men in the nineteenth century, but it is not clear whether this pattern would hold for earlier centuries. In addition, because wills did not always report all surviving children and did not always provide complete wealth information, i.e. real estate value, land area or personal wealth were missing, Clark and Cummins had to impute these values for a large share of their data. Their net fertility estimates are very high relative to the net reproduction rates estimated by Wrigley et al. (1997) for early modern England.[1] Their will database is also restricted to three counties in southeastern England: Suffolk, Essex and Surrey. These areas were more highly developed in the middle ages and early modern period and were far from the centres of the Industrial Revolution in the Northwest and West Midlands. Thus, the Wrigley et al. (1997) data provides substantially better geographic coverage than the wills database (see Figure B.1). We also believe that our more comprehensive measure of class, discussed in detail below, will better capture the fertility behaviour of the lower classes than the will data. These weaknesses, mostly a product of the historical sources that survive, do not in themselves invalidate these authors' earlier results, but they do highlight the value of testing their hypothesis with a different dataset.

In addition to the papers based on probate evidence, Boberg-Fazlic, Sharp, and Weisdorf (2011) used the same family reconstitution dataset used in the current study to test the relationship between wealth and fertility. However, rather than measuring wealth directly, Boberg *et al.* used occupational titles to sort individuals into social groups that were roughly correlated with wealth. Boberg *et al.* also found that the wealthiest occupational groups had higher fertility up until *c.* 1800 after which they were outperformed by the poorest groups. This shift in evolutionary advantage might have emerged with the ending of the Malthusian era (Clark 2007, Galor 2012) and with the rich beginning the demographic transition (fertility decline) earlier than the poor (Livi-Bacci 1986, de la Croix 2012 in Chapter 3). However, Boberg *et al.* measure fertility conditional on engaging in marriage and giving birth, so they ignore the extensive margin of fertility, which we find to be so important in this paper. Thus, the potential weaknesses in the will data and methodological problems with the earlier parish reconstitution studies provide justification for revisiting the survival of the richest question in this study.

2.5 Previous Literature on the European Marriage Pattern

Finally, we contribute to the literature on the European Marriage Pattern (EMP), which is one of the strongest models explaining how demographic behaviour influenced economic devel-

---

[1]The Wrigley et al. (1997) net reproduction rates have to be multiplied by two to capture both males and females, but they average 2.37 from 1536-1780 far lower than the values coming from the Clark and Cummins (2015) will database.

opment in the pre-Industrial era. Hajnal (1965) first described the pattern and it has been expanded since then. Hajnal argued that there were key differences between the demographic and social context in Northwestern Europe and elsewhere in the world. In Northwestern Europe, high rates of celibacy, late ages of marriage and the formation of new, nuclear households at marriage created an economic system that prevented population growth and fostered economic development.

There are two explanations of the origins of the EMP. De Moor and Van Zanden (2010) argue that norms of consent among marriage partners which emerged in Catholic doctrine in the ninth century led to more respect for and empowerment of women, producing the EMP, whereas Voigtländer and Voth (2013) argue that the population shock of the Black Death and women's comparative advantage in the pastoral sector improved women's labour market opportunities and led to later marriage. In the more traditional argument, the EMP influenced economic development by reducing fertility and preventing population growth from capturing all of the growth achieved through improvements in technology or extensive expansion in the area under agriculture (Voigtländer and Voth 2013).

However, the schematic and somewhat simplified view of the EMP in England has recently been challenged on a number of fronts. First, Dennison and Ogilvie (2014) have conducted a meta-analysis of the demographic characteristics of a number of European countries on both sides of Hajnal's line. They collected systematic information on female age at first marriage, celibacy rates and household complexity and found no correlation between key elements of Hajnal's European Marriage Pattern and economic development. In addition, two other papers have shown that people living in early modern England may have had a larger degree of fertility control than had been assumed, a fundamental breach of the traditional Malthusian population model. Kelly and Ó Gráda (2012), studying counts of births in 404 English parishes, found evidence that births responded to economic conditions in early modern England with parishes having fewer births when economic conditions were bad. Cinnirella, Klemp, and Weisdorf (2017) also found that the probability of having a child was lower when conditions were bad analysing the individual-level family reconstitution data employed in this paper. If families were able to limit their fertility within marriage, then the rates of celibacy and age at marriage might have been less important than the proponents of the EMP have thought.

Our contribution to this literature is to analyse the workings of the EMP within wealth groups of one country to give an idea of how the EMP was practiced within pre-Industrial English society. This analysis will reveal the extent to which the EMP factors were heterogeneous across groups.

3 Database

3.1 The Data

We use the family reconstitution data collected by the *Cambridge Group for the History of Population and Social Structure* and described in detail by Wrigley et al. (1997). The full dataset includes over 300,000 individuals recorded in registers coming from a total of 26 provincial, English parishes (see Figure B.1). The data are drawn from baptism, marriage and burial registers kept by the local parish priest, which have been systematically linked so that individuals can be tracked across their lifetime from birth, to marriage, to the birth (and potentially burial) of their children, to their own burial (Wrigley 1966). In addition to the demographic events present in the registers, male occupations were often recorded upon a man's marriage or burial or when his children were baptised. It is from these occupations that we construct our social groups.
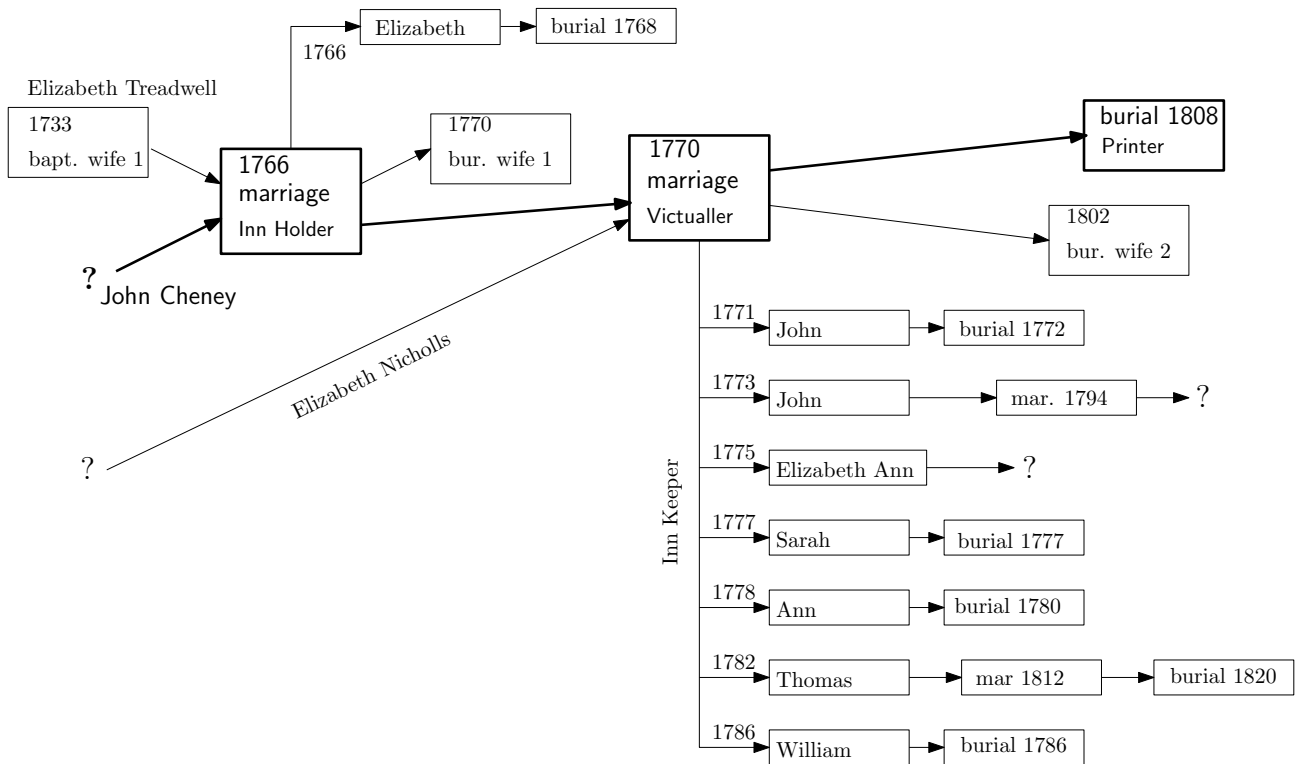
Figure 2: The Reconstitution Data in Miniature - The life of John Cheney from Banbury

To provide an idea of how the reconstitution data work, Figure 2 shows the example of John Cheney from the parish of Banbury. John was born outside the parish, so we do not observe his birth date in the baptism register. He appears in our dataset when he marries Elizabeth Treadwell in Banbury in 1766. Unfortunately, Elizabeth and their only child died by 1770,

11

so in the same year, John married Elizabeth Nicholls, with whom he went on to have seven children. Among the children from his second marriage, four died within two years of birth. The youngest child, Thomas, married and died at age 38 in the parish. We assume Elizabeth Ann migrated out of the parish since we do not observe her marriage or burial. Finally, John's son John (the surviving one) married in the parish and then migrated since we do not observe his burial. John's second wife died in 1802, and he passed away in 1808. We can also track John (the father)'s occupation over time. His occupation is initially recorded at his first marriage as inn holder. Then, by his second marriage in 1770 he has become a victualler, though he is listed as an inn keeper in the baptism register in relation to the birth of his children (1771-86). At his burial in 1808, John was listed as being a printer, suggesting substantial upward social mobility in his lifetime. John's upward mobility is corroborated by the fact that the Cheney family printers published a special book in 1967 (Cheneys 1967) to celebrate their 200-year publishing history in Banbury beginning with John Cheney. Clearly, the family reconstitutions provide a wealth of data on the early modern English society and economy.

The 26 parishes were chosen by the Cambridge Group because they were geographically dispersed across England (see map in Appendix B Figure B.1) and had high data quality. They are representative of rural areas and small market towns in the early modern England but would not capture the effects of the largest urban areas such as London. While the full data cover 1541 to 1851, not all of the 26 parishes span the entire time period: some parishes only contain records for earlier or later periods. Moreover, there was a rise in non-conformism in England at the end of the eighteenth and first half of the nineteenth century, which meant that in some parishes where non-conformism was important, a lot of vital events went unreported in Church of England registers. Thus, the years for which the data are reliable are different from the years for which reconstitution data exist.

Fortunately, Wrigley et al. (1997) conducted a careful analysis of the reliability of the registers and report the corresponding dates for each parish (p. 22-23). We accordingly restrict our analysis based on the reliable start and stop year for each parish. We require the marriage date to be 15 years before the end of the reliable period. This strategy ensures that we observe nearly all of a woman's potential children, but we would not lose observations because we did not observe the mother's death date until after the records became less reliable. The restriction to the data leads to a loss of individual observations of 21.6 per cent.

This restriction is necessary because we only know whether an individual was present in a parish when we observe a related baptism, marriage or burial. Imagine we observe a woman getting married five years before the end of the reliable period, but we do not observe her children's births or her death, which presumably occurred outside the reliable period. We cannot include the exposure of this woman up to the end of the reliable period by censoring at the end of the

reliable period formally in a hazard model because we cannot know whether she remained in the parish childless and her death went unrecorded or whether she migrated outside the parish and had children. This restriction ensures that we exclude incomplete data at the edge of the reliable period.

3.2   Social Groups

In order to measure the four margins of fertility across a social gradient, we must first sort occupations in our dataset into meaningful social groups. There are a number of ways that we could sort the occupations, including by HISCLASS (Van Leeuwen and Maas 2011) or the nineteenth-century British professional groups (Armstrong, 1972, Long, 2013), but we have opted to follow the previous literature and use Clark and Hamilton (2006)'s wealth categories which were also employed by Boberg-Fazlic, Sharp, and Weisdorf (2011) and Clark and Cummins (2015). These reflect occupations that were frequently recorded in wills by wealth and social status. However, one limitation of our study is that we cannot observe wealth directly as Clark and Hamilton (2006) and Clark and Cummins (2015) do. Some farmers in our data may have been wealthier than some of the merchants and lower gentry. Thus, these occupational groups reflect a combination of social status and wealth.

Group one consisted of mainly labourers and servants. Group two contains small farmers (husbandmen) and rural, low-skill manufacturing workers such as weavers. Group three was mainly made up of craftsmen such as tailors and skilled construction workers, whereas group four were traders and lower working proprietors. Group five consisted of farmers, which unlike the poorer husbandmen, had large amounts of agricultural lands and employed wage labor to work their farms. Thus, they were of a higher wealth and social group than the small-scale farmers and labourers in groups one and two.[2] Group six contains merchants and professionals such as clerks, clerics and higher working proprietors. Finally, group seven represents the landed gentry who often had occupations listed as gentleman or esquire.

There was a small degree of occupational mobility within an individual's lifetime. We discuss this at length in Appendix C. There is also a concern with our assumption that the husband's occupation can be taken as a useful proxy of the household occupational status, even when we are calculating net reproduction for women. From a historical perspective, this is justified. Early modern England was a patriarchal society where men had many more opportunities and advantages than women, making the husband's occupational status particularly salient

---

[2]As is clear in our demographic measures below, farmers are often an outlier group. We suspect that farmers organised themselves differently with regards to family planning than other segments of society for two reasons. They were much more closely affected by inheritance rules and the division of farm land between offspring, and as food suppliers they responded in the opposite manner to variations in food prices than all other groups who were food demanders.

for fertility decisions. For example, from the 13th or 14th century to the late 19th century, married women were subject to the laws of coverture, which severely limited their legal and economic status (Blackstone and Kerr 1896, Holcombe 1983). However, it is true that if people regularly married outside their occupational group, defined by the occupation of their father, this could attenuate differences in net reproduction across groups. Accordingly, we attempted to measure marital homogamy in our data. We were only able to find both the groom's father's occupation and the bride's father's occupation for 4.5 per cent of married couples because we cannot observe these occupations for in-migrants. 42.7 per cent of these couples married within their own social group. However, this figure is a severe underestimate of the degree of assortative matching since these couples chose from the smallest marriage pool possible: only people born and never migrating from one parish. Therefore, there are likely many individuals who migrated to find a partner in their own group or imported one into their own parish, especially among the wealthier classes. We see this to be true when looking at emigrants and immigrants below (see Tables 3 and 4).

In our final analysis of fertility, we have chosen to combine group six, merchants and professionals, and group seven, gentry. We do this because our sample size for group seven is very small on certain margins and the demographic behaviour of the two groups was very similar. Combining the two groups also adds analytical clarity since both would have been the wealthiest in society at the time. Thus, we are left with six occupational groups, across which we can test for differences in marriage rates, childlessness, marital fertility and child mortality. However, we still include the people for whom the father's occupation is unknown (59.8 per cent) in the regressions, as these contribute to identifying the baseline hazard rate and the parish and time fixed effects. Our results are robust to excluding unknown occupations (not reported).

3.3 Human Capital and Literacy

In an ideal world, to test Galor and Moav (2002)'s theory that quality-oriented families who invested in human capital had a survival advantage, we might want to study fertility differences between groups stratified by levels of human capital. The only direct measure of human capital in our dataset is signature literacy, but this is unfortunately only available in a sub-sample of parishes after 1754 when it became a legal requirement for both bride and groom to sign the marriage register (Schofield, 1981, p. 201). We do not incorporate literacy into our analysis for two reasons. First, by 1754, male literacy was no longer a strong signal of human capital. Male literacy had reached 56% by 1750 in England (de Pleijt 2018) and was substantially higher for the middle and upper classes (Cressy 1981; Schofield 1981). Literacy was a precursor to retaining productive knowledge (de la Croix, Doepke, and Mokyr 2018), but its crude, binary character as a variable means that it cannot capture how knowledge was instrumentalised, nor

can it distinguish quality of education. It is not a good proxy for upper-tail human capital (Mokyr 2005), which was important for industrial development in France (Squicciarini and Voigtländer 2015). Second, most of the growth in male literacy occurred before 1754, so any differences in net fertility after 1750 may not be instructive. This is especially true since several studies have found declining fertility among the wealthy after 1780 (Boberg-Fazlic, Sharp, and Weisdorf 2011; Clark and Cummins 2015). For these reasons, we believe that our occupational groups are better at capturing differences in human capital (and human capital investment) than a simple literacy variable. The groups capture how human capital was instrumentalised since the middle classes would use their literacy differently than the lower and upper classes and capture differences in schooling quality. This corresponds well to other studies which have found that occupational titles are a far better proxy for productive knowledge and skill than literacy (de Pleijt, Nuvolari, and Weisdorf 2018).

4   Measuring the Four Margins of Fertility

We can now turn to measuring the four margins of reproductive success laid out in the theory above. Recall that net reproduction $n(\cdot)$ as a function of social class c is:

$$n(c) = m(c)(1 - z(c))b(c)(1 - d(c)),$$

where $m(c)$ is the percentage of women who marry, $z(c)$ is the fraction of childless married women, $b(c)$ is the number of births conditional on being married and having children, and $d(c)$ is the mortality rate of children aged 0 to 15. When studying childlessness (or fertility at other parities), prior studies have usually relied on so-called completed marriages, marriages where both spouses survive until the woman reached the age of 50 (Wrigley et al. 1997). The rates childlessness ($z(c)$), marital fertility ($b(c)$) and child mortality ($d(c)$) can then be estimated using a sample of women who reached the end of their reproductive period, say aged 50+. However, as we are interested in the reproductive success of different social groups, we should not neglect the contribution of women who died prematurely, i.e. before age 50, to reproduction. Although these women will have fewer children than those with completed fertility, their children still contribute to reproductive success. We therefore use Cox Proportional Hazard models for censored data to estimate the four margins, $m(c)$, $z(c)$, $b(c)$ and $d(c)$. Following the Cox model, the risk $h(t, x, \beta)$ (of getting married, of having a child, of dying) of an individual with characteristics x changes with the survival time t according to

$$h(t, x, \beta) = h_0(t) \exp[x\beta],$$

where $h_0(t)$ is the baseline hazard. As regressors x we use the six occupational dummies defined above plus a dummy for individuals of unknown occupation, and we control for parish and time fixed effects by including twenty-six parish dummies and four time dummies (one for each quartile of the data ranked by year of marriage: 1538-1649, 1650-1718, 1719-1769, 1769-1836) in the regressions.[3] The time dummies refer to the parents' marriage year for $m(c)$ and $d(c)$ and to a couple's marriage year for $z(c)$ and $b(c)$. The reference category is a Labourer/Servant married in 1650-1718 in the parish of Alcester.

The estimated vector of parameters $\hat{\beta}$ is found by maximising the partial likelihood, which does not depend on how the baseline hazard $h_0(t)$ is specified. However, to compute $m(c)$, $z(c)$, $b(c)$ and $d(c)$, we need to estimate the cumulative hazard function and the corresponding survival. The baseline hazard, required in this computation, is obtained using the Nelson-Aalen estimator. Denoting t_1, t_2, ... the distinct event times, the expected number of events at t_i can be obtained by summing the hazards over the population at risk:

$$E(d_i) = \sum_{j:t_j > t_i} \hat{h}_0(t_i) \exp[x_j \hat{\beta}].$$

where d_i is the number of events at t_i and the sum is taken over the individuals j that are still at risk at t_i. Replacing the expected number of events $E(d_i)$ by their actual number d_i gives an estimate of $\hat{h}_0(t_i)$.

4.1   Marriage Rate $m(c)$

Although family reconstitution data, which are centred around marriages, may seem at first to be less than ideal for analysing people who never married, we can study lifetime celibacy by analysing the life courses of the children from each marriage. By looking at cohorts of births, we are able to compute the share of unmarried individuals by their age of death, providing a full picture of celibacy across the entire lifecycle.

Rather than following the literature so far by measuring celibacy among women surviving past a certain age (say 40 or 50) (Weir 1984; Schofield 1985), we use Cox proportional hazard models to measure the risk of marriage across the life course. This method thus accounts for women who died before the cut off age, married or not. Women's risk of marriage begins when they turn 16, the age at which women were allowed to marry according to Anglican tradition, and was only censored at the upper end when the woman died. Thus, in addition to restricting the data to the reliable periods as described in Section 3.1, we exclude women who died below the age of 16 and were never at risk of marriage.

---

[3]We also ran other models with a greater number of time periods, one for each decade, but this did not influence our results.

To compute our indicator of celibacy, we take the predicted marriage rate from the estimated survival function at the age of 45, which is $m(c)$. We choose age 45 because very few marriages occurred after age 45 and very few women were able to have children after the age of 45, so marriages beyond 45 were unlikely to influence net fertility across social groups. Because we do not observe women's occupations directly nor husband's occupations for women who never married, we use the occupation of a woman's father to place her into one of the social groups when estimating the risk of marriage. In addition, we impose the following requirements on the data to construct the sample used to predict the survival curves: we count women in the data as married if they (i) were baptised and buried in the sampled parish; (ii) survived beyond the age of 16 as mentioned above; and (iii) were registered as being married or having baptised a child in the parish. Lifetime celibate women satisfied conditions (i) and (ii) but were not registered as married or baptising children in the parish.

Table A.2 in Appendix A provides information on the sample of women for whom we calculate the marriage rate. There are a total of 9,007 women satisfying (i) and (ii), and 58.9 per cent of them are registered as married, hence satisfying (i), (ii), and (iii). This initial marriage rate is low because it includes many women who died before getting married, so to account for this, we estimate the risk of marriage using a Cox proportional hazard model.

Column (1) in Table 1 displays the results. Here we have reported coefficients, but to interpret the Cox model, it is often helpful to calculate a hazard ratio by exponentiating the coefficient, similar to an odds ratio for logistic regressions. The upper classes have a hazard ratio of 0.733 ($e^{-0.310}$) which means that they have a 36.3% lower risk of being married by the end of the period at risk than the reference group of labourers and servants. Thus, daughters of the upper classes had substantially lower risks of marriage, but daughters of craftsmen, traders and the unknown occupations were also less likely to marry than the daughters of laborers and servants.[4] These results are robust to excluding unknown occupations (not reported). The top left panel of Figure 3 shows the marriage rate across the social groups predicted from the survival curve at age 45 with average levels imputed for all parish and time dummies. As was clear from the regressions, there is a social gradient in marriage rates. Poorer social groups were much more likely to marry with marriage rates around 80%, but the merchants/professionals and gentry had much lower rates of marriage around 68%. The marriage rates are similar to other

---

[4]The Cox proportional hazard model assumes, as its name indicates, that the hazard rate is shifted proportionately by the control variables. It implies that the survival functions for different social classes change proportionately and do not, for instance, cross each other. One test of the proportionality assumption (by Grambsch and Therneau (1994)) is obtained by computing, for each control variable, the scaled Schoenfeld residual, and by correlating it with a transformation of time (Hosmer and Lemeshow, 1999, pp. 197-205; Schoenfeld, 1982); proportionality is rejected if the correlation is statistically significant. In the case of our regression, the correlations are respectively 0.01 (Husbandmen), 0.01 (Craftsmen), 0.00 (Traders), -0.02 (Farmers), -0.00 (Upper class), and none of them is statistically different from zero.

	(1)	(2)	(3)
	risk of	risk of	risk of
	marriage	first birth	child death
Occupations	$m(c)$	$z(c)$	$d(c)$
Husbandmen	−0.067	−0.030	−0.001
	(0.065)	(0.041)	(0.039)
Craftsmen	−0.204***	−0.013	0.039
	(0.063)	(0.039)	(0.036)
Traders	−0.162**	−0.038	0.067
	(0.079)	(0.051)	(0.046)
Farmers	0.119	−0.203***	0.004
	(0.091)	(0.062)	(0.062)
Upper Class	−0.310***	−0.204***	0.071
	(0.086)	(0.053)	(0.048)
Unknown Occupation	−0.097*	−0.284***	0.001
	(0.052)	(0.032)	(0.032)
Observations	9,007	14,730	48,514

Note: Reference category is Labourers/Servants. 26 parish dummies and 4 period dummies are included. *p<0.1; **p<0.05; ***p<0.01. Data are restricted to the reliable periods determined by Wrigley et al. (1997, p. 22-23) as described in Section 3.1 before running the models. We censor formally within the Cox models at the woman/mother's death in the case of $m(c)$ and $z(c)$ and at age 15 for children in $d(c)$. See text for more detail.

Table 1: Cox Proportional Hazard Model. Analysis of $m(c)$, $z(c)$, $d(c)$

estimates of marriage rates from historical sources. Weir (1984) and Schofield (1985) find only slightly higher aggregate marriage rates, and (Froide 2005) found similar rates in urban areas. The farmers were a bit of an outlier (as they are in some of the other calculations as well). Thus, we observe a strong social gradient in celibacy with the richest experiencing much higher rates than the poorest.[5]

4.2 Childlessness rate $z(c)$

In order to compute the rate of childlessness (i.e. the share of married couples who never have a child), we again use Cox proportional hazard models to estimate the risk of a married couple having a first birth. Couples become at risk of a first birth upon marriage and censoring occurs

---

[5]We tested for regional variation in the marriage rates by grouping our parishes into three regions: North, Midlands and South. The upper classes had lower risks of marriage in all regions and statistically signicant differences in the Midlands and South. However, the gradient appears to hold across all regions (not reported).

when the mother dies.[6] We take the predicted childlessness rate for each social group after ten years as our indicator of childlessness in the net reproduction calculations.

We limit our sample to couples whose burial dates are known to ensure that we do not attribute childlessness to couples that migrated out of the parish. As mentioned above, we also restrict the sample to first marriages in the calculations of $z(c)$ and $b(c)$ for a number of reasons. First, we want to avoid double counting of mothers and fathers who may have been childless in the second marriage but not the first. Second, we want to eliminate age effects on sterility in second marriages since the average age of women in second marriages was substantially higher: the intervals between first and second marriages for women were 43.7 months across the period (Wrigley et al., 1997, p. 172). There is also a lot of uncertainty about the marriage order of many people in our dataset since we cannot know whether an individual who migrated into a parish was married previously. The data often only state that the marriage order was greater than or equal to the first or greater than or equal to the second.

Finally, Wrigley et al. (1997) argue that it is impossible to know the precise rates of remarriage from family reconstitution data because second marriages can only be identified by linking to a birth record. Thus, the differential propensity to link different types of individuals leads to bachelor/spinster marriages being overrepresented and the other marriage types being under-represented (Wrigley et al., 1997, p. 164). It is impossible to correct for this without a full census of the population with marriage orders, which does not exist for our data. To test the prevalence of remarriage across the social groups, we assume that the marriage order was the lowest possible order listed, i.e. we assume that in-migrants' marriages were first marriages, and look at the mean marriage order of men and women across our occupational groups (see Table B.4). Although men have higher marriage orders than women, the mean marriage order is very stable across the occupational groups and is unlikely to affect our results for calculating $z(c)$ or $b(c)$.

Table A.3 shows that out of 14,730 women in the sample, 15 per cent remained childless. This is a very high number compared to the usual estimations of natural sterility,[7] but below the childlessness rate of the nobles reported in Hollingsworth (1965).

Specification (2) in Table 1 shows the results from the Cox proportional hazard model estimating the risk of a married couple having a first child using the husband's occupation to determine the social group. A social gradient is again clear with richer social groups having significantly

---

[6]Note that childbirth outside of wedlock was banned by the Church of England, and that illegitimate births, therefore, were rare ranging from 1.2 per cent to 6.0 per cent of births across our period (Wrigley et al., 1997, p. 224).

[7]Leridon (2008) measures natural sterility in the Henry database for rural eighteeth-century France when fertility control was ineffective. Restricting the sample to couples where the husband and the wife were still living together at age 50, 3.7 per cent of women who married at age 20-24 remained childless.

Marriage Rate $m(c)$

Childlessness Rate $z(c)$
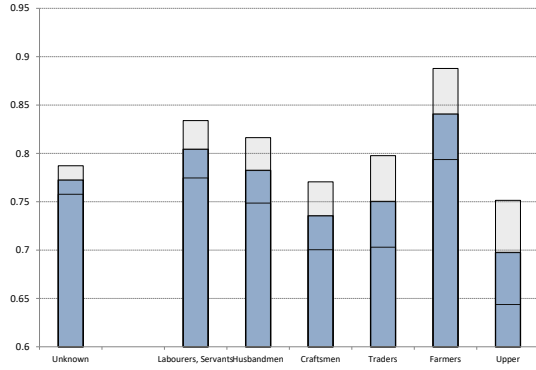
Note: Estimated survival $1 - m(c)$ computed at age 45, with 95% confidence intervals
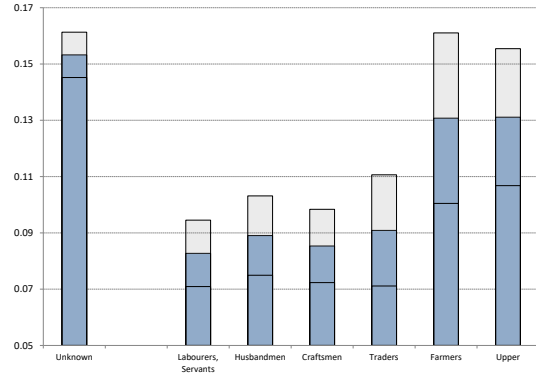
Note: Estimated survival $z(c)$ computed 10 years after marriage, with 95% confidence intervals

Births per Woman $b(c)$

Child Mortality Rate $d(c)$
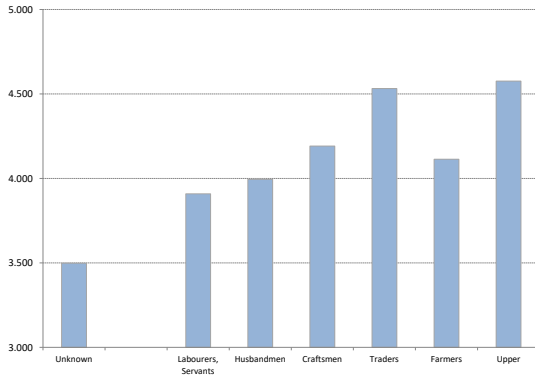
Note: Birth $b(c)$ computed from parity progression ratios using Equation (5)
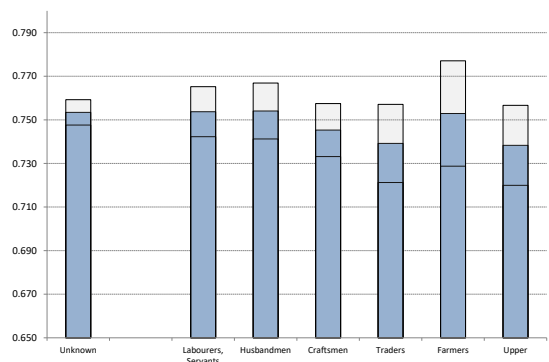
Note: Estimated infant survival rate $1-d(c)$ computed at age 15, with 95% confidence intervals

Figure 3: Computed Margins Across Social Groups

lower risks of a first birth, which implies higher rates of childlessness.[8] Again, these results are robust to excluding unknown occupations (not reported).

The top right panel of Figure 3 shows the predicted childlessness rate at ten years after marriage across social groups (with average levels for parish and time period dummies) and highlights again the strong social gradient in childlessness similar to celibacy. Marriages in wealthier social groups were more likely to remain childless relative to the poor. These rates of childlessness

---

[8]We also test for proportionality here. The correlation between the scaled Schoenfeld residual and transformation of time are respectively 0.01 (Husbandmen), 0.00 (Craftsmen), 0.01 (Traders), 0.00 (Farmers), 0.03 (Upper class), and none of them is statistically different from zero, but for the Upper class. In addition to the proportionality assumption, it is also possible to test the assumption of the Cox model that censoring is independent from survival time. We have reasons to believe that it might be violated in the case of childlessness, as sterility and death have common unobserved determinants. We consider this issue in Appendix I, using the method proposed by Jackson et al. (2014) and find that it is not important for our results.

are rather high, so they are discussed in more detail in Section 6.[9]

4.3 Marital fertility rate $b(c)$

We calculate marital fertility conditional on having a child by again using Cox proportional hazard models to predict parity progression ratios, P_n. P_n is the risk that a couple has an additional child conditional on having reached a given parity n and is given as follows:

$$P_n = P(n+1|n) \tag{3}$$

At parity zero, this is the opposite to the childless, $z(c)$, calculated above. At parity one, we measure the risk that a couple will have a second child conditional on them having had a first child. The period at risk begins at the birth of the first child and the event is closed when a second child is born. Right censoring occurs when the mother dies. Again we limit the sample to couples where the burial dates are observed for both the mother and father and to first marriages. We predict the parity progression ratios from the survival curves ten years after the previous birth. Table A.5 shows the characteristics of the sample used to predict the parity progression ratios. The first line shows the number of couples at risk for a second birth conditional on having a first birth. Of the 12,519 at risk of a second birth, 85.8 per cent had a second birth. The percentage of couples having an additional child declines with parity as we would expect.

Table 2 shows the coefficients of Cox regressions modelling the probability to move to the next parity conditionally on having one to nine children, using the husband's occupation to classify the couples into the social groups. We stop the estimation at $P(10|9)$ because after parity ten the sample size becomes too small. Here we see that the upper classes have higher probabilities to have an additional child, in particular when they already have three.[10]

Table B.2 shows the predicted parity progression ratios for various social groups from the regressions of Table 2 assuming average levels for parish and time period dummies. Lower parity progression ratios across the parities will produce a lower gross fertility rate for the social group. Thus, at first glance there appears to be a social gradient with richer groups

---

[9]Again, we tested for regional variation in the childlessness rates by grouping our parishes into three regions: North, Midlands and South. The upper classes had lower risks of first birth in all regions but the differences were only statistically significant in the South. However, the gradient appears to hold across all regions (not reported).

[10]The proportionality of the hazard rates is less clearly supported for the birth estimations than for the two extensive margins. Among the 63 coefficients in Table 4, proportionality is rejected in 26 cases (at the 5% level). However, this non-proportionality is not overly concerning because we are estimating the parity progression ratios over a short period of time (the cumulated hazard over ten years rather than 30 years for marriage rates), and non-proportionality is less likely to be an issue over short time intervals (Bellera et al. 2010).

	(1)	(2)	(3)	(4)	(5)
	risk of	risk of	risk of	risk of	risk of
Occupations	second child	third child	fourth child	fifth child	sixth child
Husbandmen	0.099**	0.070	0.051	0.087	0.049
	(0.044)	(0.048)	(0.053)	(0.060)	(0.070)
Craftsmen	0.126***	0.047	0.100**	0.166***	0.149**
	(0.042)	(0.046)	(0.050)	(0.056)	(0.064)
Traders	0.182***	0.166***	0.231***	0.280***	−0.014
	(0.055)	(0.059)	(0.064)	(0.072)	(0.084)
Farmers	0.196***	0.117*	0.164**	0.043	−0.032
	(0.065)	(0.069)	(0.075)	(0.085)	(0.100)
Upper Class	0.197***	0.180***	0.254***	0.183**	0.200**
	(0.057)	(0.061)	(0.066)	(0.074)	(0.085)
Unknown Occupation	−0.070**	−0.102***	−0.056	−0.018	−0.029
	(0.035)	(0.038)	(0.042)	(0.048)	(0.056)
Observations	12,519	10,738	8,993	7,280	5,679

	(6)	(7)	(8)	(9)	
	risk of	risk of	risk of	risk of	
Occupations	seventh child	eighth child	nineth child	tenth child	
Husbandmen	−0.067	0.103	0.216	−0.049	
	(0.083)	(0.100)	(0.132)	(0.186)	
Craftsmen	0.023	0.007	0.270**	0.049	
	(0.075)	(0.092)	(0.119)	(0.159)	
Traders	0.234**	−0.082	0.308**	0.325*	
	(0.096)	(0.122)	(0.153)	(0.197)	
Farmers	0.110	0.069	−0.055	−0.288	
	(0.115)	(0.142)	(0.198)	(0.294)	
Upper Class	0.265***	0.300**	0.551***	−0.070	
	(0.098)	(0.117)	(0.143)	(0.197)	
Unknown Occupation	−0.068	0.076	0.213**	−0.116	
	(0.067)	(0.081)	(0.106)	(0.146)	
Observations	4,201	2,959	1,932	1,154	

Note: Reference category is Labourers/Servants. 26 parish dummies and 4 period dummies are included. *p<0.1; **p<0.05; ***p<0.01. Data are restricted to the reliable periods determined by Wrigley et al. (1997, p. 22-23) as described in Section 3.1 before running the models. We censor formally within the Cox models at the mother's death at each parity. See text for more detail.

Table 2: Cox Analysis of Parity Progression Across Social Groups

having higher fertility. Following Henry (1953, chapter 14) and Feeney (1986), these parity progression ratios can be used to calculate the completed fertility using the following formula:

$$\text{Mothers' Completed Fertility} = 1 + P_1\left(1 + P_2\left(1 + P_3\ldots\left(1 + \frac{P_n}{1 - P_n}\right)\right)\right) \quad (4)$$

However, we are interested in the fertility of incomplete families as well, so we have added survival probabilities to equation (4). In order to do this, we use logistic regression to estimate the probability that a mother will die before the birth of her next child (at parity $n + 1$) conditional on her having had a child of parity n, which we take to be within five years of the previous birth. There are not significant differences across social groups in this measure (see Table B.1 in Appendix B), but we feel that it is necessary for completeness. Thus, we can add the survival probabilities s_n to the completed fertility equation above to calculate the gross fertility of women in the sample using the following equation:

$$b(c) = s_0\left(1 + s_1 P_1\left(1 + s_2 P_2\left(1 + s_3 P_3\ldots\left(1 + \frac{s_n P_n}{1 - s_n P_n}\right)\right)\right)\right) \quad (5)$$

The bottom left panel of Figure 3 shows the final values of $b(c)$ for each social group calculated from equation 5. There is again a social gradient, which in this case mirrors the survival of the richest results found in the previous literature. The richer social groups had substantially higher fertility than the poorer ones, excepting the farmers. This could be driven by a number of factors, but the more common use of wet nurses among the upper classes may explain some of the differential (Fildes 1986, pp. 99-100). Thus, the intensive and extensive margin of fertility seem to be influencing net fertility in opposite directions across social groups.

4.4   Child mortality rate $d(c)$

Child mortality is also calculated using a Cox proportional hazard model. Exposure for children begins from birth. Events are closed if a child dies and right censoring occurs after age 15 since we are only interested in survival to marriage age. To estimate the model, we must observe the child's birth year and their mother's year of death. In addition, we restrict the data to observations where the mother's year of death is 15 years after the child's birth. This requirement ensures that all of the child's mortality exposure was in the observed parish since the father may have moved the family after the mother's death. Children who we do not observe dying are assumed to have survived to age 15. We take the predicted mortality rate from the survival curve at age 15 and use the father's occupation to place children into social groups. Table A.4 shows the characteristics of the sample used to estimate child deaths. Of the 48,514 children at risk in the sample, we have a death rate of 26.0 per cent.

Specification (3) in Table 1 shows that there is no significant social gradient in child mortality. All of the coefficients are insignificant and very close to zero.[11] Thus, differences in net fertility across groups do not seem to be related to differences in relative survival of the rich and poor.

4.5 Social Gradients in Reproductive Success

Having estimated class gradients across the four margins of fertility, we now combine them and calculate the net reproduction for each social group. We calculate two measures of net fertility. n incorporates all four margins of fertility whereas $\bar{n}$ applies the average values of the extensive margin to calculate the net reproduction rate as shown in the following equation:

$$\bar{n}(c) = \bar{m} \, (1 - \bar{z}) \, b(c) \, (1 - d(c)) \tag{6}$$
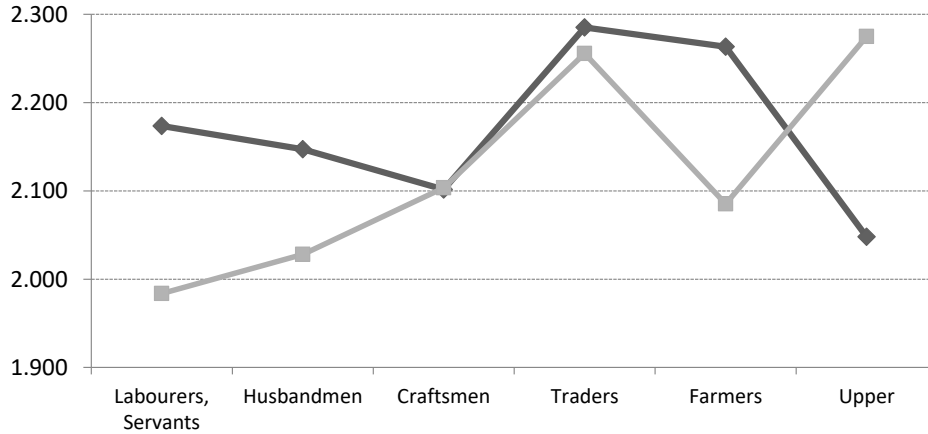
Note: dark grey line: $n(c)$. light grey line: $\bar{n}(c)$. Data from Table F.1.

Figure 4: Net reproduction rates Across Social Groups

Figure 4 plots $n(c)$ and $\bar{n}(c)$ across the social groups. Looking first at $\bar{n}(c)$, the light grey line, we see the pattern commonly reported in the literature. Richer groups have substantially higher net reproduction rates than poorer groups, confirming the survival of the richest story. Farmers had lower reproduction rates, but otherwise the social gradient is clear. However, when we include the extensive margins of celibacy and childlessness (the dark grey line), we see that the gradient changes dramatically. The net fertility rates of the upper classes decrease substantially because they have high rates of celibacy whereas the net fertility of the poor increases substantially because they have low rates childlessness.[12] The differences between

---

[11]Moreover, the proportionality of hazards is never rejected.

[12]We break the margins down further in Appendix D. This gradient also matches what one would find from

$n(c)$ and $\bar{n}(c)$ are statistically significant at the five per cent level for the upper classes and labourers and servants. The differences for farmers and husbandmen just barely miss the ten per cent threshold, whereas there are no significant differences between $n(c)$ and $\bar{n}(c)$ for craftsmen and traders as one might expect.[13] Thus, rather than seeing a more or less smooth increasing gradient in net reproduction across social groups, we see a hump-shaped pattern with the middle classes of traders and farmers having the highest net reproduction rates. This raises serious doubts about previous claims attributing an evolutionary advantage to the wealthiest classes in early modern English society.

Appendix G conducts a simple back-of-the-envelope calculation to understand whether the net fertility differences across social groups $n(c)$ had a quantitatively meaningful effect on the relative representation of each group in the English economy. We calculate relative representation across ten generations, the amount of time from the beginning to the end of our dataset, 1541-1851. The share of the middle classes increases from 13.7% to 23.6% over the ten generations, showing that the net fertility differences are quantitatively meaningful and could have led to the expansion of *quality* families as Galor and Moav (2002) suggested.

4.6   Uncertainty Surrounding the Estimates

The uncertainty surrounding each margin, $m(c)$, $z(c)$ and $d(c)$, is given by the standard errors of the estimated coefficients of the Cox regressions provided in Tables 1 and 2. To show the uncertainty around the more complex measures $b(c)$ and $n(c)$, given by Equations (1) and (5) respectively, we need to rely on simulations. We therefore generate a sample of moments $m(c)$, $z(c)$, $d(c)$ and parity progressions P_n, by drawing them from their estimated distribution (using the mean and standard error of the estimated survival, assuming they are normally distributed and uncorrelated). With these moments we can compute a sample of mothers' fertility $b(c)$ and, finally, a sample of reproduction rates $n(c)$. More details are provided in Appendix F.

These samples allow us to test whether net reproduction rates are statistically significant between social classes. We have aggregated the social classes as follows: the lower class includes Labourers, Servants, Husbandmen, and Craftsmen and the middle class includes Traders and Farmers. We find that the middle classes had a statistically signficant (one-tailed p-value=0.04) advantage in net reproduction of 0.23 children over the upper classes. However, the differences between the other groups were statistically less clear. The middle classes had 0.14 higher net reproduction than the lower classes, but this difference was not statistically significant at con-

---

a simple count of the number of children surviving to women sorted by the woman's father's occupation (see Appendix E.

[13]We use the imputation method described in the next section (section 4.6) to aggregate the uncertainty around the estimates for the four margins estimated above.

ventional levels (one-tailed p-value=0.11). The difference between the upper and lower classes was even smaller, with the lower classes having 0.09 more children, but again this difference was not statistically significant (one-tailed p-value=0.21). Thus, although we cannot be totally certain of all of the mean differences, the net reproduction of the middle classes was very likely substantially higher than the upper classes.[14] This finding clearly differs from the earlier literature.

These results could be biased if fertility of the upper classes began declining at the end of the eighteenth century as both Boberg-Fazlic, Sharp, and Weisdorf (2011) and Clark and Cummins (2015) found in earlier studies. Although we include time dummies in our regressions, these are not interacted with the social groups, so they cannot capture differences in the margins across time for each social group. Thus, to see whether our baseline estimation is sensitive to including these marriages late in our period, and whether Clark and Cummins (2015) results can be seen as one possible realization of our model, we have redone our entire analysis excluding marriages that occurred after 1780.[15] Appendix H reports these results and shows that they have little effect on the social gradient of female net reproduction. In fact, the differences between the groups become more significant. Despite the fact that the standard errors increase with the restricted sample, farmers have higher net reproduction driving the middle-class level upwards. Thus, in this robustness check, the differences in net reproduction between the middle class and upper class on the one hand and middle class and lower class on the other are both statistically significant with one-tailed p-values of 0.03 and 0.05 respectively. However, there is still no significant difference between the upper and lower classes (p-value=0.28). Again, we are not able to confirm the gradient produced in the survival of the richest literature.

5 Robustness Checks

5.1 Migration and Selection

Any analysis of family reconstitution data faces the problem that people who migrate to or from observed parishes are censored with regards to certain information. In the English data

---

[14]One might also wonder what would happen if those with unknown occupations were able to be assigned to their correct occupational categories. The unknown category has lower net reproduction than the other groups. Part of this is driven by the sources of occupations, which leads us to underestimate births and overestimate childlessness for that group, and is discussed at length in Section 5.3 and Appendix L. If we assume that the vast majority of the unknown were lower classes and reallocate all of the unknown into the lower class category, the net reproduction of the lower classes becomes statistically lower than both the middle and upper classes. However, the middle classes would still have far higher fertility than the upper classes, which is what is important for our story. The hump-shaped gradient would only become more accentuated.

[15]Note that we do not lose any parishes in this robustness check because the parishes with data after 1780 all have starting dates before 1680. Thus, there are 26 parishes in all of the analysis.

there is no foolproof way of measuring migration since the English parish register system did not include migration lists like the ones kept in Sweden (Jeub 1993). This makes it difficult to calculate precise periods of exposure for each individual in the parish and is a limitation that applies equally to the fertility and mortality calculations in Wrigley et al. (1997).[16] However, we do observe migration of a type because we know whether someone was baptised, married or buried in the parish. Thus, we might not observe the baptism date of in-migrants, but we are likely to observe their marriage, the baptism of their children or their burial. For out-migrants, we would observe their baptism and perhaps their marriage but not their burial. These three vital events provide important information that can be incorporated in defining exposure periods. For instance, there is no problem in measuring marital fertility if we observe a woman's marriage date, the baptism dates of her children and her burial date since she is very likely to have spent her reproductive years in the parish. Childlessness and child mortality are similarly easy to measure. However, it is important to note that we can only calculate these figures for non-migrants and in-migrants and must exclude out-migrants since we cannot be sure they baptised all of their children in the parish.

We run into more trouble when trying to measure celibacy rates in the data. As mentioned above, we observe never-married (celibate) individuals based on three criteria: they were baptised and buried in the parish; they survived past the age of 15; and they never married nor baptised or buried any children in the parish. These three components ensure that we have identified someone who never married and did not have any illegitimate children. Because we need to observe both the baptism and burial date of an individual, to measure celibacy we have to exclude both out-migrants and in-migrants from the calculations. Thus, because we must exclude some form of migrants from all of the calculations, it is important to try to establish that the selection bias of migrants is not driving our primary results, i.e. class-gradients in the four margins.

Before conducting some empirical tests, it might be helpful to describe who the out- and in-migrants were. Migrants in early modern England could have been one of three types. The first type were people who did not move very far, for instance to a nearby parish. These migrants were likely taking advantage of new economic opportunities such as employment as a servant in a different parish but did not substantially vary their cultural practices (Anderson, 1990, p. 11). These migrants were not insubstantial. Clark (1979) studied the migration patterns of rural witnesses in diocesan courts between 1660 and 1730 finding that 70% had moved at least once in their life. However, only 14% had moved beyond their county of origin. Likewise, Anderson (1990) found that two-thirds of rural migrants in 1851 moved less than 25 km. This

---

[16]Ruggles (1992) and Wrigley (1994) have also discussed potential migration biases to mean ages at marriage and death in reconstitution studies. These are discussed in more detail in Appendix J.

evidence from 1851 is especially striking since migration had increased substantially by this period driven by the growth of industrial cities (Shaw-Taylor and Wrigley 2014). In our data, we can capture the demographic behaviour of this type of migrant by looking at in-migrants to our own parishes, who are likely similar to out-migrants from our parishes who also moved to similar parishes.

The second type of migrants were people who left home to become apprentices and learn a trade. Minns and Wallis (2012) estimate that in 1700 around 9 per cent of English young men were apprenticed in London alone and 60 per cent of these apprentices were in-migrants to London. The London apprenticeship market was remarkably open with the majority of apprentices being unrelated to their master by kinship or place of origin and drawn from nearly all of England (de la Croix, Doepke, and Mokyr 2018), though with a clear gradient based on distance to London. Apprentices came from all social backgrounds but the sons of gentlemen, professionals, traders and craftsmen were over-represented among London apprentices relative to labourers, servants, husbandmen and farmers (Leunig, Minns, and Wallis 2011). Thus, apprentices may account for more migration among the higher and middle classes compared to the lower classes, excluding farmers. One should also note that a large minority of these apprentices returned to their original parish after their apprenticeship, so not all of these migrants were lost from observation forever (Klemp et al. 2013).

The third type of migrants were people who moved from our provincial parishes to towns or larger population centres, London being the largest draw, but did not take part in formal apprenticeships. We do not have precise information on rural to urban migration until the mid-nineteenth century with the census of 1851. This data is an imperfect proxy for our earlier period because of the enormous population growth and urbanisation that occurred in England along with the Industrial Revolution, so we must interpret the findings cautiously. According to a national sample analysed by Anderson (1990), 54% of the population lived more than 2 km from their place of birth. While only one-third of rural migrants moved more than 25 km, the percentages urban migrants moving more than 25 km were much higher: around 60% of urban migrants to provincial cities and 80% of migrants to London (Anderson, 1990, pp. 11-12). Thus, as expected urban migrants travelled longer distances than rural migrants, though this is for a much later period. Another long-standing feature of urban migration was the extremely high death rates of migrants in their new urban environment. Death rates in London were astoundingly high in the seventeenth century even as the population of London doubled across the century (Landers, 1993, pp. 193-195; Shaw-Taylor and Wrigley, 2014, p. 76). High mortality meant that many of the urban migrants would not have survived to reproduce anyway.

Keeping these three types of migrants in mind, we can now discuss how selective migration might influence our results. We will begin by discussing our celibacy calculation from which we

have to exclude both in- and out-migrants. It is highly possible that some women permanently left their parish of origin and never married or had children. However, this would only distort our estimation of the celibacy margin across social groups if there were (i) different migration rates for women across the social groups or if (ii) the women migrating from a particular social group were more likely to remain celibate than women migrating from another group. On the first point, we test whether there were significant differences in out-migration rates by class. To do this, we have to categorise individuals by their father's occupation since we are unable to observe the occupation of out-migrants at marriage. Table 3 illustrates two main features of migration in the reconstitution data: women were slightly more likely to emigrate than men and the social gradient was largely flat for women and slightly decreasing for men.[17] These gradients are confirmed by linear probability models reported in the last two columns of Table 3. The results suggest that the propensity to emigrate was statistically higher for male offspring born to labourers and servants than for offspring of more well-off families, but there were no statistically significant differences in emigration rates of female offspring (with the exception of the less mobile daughters of craftsmen). This conclusion is reassuring for our analysis of female fertility but is a constraint for our analysis of men's net reproduction in Appendix K.

Unfortunately, it is much more difficult to deal with the second issue above. We have no way of formally testing whether migrant women from different classes had the same or different propensities to be celibate or marry. We simply have to assume that celibate women's propensity to migrate was equal (or proportionally equal) to the propensities of non-migrants and conduct the analysis as such. This is a strong assumption, though, so we provide additional empirical and historical analysis to support it.

The largest potential problem for migration on our marriage rates would be if migrants from the upper classes, which have low marriage rates, had a higher propensity to marry than non-migrants from the same group. One might assume that wealthier women were more likely to migrate for marriage since they had a smaller pool of eligible bachelors locally, had more money to search for potential partners and would have benefited more from their husband's inheritance. This would mean that our low marriage rates for the upper classes are potentially under-estimates. We can apply an approximate test for this by looking at in-migrants into our own parishes who would have been similar to migrants to rural areas and market towns in the same period. If wealthy women were more likely to migrate and marry at the same time, we would expect to observe that the wealthy had higher shares of marriages between

---

[17]We are not surprised that women have a slightly higher emigration rate than men in this case. We are mostly measuring migration over very small distances, given that a migrant is simply someone who leaves their parish of birth. We would not be surprised, then, if relatively few women migrated long distances and most women migrants simply moved into their husband's parish nearby. As such, our findings do not challenge the common understanding that men have a higher propensity to emigrate.

	Emigration rates		Linear Probability Model	
	Son of	Daughter of	Son of	Daughter of
Labourers & Servants	58.5%	60.4%	Reference group	
Husbandmen	53.1%	58.1%	−0.021**	−0.001
			(0.008)	(0.009)
Craftsmen	52.8%	57.5%	−0.043***	−0.014*
			(0.008)	(0.008)
Traders	56.1%	60.0%	−0.033***	−0.013
			(0.010)	(0.010)
Farmers	54.2%	61.6%	−0.038***	0.001
			(0.013)	(0.013)
Merchants & Professionals	53.6%	57.0%	−0.034***	−0.019
			(0.013)	(0.013)
Gentry	46.7%	56.8%	−0.054***	0.008
			(0.018)	(0.018)
Unknown Occupation	53.5%	57.0%	−0.011	−0.011
			(0.007)	(0.007)
Observations			65,183	62,453

Note: The linear probability model includes 26 parish dummies and 4 period dummies.
*p<0.1; **p<0.05; ***p<0.01.

Table 3: Emigration by Social Group

female in-migrants and male non-migrants than other social groups. In addition, the share of marriages between female in-migrants and male non-migrants would be higher than the share of marriages between male in-migrants and female non-migrants in these wealthier groups.

However, as Table 4 shows, we do not observe this pattern at all. Merchants and professionals had the lowest share of female in-migrant marriages, and although the gentry had a higher share, it was below that of husbandmen, traders and farmers. The share of gentry marriages with in-migrant mothers was only slightly higher than the share of marriages with in-migrant fathers. For merchants and professionals in-migrant fathers made up a larger share of marriages. Thus, it does not seem that upper class women had a higher propensity to marry when migrating. The largest difference between the share of in-migrant mother and in-migrant father marriages, at 16 percentage points, was among farmers who had the lowest share of in-migrant fathers and highest share of in-migrant mothers. These figures suggest that it is possible that out-migrant farmers' daughters could have had a higher propensity to marry than those who did not migrate. However, farmers had the highest marriage rates and second highest net reproduction

rates in our sample. Thus, this bias would only increase their advantage over the merchants, professionals and gentry rather than eliminating our social gradient in net reproduction.

Occupation	Share of marriages betw. in-migrant Mothers and non-migrant Fathers	Share of marriages betw. non-migrant Mothers and in-migrant Fathers
Labourers and Servants	0.190	0.180
Husbandmen	0.252	0.146
Craftsmen	0.221	0.188
Traders	0.233	0.153
Farmers	0.271	0.111
Merchants and Professionals	0.134	0.182
Gentry	0.232	0.205
Unknown	0.218	0.184
Total	0.218	0.177

Note: Household occupation is the husband's occupation.

Table 4: Share of marriages by migrant status

In addition to our analysis of in-migrants, historical research on never-married women highlights the importance of kinship ties to single women's welfare across all classes in society, which may suggest that never-married women had a lower propensity to migrate than their married counterparts (Froide, 2005, p. 49). Thus, in the end we believe it is reasonable to analyse celibacy rates across the social groups from the reconstitution data.

When thinking about the other margins of reproductive success, we can begin to understand how migration would influence our results by comparing differences in the other margins between in-migrants and non-migrants. This comparison helps us understand the demographic characteristics of the first type of migrants, migrants who did not move very far, highlighted above. Appendix J provides the details of these calculations, which follow a similar format to the general estimations conducted above. We test for differences between families with in-migrant mothers, in-migrant fathers and where both parents were in-migrants. The relative prevalence of these types of marriages across the social groups is given in Table B.5. We find some relatively minor differences between in-migrants and non-migrants, but there was no significant social gradient in these effects. Thus, excluding migrants from our analysis is unlikely to influence our main results.

5.2 Net Reproduction for Men

As mentioned above, our baseline fertility calculations centred on women's fertility because this is the most common way fertility is measured among demographers. However, this may limit the comparability of our study to Clark and Hamilton (2006) and Clark and Cummins (2015) since these papers focused on the fertility recorded in wills, which mainly refer to men. Thus, we provide a robustness check in Appendix K that presents our results recalculated for male fertility. Male net reproduction was higher than female net reproduction across all social groups largely because $b(c)$ was higher for men in general (Figure K.5). When holding the extensive margin constant at the average level across the various social groups, we again see that the rich had much higher net reproduction than the poor (see Figure K.4), though the differences between rich and poor are much smaller than those found by Clark and Hamilton (2006). However, once the extensive margin of fertility is allowed to vary across social groups, the inverse u-shaped pattern that we found for women remains. Upper class men had a rate of net reproduction far below that of their middle and even lower-class counterparts. The result among the upper classes was mainly driven by lower marriage rates and higher rates of childlessness. The differences in net reproduction between the upper class and middle class and upper class and lower class are statistically significant (see Figure K.6 and Table K.3).

5.3 Sources of Male Occupations

Another potential source of error for our childlessness and marital fertility rates comes from the sources of male occupations. These rates could be skewed if the sources of occupations, from marriages, baptisms, or deaths, were significantly different across social groups. In fact, a large proportion of our occupational data comes from baptism records, which would only be recorded for individuals who were not childless. This may explain the significantly higher rate of childlessness among those with an unknown occupation. In order to test whether this was the case, we restricted our baseline analysis to father's occupations that were drawn from marriage or burial registers and moved those from baptism registers into the unknown occupation category. The results for childlessness are reported in Appendix L, Table L.2. Removing the baptism occupations, reduces the slope of the gradient of childlessness across social groups and the differences between groups become insignificant because of the smaller sample size, but the higher levels of childlessness among farmers and the upper classes remain. We perform a similar robustness check for marital fertility rates and find that the slightly lower fertility of the unknown occupation group was also partly a product of the greater propensity to observe a father's occupation if he had more children (see Figure L.2). We recalculate the final net reproduction rates in Figure L.3. This adjustment reduces the net reproduction rate for the

lower classes by increasing childlessness and decreasing marital fertility, but the general pattern that we find in our baseline holds. The upper classes did not have higher net reproduction rates than the lower classes. It was the middle classes of traders and farmers that had the highest net reproduction.

6   Discussion of the results

As was reported earlier, we find that incorporating the extensive margins of fertility shifts our understanding of the reproductive success of different socio-economic groups significantly. Our analysis showed that the lower, middle and upper classes operated by means of distinctly different demographic regimes. Although the families of all three classes married comparatively late in life on average (Table B.3 in Appendix B), the upper classes were more often unmarried or childless than the rest (top panels of Figure 3). This resulted in comparatively modest rates of reproduction among the upper classes (Figure 4). Similarly, the lower classes had lower rates of reproduction than their middle-class counterparts, but their lower net reproduction was caused by fewer births within marriage rather than fertility control on the extensive margin (Bottom left panel of Figure 3). Among middle-class families, however, birth limitation did not occur to the same extent as in upper- and lower-class families. Farmers tended to marry at higher rates, although they had fewer births on average than the upper classes. Traders had as many births on average as the upper classes, but had lower rates of childlessness.[18] These socio-specific combinations ultimately gave the middle class the upper hand in the grand scheme of reproductive success (Figure 4).

This general overview raises two sets of questions about our findings. First, what explains class differentials in childlessness? For many of the social gradients above, it is fairly easy to understand how they could emerge across the different groups. Marriage rates could be influenced by cultural practices and requirements to save enough resources before marriage. Sibship size could be influenced by differing mean ages at marriage and birth spacing across the groups. However, it is slightly more difficult to explain variations in childlessness across groups. In a time before effective contraception, it was difficult for couples who were sexually active to avoid pregnancy over a long marriage. Thus, we will look at potential mechanisms behind our childlessness findings to help explain variations in this margin. Second, how do these findings relate to earlier studies attributing pre-Industrial economic development in North-West Europe to the European Marriage Pattern? Our study allows us to look at the functioning of the EMP within a society and speculate about how these structures may have influenced economic growth.

---

[18]See Appendix D and Figure D.1 for more detail.

6.1 Possible Reasons for High Childlessness

One of the more puzzling findings presented above was the strong variations in childlessness by social group. Thus, in this section we will consider a number of potential explanations for the social gradient in childlessness in turn, though our analysis does not allow us to fully identify the causal pathway running from higher social status to higher childlessness.

First, Wrigley et al. (1997, pp. 394-8) argue that the most important variable influencing entry sterility (or the inability to produce any child in a marriage) was the mother's age at marriage. Thus, if women in higher social groups married systematically later than their lower social group peers, then we would expect to see a higher rate of childlessness. However, this cannot explain the higher childlessness rates among merchants and professionals and the gentry since they had similar mean ages of marriage to some of the other social groups (Table B.3).

Second, Gobbi and Goñi (2016) suggest that inheritance practices may have influenced whether couples remained childless. Before the twentieth century, the English nobility used marriage settlements to pass land and wealth to their heir. Marriage settlements set up a trust where the bride and groom were only life tenants to their estate and the estate was passed on unbroken to the heir of their marriage. Marriage settlements were instigated at the marriage of an heir (alternatively, settlements could be signed when the heir turn 21) and provided for the heir's younger brothers and sisters, effectively preventing the estate from being broken apart. They had to be renewed every generation. Studying the British peerage between 1603 and 1938, Gobbi and Goñi found that the rates of childlessness among heirs without marriage settlements were 8 to 13 per cent higher than rates among heirs with marriage settlements. Because the estates of heirs without marriage settlements could be sold on the market, these heirs chose to remain childless rather than pass on a smaller estate to their son and heir. An heir would not have a marriage settlement if his father died before his marriage, which occurred in 23.6 per cent of cases in Gobbi and Goñi's dataset. Thus, this inheritance structure, which would have influenced the gentry most strongly, could have contributed to higher childlessness rates among higher wealth groups in our data.

It is also possible that higher rates of venereal diseases such as gonorrhoea, chlamydia and syphilis among the upper classes led to higher rates of sterility. Gonorrhoea and chlamydia lead to infertility by preventing conception since both diseases if untreated lead to blockages of the Fallopian tubes and abnormal sperm, whereas syphilis causes higher rates of miscarriages and stillbirths (Belsey, 1976, pp. 327-29). There is anecdotal evidence that some of these diseases were prevalent among the gentry and aristocracy, e.g. Henry VIII had syphilis (Parascandola 2008) and James Boswell, a high-born Scottish lawyer, had at least 17 infections of gonorrhoea in his life (Stone, 1979, p. 350). However, it is much more difficult to find prevalence estimates

for the general population. Szreter (2017) found that in the 1770s 8% of residents in Cheshire had contracted syphilis by the age of 35, but the prevalence rate was much lower at 1% in Cheshire's rural hinterlands similar to most of the parishes in the reconstitution dataset. There is no systematic information about class gradients in syphilis prevalence until the early twentieth century when syphilis was most common among unskilled working-class men (Szreter, 2014) and therefore would not explain higher rates of childlessness among the upper classes. Prevalence rates of gonorrhoea and chlamydia were likely higher than for syphilis, but again we have no way of knowing whether these also produced a class gradient.

Goñi (2015) has also argued that consanguinity could have influenced childlessness. Again, studying the British peerage, now in the nineteenth century, he found that children of consanguineous marriages (between second cousins or closer) were 50 per cent more likely to be childless than children of peers who did not marry in the same family. Thus, if members of the gentry were more likely to marry individuals closely related to them, this could explain part of the difference in childlessness that we observe.

Finally, it is possible that the unique cultural context of marriage among the gentry led to higher levels of childlessness. Since members of the wealthier classes had less control over who they married, they may have ended up in loveless marriages at a higher rate than their counterparts down the social ladder (Stone, 1979, p. 308). If their rates of sexual activity were lower, then this might increase rates of childlessness enough to reflect the differences we have observed.

Other hints for the causes of class differences in childlessness comes from a comparison with France. The social gradient found in childlessness rates in England is also present in some French data. Bardet (1983), who reconstituted 5889 complete family histories over the seventeenth and eighteenth centuries in Rouen (second most populated city in France), finds that childlessness was more widespread among nobles and shopkeepers than among workers, and there was a positive time trend in the data. de la Croix and Brée (2019) consider the rise in materialism as a possible explanation. The increase in materialism means an increasing availability, variety, affordability, and quality of consumer goods, implying a rising demand for them. Such a trend might have affected the upper classes more intensely than the laborers. Did certain kinds of new (luxury) goods become accessible or desirable, and was being childfree a requirement to procure them? Van de Walle and Van de Walle (1972) identify major arguments that were advanced by authors of the 18th and 19th centuries to explain why women did not want children. These include the idea that women did not want to experience the physical changes of pregnancy and wanted to preserve their place in the salons of high society. This would then only have concerned the elite.

In the end, the increased childlessness among the gentry is likely caused by a number of these

factors combined. Thus, the 5 percentage point difference in childlessness rates between the wealthy and poor could have been driven by slight differences in marriage age, inheritance patterns, venereal infections, consanguinity and sexual activity that accumulated to the final effect.

6.2 The European Marriage Pattern and Economic Growth

Our study makes two important contributions to the debate on the foundations of North-West European economic success. First, it suggests that voluntary childlessness may have been a mechanism within the European Marriage pattern. Second, we argue that the class-specific demographic patterns in England were particularly conducive to its industrial development during and after the classical years of the Industrial Revolution. Rather than survival of the richest, the higher reproductive success of the middle classes ensured that those most likely to invest in the human capital of their children outgrew other groups.

Traditionally, scholars held that births could only be limited by raising the age of marriage or increasing the celibacy rate (Wrigley and Schofield, 1981; Weir, 1984). We depart from the view inherent in these previous approaches by highlighting the possibility of voluntary childlessness as a fertility limitation strategy. Our data suggest that this fertility mechanism was more prominent among the upper classes in early modern England with the lower classes managing their fertility by adjusting birth spacing (c.f. Cinnirella, Klemp, and Weisdorf 2017; Bengtsson and Dribe 2006). Thus, future studies of the European Marriage pattern should analyse the mechanisms behind class differentials in childlessness and work to incorporate childlessness as a potential mechanism through which population can adjust to changes in living standards. However, it is important to note that because the extensive margin was most influential for the fertility of the elite, we cannot rule out that other groups were still subject to Malthusian forces.

Our results also highlight how the class-specific demographic patterns would be helpful for England's economic growth. If middle-class families were the most reproductively successful and also more prone to invest in growth-promoting education of their offspring, then their reproductive success could create a growing pool of human capital in a society. Lower-class families were rarely able to afford even primary education let alone advanced education such as apprenticeships (Minns and Wallis, 2012). On the other hand, while upper-class families were able to afford to pay for growth-enhancing education and sent some of their younger sons into apprenticeships, elder sons were mainly trained to take over their family's estates and did not need growth-promoting educational training: they relied on the rental income of their estates and had no incentive to participate in activities that fostered economic growth (Doepke and

Zilibotti 2008; Wallis and Webb 2011). Crouzet's famous study of the social background of the emerging class of English industrialists confirms this view, showing that less than three per cent of the new industrialists had an upper-class origin (Crouzet 1985). Rather, 85 per cent of industrialists came from a middle-class background. More than 40 per cent were involved in an industrial trade that was similar to that of their father and 22 per cent had a father who was a farmer. It was the children of the middle-classes that helped drive England's industrial development, not those of the upper or lower classes.

It is important to note that this diffusion of growth-promoting education does not rely on low social mobility either. Although the fertility differences between social groups were large enough to produce quantitatively important differences in their relative representation over time (see Appendix G), our story would hold if there were downward social mobility from the middle classes. These children who had been raised in middle-class families would spread their values for education to other social groups widening the reach of these values. This story matches the history of the spread of literacy in pre-Industrial England. The gentry and professionals were first to gain high levels of literacy, and they were followed by the middle classes in the seventeenth century (Cressy 1981). The growth of literacy among the lower classes only occurred in the early eighteenth century reaching around 50% by 1750 (Schofield 1981). Thus, there was a diffusion of literacy and presumably other values for education from the middle classes to the lower classes during the period of our study.

This middle-class pattern of human capital investment emphasises that the European Marriage Pattern was conducive to economic growth in England not because all English families universally strove to limit births, but because those classes that invested comparatively little in growth-promoting education were more inclined to engage in birth restricting practices. In the English case, the survival of the richest was therefore not a growth-promoting strategy: the survival of the middle-classes was.

7 Conclusion

This paper has shown that when accounting for both the extensive and intensive margins of fertility, the previous findings that the rich in early modern England had the highest reproductive success disappears. Although merchants, professionals and gentry had the highest net marital fertility, as a group they were more likely to remain celibate or childless once they were married. After accounting for the extensive margin, we find that the middle classes had the reproductive advantage. They had lower rates of celibacy and childlessness and therefore produced more offspring as a whole. Their reproductive success likely influenced England's own economic success in the early modern period by boosting investments in human capital. As

mentioned above, our social groups are imperfectly correlated with wealth, limiting our ability to fully refute the earlier survival of the richest literature. However, we argue that our social groups better capture groups within society who would invest in the human capital of their children. This was the middle classes who were usually not the wealthiest in society.

This research has two sets of important implications. First, it raises questions about simple models of the European Marriage Pattern. According to our findings different social groups within early modern societies had substantially different family and demographic patterns: celibacy and childlessness were the important fertility margins for the rich, and marital fertility was the important margin for the poor. These class differences raise questions about the usefulness of a simple model that classifies entire countries into those with strong or weak European marriage patterns (Dennison and Ogilvie 2014). Thus, we must develop more dynamic models of the European marriage pattern that allow for intra-country variation in the pattern and that more properly account for change over time.

Second, our paper highlights the importance of the extensive margin of fertility in early modern societies and corroborates more modern studies that have shown that fertility decisions on the extensive margin have significantly influenced overall fertility (Aaronson, Lange, and Mazumder (2014); Baudin, de la Croix, and Gobbi (2015)). Future studies should focus more closely on the dynamics of childlessness and celibacy. For instance, there is no study to date that addresses all potential causes of childlessness in pre-Industrial societies including the voluntary aspects highlighted in the most recent literature. Scholars have already challenged some of the other fundamental assumptions underpinning the Malthusian preventive check: there is now ample evidence that pre-Industrial couples could control their fertility within marriage and that infanticide could operate as a separate control on fertility (Cinnirella, Klemp, and Weisdorf 2017; Bengtsson and Dribe 2006; Lee, Feng, and Li 2000). We must now explore childlessness in a similar manner to see whether it is responsive to changing economic conditions and how these dynamics varied around the world. Celibacy also deserves more careful scrutiny moving away from aggregate analyses of the past toward more individual-level analyses. These margins of fertility mattered in pre-industrial societies and deserve a stronger place in our theories of premodern demographic change and economic development.

References

Aaronson, Daniel, Fabian Lange, and Bhashkar Mazumder. 2014. "Fertility Transitions along the Extensive and Intensive Margins." *American Economic Review* 104 (11): 3701–24.

Althaus, Paul G. 1980. "Differential Fertility and Economic Growth." *Zeitschrift für die gesamte Staatswissenschaft / Journal of Institutional and Theoretical Economics* 136 (2): 309–326.

Anderson, Michael. 1990, June. "The social implications of demographic change." In *The Cambridge Social History of Britain, 1750–1950*, edited by F. M. L. Thompson, Volume 2, 1–70. Cambridge University Press.

Armstrong, W Alan. 1972. "The use of information about occupation." In *Nineteenth-century society: essays in the use of quantitative methods for the study of social data*, edited by Edward Anthony Wrigley, 191–310. Cambridge University Press.

Ashraf, Quamrul, and Oded Galor. 2011. "Dynamics and Stagnation in the Malthusian Epoch." *American Economic Review* 101 (5): 2003–2041 (August).

Bardet, Jean-Pierre. 1983. *Rouen au XVIIe et XVIIIe siècles*. Paris: SEDES.

Baudin, Thomas, David de la Croix, and Paula E. Gobbi. 2015. "Fertility and Childlessness in the US." *American Economic Review* 105:1852–1882.

———. 2020. "Endogenous Childlessness and Stages of Development." *Journal of the European Economic Association.*

Bellera, Carine A, Gaëtan MacGrogan, Marc Debled, Christine Tunon de Lara, Véronique Brouste, and Simone Mathoulin-Pélissier. 2010. "Variables with time-varying effects and the Cox model: some statistical concepts illustrated with a prognostic factor study in breast cancer." *BMC medical research methodology* 10 (1): 20.

Belsey, Mark A. 1976. "The epidemiology of infertility: a review with particular reference to sub-Saharan Africa." *Bulletin of the World Health Organization (WHO)* 54 (3): 319–341.

Bengtsson, Tommy, and Martin Dribe. 2006. "Deliberate Control in a Natural Fertility Population: Southern Sweden, 1766-1864." *Demography* 43 (4): 727–746.

Blackstone, William, and Robert Malcolm Napier Kerr. 1896. *The Student's Blackstone: Being the Commentaries On the Laws of England of Sir William Blackstone.* London: Reeves and Turner.

Boberg-Fazlic, Nina, Paul Sharp, and Jacob Weisdorf. 2011. "Survival of the richest? Social status, fertility and social mobility in England 1541-1824." *European Review of Economic History* 15:365–392.

Bowles, Samuel. 2007. "Genetically Capitalist?" *Science* 318 (5849): 394–396.

Burkoff, Nikolas S., Paul Metcalfe, Jonathan Bartlett, and David Ruau. 2016. "Gamma imputation tutorial." mimeo.

Cheneys. 1967. *Cheneys of Banbury, 1767–1967.* Cheneys & Sons.

Cinnirella, Francesco, Marc Klemp, and Jacob Weisdorf. 2017. "Malthus in the Bedroom: Birth Spacing as Birth Control in Pre-Transition England." *Demography* 54 (2): 413–436 (March).

Clark, Gregory. 2007. *A Farewell to Alms: A Brief Economic History of the World.* Princeton: Princeton University Press.

Clark, Gregory, and Neil Cummins. 2015. "Malthus to modernity: wealth, status, and fertility in England, 1500–1879." *Journal of Population Economics* 28 (1): 3–29.

Clark, Gregory, and Gillian Hamilton. 2006. "Survival of the Richest: The Malthusian Mechanism in Pre-Industrial England." *The Journal of Economic History* 66 (3): 1–30.

Clark, Peter. 1979. "Migration in England during the Late Seventeenth and Early Eighteenth Centuries." *Past & Present*, no. 83:57–90.

Cressy, David. 1981. "Levels of Illiteracy in England, 1530-1730." In *Literacy and Social Development in the West: A Reader*, edited by Harvey A. Graff, 105–124. Cambridge: Cambridge University Press.

Crouzet, François. 1985. *The First Industrialists.* Cambridge: Cambridge University Press.

De Moor, Tine, and Jan Luiten Van Zanden. 2010. "Girl power: the European marriage pattern and labour markets in the North Sea region in the late medieval and early modern period." *Economic History Review* 63 (1): 1–33.

de la Croix, David. 2012. *Fertility, Education, Growth, and Sustainability.* Cambridge UK: Cambridge University Press.

de la Croix, David, and Sandra Brée. 2019. "Key Forces Behind the Decline of Fertility - Lessons from Childlessness in Rouen before the Industrial Revolution." *Cliometrica* 13 (1): 25–54.

de la Croix, David, and Matthias Doepke. 2003. "Inequality and Growth: Why Differential Fertility Matters." *American Economic Review* 93 (4): 1091–1103.

de la Croix, David, Matthias Doepke, and Joel Mokyr. 2018. "Clans, Guilds, and Markets: Apprenticeship Institutions and Growth in the Pre-Industrial Economy." *Quarterly Journal of Economics* 133:1–70.

Dennison, Tracy, and Sheilagh Ogilvie. 2014. "Does the European Marriage Pattern Explain Economic Growth?" *The Journal of Economic History* 74 (03): 651–693.

de Pleijt, Alexandra M. 2018. "Human capital formation in the long run: evidence from average years of schooling in England, 1300–1900." *Cliometrica* 12 (1): 99–126 (December).

de Pleijt, Alexandra M., Alessandro Nuvolari, and Jacob Weisdorf. 2018. "Human Capital Formation during the First Industrial Revolution: Evidence from the Use of Steam Engines." Journal of the European Economic Association, forthcoming.

Doepke, Matthias, and Fabrizio Zilibotti. 2008. "Occupational Choice and the Spirit of Capitalism." *The Quarterly Journal of Economics* 123 (2): 747–793.

Feeney, Griffith. 1986. "Period Parity Progression Measures of Fertility in Japan." NUPRI Research Paper Series 35.

Fildes, Valeria A. 1986. *Breasts, Bottles and Babies: A History of Infant Feeding.* Edinburgh: Edinburgh University Press.

Foreman-Peck, James. 2011. "The Western European marriage pattern and economic development." *Explorations in Economic History* 48 (2): 292–309.

Froide, Amy M. 2005. *Never married: singlewomen in early modern England.* Oxford University Press.

Galor, Oded. 2012. *Unified Growth Theory.* Princeton: Princeton University Press.

Galor, Oded, and Marc Klemp. 2014. "The Biocultural Origins of Human Capital Formation." Working papers 2014-6, Brown University, Department of Economics.

Galor, Oded, and Stelios Michalopoulos. 2012. "Evolution and the growth process: Natural selection of entrepreneurial traits." *Journal of Economic Theory* 147 (2): 759–780.

Galor, Oded, and Omer Moav. 2001. "Evolution and growth." *European economic review* 45 (4-6): 718–729 (May).

———. 2002. "Natural Selection And The Origin Of Economic Growth." *The Quarterly Journal of Economics* 117 (4): 1133–1191.

Gobbi, Paula E. 2013. "A model of voluntary childlessness." *Journal of Population Economics* 26 (3): 963–982.

Gobbi, Paula E., and Marc Goñi. 2016. "Childless Aristocrats. Fertility, Inheritance, and Persistent Inequality in Britain (1550-1950)." unpublished.

Goñi, Marc. 2015. "Kissing cousins. Estimating the causal effect of consanguinity on fertility using evidence from the London Season." unpublished.

Grambsch, Patricia M., and Terry M. Therneau. 1994. "Proportional Hazards Tests and Diagnostics Based on Weighted Residuals." *Biometrika* 81 (3): 515–526.

Hajnal, John. 1965. "European marriage patterns in perspective." In *Population in history: essays in historical demography*, edited by Eversley DE Glass DV, 101–43. Chicago Illinois Aldine Publishing Company.

Henry, Louis. 1953. *Fécondité des Mariages – Nouvelle méthode de mesure.* Paris: Presses Universitaire de France.

Holcombe, Lee. 1983. *Wives and Property.* Toronto: University of Toronto Press.

Hollingsworth, Thomas H. 1965. *The Demography of the British Peerage.* London: Population Investigation Committee, LSE.

Hosmer, David W, and Stanley Lemeshow. 1999. *Applied survival analysis: Regression modelling of time to event data.* New York: Wiley.

Jackson, Dan, Ian White, Shaun Seaman, Hannah Evans, Kathy Baisleyb, and James Carpenter. 2014. "Relaxing the independent censoring assumption in the Cox proportional hazards model using multiple imputation." *Statistics in Medicine* 33:4681–4694.

Jeub, Ulla Nilsdotter. 1993. *Parish records: 19th century ecclesiastical registers.* Umeå: Demografiska databasen, Univ.

Kelly, Morgan, and Cormac Ó Gráda. 2012. "The Preventive Check in Medieval and Preindustrial England." *The Journal of Economic History* 72 (04): 1015–1035.

Klemp, Marc, Chris Minns, Patrick Wallis, and Jacob Weisdorf. 2013. "Picking winners? The effect of birth order and migration on parental human capital investments in pre-modern England." *European Review of Economic History* 17 (2): 210–232.

Landers, John. 1993. *Death and the metropolis: studies in the demographic history of London, 1670-1830.* Volume 20. Cambridge University Press.

Lee, James, Wang Feng, and Bozhong Li. 2000. "Population, poverty, and subsistence in China, 1700-2000." In *Population and Economy: From Hunger to Modern Economic Growth*, edited by Tommy Bengtsson and Osamu Saito, 73–109. Oxford: Oxford University Press.

Leridon, Henri. 2008. "A New Estimate of Permanent Sterility by Age: Sterility Defined as the Inability to Conceive." *Population Studies* 62 (1): 15–24.

Leunig, Tim, Chris Minns, and Patrick Wallis. 2011. "Networks in the Premodern Economy: The Market for London Apprenticeships, 1600–1749." *The Journal of Economic History* 71 (2): 413–443 (June).

Livi-Bacci, Massimo. 1986. "Social Group Forerunners of Fertility Control in Europe." In *The Decline of Fertility in Europe*, edited by Ansley Coale and Susan Cotts Watkins, 182–200. Princeton: Princeton University Press.

Long, Jason. 2013. "The surprising social mobility of Victorian Britain." *European Review of Economic History* 17 (1): 1–23.

Malthus, Thomas. 1798. *An Essay on the Principle of Population*. London: Johnson.

Minns, Chris, and Patrick Wallis. 2012. "Rules and reality: quantifying the practice of apprenticeship in early modern England." *The Economic History Review* 65 (2): 556–579.

Mokyr, Joel. 2005. "Long-term economic growth and the history of technology." In *Handbook of economic growth*, edited by Philippe Aghion and Steven N. Durlauf, Volume 1, 1113–1180. Elsevier.

Parascandola, John. 2008. *Sex, sin, and science: a history of syphilis in America*. Greenwood Publishing Group.

Ruggles, Steven. 1992. "Migration, Marriage, and Mortality: Correcting Sources of Bias in English Family Reconstitutions." *Population Studies* 46 (3): 507–522.

Schoenfeld, David. 1982. "Partial Residuals for The Proportional Hazards Regression Model." *Biometrika* 69 (1): 239–241.

Schofield, Roger. 1985. "English marriage patterns revisited." *Journal of Family History* 10:2–20.

Schofield, R S. 1981. "Dimensions of Illiteracy in England, 1750-1850." In *Literacy and Social Development in the West: A Reader*, edited by Harvey A. Graff, 201–213. Cambridge: Cambridge University Press.

Shaw-Taylor, Leigh, and Edward Anthony Wrigley. 2014. "Occupational structure and population change." In *The Cambridge Economic History of Modern Britain: Volume 1, Industrialisation, 1700–1870*, edited by Jane Humphries Roderick Floud and Paul Johnson, 53–88. Cambridge University Press.

Skirbekk, Vegard, et al. 2008. "Fertility trends by social status." *Demographic Research* 18 (5): 145–180.

Squicciarini, Mara P., and Nico Voigtländer. 2015. "Human Capital and Industrialization: Evidence from the Age of Enlightenment." *The Quarterly Journal of Economics* 130 (4): 1825–1883.

Stone, Lawrence. 1979. *The family sex and marriage: in England 1500-1800*. New York: Harper and Row.

Szreter, Simon. 2014. "The Prevalence of Syphilis in England and Wales on the Eve of the Great War: Re-visiting the Estimates of the Royal Commission on Venereal Diseases 1913–1916." *Social History of Medicine* 27 (3): 508–529.

———. 2017. "Treatment rates for the pox in early modern England: a comparative estimate of the prevalence of syphilis in the city of Chester and its rural vicinity in the 1770s." *Continuity and Change* 32 (2): 183–223.

Van de Walle, Etienne, and Francine Van de Walle. 1972. "Allaitement, stérilité et contraception : les opinions jusqu'au XIXe siècle." *Population* 27 (4): 685–701.

Van Leeuwen, Marco HD, and Ineke Maas. 2011. *HISCLASS: A historical international social class scheme.* Universitaire Pers Leuven.

Vogl, Tom. 2016. "Differential Fertility, Human Capital, and Development." *Review of Economic Studies* 83 (January): 365–401.

Voigtländer, Nico, and Hans-Joachim Voth. 2013. "The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe." *Review of Economic Studies* 80 (2): 774–811.

Wallis, Patrick, and Cliff Webb. 2011. "The education and training of gentry sons in early modern England." *Social History* 36 (1): 36–53.

Weir, David R. 1984. "Rather never than late: celibacy and age at marriage in English cohort fertility, 1541-1871." *Journal of Family History* 9 (04): 340–54.

Wrigley, Edward Anthony. 1966. "Family reconstitution." In *An Introduction to English Historical Demography: From the Sixteenth to the Nineteenth Century*, edited by David Edward Charles Eversley, Peter Laslett, and Edward Anthony Wrigley, 96–159. London: Weidenfeld and Nicolson.

———. 1994. "The Effect of Migration on the Estimation of Marriage Age in Family Reconstitution Studies." *Population Studies* 48 (1): 81–97.

Wrigley, Edward Anthony, R S Davies, James Oeppen, and Roger Schofield. 1997. *English Population History from Family Reconstitution: 1580–1837.* Cambridge: Cambridge University Press.

Wrigley, Edward Anthony, and Roger Schofield. 1981. *The population history of England 1541-1871: A Reconstruction.* Cambridge University Press.

Appendices

A   Sample

1	Labourers/Servants	incl. seamen
2	Husbandmen	small farmers, weavers
3	Craftsmen	tailors, carpenters
4	Traders	innkeepers, butchers, bakers
5	Farmers	
6	Merchants/Professionals	clerks, clergy, medical
7	Gentry	gentlemen, esquire

Table A.1: Social Groups Employed in the Analysis

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Died Age	9,007	45.932	21.136	16	27	43	64	104
Married	9,007	0.589	0.492	0	0	1	1	1
Labourers & Serv.	9,007	0.110	0.313	0	0	0	0	1
Husbandmen	9,007	0.086	0.280	0	0	0	0	1
Craftsmen	9,007	0.096	0.295	0	0	0	0	1
Traders	9,007	0.045	0.208	0	0	0	0	1
Farmers	9,007	0.029	0.169	0	0	0	0	1
Upper class	9,007	0.041	0.197	0	0	0	0	1
Unknown Occupat.	9,007	0.593	0.491	0	0	1	1	1

Note: the means on the father's occupational dummies show the proportion of the sample falling into each occupational group. The largest group is the Labourers and Servants (11.1 per cent) while the smallest one is the Farmers (2.9 per cent).

Table A.2: Sample to Compute $m(c)$

Statistic	N	Mean	St. Dev.	Min	Median	Max
At Risk for First Birth						
I(First Birth)	14,730	0.850	0.357	0	1	1
Spacing marriage to 1st birth	12,517	1.373	1.354	0.000	0.940	18.014

Table A.3: Sample to Compute $z(c)$

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Died Age	24,260	24.964	27.373	0	1	13	47	103
Died before 15	48,514	0.256	0.436	0	0	0	1	1

Table A.4: Sample to Compute $d(c)$

Statistic	N	Mean	St. Dev.	Min	Median	Max
At Risk for Second Birth						
I(Second Birth)	12,519	0.858	0.349	0	1	1
Spacing 1st to 2nd	10,739	2.362	1.317	0.000	2.071	11.444
At Risk for Third Birth						
I(Third Birth)	10,738	0.838	0.369	0	1	1
Spacing 2nd to 3rd	8,994	2.561	1.335	0.000	2.301	11.471
At Risk for Fourth Birth						
I(Fourth Birth)	8,993	0.810	0.392	0	1	1
Spacing 3rd to 4th	7,283	2.632	1.339	0.000	2.389	11.808
At Risk for Fifth Birth						
I(Fifth Birth)	7,280	0.780	0.414	0	1	1
Spacing 4th to 5th	5,682	2.600	1.286	0.000	2.397	9.929
At Risk for Sixth Birth						
I(Sixth Birth)	5,679	0.740	0.439	0	1	1
Spacing 5th to 6th	4,204	2.594	1.265	0.000	2.400	10.203
At Risk for Seventh Birth						
I(Seventh Birth)	4,201	0.704	0.456	0	1	1
Spacing 6th to 7th	2,959	2.555	1.252	0.000	2.384	9.748
At Risk for Eighth Birth						
I(Eight Birth)	2,959	0.653	0.476	0	1	1
Spacing 7th to 8th	1,932	2.530	1.202	0.000	2.397	10.222
At Risk for Nineth Birth						
I(Nine Birth)	1,932	0.597	0.491	0	1	1
Spacing 8th to 9th	1,154	2.475	1.244	0.000	2.286	9.463
At Risk for Tenth Birth						
I(Tenth Birth)	1,154	0.545	0.498	0	1	1
Spacing 9th to 10th	629	2.303	1.113	0.000	2.162	9.841

Note: Spacing between the n and $n+1$ birth displays descriptive statistics on the time in years between the births. Spacing values of zero are possible when twins were born.

The sample size of the spacing variables and children at risk in the line below do not always agree because we do not have precise birth dates for a small sample of children, which prevents us of calculating spacing between two children. These have been checked and confirmed.

Table A.5: Sample to Compute $b(c)$

B   Additional Data

Figure B.1: The 26 parishes of Wrigley et al. (1997) and the 3 counties of Clark and Cummins (2015)

	Mean	Unknown	1	2	3	4	5	6	7
					Social groups				
0	0.974	0.969	0.981	0.979	0.979	0.983	0.982	0.982	0.974
1	0.916	0.903	0.933	0.926	0.934	0.933	0.944	0.925	0.917
2	0.923	0.916	0.939	0.925	0.929	0.941	0.921	0.939	0.917
3	0.914	0.909	0.923	0.916	0.921	0.938	0.905	0.919	0.904
4	0.919	0.913	0.938	0.907	0.920	0.946	0.916	0.927	0.905
5	0.908	0.901	0.917	0.907	0.927	0.893	0.883	0.944	0.948
6	0.910	0.895	0.930	0.921	0.926	0.915	0.889	0.928	0.973
7	0.916	0.910	0.924	0.912	0.932	0.908	0.854	0.952	0.952
8	0.924	0.929	0.902	0.920	0.924	0.913	0.886	0.962	0.916
9	0.941	0.930	0.956	0.945	0.932	0.957	0.978	0.964	0.898

Note: estimated from logistic regressions

Table B.1: 5-year Survival Probabilities of Mothers by parity n and social group

Occupations	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9
Labourers, Servants	0.89	0.86	0.82	0.77	0.74	0.71	0.63	0.52	0.56
Husbandmen	0.92	0.88	0.83	0.80	0.76	0.69	0.67	0.60	0.54
Craftsmen	0.92	0.88	0.85	0.83	0.79	0.72	0.63	0.62	0.58
Traders	0.93	0.90	0.88	0.86	0.73	0.79	0.60	0.63	0.68
Farmers	0.93	0.89	0.86	0.79	0.73	0.75	0.66	0.50	0.46
Upper Class	0.94	0.91	0.87	0.82	0.78	0.79	0.72	0.70	0.56
Unknown	0.88	0.83	0.80	0.77	0.73	0.69	0.66	0.60	0.52

Note: Parity progression ratios computed from estimated survival functions 10 years after previous birth, and assuming average levels for parish and time period dummies

Table B.2: Estimated Parity Progression Ratios Across Social Groups

	Mean age at marriage	95% conf. interval	
Labourers, Servants	24.63	24.18	25.08
Husbandmen	25.45	24.95	25.95
Craftsmen	25.62	25.09	26.16
Traders	25.40	24.61	26.18
Farmers	23.83	23.06	24.60
Merchants, Professionals	25.22	24.18	26.27
Gentry	25.29	23.79	27.79
Unknown Occupation	25.67	25.47	25.86

Table B.3: Mean female age at marriage by social group

Occupation	Mean number of marriages	
	Women	Men
Labourers, Servants	1.061	1.126
Husbandmen	1.061	1,128
Craftsmen	1.063	1,124
Traders	1.064	1.135
Farmers	1.055	1.129
Merchants, Professionals	1.052	1.110
Gentry	1.048	1.123
Unknown Occupation	1.047	1.116
All	1.053	1.120

Table B.4: Mean Number of Marriages per Married Woman and Men and per occupation

Occupation	Both Non-migrant	Mother In-Migrant	Father In-Migrant	Both In-Migrant	Total
Labourers and Servants	547	277	262	371	1457
Husbandmen	372	271	157	276	1076
Craftsmen	451	263	223	252	1189
Traders	202	122	80	119	523
Farmers	144	88	36	57	325
Merchants and Professionals	174	47	64	67	352
Gentry	55	35	31	30	151
Unknown	2349	1623	1368	2104	7444
Total	4294	2726	2221	3276	12517

Table B.5: Number of marriages by migrant status

C Individual Occupational Mobility

As mentioned in the main text, the occupations in the reconstitution data were recorded at several points in a man's life: at his marriage, at the baptisms of his children and at his burial. Thus, it is possible to quantify men who moved up or down in the wealth categories during their lifetime. In total 9 per cent of the father's occupations in our dataset changed over the father's life course. Thus, we needed to figure out a way to categorise those fathers who experienced mobility across social groups. Table C.1 shows the movement of father's occupations between the Clark and Cummins social groups during the father's lifetime. The rows in the table represent the minimum social group of the father, which could occur at marriage, baptism of a child or burial, whereas the columns represent the maximum social group of the father. Thus, Table C.1 does not reveal whether there was upward or downward mobility.

There were 17,666 children whose father's minimum and maximum social group for their occupation was one, and there were 1,087 children whose father's minimum social group was one but also had a group two occupation listed. In order to simplify our analysis and improve our sample size among the occupational groups, we reassigned some groups experiencing mobility to their maximum social group. However, we only did this for social groups that could have easily been fluid to minimise measurement error in the analysis. Thus, we assign the 1,087 observations who were mobile between groups one and two to group two since the line between a labourer and a husbandmen was rather small. We also assigned those mobile between groups two and three to group three (463 obs.) and those mobile between groups three and four to group four (368 obs.). For groups five, six and seven, a more nuanced approach was required. For the farmers, group five, it made more sense to merge them with men who were mobile between group two, husbandmen, and five since both were agricultural occupations (230 obs.). Likewise, assigning those mobile between groups four and six to group six made sense because there would obviously be mobility between traders and merchants and small scale proprietors and larger scale proprietors (578 obs.). Finally, we assigned men mobile between group five and seven and group six and seven to group seven. There would obviously be some crossover between farmers, professionals and gentry, but if a person was ever ascribed an occupation of being a gentleman or esquire, they were likely to be a member of the gentry whatever their occupation recorded elsewhere in their life (91 and 232 obs.). All other individuals experiencing mobility were placed in the unknown occupation category.

Highest occupation of father

	1	2	3	4	5	6	7
1	17666	1087	624	330	187	229	37
	87.6%	5.4%	3.1%	1.6%	0.9%	1.1%	0.2%
2	0	11827	463	168	230	139	28
		92.0%	3.6%	1.3%	1.8%	1.1%	0.2%
3	0	0	14440	368	38	236	43
			95.5%	2.4%	0.3%	1.6%	0.3%
4	0	0	0	7904	129	578	96
				90.8%	1.5%	6.6%	1.1%
5	0	0	0	0	3528	78	91
					95.4%	2.1%	2.5%
6	0	0	0	0	0	3516	232
						93.8%	6.2%
7	0	0	0	0	0	0	1581
							100%

Lowest occupation of father (row axis label)

Note: The individuals in the table above do not represent individual men. Instead, they represent the occupations of fathers in the dataset. Thus, a farmer who remained a farmer across his life cycle and had six children is counted in the table six times. Gray areas represent the persons to whom we actually assign an occupation. White areas are assigned as unknown occupation.

Table C.1: Reassigning Fathers Mobile Between Social Groups

D   Decomposing the Extensive Margin

We can decompose the influence of the extensive margin on fertility differences across social groups more precisely by testing the influence of childlessness and celibacy separately. Figure D.1 replicates the gray reproduction rates produced in Figure 4 showing the difference between holding the extensive margin constant and allowing it to vary. In addition to these two lines, we have now added a line that holds the marriage rate constant while allowing all other margins to vary (orange line) and a line that holds the childlessness rate constant while allowing all other margins to vary. This figure suggests that allowing childlessness to vary increases the fertility of the lower classes but does not substantially influence the net reproduction of farmers and the upper classes. Allowing marriage rates to vary has a greater influence on the social gradient in net reproduction and is responsible for the much lower net reproduction of the upper classes. Thus, marriage, rather than childlessness, drives social differences in net fertility.
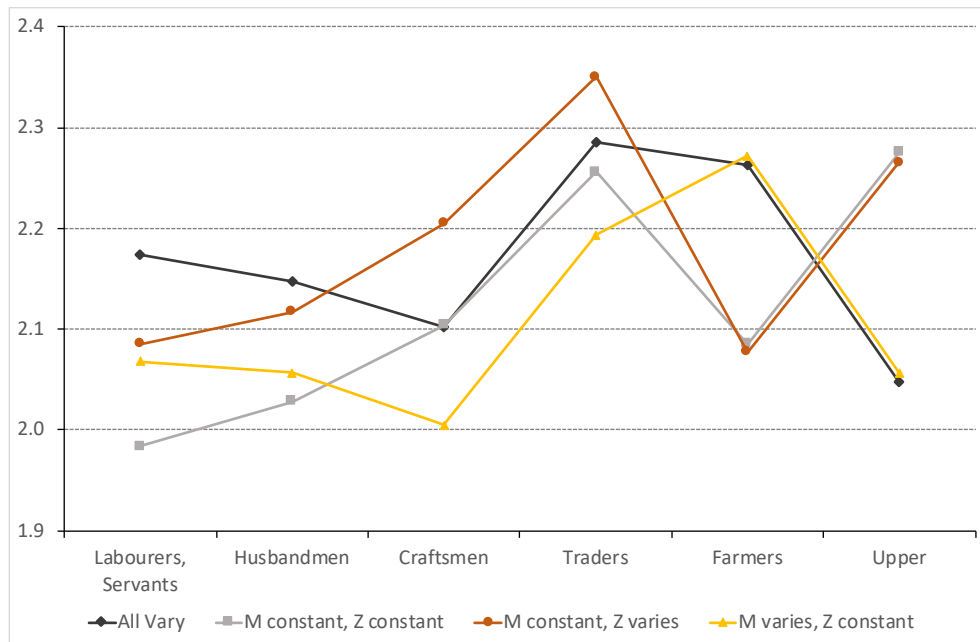
Note: Data from Table F.1.

Figure D.1: Decomposing the Extensive Margin: Net reproduction rates Across Social Groups

E   Simple Count of Surviving Children by Social Group

One of the key features of our analysis is to break down net fertility into its four margins (see Section 2). However, one might wonder whether the net fertility gradient that we estimate is similar to what would arise simply from counting the number of surviving children of all women born to fathers in a given social group. This would incorporate the intensive and extensive margin because single and married women would be counted alike. This appendix performs this simple count. However, the calculation is quite restrictive because, as mentioned in the main text, we can only observe life-long single women among non-migrant women, i.e. those who are born and die in the same parish. Thus, the starting point for this calculation is the subsample of women in the $m(c)$ calculation as reported in Table A.2. For these women, we count the number of children surviving to 15 that each has. Single and childless women have zero children. We then report the mean number of children per woman by the woman's father's occupation rather than the husband's occupation which we use to compute $z(c)$, $b(c)$ and $d(c)$ above. We have to use the father's occupation because life-long single women do not have a husband's occupation.

The results are presented in Table E.1. We find a very similar gradient to the one reported in Figure 4 and Table F.1. The upper classes have the lowest net fertility rate; the middle classes of traders and farmers have the highest; and the lower classes were somewhere in between. These figures are lower than our estimates of $n(c)$ for a couple of reasons. Most importantly, restricting the calculations to non-migrants may mean that single women dying before marriage are over-represented in this sample since women who migrated short distances for marriage are excluded. This helps explain the very large share of women in the sample that never have children. However, in this case, we are more interested in the comparative level of fertility across groups rather than the overall fertility level, so this is not a problem for our analysis.

Occupation	Number of observations	Average number of children per women
Labourers / Servants	988	1.393
Husbandmen	771	1.319
Craftsmen	869	1.224
Traders	409	1.555
Farmers	264	1.462
Upper class	366	0.839
Unknown occupations	5340	1.217
All	9007	1.253

Note: Mean number of children surviving to age 15 born to women whose fathers were in the reported social group.

Table E.1: Simple Count of Surviving Children by Social Group

# F Uncertainty Surrounding the Estimates - Details

In order to measure the uncertainty surrounding our estimates of net reproduction for the various class groups, we simulate and accumulate the errors for the four margins using the estimates and standard errors from the regressions as described in Section 4.6 in the main text. Table F.1 summarizes the various moments involved. Figure F.1 plots smooth kernel densities for the sampling distribution for $n(c)$ for each occupational group. As noted in the text, the unknown occupational group has much lower net reproduction partially because of biases that arise from the sources of occupations discussed at length in Section 5.3 in the main text and in Appendix L. It has a narrower distribution because of the much larger sample size for this groupt than the others.

We next aggregate the social classes into three larger groups as follows: the lower class includes Labourers, Servants, Husbandmen, and Craftsmen; the middle class includes Traders and Farmers and the upper class includes Merchants, Professional and Gentry. We exclude the group with unknown occupations. Figure F.2 plots the corresponding smooth kernel histograms of the values for $n(c)$. Table F.2 reports one-tailed mean difference tests of statistical signifance between the three groups. We find that the difference in net reproduction between the middle-class and upper-class group was statistically significant, but the other differences between groups were not quite statistically significant.

	$m(c)$	$1 - z(c)$	$b(c)$	$1 - d(c)$	$n(c)$	$\bar{n}(c)$
Mean	0.77	0.87	3.76	0.75	1.90	1.90
Labourers, Servants	0.80	0.92	3.91	0.75	2.17	1.98
Husbandmen	0.78	0.91	3.99	0.75	2.15	2.03
Craftsmen	0.74	0.91	4.19	0.75	2.10	2.10
Traders	0.75	0.91	4.53	0.74	2.29	2.26
Farmers	0.84	0.87	4.11	0.75	2.26	2.09
Upper class	0.70	0.87	4.58	0.74	2.05	2.27
Unknown	0.77	0.85	3.50	0.75	1.72	1.77

Note: $n(c)$ and $\bar{n}(c)$ computed from Equations (1) and (6) respectively.

Table F.1: Computed Reproductive Rates - n and $\bar{n}$
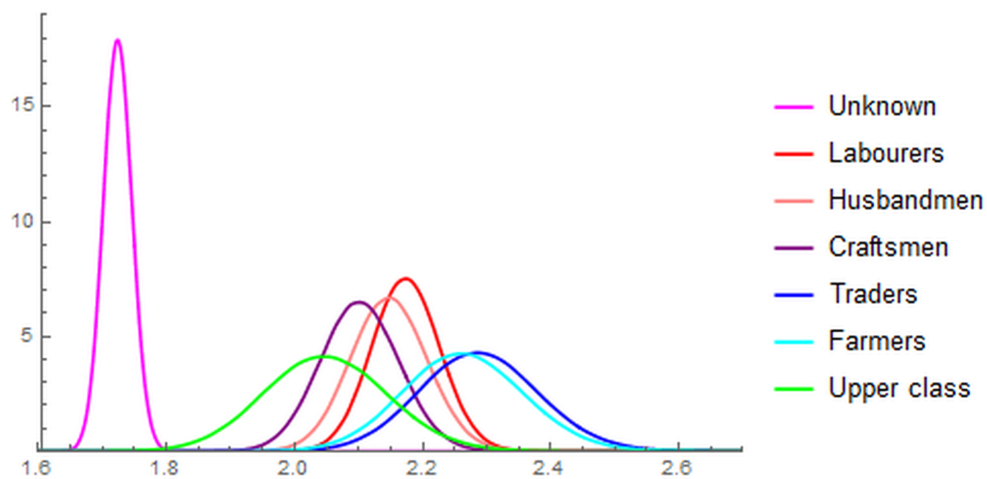
Note: Lower class includes Labourers, Servants, Husbandmen, and Craftsmen.
Middle class includes Traders and Farmers.

Figure F.1: Sampling distributions of $n(c)$ by social class
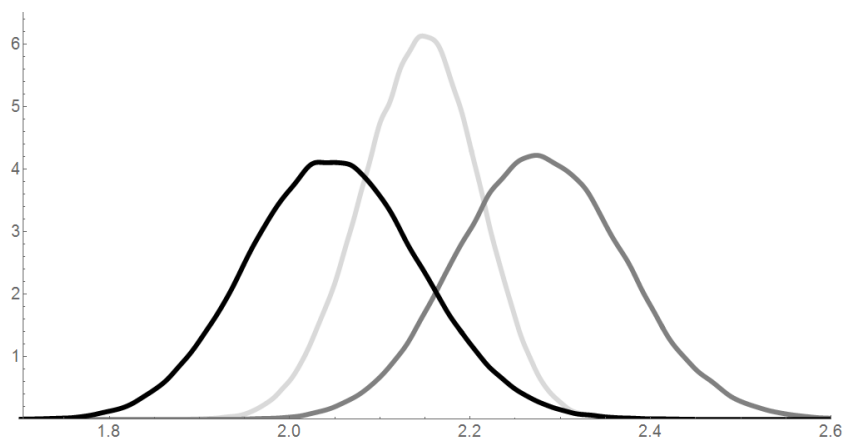
Note: Lower class includes Labourers, Servants, Husbandmen, and
Craftsmen. Middle class includes Traders and Farmers.

Figure F.2: Sampling distributions of $n(c)$ by social class (black=upper, gray=middle, light gray=lower)

	Lower Classes	Middle Classes	Upper Classes
Lower Classes		0.14	0.09
		(0.11)	(0.21)
Middle Classes			0.23
			(0.04)

Note: Mean net reproduction gap between groups from the simulation with one-tailed p-value on mean difference in parentheses.

Table F.2: Tests of statistical significance on mean differences between groups

G Fertility Differences across Social Groups and the Composition of the Population

One might question whether fertility differentials between groups were really large enough to have important quantitative effects on the relative representation of the groups in the economy over time. To test whether this was the case, we do some simple back-of-the-envelope calculations to see how fertility differences might have mattered. We look at the shares of children in each successive generation for ten generations. This is based on the assumption that the average generation is 30 years in length and our data cover just over 300 years, 1541-1851. We use the shares of individuals with known occupations from Table A.2 times one thousand as the notional population sizes for the fifth generation so that the shares in the reconstitution data reflect the mean share across all ten generations. We then assume that half the population of each group are women and multiply the number of women by the net reproduction rate $n(c)$ from Table F.1, producing the number of individuals in the subsequent generation, i.e.

$$Pop_g = \frac{Pop_{g-1}}{2} \times n(c)$$

The notional population levels and relative representation of each social group are presented in Tables G.1 and G.2. The relative representation of husbandmen, craftsmen and the upper classes fall across the generations, while the relative representation of labourers and servants remains fairly stable. However, the relative representation of traders and farmers increases substantially. As Table G.3 shows, this leads the middle classes to increase their share from 13.7% in generation zero to 23.6% in generation ten. This is a very large increase for a pre-Industrial, agricultural economy, as England was in 1541. The lower and upper classes decline as a whole. Thus, the differences in fertility that we measure could have strong effects on the relative representation of various groups in society over time and can explain in part the rising share of people working in middle class occupations over time.

Generation	Labourers, Servants	Husbandmen	Craftsmen	Traders	Farmers	Upper
0	73	60	75	23	16	36
1	79	65	79	26	18	37
2	86	70	83	30	20	38
3	93	75	87	34	23	39
4	101	80	91	39	26	40
5	110	86	96	45	29	41
6	120	92	101	51	33	42
7	130	99	106	59	37	43
8	141	106	111	67	42	44
9	153	114	117	77	48	45
10	167	123	123	88	54	46

Note: We set the average shares in our data as the shares in the fifth generation so that the average share across generations is equal to the average in the reconstitution dataset. We then calculate the population of each social group by assuming half the population of each group are women and ascribing our predicted net reproduction rates $n(c)$ to each group.

Table G.1: Notional Population of the Social Groups across 10 generations

Generation	Labourers, Servants	Husbandmen	Craftsmen	Traders	Farmers	Upper
0	25.6	21.3	26.5	8.2	5.5	12.9
1	26.0	21.3	25.9	8.7	5.8	12.3
2	26.3	21.3	25.4	9.2	6.1	11.7
3	26.5	21.3	24.8	9.8	6.5	11.1
4	26.8	21.2	24.2	10.4	6.8	10.6
5	27.0	21.1	23.6	11.1	7.1	10.1
6	27.2	21.0	23.0	11.7	7.5	9.6
7	27.4	20.9	22.4	12.4	7.8	9.1
8	27.6	20.8	21.7	13.1	8.2	8.6
9	27.7	20.6	21.1	13.8	8.6	8.1
10	27.8	20.4	20.5	14.6	9.0	7.7

Note: See Table G.1 and text for details of the calculations.

Table G.2: Predicted Relative Representation (%) of Each Social Group Across 10 Generations

Generation	Lower	Middle	Upper
0	73.4	13.7	12.9
1	73.2	14.5	12.3
2	72.9	15.4	11.7
3	72.6	16.3	11.1
4	72.2	17.2	10.6
5	71.7	18.2	10.1
6	71.2	19.2	9.6
7	70.7	20.2	9.1
8	70.1	21.3	8.6
9	69.4	22.4	8.1
10	68.7	23.6	7.7

Note: See Table G.1 and text for details of the calculations.

Table G.3: Predicted Relative Representation (%) of the Lower, Middle and Upper Classes Across 10 Generations

H   Removing Late Marriages from the Sample

One of the interesting findings of Clark and Cummins (2015) and Boberg-Fazlic, Sharp, and Weisdorf (2011) was that the survival advantage of the rich began to decline after the 1780s as the rich began to control their fertility and opt for smaller families. Although we have included time dummies in our analysis, one might wonder whether part of our low net reproduction result for the upper classes could be related to this early decline in fertility among the upper classes. To assuage fears of this, we have redone all of the regressions and calculations for our baseline excluding marriages after 1780. The social gradient that we found is very similar suggesting that our results are not being exclusively driven by reductions in fertility among the wealthy in this later period (see Table H.1 and Figure H.2). In fact the hump-shaped gradient is even more accentuated.

We also performed the simulations to understand the uncertainty around these estimates. Figure H.2 plots the sampling distributions of the net reproduction for each of the three classes. The standard errors on the estimates did increase since the sample size was reduced by removing the later marriages. However, the higher net reproduction point estimate for farmers shifts the middle class distribution upward. Now the difference in net reproduction between the middle class and upper class on the one hand and between the middle class and lower class are both statistically significant on a one-sided test of mean differences (see Table H.2). The upper class and lower class were not statistically different.

	$m(c)$	$1-z(c)$	$b(c)$	$1-d(c)$	$n(c)$	$\bar{n}(c)$
Mean	0.77	0.87	3.73	0.74	1.87	1.87
Labourers, Servants	0.79	0.91	3.84	0.74	2.07	1.92
Husbandmen	0.79	0.91	4.03	0.75	2.15	2.04
Craftsmen	0.76	0.92	4.19	0.74	2.14	2.09
Traders	0.76	0.91	4.54	0.73	2.29	2.25
Farmers	0.86	0.88	4.12	0.74	2.33	2.07
Upper Class	0.70	0.87	4.58	0.73	2.05	2.25
Unknown	0.77	0.85	3.46	0.75	1.70	1.75

Note: $n(c)$ and $\bar{n}(c)$ computed from Equations (1) and (6) respectively.

Table H.1: Computed Reproductive Rates - n and $\bar{n}$ - Restricted Sample
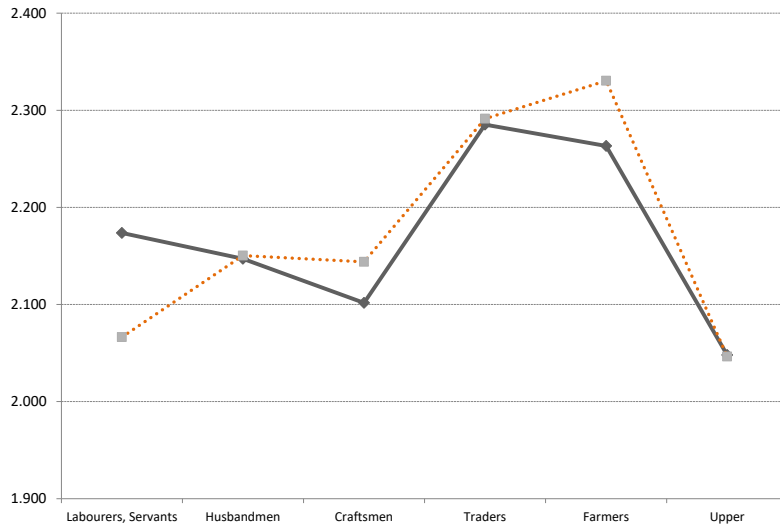
Note: solid, benchmark, dotted, excluding marriages after 1780.

Figure H.1: Net reproduction rates $n(c)$ with and without late marriages
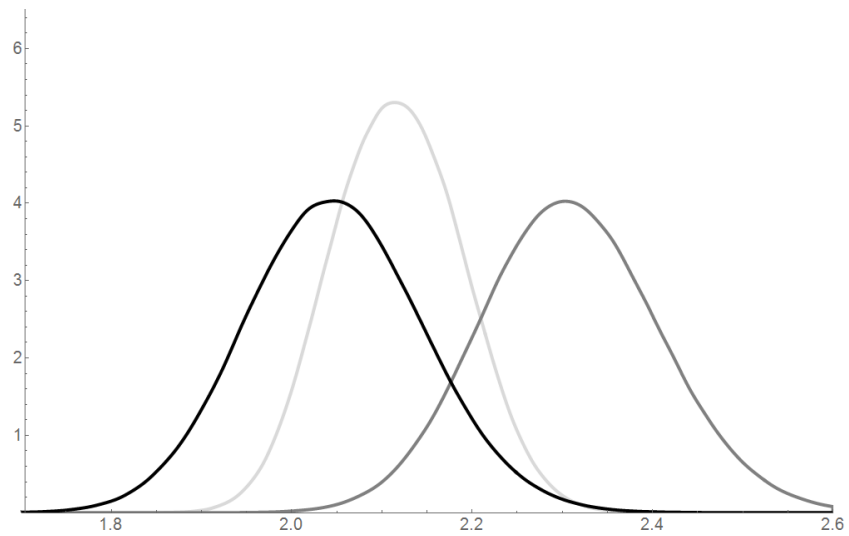
Note: Lower class includes Labourers, Servants, Husbandmen, and Craftsmen. Middle class includes Traders and Farmers. Sample restricted to marriages before 1780.

Figure H.2: Sampling distributions of $n(c)$ by social class (black=upper, gray=middle, light gray=lower)

	Lower Classes	Middle Classes	Upper Classes
Lower Classes		0.19	0.07
		(0.05)	(0.28)
Middle Classes			0.26
			(0.03)

Note: Mean net reproduction gap between groups from the simulation with one-tailed p-value on mean difference in parentheses.

Table H.2: Tests of statistical significance on mean differences between groups excluding marriages after 1780

I   Robustness: Non-Random Censoring

The Cox model assumes that censoring is independent of survival time. It might be difficult to defend this assumption in the context of childlessness: it is reasonable to assume that the probability of the mother dying (censoring) is dependent on the time she remains without a child (survival time). One can easily think of medical conditions that affect both.

We relax the independent censoring assumption in the Cox proportional hazard model using the γ-imputation method proposed by Jackson et al. (2014). The procedure generates hypothetical datasets where individuals who were censored now have a hypothetical survival, under the assumption that at the point of censoring the hazard function jumps by a constant γ. If $\gamma < 0$, which is the relevant case for our problem, there is decreased risk (of having one child) after censoring. By varying the size and magnitude of this constant γ, and reestimating the model under this condition, sensitivity analyses can be performed. Table I.1 reports results obtained with the `InformativeCensoring R` library described in Burkoff et al. (2016). Obviously, the coefficients of interest are barely affected by the choice of γ because of the small number of observations which are effectively subject to censoring because of early death. Thus, the fact that censoring may not be independent of the survival time does not influence our results in practice.

	(1) benchmark	(2) $\gamma = -5$	(3) $\gamma = -0.5$	(4) $\gamma = -0.1$
Husbandmen	-0.030	-0.033	-0.030	-0.030
	(0.051)	(0.040)	(0.041)	(0.041)
Craftsmen	-0.013	-0.013	-0.013	-0.013
	(0.050)	(0.039)	(0.040)	(0.040)
Traders	-0.038	-0.024	-0.034	-0.038
	(0.064)	(0.051)	(0.052)	(0.051)
Farmers	-0.203^{***}	-0.200^{***}	-0.204^{***}	-0.205^{***}
	(0.079)	(0.053)	(0.062)	(0.062)
Upper Class	-0.204^{***}	-0.212^{***}	-0.206^{***}	-0.204^{***}
	(0.060)	(0.053)	(0.053)	(0.053)
Unknown Occupation	-0.284^{***}	-0.294^{***}	-0.287^{***}	-0.285^{***}
	(0.038)	(0.032)	(0.032)	(0.032)
Observations	14,730	14,730	14,730	14,730

Note: Dependent variable is risk of first birth. Reference category is Labourers/Servants. 26 parish dummies and 4 period dummies are included. *p<0.1; **p<0.05; ***p<0.01

Table I.1: $z(c)$. Relaxing independent censoring with γ-imputation method

J Robustness: In-migrants vs. Non-migrants

In this appendix, we explore whether migration poses a potential threat to the social gradient in fertility that we found in our main results. As mentioned in the text, the only migrants that we can observe demographic behaviour for in our sample are those who move into our parish upon marriage, i.e. those individuals with missing baptism dates, but known marriage and burial dates. However, aside from the urban migrants, these migrants are likely similar to those leaving our parishes and heading to similar parishes. To this end, we examine how the social gradients of three of our margins of female fertility looked for in-migrants compared to non-migrants. We can measure differences across three of our margins, but unfortunately we cannot measure celibacy for in-migrants because we cannot be sure that individuals moving into the parish were not married before migrating and we would not observe their father's occupation.

We break our migrants into three groups: a mother in-migrant father non-migrant group; a mother non-migrant father in-migrant group; and a final group where both husband and wife were in-migrants. To test whether migration influenced demographic behaviour, we introduce a dummy variable for each migrant group and interact that dummy variable with the social groups dummies in the regressions to test for differences between migrants and non-migrants across the social groups.

Table J.1 presents the results for childlessness. Column 1 shows the baseline results from table 1 above. Columns 2, 3 and 4 show the results for in-migrant mothers, in-migrant fathers and both husband and wife in-migrant respectively. Controlling for various types of in-migrants did not substantially change the coefficients on the initial social groups. There were significant differences in childlessness for certain groups of migrants. For instance, in-migrant fathers were less likely to be childless, and families where both husband and wife were in-migrants were more likely to be childless. However, on the whole there were no significant differences between the social gradient of non-migrants and in-migrants. There are a few possible explanations for why joint in-migrant couples might have had higher rates of childlessness. First, we limit the data to first marriages only in an attempt to limit the maternal age and parity effects that could come with second marriages. Second, it is possible that in-migrants married at later ages than their non-migrant counterparts. Since female sterility is strongly linked to age, this would naturally increase childlessness rates. Unfortunately, we do not know the ages of in-migrant mothers because we do not observe their baptism date, so it is impossible to test this directly. However, as Ruggles (1992) notes, out-migration in reconstitution data can bias certain estimates such as mean age at marriage and life expectancy even if the characteristics and demographic rates of out-migrants and non-migrants were the same. This is because if the probability of migrating is renewed each year for all individuals at the same rate before and after marriage, then the overall probability of migrating out of the parish before marriage is greater for people who marry at older ages. Thus, the mean age of marriage for people in a parish may be an underestimate of the mean age including all out-migrants. This means that it is possible that our in-migrants (out-migrants from elsewhere) were older than the non-migrants who married in the parish. Ruggles showed that this effect could be large, but Wrigley (1994) conducted some additional tests with the actual reconstitution data and found that the bias of out-migration on marriage ages was very small. In the end, this could explain our higher childlessness rates and lower marital fertility (seen below), but we cannot be sure.

Tables J.2 and J.3 present the results of the same exercise for the parity progression ratio from one to two children and two to three children. Again, there are very few substantial differences in the social gradient between migrants and non-migrants and controlling for the various types of migrants does not alter the coefficients for non-migrants very much.

Table J.4 presents the results for the mortality rate of children under the age of 15. In-migrants couples had moderately higher mortality levels for children, but there were no statistically significant differences between non-migrants and in-migrants of each specific social group. Thus, our social gradient remains intact.

In conclusion, there were some relatively minor differences in the margins of female fertility between in-migrants and non-migrants. However, there is no significant social gradient to these effects, so not accounting for migration does not seem to strongly influence at least these three margins of reproductive success.

Dependent variable: Risk of first birth	(1) benchmark	(2) mother only	(3) father only	(4) both parents
Unknown Occupation	−0.284***	−0.265***	−0.275***	−0.305***
	(0.032)	(0.035)	(0.035)	(0.037)
Husbandmen	−0.030	−0.027	−0.002	−0.047
	(0.041)	(0.047)	(0.045)	(0.048)
Craftsmen	−0.013	−0.013	−0.008	−0.050
	(0.039)	(0.044)	(0.044)	(0.045)
Traders	−0.038	−0.014	−0.055	−0.069
	(0.051)	(0.058)	(0.056)	(0.059)
Farmers	−0.203***	−0.203***	−0.204***	−0.249***
	(0.062)	(0.072)	(0.066)	(0.069)
Upper Class	−0.204***	−0.190***	−0.161***	−0.269***
	(0.053)	(0.058)	(0.058)	(0.059)
Mother Migrant		0.040		
		(0.067)		
Unknown Occupation:Mother Migrant		−0.098		
		(0.073)		
Husbandmen:Mother Migrant		−0.024		
		(0.097)		
Craftsmen:Mother Migrant		−0.005		
		(0.097)		
Traders:Mother Migrant		−0.114		
		(0.123)		
Farmers:Mother Migrant		−0.014		
		(0.142)		
Upper Class:Mother Migrant		−0.088		
		(0.138)		
Father Migrant			0.131*	
			(0.069)	
Unknown Occupation:Father Migrant			−0.048	
			(0.075)	
Husbandmen:Father Migrant			−0.156	
			(0.110)	
Craftsmen:Father Migrant			−0.028	
			(0.101)	
Traders:Father Migrant			0.171	
			(0.140)	
Farmers:Father Migrant			0.096	
			(0.190)	
Upper Class:Father Migrant			−0.228*	
			(0.133)	
Both Migrant				−0.275***
				(0.062)
Unknown Occupation:Both Migrant				0.070
				(0.066)
Husbandmen:Both Migrant				0.041
				(0.093)
Craftsmen:Both Migrant				0.103
				(0.093)
Traders:Both Migrant				0.076
				(0.121)
Farmers:Both Migrant				0.130
				(0.158)
Upper Class:Both Migrant				0.231*
				(0.129)
Observations	14,730	14,730	14,730	14,730

Note: Dependent variable is risk of first birth. Reference category is Labourers/Servants. 26 parish dummies and 4 period dummies are included. *p<0.1; **p<0.05; ***p<0.01

Table J.1: Cox Model. Risk of first birth ($z(c)$): baseline vs. in-migrant mothers, in-migrant fathers and in-migrant couples interactions

Dependent variable: Risk of second birth	(1) benchmark	(2) mother only	(3) father only	(4) both parents
Unknown Occupation	−0.070**	−0.021	−0.069*	−0.030
	(0.035)	(0.045)	(0.038)	(0.040)
Husbandmen	0.099**	0.157***	0.101**	0.125**
	(0.044)	(0.060)	(0.048)	(0.051)
Craftsmen	0.126***	0.181***	0.105**	0.170***
	(0.042)	(0.056)	(0.047)	(0.048)
Traders	0.182***	0.265***	0.160***	0.236***
	(0.055)	(0.074)	(0.060)	(0.062)
Farmers	0.196***	0.165*	0.207***	0.181**
	(0.065)	(0.086)	(0.069)	(0.072)
Upper Class	0.197***	0.120*	0.210***	0.158**
	(0.057)	(0.072)	(0.062)	(0.064)
Mother Migrant		−0.022		
		(0.058)		
Unknown Occupation:Mother Migrant		−0.110*		
		(0.063)		
Husbandmen:Mother Migrant		−0.118		
		(0.087)		
Craftsmen:Mother Migrant		−0.133		
		(0.085)		
Traders:Mother Migrant		−0.185*		
		(0.110)		
Farmers:Mother Migrant		0.074		
		(0.129)		
Upper Class:Mother Migrant		0.218*		
		(0.116)		
Father Migrant			−0.024	
			(0.074)	
Unknown Occupation:Father Migrant			−0.001	
			(0.081)	
Husbandmen:Father Migrant			−0.021	
			(0.120)	
Craftsmen:Father Migrant			0.118	
			(0.109)	
Traders:Father Migrant			0.145	
			(0.148)	
Farmers:Father Migrant			−0.106	
			(0.199)	
Upper Class:Father Migrant			−0.072	
			(0.144)	
Both Migrant				−0.019
				(0.066)
Unknown Occupation:Both Migrant				−0.149**
				(0.072)
Husbandmen:Both Migrant				−0.113
				(0.100)
Craftsmen:Both Migrant				−0.210**
				(0.100)
Traders:Both Migrant				−0.244*
				(0.131)
Farmers:Both Migrant				0.095
				(0.164)
Upper Class:Both Migrant				0.227
				(0.138)
Observations	12,519	12,519	12,519	12,519

Note: Dependent variable is risk of second birth. Reference category is Labourers/Servants. 26 parish dummies and 4 period dummies are included. *p<0.1; **p<0.05; ***p<0.01

Table J.2: Cox Model. Risk of second birth: baseline vs. in-migrant mothers, in-migrant fathers and in-migrant couples interactions

Dependent variable: Risk of third birth	(1) benchmark	(2) mother only	(3) father only	(4) both parents
Unknown Occupation	−0.102***	−0.108***	−0.088**	−0.126***
	(0.038)	(0.041)	(0.041)	(0.043)
Husbandmen	0.070	0.103*	0.056	0.062
	(0.048)	(0.054)	(0.052)	(0.054)
Craftsmen	0.047	0.046	0.065	0.036
	(0.046)	(0.051)	(0.050)	(0.051)
Traders	0.166***	0.214***	0.172***	0.130*
	(0.059)	(0.066)	(0.064)	(0.067)
Farmers	0.117*	0.113	0.114	0.077
	(0.069)	(0.080)	(0.074)	(0.077)
Upper Class	0.180***	0.120*	0.192***	0.147**
	(0.061)	(0.067)	(0.067)	(0.069)
Mother Migrant		−0.005		
		(0.078)		
Unknown Occupation:Mother Migrant		0.027		
		(0.085)		
Husbandmen:Mother Migrant		−0.129		
		(0.113)		
Craftsmen:Mother Migrant		0.004		
		(0.112)		
Traders:Mother Migrant		−0.204		
		(0.144)		
Farmers:Mother Migrant		0.013		
		(0.159)		
Upper Class:Mother Migrant		0.404**		
		(0.158)		
Father Migrant			0.060	
			(0.080)	
Unknown Occupation:Father Migrant			−0.083	
			(0.088)	
Husbandmen:Father Migrant			0.123	
			(0.126)	
Craftsmen:Father Migrant			−0.093	
			(0.117)	
Traders:Father Migrant			−0.031	
			(0.160)	
Farmers:Father Migrant			0.060	
			(0.213)	
Upper Class:Father Migrant			−0.065	
			(0.157)	
Both Migrant				−0.192***
				(0.073)
Unknown Occupation:Both Migrant				0.081
				(0.079)
Husbandmen:Both Migrant				0.008
				(0.110)
Craftsmen:Both Migrant				0.006
				(0.111)
Traders:Both Migrant				0.122
				(0.142)
Farmers:Both Migrant				0.156
				(0.175)
Upper Class:Both Migrant				0.116
				(0.150)
Observations	10,738	10,738	10,738	10,738

Note: Dependent variable is risk of third birth. Reference category is Labourers/Servants. 26 parish dummies and 4 period dummies are included. *p<0.1; **p<0.05; ***p<0.01

Table J.3: Cox Model. Risk of third birth: baseline vs. in-migrant mothers, in-migrant fathers and in-migrant couples interactions

Dependent variable: Risk of death	(1) benchmark	(2) mother only	(3) father only	(4) both parents
Unknown Occupation	0.001	0.002	−0.003	0.030
	(0.032)	(0.034)	(0.035)	(0.036)
Husbandmen	−0.001	0.008	−0.022	0.034
	(0.039)	(0.043)	(0.044)	(0.044)
Craftsmen	0.039	0.042	0.018	0.062
	(0.036)	(0.040)	(0.042)	(0.042)
Traders	0.067	0.116**	0.028	0.106**
	(0.046)	(0.050)	(0.053)	(0.053)
Farmers	0.004	−0.008	0.002	−0.012
	(0.062)	(0.071)	(0.068)	(0.069)
Upper Class	0.071	0.098*	0.047	0.078
	(0.048)	(0.051)	(0.055)	(0.055)
Mother Migrant		−0.025		
		(0.068)		
Unknown Occupation:Mother Migrant		−0.011		
		(0.075)		
Husbandmen:Mother Migrant		−0.044		
		(0.098)		
Craftsmen:Mother Migrant		−0.019		
		(0.096)		
Traders:Mother Migrant		−0.330**		
		(0.133)		
Farmers:Mother Migrant		0.053		
		(0.146)		
Upper Class:Mother Migrant		−0.252*		
		(0.148)		
Father Migrant			0.001	
			(0.058)	
Unknown Occupation:Father Migrant			0.017	
			(0.066)	
Husbandmen:Father Migrant			0.109	
			(0.091)	
Craftsmen:Father Migrant			0.105	
			(0.086)	
Traders:Father Migrant			0.207*	
			(0.112)	
Farmers:Father Migrant			0.007	
			(0.171)	
Upper Class:Father Migrant			0.112	
			(0.113)	
Both Migrant				0.122**
				(0.059)
Unknown Occupation:Both Migrant				−0.115*
				(0.066)
Husbandmen:Both Migrant				−0.148
				(0.091)
Craftsmen:Both Migrant				−0.093
				(0.086)
Traders:Both Migrant				−0.169
				(0.112)
Farmers:Both Migrant				0.160
				(0.159)
Upper Class:Both Migrant				−0.018
				(0.113)
Observations	48,514	48,514	48,514	48,514

Note: Dependent variable is risk of death. Reference category is Labourers/Servants. 26 parish dummies and 4 period dummies are included. *p<0.1; **p<0.05; ***p<0.01

Table J.4: Cox Model. Risk of child death ($d(c)$): baseline vs. in-migrant mothers, in-migrant fathers and in-migrant couples interactions

K Men's Reproductive Success

Although we have presented our baseline results in the main paper from the perspective of women, we are also able to produce the same statistics for men, and there is some reason to believe that these may actually be more accurate in reflecting fertility differences by occupational group. Because female occupations were very rarely recorded in parish records, we have to assign a woman's occupational class in relation to the men in her life. When calculating the marriage rate, we use her father's occupation, and for the other three margins, we use her husband's occupation. If there was not strong marital homogamy, there could be measurement error in the occupational groups, which might lead to a different pattern across groups.

Thus, we present the equivalent regression tables and graphs for men that we included in the original paper for women. In the case of men, there were a few additional issues that needed to be dealt with to estimate the results. First, men were nearly twenty times more likely to baptise children in a parish without having been married in the parish, i.e. we do not observe the father's marriage date. It seems that men would occasionally get married in their spouse's parish, but then return to their home parish to live. Thus, we count all men who returned to their parish of birth to baptise children as being married even if we do not observe their marriage date. However, we need a marriage date to estimate the risk of marriage, so we imputed their marriage date to be one year before the birth of their first child. We do not use these individuals in calculating $z(c)$ and $b(c)$ though because the results would be sensitive to our imputed marriage dates.

Another potential issue is remarriage since men could remarry and continue to produce offspring at much later ages than women. Remarriage would obviously not influence the celibacy, childlessness or child mortality margins, but it could potentially influence the marital fertility of wealth groups. However, given the data constraints in determining the marriage order of individuals and the minuscule differences in marriage order across wealth groups, described in Section 4.2 of the main text and in Table B.4, we proceed just using the births from first marriages. Excluding remarriage likely leads to an overstatement of childlessness rates in men since some men who had childless first marriages would go on to have children in a second marriage and underestimate marital fertility since the children in second marriages would not be counted.

Finally, we also update all factors that were specific to women for men, so we adjust the parity progression ratios for male mortality rather than female mortality. This is especially important since men would not have suffered the high death rates in childbirth that women did.

Table K.1 reports the regressions for $m(c)$, $z(c)$ and $d(c)$, the equivalent of table 1 above, and the predicted marriage rates and childless rates for men are presented next to the rates for women in figures K.1 and K.2. The greater celibacy and childlessness of the upper classes holds when looking at fertility for men. In fact, the gradients across classes are even larger. We also see that $b(c)$ for men has been shifted upward relative to women in figure K.3 although the gradient is more or less the same. This higher level of fertility then produces a higher net reproduction rate for men relative to women (figure K.5) when all of the calculations are completed. Overall, the general patterns of net reproduction reamin. When the extensive margin is held at average levels, we find that the upper classes had higher fertility than the middle and lower classes (see Table K.2 and Figure K.5). However, when we allow the extensive margin to vary across

social groups, the upper classes have the lowest level of fertility by a large margin. Their net reproduction rate is far below the lower classes, and the middle classes maintain the highest net reproduction.

Similarly to what is done in Section F, we can assess the significativity of the results by simulation. Figure K.6 plots smooth kernel densities for the sampling distribution for $n(c)$ for the three groups (Low, Middle, and Upper social classes). It can be compared to Figure F.2. Table K.3 reports one-tailed mean difference tests of statistical significance between the three groups (comparable to Table F.2 for women). We find that the difference in net reproduction between the middle-class and upper-class group was statistically significant, as well as the gap between upper and lower classes.

|	(1) risk of marriage $m(c)$	(2) risk of first birth $z(c)$	(3) risk of child death $d(c)$
Husbandmen	−0.089	−0.007	0.003
	(0.057)	(0.043)	(0.040)
Craftsmen	−0.044	0.003	0.038
	(0.055)	(0.041)	(0.038)
Traders	−0.201***	−0.015	0.093**
	(0.069)	(0.053)	(0.047)
Farmers	−0.139*	−0.237***	0.025
	(0.076)	(0.064)	(0.063)
Upper Class	−0.446***	−0.252***	0.088*
	(0.075)	(0.055)	(0.050)
Unknown Occupation	−0.031	−0.250***	−0.003
	(0.046)	(0.033)	(0.033)
Observations	11,080	13,292	45,887

Note: Reference category is Labourers/Servants. 26 parish dummies and 4 period dummies are included. *p<0.1; **p<0.05; ***p<0.01

Table K.1: Cox Model. Analysis of $m(c)$, $z(c)$, $d(c)$ for men
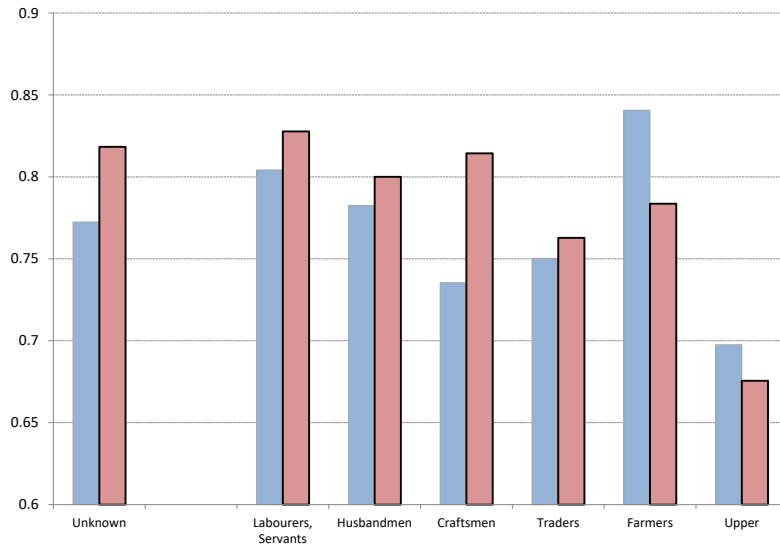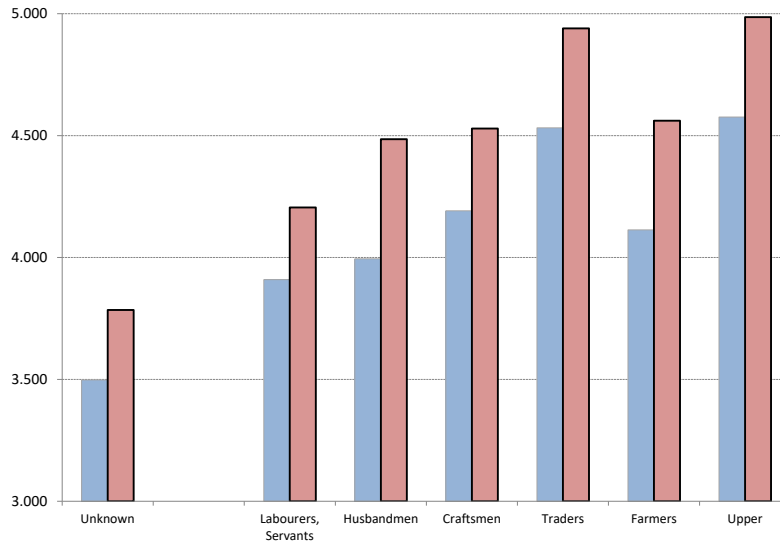
Note: filled blue, women; filled red with borders, men.

Figure K.1: Computed Marriage rates: the point of view of men
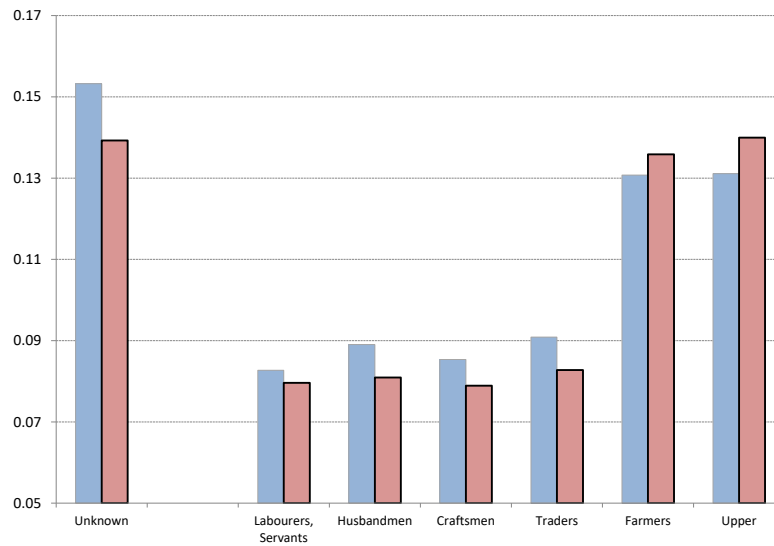
Note: filled blue, women; filled red with borders, men.

Figure K.2: Computed Childlessness Rate: the point of view of men

Note: filled blue, women; filled red with borders, men.

Figure K.3: Computed Births: the point of view of men

	$m(c)$	$1 - z(c)$	$b(c)$	$1 - d(c)$	$n(c)$	$\bar{n}(c)$
Mean	0.81	0.88	4.10	0.75	2.20	2.20
Labourers, Servants	0.83	0.92	4.21	0.75	2.41	2.26
Husbandmen	0.80	0.92	4.49	0.75	2.49	2.41
Craftsmen	0.81	0.92	4.53	0.75	2.53	2.41
Traders	0.76	0.92	4.94	0.74	2.55	2.60
Farmers	0.78	0.86	4.56	0.75	2.33	2.45
Upper	0.68	0.86	4.99	0.74	2.14	2.62
Unknown	0.82	0.86	3.78	0.75	2.01	2.03

Note: $n(c)$ and $\bar{n}(c)$ computed from the point of view of men.

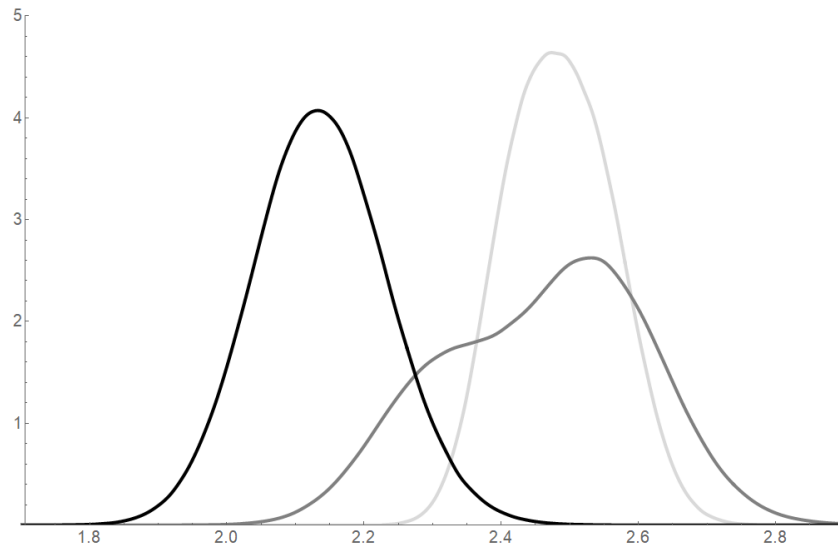Table K.2: Computed Reproductive Rates - n and $\bar{n}$
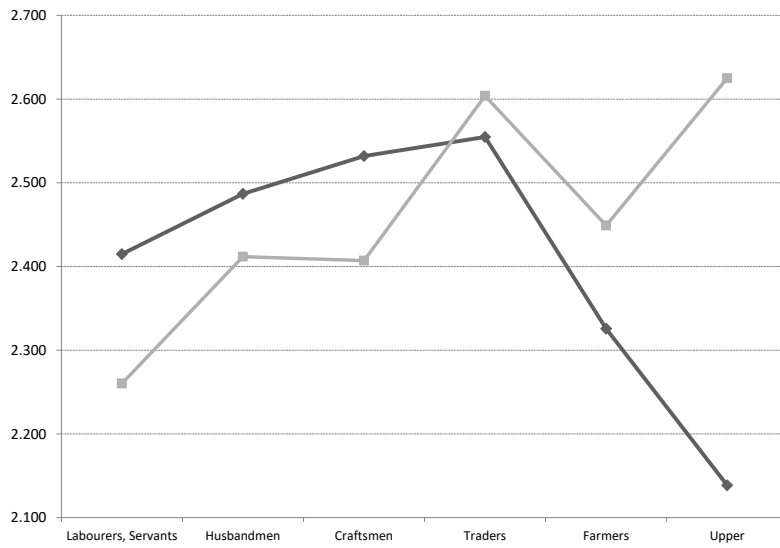
Notes: dark grey line: $n(c)$. light grey line: $\bar{n}(c)$. Data from Table K.2.

Figure K.4: Net reproduction rates Across Social Groups
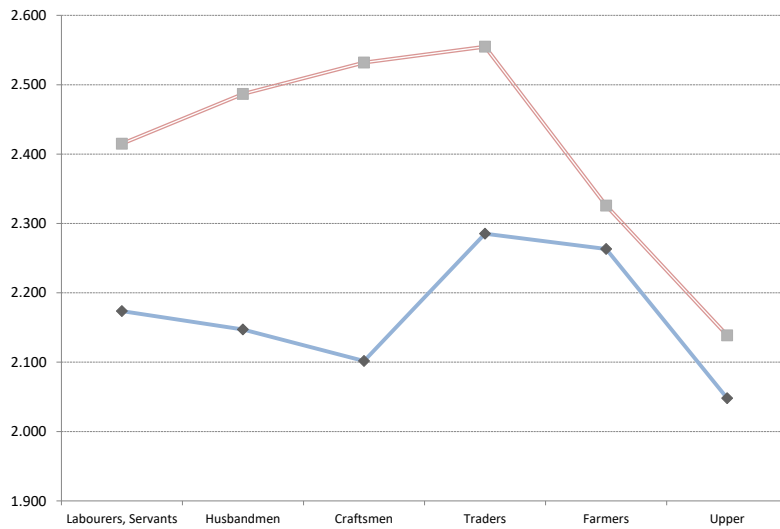
Notes: blue, women; red, men.

Figure K.5: Net reproduction rates $n(c)$: the point of view of men

Note: Lower class includes Labourers, Servants, Husbandmen, and Craftsmen. Middle class includes Traders and Farmers.

Figure K.6: Sampling distributions of $n(c)$ for men by social class (black=upper, gray=middle, light gray=lower)

	Lower Classes	Middle Classes	Upper Classes
Lower Classes		-0.02	-0.035
		(0.45)	(0.00)
Middle Classes			0.32
			(0.03)

Note: Mean net reproduction gap between groups from the simulation with one-tailed p-value in parentheses.

Table K.3: Tests of statistical significance on mean differences between groups

L   Robustness: Source of Occupations

As mentioned before, the occupations in our data could have been recorded at three types of events: a marriage, a baptism of a child, and a death. For our general analysis, we did not distinguish between these three, but for the childlessness rates, we worried that the pattern across social groups could be skewed if we were more likely to observe an individual's occupation if they had children. This would be true if an individual's occupation was observed in a baptism register. Thus, in this robustness check, we limit the occupations to those given at marriage or burial, moving the other individuals into the unknown occupation category. Table L.1 shows what this implies for the sample size of the different groups.

The results of the Cox regression for the probability to move from zero to one child are presented in Table L.2. The first column is the one reported in the main text, while the second column is based on the alternative way of defining occupations. After removing the occupations obtained from baptism registers, all of the occupation groups become insignificant, though the basic gradient in childlessness remains across all classes. Farmers and the upper classes have much higher rates of childlessness although these rates are insignificant because of the smaller sample size under the restricted occupations. We also see a sharp decrease in the childlessness rate of couples of unknown occupation, which confirms our suspicion that the high rates of childlessness in this group were largely driven by this issue. We do not use this as our primary specification because the sample size for each occupational category shrinks a lot when restricting occupations, increasing the standard errors. Figure L.1 shows the average levels of childlessness between the two specifications and confirms these findings.

We also re-estimated all the parity progression ratios with the alternative definition of occupations and compute the gross fertility of women using Equation (5). As Figure L.2 shows, the source of occupations does not strongly influence the gross marital fertility of women in our dataset. The general class gradient remains the same or even becomes slightly more pronounced.

Figure L.3 displays the female net reproduction rates recalculated from all of the various regressions. The net reproduction of the lower classes decreases a bit since removing the baptism occupations increased their childlessness and decreased their fertility. However, the general gradient holds. The middle classes have higher net reproduction rates than the lower and upper classes. Thus, the source of occupations does not explain away the social gradient found in our baseline results.

	benchmark	excluding occupations from baptism registry
Labourers, servants	1665	996
Husbandmen	1222	846
Craftsmen	1371	919
Traders	603	415
Farmers	379	241
Merchants, Professionals	406	359
Gentry	183	149
Unknown Occupation	8901	10805

Table L.1: Number of married women with known death date

|	(1)	(2)
Dependent Variable: Risk of first birth	benchmark	excluding occupations from baptism registry
Husbandmen	−0.030	−0.008
	(0.041)	(0.051)
Craftsmen	−0.013	0.0278
	(0.039)	(0.050)
Traders	−0.038	−0.013
	(0.051)	(0.064)
Farmers	−0.203***	−0.107
	(0.062)	(0.079)
Upper Class	−0.204***	−0.092
	(0.053)	(0.060)
Unknown Occupation	−0.284***	−0.083*
	(0.032)	(0.038)
Observations	14,730	14,730

Note: Dependent variable is risk of first birth. Reference category is Labourers/Servants. 26 parish dummies and 4 period dummies are included. *p<0.1; **p<0.05; ***p<0.01

Table L.2: With and without Recording Occupations from Baptism Registry: $z(c)$
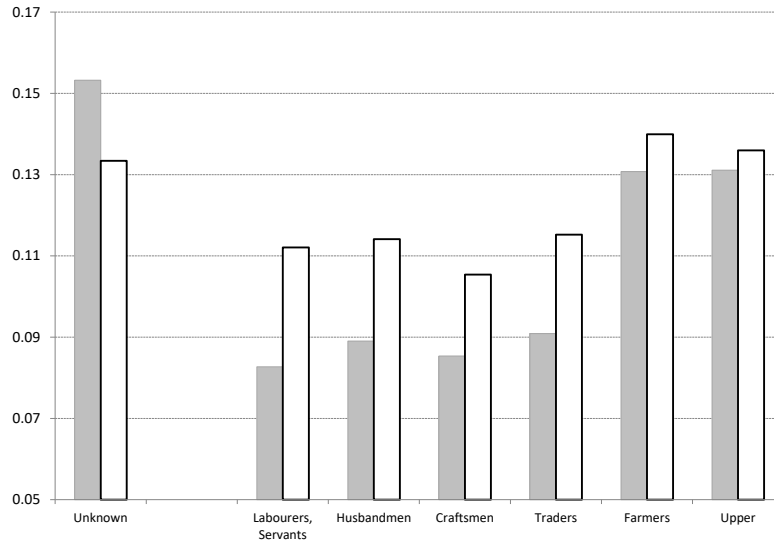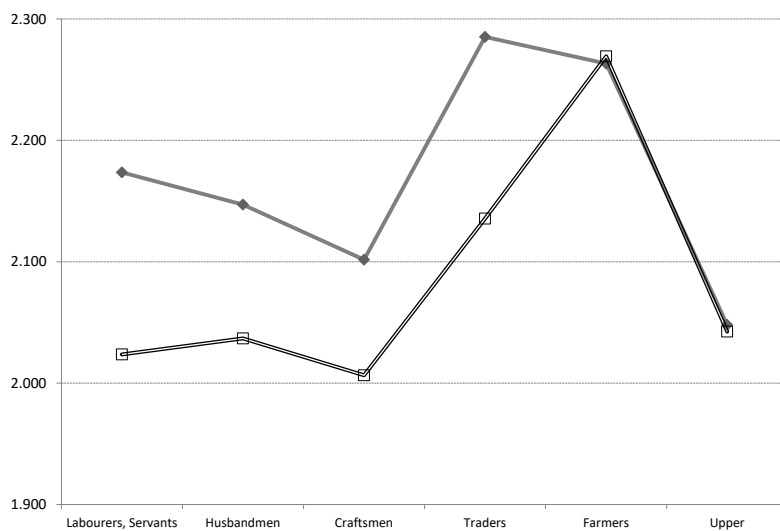
Note: filled grey, benchmark; solid white, alternative.

Figure L.1: Computed Childlessness Rate with Alternative Definition of Occupations
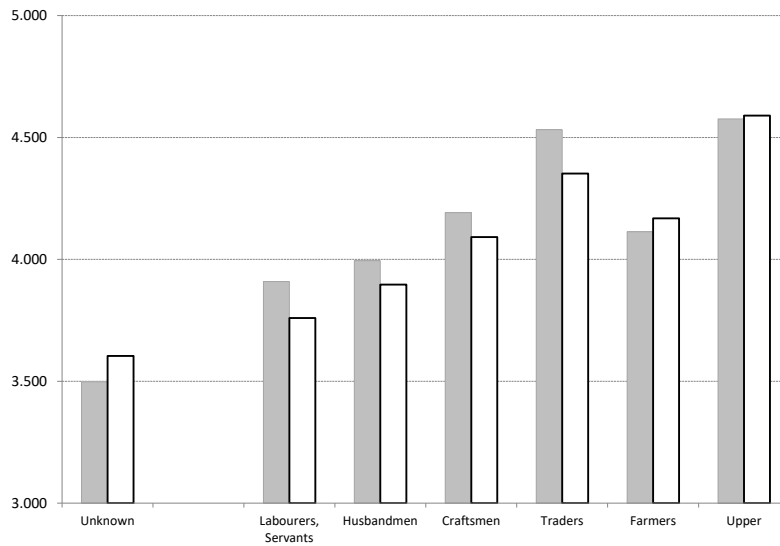
Note: filled grey, benchmark; solid white, alternative.

Figure L.2: Computed Births with Alternative Definition of Occupations

79

Note: filled grey, benchmark; solid white, alternative.

Figure L.3: Net reproduction rates $n(c)$ across Social Groups of the Baseline compared with the Sample Excluding Occupations from Baptism Registers