WARWICK
THE UNIVERSITY OF WARWICK

Aug 2016 No.296

Fecundity, Fertility and the Formation of Human Capital

Marc Klemp and Jacob Weisdorf

WORKING PAPER SERIES

Centre for Competitive Advantage in the Global Economy

Department of Economics

CAGE
Competitive Advantage in the Global Economy

E·S·R·C
ECONOMIC & SOCIAL RESEARCH COUNCIL

Fecundity, Fertility and the Formation of Human Capital*†

Marc Klemp Jacob Weisdorf

July 16, 2016

Abstract This research explores a fundamental cause of variation in human capital formation across families in the pre-modern period, as well as the mitigating effects of family-level economic prosperity. Exploiting a vast genealogy of English individuals in the 17^{th} to the 19^{th} centuries, the study proposes and tests the hypothesis that lower parental reproductive capacity positively affected the socioeconomic achievements of offspring. In particular, the research establishes an effect of reproductive capacity on offspring human capital in the pre-modern era. Using the time interval between the date of marriage and the first birth as a measure of reproductive capacity, the research establishes that children of parents with lower fecundity were more likely to become literate and employed in skilled and high-wealth professions. The analysis finds that parental fecundity significantly affected the number of siblings, indicating that a trade-off between child quantity and quality was present in England during the industrial revolution and supporting leading theories of the origins of modern economic growth. Furthermore, it finds that the effect was weaker for the socioeconomic elite, who could offset the cost of additional children by raising total investment in offspring human capital.

Keywords Human Capital Formation, Child Quantity-Quality Trade-Off, Reproductive Capacity, Fecundity, Demographic Transition, Long-Run Economic Growth

JEL Classification Codes J13, N30, O10.

*We thank the seminar and workshop audience at the universities of Brown, Cambridge, Copenhagen, Lund, Madrid (Carlos III), Odense, Oxford, Pisa, Utrecht, Valencia, as well as at the European University Institute and the London School of Economics. For specific comments and suggestions we thank Thomas Barnebeck, Sascha Becker, Tommy Bengtsson, Steve Broadberry, Davide Cantoni, Mario Carillo, Matteo Cervellati, Francesco Cinnirella, Gregory Clark, Carl-Johan Dalgaard, Romola Davenport, Matthias Doepke, Martin Dribe, Véronique Gille, Cormac Ó Gráda, Lothar Grall, Knick Harley, Tim Hatton, Jane Humphries, David Jacks, Nicolai Kaarsen, Tim Leunig, Anastasia Litina, Chris Minns, Steven Nafziger, Martin Nielsen, Alessandro Nuvolari, Karl Gunnar Persson, Alice Reid, Leigh Shaw-Taylor, Richard Smith, Simon Szreter, Ludger Voessmann, Joachim Voth, Jan de Vries, Patrick Wallis, David Weil and Tony Wrigley, and particularly Oded Galor. We thank Paul Sharp for research assistance; Kerstin Enflo for her contribution at an early stage. We also thank the Cambridge Group for making their data available, and Gill Newton for help with clarifying questions. The research of Klemp is funded by the Carlsberg Foundation and by the Danish Research Council reference no. 1329-00093 and reference no. 1327-00245. Jacob Weisdorf thanks the Robert Schumann Centre of Advanced Studies at the European University Institute in Florence for its financial support through a Jean Monnet Fellowship, as well as the financial support provided by Tine de Moor through her ERC Grant "Unified We Stand" (Grant no. 240928). Marc Klemp (marc_klemp@brown.edu): Department of Economics and Population Studies & Training Center, Brown University, 64 Waterman St., Providence, RI 02912, USA; Department of Economics, University of Copenhagen, ster Farimagsgade 5, building 26, DK-1353 Copenhagen K, Denmark. Jacob Weisdorf (jacobw@sam.sdu.dk): Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark; CEPR; and CAGE.

†A working paper version of this paper was circulated with the title "Fecundity, Fertility and Family Reconstitutions Data: The Child Quantity Quality Trade-off Revisited".

1 Introduction

The demographic transition and the entailing rise in the global level of human capital and living standards over the past two centuries has caused one of the most significant transformations of human societies in the course of human history. While living standards in the world economy stagnated during the millennia preceding the demographic transition and the industrial revolution, the fall in fertility rates over the past two centuries is associated with rising levels of human capital and a twelvefold increase in income per capita.[1]

Long-run growth theories posit that technological progress during the pre-industrial period raised the incentive to reduce fertility in order to increase investment in the average level of human capital of offspring, and that this helped cause the demographic transition and drove the industrial revolution (Galor, 2011). Similarly, evolutionary growth theories hypothesize that individuals with a bias towards low fertility improved the income potential and hence the reproductive success of their lineage, gradually increasing the representation of this growth-enhancing trait in the population, contributing to the process of development and the take-off from stagnation to growth (Galor and Moav, 2002; Galor and Klemp, 2016).[2]

This research proposes and tests the hypothesis that human fecundity – the individually variable capacity to reproduce – has played a significant role in the formation of human capital in the pre-industrial world. In particular, by exploiting a vast genealogy of English individuals living in the 17$^{\text{th}}$ to the 19$^{\text{th}}$ centuries, the study finds that lower parental reproductive capacity positively affected the socioeconomic achievements of offspring in the pre-industrial world. Using the time interval between the date of marriage and the first birth as a measure of reproductive capacity, and hence as a novel source of unplanned, exogenous variation in family size, the research establishes that children of parents with lower fecundity were more likely to obtain literacy, a skilled profession, and high occupational wealth. The analysis further establishes that parental fecundity negatively affected the number of surviving offspring, indicating that a trade-off between child quantity and child quality was present in England before and during the industrial revolution.

Furthermore, the results highlight the importance of biocultural factors in the economic development process. Over most of human existence, Malthusian pressures has governed the evolution of the size of the population and has conceivably shaped the biological and cultural composition of the population, contributing to the process of development and the take-off from stagnation to growth. The present results provide the first evidence that lower fecundity historically raised both the level of education and the occupational status of the next generation, suggesting that

[1]The transition from stagnation to growth and the associated divergence of income per capita across the globe have been the subject of intensive research in the growth literature in recent years (Galor and Weil, 1999, 2000; Galor and Moav, 2002; Hansen and Prescott, 2002; Lucas, 2002; Lagerlöf, 2003; Doepke, 2004; Galor, 2005; Strulik and Weisdorf, 2008; Dalgaard and Strulik, 2010; O'Rourke et al., 2013).

[2]Originating in the field of evolutionary biology, the notion of a trade-off between the quantity and quality of offspring found some of its first uses in economics by Becker (1960), Becker and Lewis (1973), and Becker and Tomes (1976). Furthermore, the notion plays a central role in the literature of unified growth theory (Galor and Weil, 2000; Galor, 2011).

the transition from stagnation to growth may be attributed to decreasing levels of fecundity in the population in accordance with evolutionary growth theories and empirics.[3]

In addition, the existence of a historical trade-off between the quantity and quality of children, is particularly relevant for the assessment of unified growth theories that explain the transition from millennia of economic stagnation to an era of sustained economic growth, and its connection with a demographic transition. In particular, the leading unified growth theories crucially depend on the assumption that there existed a trade-off between the number of offspring in a family and the occupational income of the family offspring. For instance, Galor and Weil (2000) have argued that technological progress during England's industrial revolution motivated parents to lower their fertility rate in order to increase their investment in the human capital of each of their offspring, hence facilitating an increase in their children's income. The current research presents the first evidence suggesting that population dynamics in England – the cradle of the industrial revolution – were indeed characterized by a child quantity-quality trade-off around the time of the industrial revolution. In particular, the analysis suggests that the number of family births negatively affected children's levels of human capital and social position. This finding therefore lends credence to the hypothesis that increasing returns to education around the time of the industrial revolution contributed to human capital formation via a child quantity-quality trade-off and to the onset of the demographic transition and the evolution of societies from an epoch of stagnation to sustained economic growth.[4]

The analysis exploits a rich set of demographic data which is built from historical church books and is regarded as the gold standard of pre-modern demographic statistics. It covers one of the most extraordinary periods in British economic history when England underwent the industrial revolution, which, according to the leading long-run growth theories, represents the key break-point in the global transition from millennia of economic stagnation to the rise of modern economic growth. Furthermore, contrary to most other historical demographic datasets covering this time period, the data are reported at the individual level making it possible for the first time to study the direct within-family effects a of couple's reproductive capacity on their children's socioeconomic achievements later in life. The data provide unusually detailed statistics about individuals' socioeconomic performance, including their literacy status, their working skills, and their occupational wealth, as

[3] In particular, Galor and Klemp (2016) have established that a parental disposition towards moderate fertility was conducive to long-run reproductive success in historical Quebec: subsequent generations of couples prone to restrained fertility had higher reproduction that more than compensated for the reduced fertility in the first generation. The leading theories explaining how fertility limitation can be a dominant reproductive strategy builds on the hypothesis of a parental trade-off between the quantity and quality of children. Parents disposed to low fertility and living in surroundings where the rewards for human capital investments are high can help their offspring proliferate by raising their occupational income by way of increasing their human capital (Galor and Moav, 2002). Furthermore, Clark and Hamilton (2006) have demonstrated, using English wills record, that that the wealth at death of male testators was positively correlated with the number of offspring (with the richest testators leaving twice as many surviving offspring as the poorest), consistent with the notion that occupational income increased fertility in the pre-demographic transition era.

[4] This finding is consistent with conclusions from studies based on aggregate (i.e., non-individual) census data from other countries in historical Europe which have illustrated that lower birth rates were associated with higher educational attainments on aggregate levels (Basso, 2012; Becker et al., 2010; Fernihough, 2011; Diebolt et al., 2016).

derived from their occupational title, not to mention an extraordinary set of family characteristics, including the occupational status and longevity of the parents. The data therefore enable a comprehensive investigation of the historical effects of parental fecundity on the formation of offspring human capital.

Through its focus on the effect of fecundity, rather than fertility itself, the research generates a methodological innovation that is widely applicable for analyses of historical genealogical data in societies where marriage marked a couple's intention to conceive. In particular, in light of the social norm observed in pre-modern England, where marriage did mark the intention to conceive, the research exploits variation in the random component of the time interval between the date of marriage and the date of first birth to capture the effect of fecundity on human capital.[5]

The study establishes that the time to first birth is unrelated to the socioeconomic conditions of the parents, including their occupational wealth, human capital attainment, and longevity, among other factors. Furthermore, the analysis demonstrates that the distribution of time to first birth in our data is virtually identical to that of modern-day populations studied by reproductive health researchers in which couples' reported that marriage marked the onset of unprotected sex. Moreover, our analysis shows that this arguably exogenous variable has a strong, statistically significant effect on the number of children born within a family over this historical time period. Therefore, in addition to enabling an investigation of the effect of fecundity on human capital formation, our identification strategy helps resolving the endogeneity problems that are inherent in estimating the trade-off between the quantity and quality of children.[6]

The analysis shows that a delay in the conception of the first child by around two years, corresponding to one less surviving offspring, on average: (i) increased the probability of each offspring acquiring a skilled profession by 7.9 percentage points, (ii) increased the probability of each offspring acquiring literacy skills by 7.3 percentage points, and (iii) increased offspring occupational wealth by 21.7 percent of a point on a seven-point scale. Furthermore, the analysis establishes that the effect was weaker for the socioeconomic elite, who could offset the cost of additional children by raising total investment in offspring human capital.

The paper proceeds as follows. The next section explains the empirical strategy. Section 3 then describes the data and the main variables. In Section 4, time to first birth is related to observable information and its usefulness as an exogenous determinant of family size is discussed. The main results of the effect of time to first birth on offspring quality is presented in Section 5. Next, Section

[5] As will become clear, neither intentional nor unintentional attempts to conceive before marriage is a threat to our identification strategy. Our sole identifying assumption is that marriage marked the intention to conceive.

[6] The study by Galor and Klemp (2016) also employs the time from the marriage to the first birth as an exogenous source of variation in fertility. Furthermore, a similar type of identification strategy is used in (Aguero and Marks, 2008, 2011) and Jensen (2012), using failure to conceive a second child despite regular intercourse without contraception as an instrument for family size among modern populations. In addition, previous related empirical analyses of data for modern economies have instrumented fertility by the occurrence of twin births or by the sex composition of the first births in the family (Rosenzweig and Zhang, 2009; Angrist et al., 2010; Black et al., 2005). These instruments are generally weak in historical populations in which fertility is much higher. Furthermore, the occurrence of twin births has been shown to be a problematic instrument due to the direct effect of twinning on child quality (Rosenzweig and Zhang, 2009).

6 discuss the use of time to first birth as an instrument for family size, and the effect of family size on human capital is estimated. Finally, Section 7 concludes.

2 Empirical Strategy

The identification of the causal effect of parental fertility (offspring quantity) on the human capital achievements of children (offspring quality) is affected by two major econometric obstacles, namely the potential for omitted variable bias and the potential for reverse causality. Omitted variable bias obscures the effect of quantity on quality and arises if correlates of both offspring quantity and quality are not controlled for. For instance, wealthier parents may be able to produce more offspring of higher quality, and failing to account for parental wealth may therefore obscure the effect of child quantity on child quality.[7] In particular, an observed negative relationship between the number of children and their human capital achievement may be mis-interpreted as indicative of the lack of a trade-off between quantity and quality.

Likewise, reverse causality from offspring's quality to the aggregate quantity of offspring may also obscure the presence of a trade-off between quantity and quality of offspring. For instance, an adverse effect of low offspring quality on the offspring survival rate may positively affect the total number of offspring born via a child replacement effect, generating a negative correlation between the human capital achievement and the quantity of offspring regardless of the presence or absence of a negative effect of fertility on child quality.

This research mitigates these obstacles by focusing on the effect of fecundity, rather than fertility, on child quality (see Galor and Klemp, 2016, for a similar empirical strategy). Furthermore, it exploits the inherent randomness of conception in order to identify the effect of fertility on the quality of offspring.[8]

In particular, in light of the social norm observed in historical England, whereby marriage marked the intention to conceive, the research exploits variation in the time interval between the marriage and the first birth (TFB) conditional on the maternal marriage age and other socioeconomic and environmental characteristics, in order to capture its effect on child quality. Since TFB proxies parental fecundity, this allow us to infer the effect of fecundity on offspring human capital over this historical time period.[9]

After their first birth, married women in pre-modern England typically continued to give birth until sterility set in at menopause, usually after age 40 (Wrigley et al., 1997). In light of the fact that premarital births were seen as an immoral act by the Church of England, and by English society

[7]Indeed, in the time period under investigation, wealthier parents simultaneously gave birth to more children and allocated more resources to the education of their offspring. See e.g. Boberg-Fazlic et al. (2011); Clark and Hamilton (2006); Leunig et al. (2011).

[8]As shown below, our measure of fecundity is not related to any observable parental characteristic in the data, controlling for maternal age. Furthermore, as will be argued below, our measure of fecundity is likely not to be affected by unobserved confounding factors.

[9]The event of couples having had unprotected intercourse before their marriage will not affect the validity of the empirical strategy because fecundity and TFB is unaffected by past sexual activity – in other words, conception is a "memoryless" process.

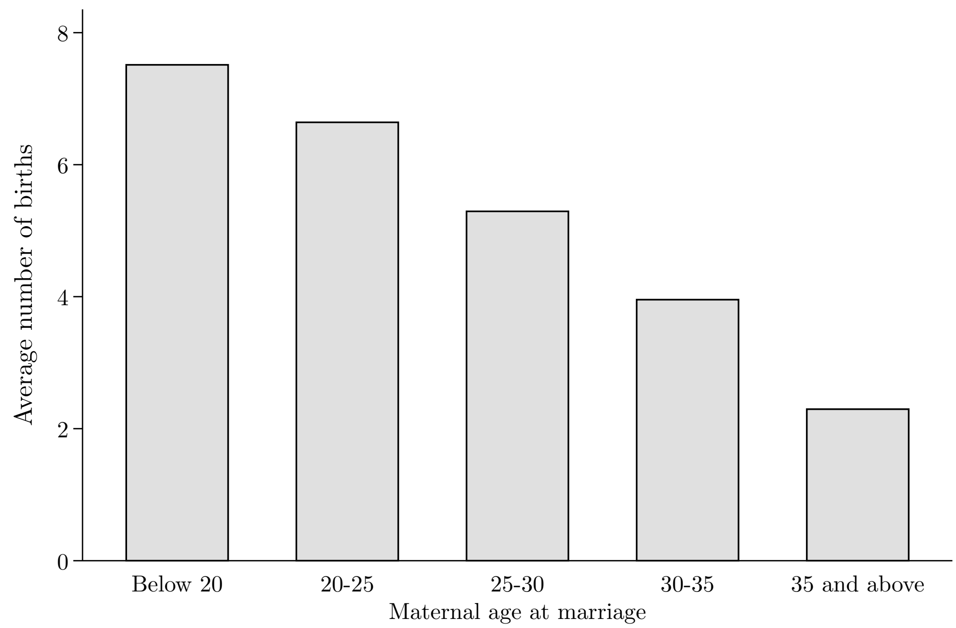


Figure 1: Number of births by the mother's age at marriage among 4,582 families in the data in which both parents survived to the date at which the mother turned 50.

as a whole, the postponement of marriage was a key form of fertility limitation.[10] With an average age at marriage of 25 years (Wrigley et al., 1997), fertility in pre-modern England was therefore high in comparison to fertility in the modern period. In particular, conditional on parental survival until the date at which the mother reached the age of 50, a married woman in the sample used in the present analysis produced on average six children.[11] Figure 1 depicts the strong correlation between the age at marriage and the total number of births over the period corresponding to the data.[12] This highlights the fact that the onset of reproduction was an important determinant of completed family size.

In an additional analysis, we therefore employ TFB as a plausibly exogenous source of variation in the number of children and examine the causal effect of an additional child on the average offspring's achieved level of human capital in adulthood.[13]

[10]Indeed, a main reason why English upper-class families were larger was because of a lower maternal marriage age (Boberg-Fazlic et al., 2011).

[11]See Table A.1 in the appendix. The average mother in the 18$^{\text{th}}$ century, including also those who died before age 50, gave birth to approximately five children (Wrigley et al., 1997).

[12]Birth control was practiced not only by means of regulating the wife's age at marriage but also within marriage by sexual abstinence, coitus interruptus, and extended breastfeeding (McLaren, 1978; Santow, 1995). Using an extended sample of the present data, Cinnirella et al. (2013) have found that pre-modern couples practiced parity-dependent birth spacing and that they responded to variations in living standards (measured by real wages and wheat prices) by increasing the time elapsed between two births when times were hard. Importantly, the time from the marriage to the first birth is not associated with variations in living standards.

[13]In high-fertility populations, TFB is a strong determinant of unplanned variation in family size. As in contemporary, post-demographic societies, the length of historical birth intervals were subject to parental decision making (Cinnirella et al., 2013), meaning that parents could delay the onset of unprotected intercourse after a birth and/or

Figure 2: The histogram depicts the durations (per two-week interval) from marriages to first marital births of 38,914 couples in England between 16$^{\text{th}}$ and the 19$^{\text{th}}$ century who gave birth between the first and 260$^{\text{th}}$ week of their marriage date.

TFB is a standard measure of fecundity in the fields of historical demography and medicine and is used to estimate the level of fecundity of historical populations.[14] This study exploits the fact that fecundity is primarily determined by biological factors, such as age, and not by unobserved socioeconomic factors, in order to estimate the effect of TFB, and by extension fecundity, on offspring human capital.

Indeed, as depicted in Figure 2, which is generated using the historical church book data described in detail below, a marriage over this period signalled a deliberate attempt by the couple to conceive. A sharp spike in births occurs starting in the 34$^{\text{th}}$ week after marriage and a quarter of all births thereafter happen within the 36$^{\text{th}}$ and 44$^{\text{th}}$ week.[15] Furthermore, the sampled couples' adherence to the existing social and religious norms is reflected in the fact that premarital conception was relatively uncommon at the time. In particular, three quarters of the sampled first births occurred after 35 weeks of marriage, with the existence of premature births suggesting that even this fraction understates the share of offspring conceived after marriage.[16]

engaged in other contraceptive practices to expand the birth spacing interval. Therefore, birth intervals are not informative about unprotected sexual activity and cannot be used to infer unplanned variation in family size.

[14]E.g. Bongaarts (1975); Olsen and Andersen (1999); and Woods (1994).

[15]Full term children are born upon 38 weeks of gestation, but since marriage may not coincide with the ovulation period, children born between weeks 36 and 44 are considered at term.

[16]It should be noted that the fraction of births occurring after 35 weeks is larger in the Canadian province Quebec in the period between the 17$^{\text{th}}$ and the 18$^{\text{th}}$ century (Galor and Klemp, 2016), potentially due to an increased prevalence of pre-marital conceptions in England after the 18$^{\text{th}}$ century. The results are robust towards excluding this later period. It should furthermore be noted that in the complete data of all 41,238 mothers, less than two percent of births occurred prior to the marriage date, while 3.6 percent of the births occurred after five years of

The analysis accounts for a range of socio-environmental factors. In particular, cultural and socioeconomic factors that are accounted for by the inclusion of control variables capturing parental human capital, occupational wealth, and family location. Additional control variables include dummies for parental marriage age, marriage year, quarter of marriage, literacy, occupational skills, as well as the gender of the individuals, and their geographic location, birth order, quarter of birth, quarter of birth of the firstborn sibling, and more.[17]

3 Data and Main Variables

The study exploits information on individuals of a sample of 26 English parishes. The information was originally recorded between 1541 and 1871 and later transcribed and collated from English church books by the *Cambridge Group for the History of Population and Social Structure* and documented in Wrigley et al. (1997).[18] The parishes were selected by the Cambridge Group on merit of data quality and have been shown to represent England as whole rather well (Wrigley et al. (1997) pages 41ff). Their locations are depicted in Figure 3. In addition to documenting baptisms, marriages, burials, and the genealogy of individuals, the data frequently contains information on individual occupations and literacy status.[19]

3.1 Sample

The study focuses on offspring from families in which the firstborn child was conceived after the day of the wedding, meaning families where the time span between the parental marriage and

marriage. Finally, it should be noted that the small hump in the histogram around week 11–16 possibly reflects the sum of the average time from conception to the realization that the woman is pregnant and the average time to arrange a "shotgun" wedding.).

[17]While the present study is, to our knowledge, the first to investigate the effect of TFB in a family on offspring literacy and occupation in adulthood, studies using data for modern populations have investigated the related phenomenon of the effects of time to pregnancy (i.e., the time from the intention to conceive until pregnancy) on the birth-related outcomes of the particular birth resulting from the pregnancy. Some of these studies have found that increased time to pregnancy is associated with adverse birth outcomes such as miscarriage, preterm delivery, caesarean delivery, neonatal death, extrauterine pregnancies, low Apgar score, low umbilical vein pH, or need for neonatal intensive care (Joffe and Li, 1994; Henriksen et al., 1997; Basso and Baird, 2003; Basso et al., 2005; Axmon and Hagmar, 2005; Raatikainen et al., 2010). Meanwhile, other studies have found that time to pregnancy is not related to some of these, as well as other, outcomes, including preterm delivery, small-for-gestational-age, sex ratio, and birthweight (Joffe and Li, 1994; Cooney et al., 2006; Joffe et al., 2007). It is important to note that a possible adverse effect of time to pregnancy on the birth outcome of the particular birth will presumably bias against finding a positive effect of TFB on offspring education when the first birth is included in the sample. Nevertheless, we find strong positive effects of TFB on offspring quality in adulthood across all siblings no matter if we control for birth order fixed effects or not. Furthermore, we find that the estimated effect of TFB is not significantly different if the sample is restricted to birth orders above one.

[18]Table A.2 in the appendix provides an example of a reconstituted family from the sample, displaying the available information about family individuals.

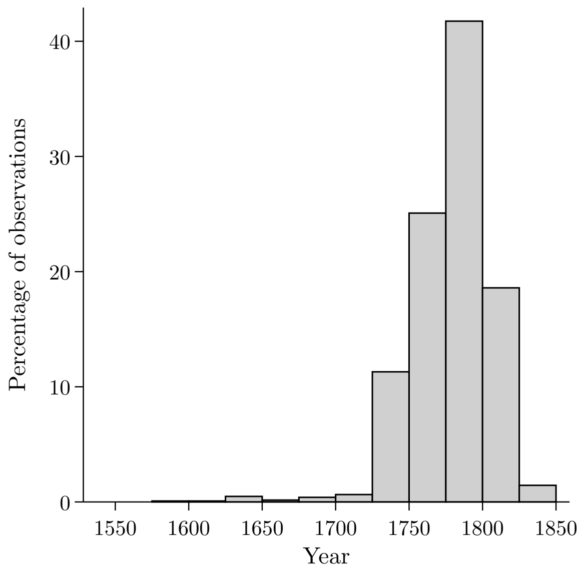
[19]For 13.2 percent of the individuals in the sample with known birth or baptism date, the date of birth is known. Given the fact that almost all children were baptized within one month of birth (Midi Berry and Schofield, 1971), the date of birth of individuals with unknown birth date is estimated to be three weeks prior to the date of baptism. Likewise, for 2.6 percent of the individuals in the sample with known death of burial date, the date of death is known. Since burials usually took place within three days of death (Schofield, 1970), the date of death of individuals with unknown death date is estimated to be the burial date.

Figure 3: Locations of the parishes (based on Schofield, 2005).

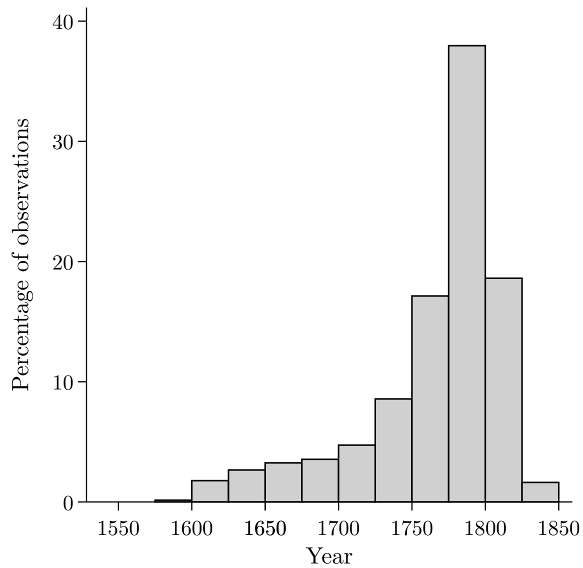
TFB was 40 weeks or more.[20] Furthermore, in order to avoid confounding the effect of parental mortality on family size and investments in offspring human capital, the study follows standard demographic procedures by excluding families where parents were unable to exploit their full reproductive potential due to the premature death of either spouse. The sample is therefore restricted to *completed families* in which both parents survived until the mother reached the age of 50 (Wrigley et al., 1997, p. 359). Because a missing birth or death date of an individual implies inter-parish migration (Souden, 1984), the restriction of the data to completed marriages (implicitly asking that the mother's birth and death dates are both known) further mitigates the possibility of unobserved births occurring outside of the sampled parishes. Families where the father's birth and death dates are missing are excluded from the sample for the same reason. In total, these restrictions result in a sample of 1,517 individuals with known literacy and/or occupation coming from 729 families.[21]

[20]The average time to ovulation is two weeks. Thus, in light of the 38 week gestation period, the average time from the onset of intercourse to the first birth in couples that conceive in the first cycle is 40 weeks. Table A.3 in the appendix establishes that the qualitative conclusions are robust to the inclusion of families in which the first birth occurred in weeks 38–40.

[21]In this sample, literacy is known for 1,248 individuals, working skills for 652 individuals and occupational wealth for 686 individuals. The summary statistics of the combined sample are presented in Table A.1 in the appendix.



(a) Observations in literacy sample.



(b) Observations in occupation sample.

Figure 4: Histograms of birth dates for the observations on literacy and occupation in the total regression sample.

Figure 4 depicts the distributions of birth dates of individuals in the sample with known occupation or literacy status. Nine tenths of the individuals were born between 1690 and 1814, spanning the majority of the classic years of the industrial revolution.

3.2 Main Variables

In the main analysis that explores the effect of reproductive capacity on offspring human capital, the dependent variable is either literacy skills, occupational skills or occupational wealth. The main independent variable is TFB. In the additional analysis, examining the mechanism through which TFB affects offspring human capital, the same dependent variables are used while the independent variable is the number of children surviving to age five, instrumented by TFB.

The study accounts for important sources of variation in living standards among the sampled families by inferring parental occupational wealth and working skills from occupational titles, in addition to parental literacy information. Furthermore, the study controls for parish-type and parish fixed effects, accounting this way for differences in living standards and other determinants of offspring human capital (e.g., access to education possibilities) linked to differences in location. Moreover, the study accounts for secular changes by controlling for the year of marriage of the parents.[22] The next four subsections explain in more detail how the main variables are derived.

[22]Table A.4 in the appendix establishes that the results are robust to inclusion of dummy variables indicating the time period in which the individual was born.

3.2.1 Dependent Variables

The empirical examination of the quantity-quality trade-off exploits the exceptionally rich information on pre-modern individual human capital attainments contained in the dataset. In particular, individual human capital attainments are inferred from two mappings of occupational titles and from a proxy of literacy skills.

Literacy. Literacy status is inferred from the existence or absence of an individual's signature on his or her marriage certificate. Whereas literate individuals would leave their signature on their marriage certificate, illiterate individuals simply left a mark. Because individuals who were able to write would, generally, be able to write their own name, the absence of a signature has a strong predictive value on the absence of literacy.

Occupational skills. This study uses the extensive HISCO and HISCLASS schemes (Leeuwen et al., 2007; Leeuwen and Maas, 2011) in order to divide the sampled individuals into skilled and unskilled workers based on the educational training required to conduct the work described by their occupational titles.[23] To this end, a standard two-step procedure is employed. An occupational title is first assigned its relevant five-digit code specified in the HISCO system. Subsequently, the code is entered into the HISCLASS system, which classifies the professional skills associated with the occupational title using a two-dimensional scheme quantifying the academic and vocational training needed to conduct the work.[24] For example, according to the HISCO scheme, an English factory worker would be classified as code number 99930, which according to the HISCLASS scheme designates an "unskilled" profession.[25]

Because the HISCLASS scheme does not designate a skill level to the titles "Paupers" and "Gentry" recorded in the data, individuals with these titles are excluded from the analysis where working skills is the outcome variable.[26]

Occupational wealth. The occupational titles are also divided according to the occupational wealth of the profession. This is done using the classification of Clark and Cummins (2010) which is based on information of wealth and occupations recorded in pre-modern English wills. In ascending order of wealth, this classification groups occupations into labourers, husbandmen, craftsmen,

[23]The HISCO system is a historical extension of the *International Standard Classification of Occupations* (ISCO) managed by the International Labour Organization (ILO). The HISCLASS system is a historical extension of the *Dictionary of Occupations* (DOT) system, which gives scores for the requirements of skills of a wide range of occupations and which was originally created in the 20th century by the US Employment Service to match job seekers to jobs (Leeuwen and Maas, 2011).

[24]Using the earliest recorded occupations of the sampled individuals, more than hundred distinct occupational titles in the data were mapped into skilled and unskilled professions. 89 percent of the occupations were derived from marriage records or the earliest ecclesiastical event thereafter (typically the baptism of firstborns). Approximately 7 percent of the occupations were derived from burial records, i.e, were recorded at the time of the individual's death. The remaining occupational titles (approximately 4%) were derived from an intermediate event, i.e., the baptism (or burial) of offspring of parity two or above.

[25]The occupational titles of our sample were coded using `http://historyofwork.iisg.nl`.

[26]As established in Table A.5 in the appendix, the findings are robust to their inclusion on the assumption that paupers were unskilled and gentry skilled.

traders, farmers, merchants, and gentry. The occupational wealth variable is given the values 1–7 according to this ordering.

It is important to note that the three measures of human capital are not perfectly correlated and thus capture different variants of human capital. In the regression sample, the correlation coefficient between literacy skills and occupational skills is 40 percent, the correlation coefficient between literacy and occupational wealth is 44 percent, and the correlation coefficient between occupational skills and occupational wealth is 63 percent; all correlations are highly statistically significant ($p < 0.0001$).

3.2.2 Independent Variables

Time to first birth (TFB). The main independent variable is the time, measured in years, from the parental marriage date to the birth date of the first sibling. The exogeneity of TFB will be discussed in Section 6.

Number of surviving siblings. On the premise that surviving children consume more parental resources than those dying during childhood, the independent variable of interest here is family size measured as the number of siblings surviving to age five.[27]

3.2.3 Control Variables on the Individual Level

Gender. Parental investment in the education of their children depended on gender. This probably reflects the different labour force participation rates and occupational wealth of men and women at the time. Parents were more likely to invest in their male offspring (see Klemp et al., 2013) which is reflected by higher literacy rates and occupational wealth among men compared to women in the data. A dummy variable indicating gender is therefore included in the models.

Birth Order. Many studies have linked birth order with human capital achievements, both in historical and in present times (see e.g., Ejrnæs and Pörtner, 2004; Black et al., 2005; Klemp et al., 2013). Due to the mechanical association between the number of siblings and birth order, and due to the fact that we are interested in the effect of TFB on human capital independently of birth order effects, we include dummies for each birth order to account for birth order fixed effects in all the main regression analyses.[28]

Non-Sunday baptism. Although the Prayer Books of the English Church prescribed that baptisms take place on Sundays, many families did not submit to this rule. Non-Sunday baptism services were possible for an additional fee, which means that non-Sunday baptisms might pos-

[27] Table A.6 in the appendix establishes that the results are robust to the inclusion of all births in the calculation of the family size rather than children surviving to age five.

[28] Appendix Table A.7 establishes that the findings are robust to excluding these dummies, as well as to controlling for birth order effects in other ways. Furthermore, that table also establishes that controlling for birth order is the most conservative approach.

itively reflect family income.[29] Meanwhile, the occurrence of a non-Sunday baptism could also indicate a perceived higher risk of infant death and hence an immediate baptism. The occurrence of a non-Sunday baptism is therefore included in the regression analysis to account for these possibly confounding factors.

3.2.4 Control Variables on the Family and Parish Level

By controlling for measures of parental human capital attainments that are similar to those of the offspring, the analysis accounts for relevant educational heterogeneity between families.[30]

Parental literacy. Literate parents may have higher income, and may therefore support larger families. Furthermore, literacy can be taught by parents at a time cost, potentially reducing the total cost of endowing offspring with literacy. Thus, dummy variables indicating maternal and paternal literacy are included in the regression analysis, along with dummy variables indicating unobserved literacy, to account for this possibly confounding factor.

Parental occupational skills. Since occupational skills can potentially be taught by educated parents at an alternative time cost, skilled parents may face different costs of endowing their offspring with occupational skills than unskilled ones. Thus, dummy variables indicating maternal and paternal skills are included in the regression analysis to account for this possibly confounding factor, along with dummy variables indicating unobserved parental skills.

Paternal occupational wealth. Fathers with occupations that are associated with higher wealth can simultaneously afford larger families and devote more resources to their offspring. Furthermore, fathers holding those occupations may have an increased propensity and ability to direct their children towards similar professions. Since occupational wealth and skills are both based on occupational titles, paternal occupational wealth is divided into two main categories, with labourers and husbandmen making up the poorest segments of the English society.[31] Thus, a dummy variable indicating paternal occupational wealth is included in the regression analysis to account for the possibly confounding factors associated with paternal wealth, along with a dummy variable indicating unobserved occupational wealth.[32]

Parental marriage time period. Parental fecundity, the technological environment, the educational environment, and the affluence of individuals may change over time. Thus, dummy variables

[29] The fact that non-Sunday baptisms were often requested by affluent families is supported by the positive associations between a non-Sunday baptism and the number of surviving siblings as well as their level of human capital, as established in the regression analyses below.

[30] The parental human capital control variables enter as dummy variables. This allows the inclusion of families with unobserved parental human capital, captured by a dummy variable indicating missing information.

[31] Table A.9 in the appendix establishes that the results are robust to inclusion of dummies for all seven paternal wealth groups.

[32] Given the low female labour participation rate, information on maternal occupational wealth is omitted.

indicating the time-period of marriage of the parents (for 20-year intervals) is included in the regression analysis to account for marriage-year fixed effects.[33]

Maternal age at marriage. Since fecundity is affected by age, the age at marriage may have a direct effect on TFB. Furthermore, the marriage age may influence family fertility through the length of the reproductive period remaining after the marriage. Moreover, since marriage age was inversely related to affluence during the time period investigated, earlier marriages may be associated with higher offspring quality due to differences in income. Thus, dummy variables indicating the age at marriage of the mother (for 5-year intervals) are introduced so as to account for marriage-age fixed effects and the confounding effects of the age at marriage on family fertility and offspring quality.

Local occupational structures. The sampled parishes range from market towns to remote rural villages and have been organised by Schofield (2005) in four groups: "agriculture", "industry", "retail and handicraft" and "other" (a mix). The local occupational structure may affect the return to different types of human capital investments and may furthermore be correlated with genetic differences determining fecundity (see e.g. Juul et al., 1999, for related evidence for modern populations). Thus, dummies capturing the four different types of occupational structure are introduced to account for these confounding effects.[34]

3.2.5 Additional Control Variables

Robustness analyses in the appendix also account for the offspring's year of birth (Table A.4); alternative birth order specifications (Table A.7); offspring gender division (Table A.14); parental longevity (Table A.15); the season of the parental marriage (Table A.16); the season of birth of the firstborn (Table A.16); the climate captured by the average yearly land surface temperatures (Table A.17); yearly crude birth- and death-rates (Table A.18); and the level of unskilled wage rates (Table A.19).

4 Time to First Birth

In this section, we explore the characteristics of TFB in various ways. It establishes that (i) TFB is a useful proxy for fecundity, (ii) TFB is unaffected by potentially confounding factors in our data, (iii) studies of modern data suggest that TFB is not affected by unobserved confounding factors, (iv) the distribution of TFB in our data is comparable to that from modern data for which the onset of regular unprotected intercourse is known.

[33] The results are robust to accounting for marriage year (i.e., 1-year interval) fixed effects (see Table A.10 in the appendix).

[34] Table A.11 in the appendix demonstrate that the results are robust to controlling for parish-level fixed effects.

4.1 Distribution of TFB and its Relation to Fecundity

Imagine a population consisting of couples with an identical level of fecundity who have unprotected sexual activity on a regular and identical basis. Despite their identical levels of fecundity, these couples will have widely different times to conception, reflecting the fact that the process of conception is a highly random process. It is easy to see that the distribution of times to conception has a high variance. In a homogenous population, the time interval from the onset of unprotected intercourse to conception follows a geometric distribution. Fecundity is measured by *fecundability* – the probability of achieving conception in a given monthly cycle (Gini, 1924). If $p \in (0, 1)$ denotes fecundability, the coefficient of variation of the time to conception, given the geometric distribution, will be $p^{-1/2} > 1$. Given a monthly probability of conception of 10 percent, the standard deviation of the time to conception in a homogeneous population will therefore be more than 3.16 times the average time to a conception.

In a heterogenous population, the proportion of the variance in the time to conception that can be attributed to differences in fecundity depends on the difference in fecundability between more and less fecund individuals. In particular, when there is a small difference in the fecundability of individuals, only a small proportion of the variance in time to conception can be attributed to the difference in fecundity. For example, assume that the population consists of 50 percent individuals with a fecundability of 17 percent per monthly cycle and 50 percent individuals with a fecundability of 15 percent per cycle. In this case, simulations show that less than 0.5 percent of the variance in time to conception can be attributed to the difference in fecundity between the two groups. Meanwhile, when there is a large difference in fecundability, a substantial amount of the variance in time to conception can be attributed to differences in fecundity. If, for example, the low fecundability is changed to 5 percent per cycle, while the high fecundability is kept at 17 percent per cycle, simulations show that almost 20 percent of the variance in time to conception can be attributed to the difference in fecundity. As will be demonstrated empirically in Section 4.4, fecundability in our data range from around 6 to 17 percent per cycle for conceptions in the first year of marriage, and we therefore expect that a substantial fraction of the variance of TFB can be attributed to differences in fecundity. Nevertheless, even the mostly random variation in TFB allow us to infer the effects of fecundity.

4.2 TFB and Observed socioeconomic Variables

The likelihood of the existence of omitted variables can be assessed by examining the association between a couple's TFB and their observed characteristics. Although the possibility of omitted variables can never be ruled out entirely, an investigation of the correlation between the TFB and the observed variables provides an indirect assessment of the exogeneity of TFB. As will be argued here, TFB, is not determined by observable parental characteristic in the data while controlling for maternal age. This finding is consistent with studies cited in the next section.

Since the human capital achievements of the couple's offspring are not needed for such an assessment, we can use a larger sample, than that of the main analysis, consisting of all completed families in the dataset. Table 1 shows the results of a set of duration models, conducted at the family level, using a Cox proportional hazards duration model. The Cox model assumes that

$$h(t) = h_0(t) \exp(\boldsymbol{\beta}' \boldsymbol{\phi})$$

where $h(t)$ the hazard of the outcome event at time t, $h_0(t)$ is a baseline, unspecified, hazard function, and $\boldsymbol{\phi}$ is a vector of explanatory variables. The outcome event in the present case is the birth of the firstborn child; $\boldsymbol{\phi}$ is a vector of family and parish level control variables; and time zero is the marriage of the couple. Positive coefficients indicate an increased probability of birth, meaning shorter time spans.

The estimations displayed in Table 1 include all of the relevant family- and parish-level geographical variables used in the main analysis further below. The results suggest an absence of any significant correlation between socioeconomic characteristics and TFB. Strikingly, the adjusted R^2 never exceed 0.2 percent. The lack of a significant correlation between TFB and important determinants of living standards and morbidity, such as parental skills, occupational wealth, literacy, and longevity, indicate that TFB is generally not associated with parental socioeconomic characteristics that may co-determine offspring human capital.

Interestingly, if the lack of significance is disregarded, and the estimated coefficients are taken at face value, they tend to indicate a *negative* correlation between the parents' socioeconomic status and their TFB. In the context of our empirical strategy, this would imply a bias (if any) *against* finding a positive association between TFB and human capital achievements and therefore also against finding evidence in favour of a quantity-quality trade-off.

4.3 Determinants of Time to Pregnancy in Modern Populations

The medical literature widely accepts that age is an important determinant of fecundity (Schwartz and Mayaux, 1982). However, there is little consensus about the effect of many other candidate determinants. Studies using data for modern populations have investigated a range of possible determinants of the time from the intention to conceive until pregnancy. These studies tend to find that time to pregnancy is mostly determined by purely biological factors (e.g., age, menstrual cycle length, parity, or the degree of oxidative damage to sperm DNA).[35] However, those biological factors that correlate with fecundity have little predictive power for time to pregnancy. In particular, Axmon et al. (2006) conclude that, although they had information on several factors, their multivariate model "explained only a small fraction of the variation in the observed time to pregnancies." This is consistent with the literature review in Ecochard (2006), stating that "most of the biological heterogeneity" in fecundity "remains unexplained". Nevertheless, biological determinants of variations in fecundity are not a concern in relation to the identification strategy of the present study, since it is the effects of the inherent biological nature of fecundity that we are interested in.

[35] See Juul et al. (1999); Loft et al. (2003); Ecochard (2006); Axmon et al. (2006); Wise et al. (2011).

Table 1: The association between the time to first birth (TFB) and observed socioeconomic variables

	Years to First Birth					
	(1)	(2)	(3)	(4)	(5)	(6)
Poor Father	-.038					-.022
	(.052)					(.064)
Skilled Father		.027				-.007
		(.063)				(.077)
Skilled Mother		.356				.347
		(.234)				(.230)
Literate Father			.088			.084
			(.074)			(.078)
Literate Mother			-.078			-.105
			(.084)			(.083)
Maternal Marriage Age (Years)				-.002		-.002
				(.003)		(.003)
Retail Location					.057	.056
					(.050)	(.051)
Industrial Location					.067	.078
					(.052)	(.053)
Agricultural Location					-.059	-.049
					(.044)	(.047)
N. of Observations	3,003	3,003	3,003	3,003	3,003	3,003
N. of Families	263	263	263	263	263	263

This table presents the results of a series of Cox Proportional Hazard regressions analyses of the time to first birth on various observable parental and locational charactaristics. All regressions account for parental marriage year and dummies indicating unknown information. The coefficient on a constant term is omitted from the table. Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Another set of factors that have been found in some cases to determine fecundity, although generally to a much lesser extent, consists of environmental exposure to chemicals. Among these, the most robust findings relate to pesticides,[36] although even these findings are often ambiguous. For example, while Curtis et al. (1999) argue that pesticides may affect fecundity, overall they found "no strong or consistent pattern of associations of pesticide exposure with time to pregnancy". Among other environmental exposures, there is less evidence for effects on time to pregnancy. For example, Joffe et al. (2003) found no effect of lead exposure on time to pregnancy. Thus, there is some evidence that exposure to certain chemicals may affect fecundity. However, it should be noted

[36] See de Cock et al. (1994); Curtis et al. (1999); Cohn et al. (2003); Axmon et al. (2006).

that these factors are not a concern in relation to the present study, given the historical period that we analyse.

Contrary to the biological and the chemical factors, most lifestyle-related and socioeconomic factors have, interestingly, been found not to determine time to pregnancy. For example, Juul et al. (1999) found that smoking, body mass index, age, and parity did not determine regional differences in fecundity in their data. Likewise, Loft et al. (2003) found that smoking did not affect time to pregnancy, and Juhl et al. (2001) did not find that moderate alcohol intake is associated with a reduction in fecundity. In addition, Joffe and Barnes (2000) found that none of a range of socioeconomic characteristics affects the time to pregnancy of the first birth in the subsequent generation. These socioeconomic characteristics included body mass index, height, smoking habits, and social class. This is particularly interesting as it suggests that grand-parental socioeconomic factors are also not related to TFB. Overall, it appears that fecundity is mainly determined by biological factors, rather than lifestyle-related factors. Indeed, Axmon et al. (2006) conclude based on their data that "female biological factors seemed more important predictors of [time to pregnancy] than lifestyle factors." These findings are reassuring, since they imply a reduced scope for omitted confounding socioeconomic or lifestyle-related factors to affect our analysis.[37]

4.4 Historical and Contemporary Distributions of TFB

In order to validate our data, we compare the distribution of TFB among the sampled couples to those of a group of newly-wed Muslim couples in rural Palestine, documented by Issa et al. (2010). There are two main reasons why the Palestinian data is appropriate for this comparison. First, in the investigated Palestinian population there are strong religious and social norms against having intercourse before the day of the wedding. Consistent with this, Issa et al. found no evidence of pre-marital pregnancies (much less co-habitation) among their sampled couples (ibid., p. 4). Second, the declared intention among the investigated Palestinian population is to become pregnant on the day of the marriage or as soon as possible thereafter. Indeed, all the sampled Palestinian couples reported that their wedding night marked the onset of unprotected sex, after which intercourse occurred frequently up until the time of pregnancy. Specifically, 16 percent of the Palestinian couples reported having had sexual intercourse between one and six times per week, while 73 percent reported having had intercourse more than seven times weekly.[38]

Any tendency, intentional or not, among the sampled English couples to delay post-marriage pregnancy would supposedly result in a lower frequency of births following marriage compared to the Palestinian couples. However, we do not find this to be the case. In fact, the English couples were slightly *more* likely to achieve pregnancy within one year of marriage, compared to

[37] This view is consistent with Aguero and Marks (2008, 2011) who show that infertility is largely unrelated to parental characteristics besides age, concluding that according to their analysis, "infertility is not correlated with 'predetermined' or background characteristics of women" (Aguero and Marks, 2011).

[38] The remaining 11 percent refused to reveal their sexual frequency.

their Palestinian counterparts.[39] The chances of conception in the most relevant control group among the Palestinians (women with less than 10 years of schooling) were 12 percent after one month; 64 percent after six months; and 76 percent after 12 months (ibid., Table 1). In the English sample, the numbers are 17 percent after one month; 57 percent after six months; and 77 percent after 12 months. Calculating the average monthly probability of conception, this translates for the Palestinian sample into 12 percent in month 0-1; 11 percent in months 2-6; and 5 percent in months 6-12; and for the English sample into 17 percent in month 0-1; 10 percent in months 2-6; and 6 percent in months 6-12. It should be noted that the fact that the probability of conception is lower for couples with longer TFB supports the notion that TFB measures fecundity.

Overall, the similarity of the distributions of TFB in our data and that of Issa et al. (2010) is reassuringly consistent with the notion that marriage in pre-modern England marked the onset of unprotected intercourse with no apparent delay of post-marriage pregnancy.[40]

5 Effect of TFB on Human Capital Achievement of Offspring

5.1 Econometric Specification

To assess the effect of TFB on the offspring's human capital attainments in adulthood, the we estimate a series models on the form

$$H_{i,j} = \beta_0 + \beta_1 TFB_j + \mathbf{\Gamma}_i \boldsymbol{\beta}_2 + \mathbf{\Phi}_j \boldsymbol{\beta}_3 + \varepsilon_{i,j}, \tag{1}$$

where $H_{i,j}$ is the measure of human capital (i.e., literacy, occupational skills, or occupational wealth) of individual i in family j; $\mathbf{\Gamma}_{i,j}$ is a vector of individual-level control variables for individual i in family j; $\mathbf{\Phi}_j$ is a vector of family and parish level control variables for family j; and $\varepsilon_{i,j}$ is an error term for individual i correlated within family j.[41] Appendix Table A.13 establishes that the results are robust to the use of logistic regression analysis.

5.2 Results

The baseline OLS estimates of the effect of parental TFB on the human capital achievements of their offspring are presented in Table 2, accounting for the age at marriage, the year of marriage, parental occupational skills and literacy, paternal wealth, family location, as well as offspring gender and birth order. Consistent with the proposed hypothesis, Table 2 establishes a highly statistically significant negative association between parental fecundity, measured by the couple's TFB, and the human capital of their offspring. We find that an increase in TFB by one year results in a 3.2 percentage point higher probability of the offspring being literate; a 3.8 percentage point higher

[39] Note that the pregnancies in our sample all lead to live births; this was not necessarily the case among the Palestinians.

[40] See also Stone (1977).

[41] We examine the robustness of the estimates to different sets of control variables. We refer to "Baseline Controls" as variables indicating gender, non-Sunday baptism, paternal poor status, and parish type.

Table 2: The Effect of the Time to First Birth (TFB) on the Human Capital of Offspring

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.032***	.032**	.032***	.040***	.039***	.038***	.119***	.120***	.113***
	(.012)	(.012)	(.011)	(.014)	(.013)	(.013)	(.042)	(.041)	(.041)
Male	.098***	.101***	.109***	-.288***	-.287***	-.280***	.387**	.387**	.384**
	(.028)	(.028)	(.027)	(.069)	(.071)	(.073)	(.179)	(.178)	(.187)
Poor Father	-.387***	-.306***	-.203***	-.398***	-.260***	-.246***	-1.585***	-1.396***	-1.258***
	(.043)	(.056)	(.057)	(.044)	(.055)	(.054)	(.153)	(.171)	(.165)
Non-Sunday Baptism	.095***	.098***	.070**	.116***	.113***	.113***	.280**	.257**	.225*
	(.030)	(.030)	(.030)	(.038)	(.037)	(.037)	(.117)	(.118)	(.120)
Skilled Father		.147**	.094		.227***	.214***		.296*	.206
		(.063)	(.063)		(.068)	(.068)		(.169)	(.162)
Skilled Mother		.306	.370*		.828***	.795***		2.569**	2.475**
		(.213)	(.212)		(.222)	(.226)		(1.029)	(1.011)
Literate Father			.191***			.058			.471***
			(.048)			(.060)			(.178)
Literate Mother			.208***			.029			.293
			(.048)			(.062)			(.226)
Agricultural Location	.093*	.089*	.092*	-.064	-.088	-.105*	-.123	-.178	-.213
	(.051)	(.050)	(.052)	(.065)	(.059)	(.061)	(.194)	(.160)	(.166)
Industrial Location	.127**	.107**	.106**	.329***	.273***	.250***	-.303	-.361*	-.439**
	(.052)	(.052)	(.051)	(.062)	(.059)	(.060)	(.210)	(.208)	(.203)
Retail Location	.196***	.196***	.142**	.034	.024	.013	-.102	-.091	-.112
	(.060)	(.059)	(.061)	(.050)	(.051)	(.052)	(.165)	(.169)	(.170)
R^2	.161	.171	.225	.279	.320	.329	.311	.333	.355
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

This table presents the results of a series of OLS regression analyses of measures of human capital achievements on the time to first birth (TFB) in the family. All regressions account for parental marriage time period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. The coefficient on a constant term is omitted from the table. Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

probability of the offspring holding a skilled occupation; and a 0.11 point higher score on the occupational wealth scale.

In particular, column 1 establishes a highly significant effect of TFB on the probability of achieving literacy. One additional year from the parent's marriage date to their first birth results in a 3.2 percentage point higher probability of achieving literacy for each child. The column also establishes that males were more likely to become literate, and that children of poorer father's were less likely to become more likely to become literate. Also, offspring who were baptized on a non-Sunday were significantly more likely to become literate. Furthermore, the column establishes that the fixed effect of living in a retail location on the probability of achieving literacy is larger than that of both agricultural locations, industrial locations, and mixed locations (the omitted category).

Column 2 establishes that the highly significant coefficient estimate of the effect of TFB on literacy is robust to controlling for parental skills, in addition to the baseline control variables. The column establishes that paternal skills are significantly positively associated with offspring literacy. Likewise, maternal skills are positively associated with offspring literacy, while the estimate is not statistically significant.

Moreover, column 3 establishes that the highly significant coefficient estimate of the effect of TFB on literacy is robust to controlling for parental literacy, in addition to the other control variables. The column establishes that both paternal and maternal literacy is significantly positively associated with offspring literacy. The coefficient estimates are all in line with our *a priori* expectations. Overall, columns 1–3 establish that TFB positively affects offspring literacy.

Furthermore, column 4 establishes a highly significant effect of TFB on the probability of achieving a skilled occupation. One additional year from the parent's marriage date to their first birth results in a 4.0 percentage point higher probability of achieving a skilled occupation for each child. The column also establishes that males were less likely to achieve a skilled occupation. At first glance, the latter finding may appear surprising. However, unskilled work was physically very demanding and working women were, therefore, usually engaged in skilled work (notably spinning and weaving). The column also establishes that children of poorer fathers were less likely to achieve a skilled occupation. Additionally, offspring who were baptized on a non-Sunday were significantly more likely to achieve a skilled occupation. Furthermore, the column establishes that the fixed effect of living in an industrial location on the probability of achieving a skilled occupation is larger than that of both agricultural locations, retail locations, and mixed locations (the omitted category).

In addition, column 5 establishes that the highly significant coefficient estimate of the effect of TFB on the probability of achieving a skilled occupation is robust to controlling for parental skills, in addition to the baseline control variables. The column establishes that both paternal and maternal skills are significantly positively associated with the offspring's probability of achieving a skilled occupation.

Furthermore, column 6 establishes that the highly significant coefficient estimate of the effect of TFB on literacy is robust to controlling for parental literacy, in addition to the other control variables. Interestingly, the column establishes that both paternal and maternal literacy is not significantly associated with the offspring's probability of achieving a skilled occupation. Overall, columns 4–6 establish that TFB positively affects offspring's probability of achieving a skilled occupation.

Column 7 establishes a highly significant effect of TFB on the probability of achieving a high-wealth occupation. One additional year from the parent's marriage date to their first birth results in a 0.12 point higher score on the occupational wealth scale for each child in the family. The column also establishes a positive correlation between male sex and a high-wealth occupation, although the correlation is not statistically significant. The lack of statistical significance is likely related to the fact that those women who were employed in an occupation outside of the household were also less likely to be engaged in skilled work, as noted above. The column furthermore establishes that children of poorer father's were less likely to achieve a high-wealth occupation. Additionally, offspring who were baptized on a non-Sunday were significantly more likely to achieve a high-wealth occupation. Furthermore, the column establishes that the different locational types are characterized by similar levels of occupational wealth, conditional on the baseline control variables.

In addition, column 8 establishes that the highly significant coefficient estimate of the effect of TFB on the probability of achieving a skilled occupation is robust to controlling for parental skills, in addition to the baseline control variables. The column establishes that both paternal and maternal skills are significantly positively associated with the offspring's probability of achieving a high-wealth occupation. In this specification, the partial correlation of industrial locations is significantly negative at the 10 percent significance level.

Finally, column 9 establishes that the highly significant coefficient estimate of the effect of TFB on literacy is robust to controlling for parental literacy, in addition to the other control variables. The column establishes that parental literacy is significantly positively associated with the offspring's probability of achieving a high-wealth occupation. In this specification, the partial correlation of industrial locations is significantly negative, indicating that industrial locations are associated with the lowest levels of occupational wealth, conditional on all the control variables. Overall, columns 7–9 establish that TFB positively affects offspring's probability of achieving a skilled occupation.

5.3 Heterogenous Effects

Having established that TFB has a significantly positive effect on the human capital achievements of offspring, we turn to the question of how this effect varies with the socioeconomic standing of families. In particular, we are interested in investigating whether skilled and educated parents, those we can describe as members of the "socioeconomic elite", face a less severe trade-off between quantity and quality of children than poorer or less educated parents. Since the budget constraint is less likely to be binding for richer parents, we hypothesize that the effect of TFB on offspring human capital achievements is smaller for the socioeconomic elite.

Table 3: Heterogeneity of the Effect of the Time to First Birth (TFB) on the Human Capital of Offspring

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.042**	.039*	.031*	.076***	.074***	.073***	.255***	.244***	.236***
	(.021)	(.021)	(.018)	(.026)	(.025)	(.025)	(.083)	(.081)	(.079)
Years to First Birth × Literate & Skilled Father	-.037	-.032	-.021	-.084**	-.078**	-.079**	-.282**	-.258**	-.258**
	(.026)	(.026)	(.024)	(.034)	(.033)	(.032)	(.119)	(.118)	(.117)
Literate & Skilled Father	.247***	.225***	.014	.277***	.203**	.191*	.931***	.894***	.572
	(.068)	(.075)	(.088)	(.080)	(.081)	(.097)	(.288)	(.291)	(.351)
Skilled Father		.054	.098		.182**	.180**		.073	.111
		(.072)	(.070)		(.071)	(.071)		(.166)	(.168)
Literate Father			.196***			.017			.371*
			(.057)			(.076)			(.223)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
Signif. of Sum of Interaction Terms	.766	.693	.590	.741	.846	.798	.755	.869	.795
R^2	.177	.182	.226	.302	.333	.338	.335	.354	.365
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

This table presents the results of a series of OLS regression analyses of measures of human capital achievements on the time to first birth (TFB) in the family, as well as TFB interacted with a dummy variable indicating if the father of the family was both literate and had a skilled profession. Except for including the interaction term between TFB and the literate and skilled father dummy variable, and the literate and skilled father dummy variable itself, the specifications correspond to those of the corresponding columns of Table 2, omitting the estimated coefficients on control variables due to space concerns. All regressions account for parental marriage year, maternal marriage age, and dummies indicating unknown information. The coefficient on a constant term is omitted from the table. Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Furthermore, as argued theoretically by Galor (2012), in sufficiently wealthy populations where intergenerational transfers take place, unexpected births are more likely to reduce future intergenerational transfers to the children as opposed to reducing child quality. If parents intend to transfer future income to their children, the optimal investment in child quality equates the rate of return on investment in human capital to the rate of return on investment in physical capital (which is constant from the viewpoint of the individual), whereby the optimal adjustment would be a reduction in intergenerational transfers, leaving investment in child quality intact. An investigation of the difference in the effect of TFB on child quality between the poorer and richer parts of society may therefore also help explain why some studies of the existence of a child quantity quality trade-off for modern, developed societies have not been able to detect a trade-off (see e.g., Black et al., 2005; Angrist et al., 2010).

We investigate whether the effect of TFB on offspring human capital achievements is smaller for the socioeconomic elite by adding an interaction term between TFB and a combined indicator of parental skills and education to the model (of course, while also including the indicator variable separately). The socioeconomic elite, defined by having skilled, literate fathers, amount to around to a little less than a fifth of the families in the combined regression sample. The estimates of the heterogeneous effects of parental TFB on the human capital achievements of their offspring are presented in Table 3, accounting for the same control variables as in Table 2.

Table 3 establishes that the effect of TFB on offspring quality is indeed smaller for the socioeconomic elite. In particular, there is no effect of TFB on offspring occupational skills or wealth for this stratum. This indicates that the socioeconomic elite offset lower TFB by increasing investment in total offspring human capital, whereas families that did not have similar economic capacity did not.

Meanwhile, the effect of TFB on literacy is not significantly smaller for the socioeconomic elite, indicating that the socioeconomic elite was more focused on offsetting the detrimental effects of lower TFB on occupational outcomes compared to literacy. However, in all specifications, the sum of the coefficients on years to first birth and its interaction is not significantly different from zero, even when considering literacy as the outcome variable.

Interestingly, the fact that the effect on the occupational outcomes is smaller for the elite than for the non-elite implies that the effect for the non-elite is even larger than that uncovered for the population as a whole in Table 2. In particular, according to column 6, one additional year from the parent's marriage date to their first birth results in a 7.3 percentage point higher probability of achieving a skilled occupation for each child in a non-elite family. Likewise, according to column 9, one additional year results in a .236 point higher score on the occupational wealth scale for each child in a non-elite family.

Table 4: The Effect of the Time to First Birth (TFB) on the Human Capital of Offspring – Robustness to Alternative TFB Specifications

	Literacy			Skilled Occupation			Occupational Wealth		
	Winzor-ised	No Long TFB's	Control-ling for Own TFB	Winzor-ised	No Long TFB's	Control-ling for Own TFB	Winzor-ised	No Long TFB's	Control-ling for Own TFB
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth (Winsorized)	.068***			.079***			.182**		
	(.022)			(.024)			(.079)		
Years to First Birth		.065**	.033***		.093***	.039***		.167*	.118***
		(.025)	(.011)		(.027)	(.013)		(.089)	(.041)
TFB of Individual $<$ 40 Weeks			-.004			.001			-.277*
			(.037)			(.046)			(.150)
TFB of Individual $\geq$ 40 Weeks and $<$ 1 Year			-.008			-.019			-.067
			(.041)			(.051)			(.154)
TFB of Individual $\geq$ 1 Year and $<$ 2 Years			-.035			.017			-.045
			(.041)			(.046)			(.155)
TFB of Individual $\geq$ 2 Years and $<$ 3 Years			.088			-.044			-.266
			(.066)			(.075)			(.170)
TFB of Individual $\geq$ 3 Years			.032			-.022			.118
			(.062)			(.069)			(.260)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Literacy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Joint Signif. of Own TFB Dummies			.583			.969			.293
R^2	.228	.215	.227	.333	.333	.330	.354	.352	.360
N. of Observations	1,248	1,180	1,248	652	616	652	686	647	686
N. of Families	571	525	571	453	422	453	468	436	468

This table presents the results of a series of alternative OLS regression analyses of measures of human capital achievements on the time to first birth (TFB) in the family. In columns 1, 4, and 7, the TFB measure is winsorized at 3 years. In columns 2, 5, and 8, observations with TFB longer than 3 years are excluded. In columns 3, 6, and 9, we control for the individual's own TFB (in the family that he or she starts) by including dummies for intervals of their own TFB as well as a dummy for unknown own TFB. All regressions account for parental marriage year, maternal marriage age, and dummies indicating unknown information. The coefficient on a constant term is omitted from the table. Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

5.4 Robustness

To gauge the robustness of our main results, we perform three main robustness checks, dealing with the occurrence of long TFB, i.e., the potential hereditary element of parental fecundity; the potential importance of additional control variables; and the use of alternative estimation methods.

5.4.1 Accounting for Potential Measurement Errors Related to Long TFB

Some of the sampled couples have TFBs exceeding three years. Even though the duration to a conception can indeed be rather extensive (as is well-known from present-day populations) we wish to ensure that the estimation results are not due to a few extensive TFBs, which could be the result of measurement error or represent unobserved heterogeneities that may affect human capital formation. Hence, the analysis is repeated while accounting for long TFB in two ways. First, we use a Winzorised version of independent variable, where any TFB exceeding three years, therefore falling outside of the 90^{th} percentile, is reset to three years. Thus, the analysis retains the useful information that the affected families were of low fecundity without giving them excessive leverage, everything else being equal. The results of the Winsorized regressions (Table 4, columns 1, 4 and 7) yields larger estimates, compared to the baseline specification (Table 2) that remain highly significant, verifying that the main findings are not driven by TFBs falling outside of the 90^{th} percentile of the distribution. Second, the study reaches the same conclusion by simply removing families with parental TFB greater than 3 years from the analysis altogether (Table 4, columns 2, 5 and 8).

5.4.2 Accounting for the Heritability of Fecundity

In light of the possible heritability of fecundity, the findings from the regression analysis may reflect a correlation between the genetic elements of fecundity and offspring quality. In order to examine if this potential channel can be ruled out, we control for the fecundity of offspring, thus accounting for variations in the hereditary components of fecundity.[42] We have included dummy variables indicating if the individual's own TFB is below 40 weeks; between 40 weeks and 1 year; between 1 and 2 years; between 2 and 3 years; or 3 years and above. The background variable is unknown TFB. Table 4, columns 3, 6 and 8 establish that the baseline results (i.e., those of Table 2) are robust to accounting for the potential channel operating through the heritability of fecundity. These regressions also establish that the offspring's own TFB is not correlated with their human capital attainments, mirroring the conclusions from Table 1.

5.4.3 Alternative Estimation Methods

The main results are also qualitatively unchanged if we use alternative estimation methods. In particular, if we use a Heckit model to account for potential issues of sample selection, we obtain

[42] The TFB of offspring is known in 71 percent of the cases.

the same qualitative conclusion (Table A.12 in the appendix). Likewise, the use a logistic model with binary outcome variables also yields the same qualitative conclusions (Table A.13).

5.4.4 Accounting for Additional Factors

We perform a wide range of additional robustness checks in the appendix. These checks establish that the main results are qualitatively unchanged if we account for offspring birth year (Table A.4), alternative offspring birth order specifications (Table A.7), alternative marriage age specifications (Table A.8), stoppage age (Table A.8), offspring gender division (Table A.14), parental longevity (Table A.15), season of birth of the firstborn (Table A.16), season of parental marriage (Table A.16), climate (average yearly land surface temperatures) (Table A.17), aggregate yearly birth and death rates (Table A.18), and a national-level wage rates (Table A.19).

5.5 Remaining Potential Confounding Factors

Despite the conclusions from Section 4, one may still be concerned that TFB is correlated with the outcome variables for reasons not directly linked to fecundity. For example, one may hypothesize that some newly-wed husbands were engaged in trades that involved their absence from the household. Such an absence would not only increase TFB, but also be related to family income and therefore offspring human capital.

To account for such a confounding mechanism, we control for various time and seasonal fixed effects. In particular, in the baseline specifications, we account for time trends by controlling for the marriage-year fixed effects. Furthermore, in Table A.16, we control for season (i.e., quarter-of-year) fixed effects in which the couple was married.

Likewise, given the possible association between temperatures and extraordinary working conditions that may entail a change in sexual activity, while at the same time potentially affecting family income and ultimately investment in offspring human capital, we control in Table A.17 for the average yearly temperature in the 20-year period surrounding (i) the birth of each child; (ii) the marriage of the parents; and (iii) the marriage of the individual.

In addition, given the possible association between vital rates and extraordinary living conditions that may entail a change in sexual activity, while at the same time potentially affecting family income and ultimately investment in offspring human capital, we control in Table A.18 for the average yearly birth and death rate in the 20-year period surrounding (i) the birth of the individual; (ii) the marriage of the parents; and (iii) the marriage of the individual.

Finally, given the possible association between wages and extraordinary working conditions (for example, temporary migration) that may entail a change in sexual activity, while at the same time potentially affecting family income and ultimately investment in offspring human capital, we control for aggregate wages in the 20-year period surrounding (i) the birth of the individual; (ii) the marriage of the parents; and (iii) the marriage of the individual as reported in Table A.19.

These robustness analyses establish that the results are robust to controlling for seasonal fixed effects as well as time-varying variables that may be associated with extraordinary working conditions.

6 Effect of Number of Siblings on Human Capital Achievement of Offspring

The previous section established that parental fecundity, measured by TFB, is positively correlated with the human capital attainments of offspring. While the analysis focuses on an important determinant of fertility, namely TFB, it is possible to exploit the highly significant effect of TFB on the number of children as an instrument for fertility to estimate the effect of the number of children on long-run reproductive success. In this section, the study therefore examines the underlying mechanism behind the observed correlation between TFB and human capital. We find that TFB significantly affects the total number of children born in the family, consistent with the notion that the positive association between TFB and offspring human capital may operate via a trade-off between child quantity and child quality. This section establishes via 2SLS regressions that the number of siblings, as instrumented by TFB, statistically significantly affects the human capital of offspring.

The identifying assumption in the 2SLS analysis is that TFB affects the three measures of human capital through family size alone, conditional on the control variables. In light of the discussion in Section 4, this assumption seems justified. Furthermore, plausible violations of the exclusion restriction, as discussed above, will presumably associate *longer* TFB with *lower* human capital of offspring and therefore bias the estimated effect of parental fertility on offspring human capital upwards and thus bias *against* finding evidence of a child quantity-quality trade-off.

6.1 Effect of TFB on Family Size

This subsection examines the relationship between parental fecundity, measured by TFB, and the number of surviving family children. It establishes theoretically as well as empirically that there is a strongly significant positive effect of parental fecundity on parental fertility. There are two reasons for expecting an effect of TFB on the number of children in the pre-demographic transition era. First, a longer TFB reduce the time available for producing children by moving the starting point of procreation. Second, to the extent that TFB reflects fecundity, a longer TFB is positively associated with the time periods between subsequent births. The empirical analysis shows that an additional year of time from marriage to the first birth results in a decrease in the family size of around half a child, conditional on a range of control variables.

6.1.1 Theoretical Effect of TFB on Family Size

The compound effect of time to first birth on marital fertility can be described formally as follows. Consider a married couple that has chosen their marriage age with the expectation of an immediate conception and with a target family size achieved through natural fertility within marriage, i.e. without fertility limitation. The remaining fertile period after the couple's first birth is $f - t$, where f denotes the reproductive (i.e. fertile) period of a couple (i.e. the time-period spanned by the marriage date and the date of sterility) and t denotes the time from the marriage to the first birth. The total number of births is then inversely related to the average birth-spacing interval, denoted $s(t)$. If x denotes the total number of births, then $x = (f - t)/s(t) + 1$. The average within-marriage birth interval can be approximated by a linear function in t, $s(t) = c + \lambda t$ where c and λ are constants, hence obtaining the expression $x = (f - t)/(c + \lambda t) + 1$. Linearizing this expression around the average time to first birth, denoted by $\bar{t}$, means that $x \approx \gamma_0 - \gamma_1 t$, where $\gamma_0 \equiv (f - \bar{t})/(c + \lambda\bar{t}) + \bar{t}(c + \lambda f)/(c + \lambda\bar{t})^2$ and $\gamma_1 \equiv (c + \lambda f)/(c + \lambda\bar{t})^2$. In the simplest case, where TFB is not correlated with average birth intervals (a counter-factual but useful approximation), $\lambda = 0$ and the average birth interval, c, is 2.5 years, an increase in TFB of one year will result in 0.4 fewer children.[43] A higher correlation between TFB and average birth intervals will increase the absolute size of the effect.[44] In light of the large variation in TFB depicted in Figure 2, this analysis indicates that TFB is a highly relevant instrument of family size in historical populations.

Now consider a couple that incorporates a buffer of time in their family planning and chooses their marriage age on the assumption that their target family size could be achieved in the presence of fertility limitation within marriage, i.e. following the first birth. In this case, TFB affects the family size only if it exceeds the length of the buffer. These discrepancies between planned and actual family sizes are presumably unrelated to parental budget constraints and may therefore affect the allocation of resources to each child in the family.

6.1.2 Empirical Effect of TFB on Family Size

The effect of TFB on the number of offspring is examined using an OLS model given by the regression equation

$$S_{i,j} = \gamma_0 + \gamma_1 TFB_{i,j} + \mathbf{\Gamma}_{i,j}\boldsymbol{\gamma}_2 + \mathbf{\Phi}_{i,j}\boldsymbol{\gamma}_3 + \mu_{i,j}, \tag{2}$$

where $S_{i,j}$ is the number of siblings of individual i in family j; where TFB_j is the time from the marriage of the parents in family j to the birth of the first sibling; $\mathbf{\Gamma}_{i,j}$ is a vector of individual-

[43]Empirically, a lower bound of the point in time when sterility sets in is given by the couple's final delivery. Using this as a proxy for the actual time of sterility, the average fertile period in the sample is $\bar{f} = 16.16$ years. The average time to first birth is $\bar{t} = 1.59$ years (cf. Table A.1 in the appendix). A simple regression of birth intervals on the time to first birth, with standard errors clustered at the family level, yields $\bar{c} = 2.48$ ($p < 0.000$) and $\bar{\lambda} = 0.08$ ($p = 0.010$). These numbers imply that $\bar{\gamma}_0 = 6.47$ and $\bar{\gamma}_1 = 0.56$. Indeed, a simple regression of family size on TFB, with standard errors clustered at the family level, provides similar estimates of ($\hat{\gamma}_0 = 7.87$ ($p < 0.000$) and ($\hat{\gamma}_1 = 0.57$ ($p < 0.000$)).

[44]It should be noted that a correlation between TFB and average birth interval is consistent with TFB reflecting fecundity.

level control variables for individual i in family j; $\mathbf{\Phi}_j$ is a vector of family and parish level control variables for family j; and $\mu_{i,j}$ is an error term for individual i correlated within family j.

The estimates of the effect of TFB on the number of offspring are presented in Table 5, accounting for the marriage age, the marriage year, parental occupational skills and literacy, paternal wealth, family location, and offspring gender and birth order. Consistent with the proposed theoretical effect of TFB on family size, Table 5 establishes a highly statistically significant negative association between parental TFB and the number of offspring.

In particular, column 1 establishes in the sample with known literacy information, i.e., the sample underlying columns 1–3 of Table 2, that an additional year of time to first birth decreases, on average, the final family size by 0.45 surviving children. Likewise, column 2 establishes in the sample with known occupational skills information, i.e. the sample underlying columns 4–6 of Table 2, that an additional year of time to first birth decreases, on average, the final family size by 0.49 surviving children. Similarly, column 3 establishes in the sample with known occupational wealth information, i.e. the sample underlying columns 7–9 of Table 2, that an additional year of time to first birth decreases, on average, the final family size by 0.52 surviving children. Finally, column 4 establishes in the sample with either known literacy, known occupational skills or known occupational wealth information, i.e. the combined sample of column 1–9 of Table 2, that an additional year to first birth decreases, on average, the final family size by 0.47 surviving children. It follows that a one-year increase in TFB reduces the number of surviving offspring by close to half a child, depending on the sample used.

Slightly less than half of the variation in family size is explained by TFB and the covariates. The covariates have virtually the same partial effects regardless of the sample used. The fact that low-wealth fathers have relatively many children is consistent with evidence showing that the poor were outcompeting the rich in terms of number of births after 1700 (Boberg-Fazlic et al., 2011). In addition, the estimates establish that literate fathers father relatively many offspring.

6.2 Econometric Specification

Under the assumption of conditional exogeneity of TFB and in light of the established effect of TFB on family size in historical England, we now estimate the effect of family size on the human capital attainments of offspring using 2SLS model with the number of siblings instrumented by the parental TFB given by Equation 2. The second stage of the 2SLS model is given by the regression equation

$$H_{i,j} = \delta_0 + \delta_1 \hat{S}_j + \mathbf{\Gamma}_{i,j}\boldsymbol{\delta}_2 + \mathbf{\Phi}_j\boldsymbol{\delta}_3 + \nu_{i,j}, \tag{3}$$

where $H_{i,j}$ is the measure of human capital (i.e., literacy, occupational skills, or occupational wealth) of individual i in family j; $\hat{S}_{i,j}$ is the predicted number of siblings of individual i in family j resulting from the estimation of equation 2; $\mathbf{\Gamma}_{i,j}$ is a vector of individual-level control variables

Table 5: The Effect of TFB on the Number of Siblings

	Number of Surviving Siblings (> 5 Years)			
	Literacy Sample	Skills Sample	Wealth Sample	Total Sample
	(1)	(2)	(3)	(4)
Years to First Birth	-.450***	-.485***	-.520***	-.467***
	(.056)	(.065)	(.064)	(.051)
Male	.009	.106	.122	.069
	(.107)	(.340)	(.282)	(.107)
Poor Father	.365	.675**	.724**	.450*
	(.281)	(.330)	(.316)	(.259)
Non-Sunday Baptism	-.092	-.211	-.225	-.148
	(.139)	(.184)	(.177)	(.128)
Skilled Father	-.260	-.168	-.124	-.187
	(.309)	(.314)	(.302)	(.288)
Literate Father	.703**	.794**	.717**	.672**
	(.279)	(.308)	(.299)	(.275)
Literate Mother	-.172	-.229	-.224	-.103
	(.288)	(.345)	(.336)	(.281)
Agricultural Location	.347	.433	.413	.291
	(.277)	(.307)	(.305)	(.261)
Industrial Location	.263	.413	.325	.157
	(.280)	(.383)	(.378)	(.267)
Retail Location	-.528*	-.381	-.512**	-.510**
	(.301)	(.261)	(.256)	(.222)
Skilled Mother	3.269*	4.092**	3.935**	3.075*
	(1.706)	(1.714)	(1.613)	(1.608)
R^2	.458	.489	.495	.456
N. of Observations	1,248	652	686	1,517
N. of Families	571	453	468	729

This table presents the results of a series of OLS regression analyses of the number of surviving siblings on the time to first birth (TFB) in the family. All regressions account for parental marriage year, maternal marriage age, and dummies indicating unknown information. The coefficient on a constant term is omitted from the table. Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

for individual i in family j; $\mathbf{\Phi}_j$ is a vector of family and parish level control variables for family j; and $\nu_{i,j}$ is an error term for individual i correlated within family j.

6.3 Results

Table 6 presents the estimates from the 2SLS estimates of the effect of the number of siblings, instrumented by TFB, on human capital achievements, accounting for parental marriage age and year, parental occupational skills and literacy, paternal wealth, family location, and offspring gender and birth order. Consistent with the notion that the influence of parental TFB (and by extension parental fecundity) on offspring human capital operates via a child quantity-quality trade-off, Table 6 establishes a highly statistically significant negative association between the number of siblings, as instrumented by TFB, and human capital achievement, regardless of the human capital measure used and conditional on a range of control variables. In particular, TFB is a strong instrument for the number of siblings, i.e., the Wald F-test statistics, based on the Kleibergen-Paap rk statistic (Kleibergen and Paap, 2006), is always above 53, well above the rule-of-thumb-value of 10 (Baum et al., 2007).

Focusing first on literacy as the outcome variable, column 1 establishes that the number of surviving siblings has a highly statistically significant and causal negative effect on the probability of becoming literate, controlling for a range of potentially confounding variables such as gender, paternal wealth, and parish type fixed effects. Furthermore, column 2 establishes that the results are robust to accounting for paternal as well as maternal skills. In particular, the coefficient of interest remains stable and highly statistically significant. In addition, column 3 establishes that the effect of surviving siblings on offspring literacy remains highly significant and stable while accounting for paternal literacy. The coefficient estimate implies that an additional sibling decreases the probability of achieving literacy by 7.1 percentage points.

Turning to occupational skills as the outcome variable, column 4 establishes that the number of surviving siblings has a highly statistically significant and causal negative effect on the probability of achieving a skilled occupation, controlling for a range of potentially confounding variables such as gender, paternal wealth, and parish type fixed effect. Furthermore, column 5 establishes that the results are robust to accounting for paternal as well as maternal skills. In particular, the coefficient of interest remains stable and highly statistically significant. In addition, column 6 establishes that the effect of surviving siblings on occupational skills remains highly significant and stable while accounting for paternal literacy. The coefficient estimate implies that an additional sibling decrease the probability of achieving a skilled occupation by 7.9 percentage points.

Turning to occupational wealth as the outcome variable, column 6 establishes that the number of surviving siblings has a causal negative effect on occupational wealth, controlling for a range of potentially confounding variables such as gender, paternal wealth, and parish type fixed effect. Furthermore, column 7 establishes that the results are robust to accounting for paternal as well as maternal skills. In particular, the coefficient of interest remains stable and becomes statistically highly significant. In addition, column 9 establishes that the the effect of surviving siblings on

Table 6: Instrumental Variables Regressions: The Effect of the Number of Siblings on Human Capital

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Surviving Siblings (> 5 Years)	-.067**	-.068***	-.071***	-.084***	-.082***	-.079***	-.230***	-.237***	-.217***
	(.026)	(.026)	(.025)	(.031)	(.029)	(.028)	(.085)	(.083)	(.081)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
R^2	.082	.096	.158	.159	.224	.243	.207	.244	.285
F (Kleibergen-Paap)	68.8	67.7	64.6	57.8	53.4	56.4	65.5	62.4	66.6
Anderson-Rubin F stat. p-value	.010	.011	.004	.004	.003	.003	.005	.003	.006
Endogeneity test p-value	.009	.011	.011	.008	.008	.009	.005	.005	.011
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

This table presents the results of a series of 2SLS regression analyses of measures of human capital achievements on the number of surviving siblings in the family, instrumented by the time to first birth (TFB). All regressions account for parental marriage year, maternal marriage age, and dummies indicating unknown information. The coefficient on a constant term is omitted from the table. Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

literacy remains highly significant and stable while accounting for paternal literacy. The coefficient estimate implies that an additional sibling decrease the probability of achieving a high-wealth occupation by 21.7 percentage points.

To gauge the robustness of our results we perform three main robustness checks, dealing with the occurrence of long TFB, the potential hereditary element of fecundity and the occurrence of long intervals between births.

6.3.1 Accounting for Potential Measurement Errors Related to Long TFB

Analogous to Section 5.4.1, the 2SLS analysis is repeated accounting for long TFBs in two ways. Again, we use a Winzorised version of the instrument, whereby any TFB exceeding three years, falling outside of the 90$^{\text{th}}$ percentile, is reset to three years. The results of the Winsorized regressions (Table 7, columns 1, 4 and 7) yield larger estimates (in absolute values) for the literacy and skills outcomes, and a smaller estimate (in absolute values) for the wealth outcome, compared to the baseline 2SLS specification (Table 7). The coefficients also remain highly significant. These results verify that the main findings are not driven by TFB falling outside of the 90$^{\text{th}}$ percentile of the distribution. Furthermore, the same conclusion is reached when excluding families with parental TFB greater than 3 years from the analysis altogether (Table 7, columns 2, 5 and 8).

6.3.2 Accounting for the Heritability of Fecundity

Analogous to Section 5.4.2, we control for TFB of offspring, accounting for variations in the hereditary components of fecundity. As above, the study includes dummy variables indicating if the individual's own TFB is below 40 weeks; between 40 weeks and 1 year; between 1 and 2 years; between 2 and 3 years; or 3 years and above. The background variable is unknown TFB. Table 4, columns 3, 6 and 8 establish that the baseline 2SLS results (Table 7) are robust to accounting for the potential channel operating through the heritability of fecundity. As previously, these regressions also establish that the offspring's own TFB are not correlated with their human capital outcome, mirroring the conclusions from Table 1.

6.3.3 Accounting for Potential Unobserved Children Related to a Long TFB

To address the possibility that long TFBs could reflect unobserved offspring, the study performs an analysis in which an extra child has been imputed wherever TFB exceeds a period of three years combined with the Winsorized TFB. This robustness check still results in highly significant estimates (Table 7, columns 4, 9, and 14), indicating that the analysis is robust to accounting for the possibility of unobserved children.

6.3.4 Accounting for Potential Measurement Errors Related to Long Birth Spacings

By confining the sample to couples who have completed their marriage (i.e. both spouses survive in marriage until the wife turns 50) we automatically exclude the possibility of permanent migration.

Table 7: Robustness to alternative IV specifications

	Literacy					Skilled Occupation					Occupational Wealth				
	Winsorized	No Long TFBs	Controlling for Own TFB	Winsorized TFB and Imputed Siblings	Imputed Siblings (All Intervals)	Winsorized	No Long TFBs	Controlling for Own TFB	Winsorized TFB and Imputed Siblings	Imputed Siblings (All Intervals)	Winsorized	No Long TFBs	Controlling for Own TFB	Winsorized TFB and Imputed Siblings	Imputed Siblings (All Intervals)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Surviving Siblings (> 5 Years)	-.104***	-.152**	-.073***			-.116***	-.247**	-.078***			-.123***	-.210*	-.104***		
	(.036)	(.072)	(.025)			(.039)	(.109)	(.028)			(.046)	(.110)	(.034)		
Surviving Siblings				-.195**					-.194**					-.200**	
(With Imputed Siblings, First Interval)				(.084)					(.078)					(.084)	
Surviving Siblings					-.073***					-.086***					-.104***
(With Imputed Siblings, Any Interval)					(.026)					(.033)					(.036)
Years to First Birth (Winsorized)															
TFB of Individual < 40 Weeks			.016					.004					-.115*		
			(.039)					(.046)					(.059)		
TFB of Individual ≥ 40 Weeks and < 1 Year			.010					-.050					-.073		
			(.044)					(.052)					(.066)		
TFB of Individual ≥ 1 Year and < 2 Years			-.036					.006					-.025		
			(.041)					(.046)					(.062)		
TFB of Individual ≥ 2 Years and < 3 Years			.098					-.034					-.101		
			(.065)					(.077)					(.083)		
TFB of Individual ≥ 3 Years			.071					-.025					.027		
			(.065)					(.069)					(.098)		
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Literacy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	.081	-.099	.158	-.285	.150	.141	-.530	.248	-.198	.205	.255	.032	.291	.068	.267
F (Kleibergen-Paap)	30.0	11.8	62.0	8.7	44.0	28.4	6.9	58.4	10.3	32.1	34.3	8.8	69.7	13.1	40.7
Anderson-Rubin F stat. p-value	.002	.010	.004	.002	.004	.001	.001	.003	.001	.003	.006	.021	.001	.006	.002
Endogeneity test p-value	.004	.012	.010	.002	.016	.002	.001	.009	.001	.008	.011	.028	.005	.008	.006
N. of Observations	1,248	1,180	1,248	1,248	1,248	652	616	652	652	652	686	647	686	686	686
N. of Families	571	525	571	571	571	453	422	453	453	453	468	436	468	468	468

This table presents the results of a series of alternative 2SLS regression analyses of measures of human capital achievements on the number of surviving siblings in the family, instrumented by the time to first birth (TFB). In columns 1, 6, and 11, we have winsorized TFB at 3 years. In columns 2, 7, and 12, we exclude observations for families with a TFB longer than three years. In columns 3, 8, and 13, we control for the individual's own TFB (in the family that he or she starts) by including dummies for intervals of their own TFB as well as a dummy for unknown own TFB. In column 4, 9, and 14, we impute an extra sibling for those families in which TFB exceeds 3 years. In columns 5, 10, and 15, we impute an extra sibling for each birth interval above 3 years. All regressions account for parental marriage year, maternal marriage age, and dummies indicating unknown information. The coefficient on a constant term is omitted from the table. Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

For that reason we steer clear of births occurring in parishes to which migrating parents moved but where we cannot observe them (see the discussion above). Nevertheless, it was not unusual for a married couple to migrate to an unobserved parish temporarily (Souden, 1984). Being away for more than a couple of years, it is not unlikely that the couple would conceive (and thus baptise) a child in their interim location. Such incidences would appear in the data in the form of an extended birth-spacing interval, and the resulting child would remain unobserved. To address this issue we impute an extra sibling for all the birth-spacing intervals exceeding three years, artificially increasing the average family size by 1.3 children. The revised estimates are reported in Table 7, columns 5, 10, and 15. The estimated effects are largely unchanged and remain significant.

6.3.5 Alternative Estimation Methods

Lastly, we also performed a Heckit IV analysis (Table A.20 in the appendix), finding again no indications that our results are driven by issues of sample selection.

7 Conclusion and Discussion

This research explores the effects of parental reproductive capacity on offspring human capital formation in pre-industrial England. Exploiting a vast genealogy of individuals living between the 17$^{\text{th}}$ and the 19$^{\text{th}}$ centuries, the study proposes and tests the hypothesis that lower parental reproductive capacity positively affects the socioeconomic achievements of the family offspring. Using the time interval between the date of the parent's marriage and the date of their first birth as a measure of their reproductive capacity, and hence as a novel source of unplanned, exogenous variation in family size, the research establishes that children of parents with lower fecundity were more likely to become literate and employed in skilled and high-wealth professions. Furthermore, the analysis establishes that the effect was weaker for the socioeconomic elite, who could potentially offset the cost of additional children by raising total investment in offspring human capital.

The analysis shows also that the parents' capacity to reproduce was positively associated with their number of surviving offspring. In combination with the finding that reproductive capacity was negatively associated with the offspring's human capital formation, this indicates that a trade-off between child quantity and child quality was operating in England during the industrial revolution. These findings therefore lend support to the quantity-quality trade-off hypothesis, and by extension to unified growth theory (Galor, 2011).

In particular, using marital fecundity as an instrument for marital fertility in a 2SLS analysis, the study found that additional siblings significantly reduced offspring human capital. This presents a further contribution of the present study, namely the use of fecundity as a source of exogenous variation in fertility. This methodological contribution can be employed in the context of a wide range of datasets, in developing countries and historical economies alike, and is a particularly useful tool for estimating the child quantity-quality trade-off effects in the growing number of family reconstructions of historical populations that are currently becoming available.

Furthermore, our results lend support to evolutionary economic growth theories which hypothesize that individuals with a bias towards low fertility improved the income potential and hence the reproductive success of their lineage, gradually increasing the representation of this growth-enhancing trait in the population, contributing to the process of development and the take-off from stagnation to growth (Galor and Moav, 2002; Galor and Klemp, 2016)).

Overall, we conclude that the individually varying trait of reproductive capacity, as measured by TFB, has affected human human capital formation in the pre-modern period and may have played a central role in the transition from an era of stagnation to one of modern economic growth.

References

Aguero, J. M. and M. S. Marks (2008): "Motherhood and Female Labor Force Participation: Evidence from Infertility Shocks," *American Economic Review*, 98, 500–504. 3, 17

Aguero, J. M. and M. S. Marks (2011): "Motherhood and Female Labor Supply in the Developing World: Evidence from Infertility Shocks," *Journal of Human Resources*, 46, 800–826. 3, 17

Angrist, J. D., V. Lavy, and A. Schlosser (2010): "Multiple Experiments for the Causal Link between the Quantity and Quality of Children," *Journal of Labour Economics*, 28, 773–824. 3, 23

Axmon, A. and L. Hagmar (2005): "Time to pregnancy and pregnancy outcome," *Fertility and sterility*, 84, 966–974. 7

Axmon, A., L. Rylander, M. Albin, and L. Hagmar (2006): "Factors affecting time to pregnancy," *Human reproduction*, 21, 1279–1284. 15, 16, 17

Basso, A. (2012): "Fertility transition and the quantity-quality trade-off: historical evidence from Spain," Tech. rep., Working Paper. Universidad de Alicante. 2

Basso, O. and D. D. Baird (2003): "Infertility and preterm delivery, birthweight, and Caesarean section: a study within the Danish National Birth Cohort," *Human Reproduction*, 18, 2478–2484. 7

Basso, O. et al. (2005): "Subfecundity and neonatal mortality: longitudinal study within the Danish national birth cohort," *BMJ*, 330, 393–394. 7

Baum, C. F., M. E. Schaffer, and S. Stillman (2007): "ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression," http://ideas.repec.org/c/boc/bocode/s425401.html. 31

Becker, G. S. (1960): "An Economic Analysis of Fertility," in *Demographic and Economic Change in Developed Countries*, ed. by A. J. Coale, Princeton, NJ: Princeton University Press, 209–40. 1

Becker, G. S. and H. G. Lewis (1973): "On the Interaction Between the Quantity and Quality of Children," *Journal of Political Economy*, 81, 279–288. 1

Becker, G. S. and N. Tomes (1976): "Child Endowments and the Quantity and Quality of Children," *Journal of Political Economy*, 84, 143–162. 1

Becker, S., F. Cinnirella, and L. Woessmann (2010): "The Trade-off between Fertility and Education: Evidence from before the Demographic Transition," *Journal of Economic Growth*, 15, 177–204. 2

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005): "The More The Merrier? The Effect Of Family Size And Birth Order On Children's Education," *Quarterly Journal of Economics*, 2, 669–700. 3, 11, 23

Boberg-Fazlic, N. P., P. R. Sharp, and J. Weisdorf (2011): "Survival of the Richest? Social Status, Fertility, and Social Mobility in England 1541–1824," *European Review of Economic History*, 15, 365–392. 4, 5, 29

Bongaarts, J. (1975): "A Method for the Estimation of Fecundability," *Demography*, 12, 645–660. 6

Broadberry, S. and B. Gupta (2006): "The early modern great divergence: wages, prices and economic development in Europe and Asia, 1500–18001," *The Economic History Review*, 59, 2–31.

Cinnirella, F., M. P. B. Klemp, and J. L. Weisdorf (2013): "Malthus in the Bedroom: Birth Spacing as a Preventive Check Mechanism in Pre-Modern England," *CAGE Online Working Paper Series*. 5

Clark, G. and N. Cummins (2010): "Malthus to Modernity: England's First Fertility Transition, 1760–1800," *University of California Davis Manuscript*. 10

Clark, G. and G. Hamilton (2006): "Survival of the Richest: The Malthusian Mechanism in Pre-Industrial England," *The Journal of Economic History*, 66, 707–736. 2, 4

Cohn, B. A., P. M. Cirillo, M. S. Wolff, P. J. Schwingl, R. D. Cohen, R. I. Sholtz, A. Ferrara, R. E. Christianson, B. J. van den Berg, and P. K. Siiteri (2003): "DDT and DDE exposure in mothers and time to pregnancy in daughters," *The Lancet*, 361, 2205–2206. 16

Cooney, M. A., B. Louis, M. Germaine, W. Sun, M. M. Rice, and M. A. Klebanoff (2006): "Is conception delay a risk factor for reduced gestation or birthweight?" *Paediatric and perinatal epidemiology*, 20, 201–209. 7

Curtis, K. M., D. A. Savitz, C. R. Weinberg, and T. E. Arbuckle (1999): "The effect of pesticide exposure on time to pregnancy." *Epidemiology*, 10, 112–117. 16

Dalgaard, C.-J. and H. Strulik (2010): "The Physiological Foundations of the Wealth of Nations," Discussion Papers 10-05, University of Copenhagen. Department of Economics. 1

de Cock, J., K. Westveer, D. Heederik, E. te Velde, and R. van Kooij (1994): "Time to pregnancy and occupational exposure to pesticides in fruit growers in The Netherlands." *Occupational and environmental medicine*, 51, 693–699. 16

Diebolt, C., A.-R. Menard, F. Perrin, et al. (2016): "Behind the Fertility-Education Nexus: What Triggered the French Development Process?" Tech. rep. 2

Doepke, M. (2004): "Accounting for fertility decline during the transition to growth," *Journal of Economic growth*, 9, 347–383. 1

Ecochard, R. (2006): "Heterogeneity in fecundability studies: issues and modelling," *Statistical methods in medical research*, 15, 141–160. 15

Ejrnæs, M. and C. C. Pörtner (2004): "Birth order and the intrahousehold allocation of time and education," *Review of Economics and Statistics*, 86, 1008–1019. 11, 49

Fernihough, A. (2011): "Human Capital and the Quantity-Quality Trade-Off during the Demographic Transition: New Evidence from Ireland," *Working Papers 201113, School of Economics, University College Dublin*. 2

Galor, O. (2005): "From stagnation to growth: unified growth theory," in *Handbook of economic growth (vol. 1A)*, ed. by P. Aghion and S. Durlauf, Amsterdam: North-Holland. 1

Galor, O. (2011): *Unified Growth Theory*, Princeton: Princeton University Press. 1, 35

Galor, O. and M. Klemp (2016): "The Biocultural Origins of Human Capital Formation," Working Papers 2014-6, Brown University, Department of Economics. 1, 2, 3, 4, 6, 36

Galor, O. and O. Moav (2002): "Natural selection and the origin of economic growth," *Quarterly Journal of Economics*, 117, 1133–1191. 1, 2, 36

Galor, O. and D. N. Weil (1999): "From Malthusian Stagnation to the Demographic Transition and Beyond," *The American Economic Review*, 89, 150–154. 1

Galor, O. and D. N. Weil (2000): "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *The American Economic Review*, 90, 806–828. 1, 2

Gini, C. (1924): "Prmieres recherches sur la fcondabilit de la femme," in *Proceedings of the Int. Math. Congress*, Toronto: North-Holland, 889–892. 14

Hansen, G. D. and E. C. Prescott (2002): "Malthus to Solow," *American Economic Review*, 92, 1205–1217. 1

Henriksen, T. B., D. D. Baird, J. Olsen, M. Hedegaard, N. J. Secher, and A. J. Wilcox (1997): "Time to pregnancy and preterm delivery," *Obstetrics & Gynecology*, 89, 594–599. 7

Issa, Y., M. Sallme, K. Nijem, E. Bjertness, and P. Kristensen (2010): "Fecundability among newly married couples in agricultural villages in Palestine: a prospective study," *Human Reproduction*, 25, 2132–2138. 17, 18

Jensen, R. (2012): "Another Mouth to Feed? The Effects of (In)Fertility on Malnutrition," *CESifo Economic Studies*, 58, 322–347. 3

Joffe, M. and I. Barnes (2000): "Do parental factors affect male and female fertility?" *Epidemiology*, 11, 700–705. 17

Joffe, M., J. Bennett, N. Best, and T. K. Jensen (2007): "Sex ratio and time to pregnancy: analysis of four large European population surveys," *BMJ*, 334, 524. 7

Joffe, M., L. Bisanti, P. Apostoli, P. Kiss, A. Dale, N. Roeleveld, M.-L. Lindbohm, M. Sallmén, M. Vanhoorne, and J. P. Bonde (2003): "Time to pregnancy and occupational lead exposure," *Occupational and environmental medicine*, 60, 752–758. 16

Joffe, M. and Z. Li (1994): "Association of time to pregnancy and the outcome of pregnancy," *Fertility and sterility*, 62, 71–75. 7

Juhl, M., A.-M. N. Andersen, M. Grønbæk, and J. Olsen (2001): "Moderate alcohol consumption and waiting time to pregnancy," *Human Reproduction*, 16, 2705–2709. 17

Juul, S., W. Karmaus, J. Olsen, E. I. S. S. Group, et al. (1999): "Regional differences in waiting time to pregnancy: pregnancy-based surveys from Denmark, France, Germany, Italy and Sweden," *Human Reproduction*, 14, 1250–1254. 13, 15, 17

Kleibergen, F. and R. Paap (2006): "Generalized reduced rank tests using the singular value decomposition," *Journal of Econometrics*, 133, 97–126. 31

Klemp, M., C. Minns, P. Wallis, and J. 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, 210–232. 11

Lagerlöf, N.-P. (2003): "From Malthus to Modern Growth: Can Epidemics Explain the Three Regimes?" *International Economic Review*, 44, 755–777. 1

Leeuwen, M. H. D. V. and I. Maas (2011): *HISCLASS: A historical international social class scheme*, Leuven: Leuven University Press. 10

Leeuwen, M. H. D. V., I. Maas, and A. Miles (2007): *HISCO: Historical International Standard Classification of Occupations*, Cornell: Cornell University Press. 10

Leunig, T., C. Minns, and P. Wallis (2011): "Networks in the Premodern Economy: the Market for London Apprenticeships, 1600–1749," *Journal of Economic History*, 71, 413–443. 4

Loft, S., T. Kold-Jensen, N. H. Hjollund, A. Giwercman, J. Gyllemborg, E. Ernst, J. Olsen, T. Scheike, H. E. Poulsen, and J. P. Bonde (2003): "Oxidative DNA damage in human sperm influences time to pregnancy," *Human Reproduction*, 18, 1265–1272. 15, 17

Lucas, R. (2002): *The Industrial Revolution: Past and Future*, Cambridge: Harvard University Press. 1

McLaren, D. (1978): "Fertility, Infant Mortality and Breast Feeding in the Seventeenth Century," *Medical History*, 22, 378–396. 5

Midi Berry, B. and R. S. Schofield (1971): "Age at Baptism in Pre-Industrial England," *Population Studies*, 25, 453–463. 7

Olsen, J. and P. K. Andersen (1999): "We Should Monitor Human Fecundity, but How? A Suggestion for a New Method That May Also Be Used to Identify Determinants of Low Fecundity," *Epidemiology*, 10, 419–421. 6

O'Rourke, K. H., A. S. Rahman, and A. M. Taylor (2013): "Luddites, the industrial revolution, and the demographic transition," *Journal of Economic Growth*, 18, 373–409. 1

Raatikainen, K., M. Harju, M. Hippeläinen, and S. Heinonen (2010): "Prolonged time to pregnancy is associated with a greater risk of adverse outcomes," *Fertility and sterility*, 94, 1148–1151. 7

Rosenzweig, M. R. and J. Zhang (2009): "Do Population Control Policies Induce More Human Capital Investment? Twins, Birthweight, and China's "One Child" Policy," *Review of Economic Studies*, 76, 1149–1174. 3

Santow, G. (1995): "Coitus Interruptus and the Control of Natural Fertility," *Population Studies*, 49, 19–43. 5

Schofield, R. S. (1970): "Perinatal mortality in Hawkshead, Lancashire, 1581–1710," *Local Population Studies*, 4, 11–16. 7

Schofield, R. S. (2005): ""Monday's child is fair of face": favoured days for baptism, marriage and burial in pre-industrial England," *Continuity and Change*, 20, 93–109. 8, 13

Schwartz, D. and M. Mayaux (1982): "Female fecundity as a function of age: results of artificial insemination in 2193 nulliparous women with azoospermic husbands." *Obstetrical & Gynecological Survey*, 37, 548–550. 15

Souden, D. (1984): "Movers and Stayers in Family Reconstitution Populations," *Local Population History*, 33, 11–28. 8, 35

Stone, L. (1977): *The family, sex and marriage in England, 1500–1800*, New Yourk: Harper & Row. 18

Strulik, H. and J. Weisdorf (2008): "Population, food, and knowledge: a simple unified growth theory," *Journal of Economic Growth*, 13, 195–216. 1

Voigtländer, N. and H.-J. Voth (2009): "Malthusian dynamism and the rise of Europe: make war, not love," *The American Economic Review*, 99, 248–254.

Wise, L. A., E. M. Mikkelsen, K. J. Rothman, A. H. Riis, H. T. Sørensen, K. F. Huybrechts, and E. E. Hatch (2011): "A prospective cohort study of menstrual characteristics and time to pregnancy," *American journal of epidemiology*, 174, 701–709. 15

Woods, J. W. (1994): *Dynamics of Human Reproduction: Biology, Biometry, Demography*, New York: Aldine de Gruyter. 6

Wooldridge, J. M. (2010): *Econometric Analysis of Cross Section and Panel Data*, London: MIT Press, second ed. 62

Wrigley, E. A., R. Davies, J. Oeppen, and R. Schofield (1997): *English Population History from Family Reconstitution*, Cambridge: Cambridge University Press. 4, 5, 7, 8

Appendix for Online Publication

Table A.1: Summary Statistics

	Mean	S.D.	Count	P10	P90
Surviving Siblings (> 5 Years)	4.83	2.51	1517	2	8
Literate	0.56	0.50	1248	0	1
Skilled	0.68	0.47	652	0	1
Years to First Birth	1.58	1.18	1517	0.81	2.99
Male	0.53	0.50	1517	0	1
Non-Sunday Baptism	0.53	0.50	1485	0	1
Skilled Father	0.69	0.46	925	0	1
Skilled Mother	0.63	0.49	35	0	1
Literate Father	0.60	0.49	969	0	1
Literate Mother	0.32	0.47	942	0	1
Longevity of Father (Years)	72.37	9.79	1517	59.0	84.1
Age at Marriage of Mother (Years)	25.08	4.68	1517	19.8	31.0
Agricultural Location	0.25	0.43	1517	0	1
Industrial Location	0.24	0.43	1517	0	1
Retail Location	0.16	0.36	1517	0	1
Centuries since 1500	2.71	0.38	1517	2.33	3.07
Observations	1,517				

Table A.2: Example Family

Parish Level Information		Family Level Information
Parish	Occupational Type	TFB
Odiham	Mixed	0.92

Individual Level Information

Family Member	Name	Birth Date	Death Date	Age at Death	Occupation	Skilled Profession	Literate	Marriage Age
Mother	Hanna Sury	21 July 1740	10 Nov. 1816	76.3	-	-	No	21.2
Father	Edward Neville	14 May 1773	3 Nov. 1816	83.5	Labourer	No	No	28.4
Daughter	Ann	8 Oct. 1762	-	-	-	-	No	
Son	John	17 Apr. 1765	13 Oct. 1850	85.5	Sawyer	Yes	No	
Son	Edward	3 Mar. 1767	8 May 1852	85.2	Baker	Yes	Yes	
Son	James	3 May 1769	14 Apr.1849	79.9	Labourer	No	No	
Son	Thomas	6 Mar. 1771	20 Mar. 1771	0.0	-	-	-	
Son	Daved	28 Mar. 1773	13 May 1858	85.1	-	-	No	
Son	Thos	23 Apr. 1775	21 Dec. 1855	80.7	Sawyer	Yes	No	
Son	Francis	8 June 1777	9 May 1780	2.9	-	-	-	
Daughter	Hannah	5 Dec. 1779	-	-	-	-	-	

"TFB" is the time from the marriage to the first birth, measured in years. "-" indicates missing information.

The table provides an example of the statistics transcribed from the church book as well as those inferred either by us or by the Cambridge Group. The record shows that in Odiham on 15 Oct. 1761 Edward Neville (baptized 14 May 1733, buried 3 Nov. 1816 at age 83) married Hannah Sury (baptized 21 July 1740, buried 10 Nov. 1816 at age 76). At the time of the marriage, husband Edward was registered in the church book as a labourer, which according to the HISCLASS is an unskilled occupation. He was recorded as being illiterate, as was his wife. Wife Hannah gave birth to a total of nine children (seven boys and two girls), two of which (Thomas and Francis) died before reaching the age of five, leaving a total of seven "surviving" children. Six of the seven survivors married in their parish of birth. James (a labourer) was unskilled, while Edward (a baker), John and Thos (both sawyers) were skilled workers. The record also shows that Edward was literate but that his siblings were all illiterate, except for lastborn Hannah who at some stage during her life moved away to a parish outside the sample (indicated by her missing death date) rendering her marriage and literacy status unknown.

Table A.3: Robustness to Cutoff at 38 Weeks (Instead of 40 Weeks)

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.025**	.026**	.025**	.036**	.035***	.035***	.097**	.100**	.091**
	(.012)	(.012)	(.011)	(.014)	(.013)	(.013)	(.042)	(.041)	(.041)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
R^2	.163	.170	.220	.266	.300	.310	.303	.322	.341
N. of Observations	1,350	1,350	1,350	719	719	719	754	754	754
N. of Families	616	616	616	497	497	497	513	513	513

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to the inclusion of families with TFB between 38 and 40 weeks, we perform the analysis on an en extended sample including these families. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the same column in Table 2. The table establishes that the qualitative conclusion is robust to this alternative sample restriction. In particular, the association between TFB and the three measures of human capital achievements remain significant in all specifications.

Table A.4: Robustness to Birth Year

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.036***	.035***	.034***	.045***	.042***	.042***	.131***	.132***	.125***
	(.013)	(.013)	(.012)	(.014)	(.013)	(.013)	(.042)	(.041)	(.040)
Birth Time Period Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
R^2	.168	.177	.230	.295	.335	.343	.322	.343	.366
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Fecundity and affluence of individuals may be affected by the socioeconomic and demographic patterns over the individuals' lifetime, as captured by their birth year. Thus, a dummy variable indicating the birth year of the individual (on the vigintennial, i.e., the 20-year, level) is included in the regression analysis to account for these confounding factors. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the full specifications in Table 2, i.e. to columns 3, 6, and 9, except for the inclusion of birth vigintennial dummy variables. The table establishes that the qualitative conclusion is robust to these alternative specifications. In particular, the association between TFB and the three measures of human capital achievements remain significant in all specifications.

Table A.5: Robustness to Inclusion of Paupers and Gentry in the Analysis of Skills

	Occupational Skills (with Paupers and Gentry)		
	(1)	(2)	(3)
Years to First Birth	.030**	.031**	.030**
	(.014)	(.013)	(.012)
Baseline Controls	Yes	Yes	Yes
Parental Skills	No	Yes	Yes
Parental Literacy	No	No	Yes
R^2	.258	.293	.302
N. of Observations	686	686	686
N. of Families	468	468	468

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to the inclusion of paupers and gentry in the occupational skills measure, we perform the analysis using an alternative occupational skills measure in which paupers are defined as unskilled (as opposed to being omitted from the analysis) and gentry are defined as skilled (as opposed to being omitted from the analysis). We update both the parental occupational skills variable and the outcome occupational skills variable. The specifications underlying each column in the table corresponds to the specifications in Table 2, i.e. to columns 7, 8, and 9. The table establishes that the qualitative conclusion is robust to this alternative occupational skills variable defined for a larger sample. In particular, the association between TFB and the measure of occupational skills remain significant in all specifications.

Table A.6: Robustness to Use of all Births (Instead of Surviving Births Only)

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Siblings (All)	-.067***	-.069***	-.072***	-.074***	-.074***	-.072***	-.209***	-.220***	-.202***
	(.026)	(.027)	(.025)	(.028)	(.027)	(.026)	(.077)	(.077)	(.075)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
R^2	.069	.088	.144	.137	.209	.226	.201	.245	.283
F (Kleibergen-Paap)	60.1	60.5	58.4	53.6	50.7	51.6	61.6	59.5	61.1
Anderson-Rubin F stat. p-value	.010	.011	.004	.004	.003	.003	.005	.003	.006
Endogeneity test p-value	.013	.018	.012	.006	.006	.007	.008	.009	.016
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the instrumental variable estimates to the inclusion of children that died between age 0 and 5 years in the measure of family size, we perform the instrumental variable analysis with the total number of siblings, rather than the number of surviving siblings, as the main explanatory (and endogenous) variable. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the same column in Table 6. The table establishes that the qualitative conclusion is robust to this alternative sample restriction. In particular, TFB remain a strong instrument, with Kleibergen-Paap F-statistics above 50, and the estimate of the effect of the number of siblings on the level of human capital remain highly significant in all specifications.

Table A.7: Robustness to Alternative Birth Order Specifications

	Literacy			Skilled Occupation			Occupational Wealth		
	No control for birth order	Control for relative birth order	Control for relative birth order	No control for birth order	Control for relative birth order	Control for relative birth order	No control for birth order	Control for relative birth order	Control for relative birth order
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.034***	.032***	.032**	.038***	.038***	.034**	.119***	.118***	.117**
	(.011)	(.011)	(.013)	(.013)	(.013)	(.014)	(.040)	(.040)	(.047)
Order		-.005			-.001			-.001	
		(.006)			(.007)			(.022)	
Relative Order			-.021			.002			-.089
			(.025)			(.030)			(.090)
Baseline Controls (Except Birth Order FE)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Literacy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	.222	.223	.225	.321	.321	.327	.348	.348	.346
N. of Observations	1,248	1,248	1,210	652	652	634	686	686	667
N. of Families	571	571	533	453	453	435	468	468	449

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to the control for birth order effects, and to shed light on the effect of birth order, we perform the analysis while not accounting for birth order, and, while including alternative birth order controls. In particular, we completely omit birth order variables (in columns 1, 4, and 7), we control for the absolute birth order of the individual (in columns 2, 5, and 8), and we control for the relative birth order of the individual (in columns 3, 6, and 9). Since birth order is naturally linked to family size, the relative birth order measure helps controlling for the effect of being later-born on human capital achievements, while avoiding the conflation of order and size inherent to the absolute birth order measure (Ejrnæs and Pörtner, 2004). While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the full specifications in Table 2, i.e. to columns 3, 6, and 9, except for the alternative ways of controlling for birth order. The table establishes that the qualitative conclusion is robust to these alternative specifications. In particular, the association between TFB and the three measures of human capital achievements remain significant in all specifications.

Table A.8: Robustness to Alternative Marriage and Stoppage Age Specifications

	Literacy			Skilled Occupation			Occupational Wealth		
	No stoppage or marriage age controls	Control for maternal stoppage age	Control for maternal stoppage and marriage age	No stoppage or marriage age controls	Control for maternal stoppage age	Control for maternal stoppage and marriage age	No stoppage or marriage age controls	Control for maternal stoppage age	Control for maternal stoppage and marriage age
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.032***	.031***	.031***	.035***	.034***	.036***	.108***	.106***	.111***
	(.011)	(.011)	(.011)	(.012)	(.012)	(.013)	(.040)	(.040)	(.041)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
R^2	.221	.228	.233	.321	.333	.340	.349	.350	.355
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to controlling for the timing of the last birth in the family as well as to omitting the marriage age control variable, we perform the analysis with alternative marriage and stoppage age specifications. In particular, we estimate the model (i) with neither the marriage age nor the stoppage age control variables, (ii) with the stoppage age control variable, and (iii) with both the marriage age and the stoppage age control variables. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the full specifications in Table 2, i.e. to columns 3, 6, and 9, except for the differences in the marriage and stoppage age specifications. The table establishes that the qualitative conclusion is robust to these alternative specifications. In particular, the association between TFB and the three measures of human capital achievements remain significant in all specifications.

Table A.9: Robustness to Accounting for Paternal Occupational Wealth

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.032**	.032**	.033***	.037***	.040***	.040***	.113***	.117***	.110***
	(.013)	(.013)	(.011)	(.014)	(.013)	(.013)	(.044)	(.042)	(.042)
Paternal Occupational Wealth Class Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
R^2	.168	.174	.229	.303	.325	.336	.338	.359	.376
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to accounting for family living standards, we perform the analysis while including control variables capturing paternal wealth. In particular, we include dummies indicating which of the seven occupational wealth classes the father's occupation belongs to, or if the father's occupation is unknown. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the same column in Table 2. The table establishes that the qualitative conclusion is robust to these alternative specifications. In particular, the association between TFB and the three measures of human capital achievements remain significant in all specifications.

Table A.10: Robustness to Fixed Effects on the Yearly Level

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.036***	.037***	.038***	.041***	.041***	.039***	.123***	.127***	.114***
	(.012)	(.012)	(.011)	(.014)	(.013)	(.013)	(.043)	(.042)	(.042)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
R^2	.189	.198	.258	.311	.349	.362	.348	.374	.398
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to the resolution of time-related dummy variables, we perform the analysis while controlling for parental marriage year fixed effects on the 1-year level (instead of the 20-year level) and maternal marriage age year fixed effects on the 1-year level (instead of on the 5-year level). While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the full specifications in Table 2, i.e. to columns 3, 6, and 9, except for alternative marriage year and marriage age dummies. The table establishes that the qualitative conclusion is robust to this alternative occupational skills variable defined for a larger sample. In particular, the association between TFB and the measure of occupational skills remain significant in all specifications.

Table A.11: Robustness to Parish Fixed Effects

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.029**	.029**	.030***	.029**	.028**	.028**	.036**	.036***	.035**
	(.012)	(.012)	(.011)	(.014)	(.013)	(.013)	(.014)	(.014)	(.014)
Parish Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
Signif. of Parish FE	.000	.000	.000	.000	.000	.000	.000	.000	.000
R^2	.185	.194	.250	.332	.362	.371	.411	.424	.434
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to accounting for parish-specific fixed effects, we perform the analysis while including dummy variables indicating the parish. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the same column in Table 2. The table establishes that the qualitative conclusion is robust to these alternative specifications. In particular, the association between TFB and the three measures of human capital achievements remain significant in all specifications.

Table A.12: Robustness to Heckit Estimation

	Literacy		Skilled Occupation		Occupational Wealth	
	(1)	(2)	(3)	(4)	(5)	(6)
Years to First Birth		.032***		.038***		.113***
		(.011)		(.013)		(.041)
No Death Date	-.593***		-.181***		-.222***	
	(.079)		(.058)		(.057)	
No Marriage Date	-3.897***		-1.770***		-1.832***	
	(.146)		(.070)		(.069)	
Inverse Mills Ratio		.020		-.004		.014
		(.033)		(.029)		(.094)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	Yes	Yes	Yes	Yes	Yes	Yes
Parental Literacy	Yes	Yes	Yes	Yes	Yes	Yes
R^2		.225		.329		.355
N. of Observations	8,619	1,248	8,619	652	8619	686
N. of Families	1631	571	1631	453	1631	468

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the possibility of sample selection bias in the reduced-form estimation of the association between human capital and TFB, we perform a series of Heckit analyses. In particular, we exploit the fact that a missing marriage and/or death date significantly predict the availability of human capital information. A missing death date indicates that the individual migrated to another parish before death, and a missing marriage date indicates that the individual either did not get married or that the individual migrated before marriage. To the extend that these events do not independently affect human capital achievements, they act as useful sample selection predictors.

Table A.12 shows that the dummies for missing marriage and death dates are both highly significant, confirming their significance in predicting a missing literacy or skill status. The inverse Mills ratio turns out to be highly insignificant in both stages of both regressions, verifying the absence of a sample selection bias. Indeed, the coefficient on the years to first birth remains highly significant for all three measures of human capital.

Table A.13: Robustness to Logit Regression

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.161**	.161**	.174***	.272***	.303***	.309***	.274***	.284***	.272***
	(.068)	(.070)	(.066)	(.105)	(.116)	(.117)	(.088)	(.089)	(.092)
Male	.464***	.480***	.540***	-1.934***	-2.049***	-2.033***	1.309***	1.319***	1.314***
	(.135)	(.136)	(.141)	(.511)	(.555)	(.583)	(.392)	(.399)	(.430)
Poor Father	-1.915***	-1.553***	-1.110***	-2.318***	-1.716***	-1.668***	-2.867***	-2.868***	-2.787***
	(.243)	(.283)	(.290)	(.298)	(.356)	(.355)	(.278)	(.340)	(.336)
Non-Sunday Baptism	.460***	.477***	.392**	.708***	.765***	.784***	.623***	.619***	.569**
	(.142)	(.144)	(.153)	(.237)	(.242)	(.246)	(.215)	(.219)	(.223)
Skilled Father		.691**	.502*		1.041***	.982***		.078	-.113
		(.293)	(.294)		(.374)	(.379)		(.403)	(.388)
Skilled Mother		1.502	2.086*		18.669***	17.165***		16.863***	16.799***
		(1.186)	(1.257)		(1.754)	(1.700)		(1.696)	(1.606)
Literate Father			.928***			.271			.819**
			(.224)			(.368)			(.350)
Literate Mother			1.197***			.379			.130
			(.274)			(.440)			(.400)
Agricultural Location	.475*	.450*	.488*	-.392	-.615	-.757*	-.722**	-.834***	-.931***
	(.245)	(.243)	(.263)	(.383)	(.374)	(.401)	(.368)	(.312)	(.325)
Industrial Location	.625**	.526**	.523**	2.535***	2.168***	2.012***	-.336	-.349	-.456
	(.246)	(.245)	(.251)	(.645)	(.645)	(.632)	(.388)	(.385)	(.381)
Retail Location	1.115***	1.113***	.935**	.075	-.027	-.167	.306	.332	.241
	(.364)	(.359)	(.376)	(.316)	(.331)	(.356)	(.307)	(.314)	(.328)
N. of Observations	1,242	1,242	1,242	640	640	640	680	680	680
N. of Families	566	566	566	444	444	444	464	464	464

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to estimation by an alternative estimator, we perform the analysis using logistic regression. The literacy and occupational skills variables are binary variables and therefore fit naturally into the logistic regression model framework. Based on the occupational wealth variable, which contains seven categories, we create a binary outcome variable indicating a high-wealth occupation. The variable is one when the occupational wealth variable is 3 or above and zero when the occupational wealth variable is below 3. This binary variable is used as the outcome in columns 7–9. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the same column in Table 2. The table establishes that the qualitative conclusion is robust to this alternative sample restriction. In particular, the association between TFB and the three measures of human capital achievements remain significant in all specifications.

Table A.14: Robustness to Gender Division

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.041***	.040**	.036**	.037**	.036***	.034**	.106**	.110**	.096**
	(.016)	(.016)	(.015)	(.014)	(.013)	(.013)	(.045)	(.045)	(.044)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
R^2	.197	.215	.258	.290	.336	.348	.320	.351	.380
N. of Observations	539	539	539	598	598	598	612	612	612
N. of Families	367	367	367	431	431	431	440	440	440

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to the focus on males, we perform the analysis on an restricted sample of males only. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the same column in Table 2. The table establishes that the qualitative conclusion is robust to this alternative sample restriction. In particular, the association between TFB and the three measures of human capital achievements remain significant in all specifications.

Table A.15: Robustness to Parental Longevity

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.032**	.035***	.035***	.049***	.041***	.041***	.113**	.101**	.093**
	(.014)	(.013)	(.012)	(.015)	(.013)	(.013)	(.051)	(.048)	(.046)
Paternal and Maternal Longevity Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
Signif. of Maternal Longevity FE	.064	.123	.137	.000	.000	.001	.030	.282	.634
Signif. of Paternal Longevity FE	.000	.000	.000	.000	.000	.001	.000	.000	.000
R^2	.108	.171	.233	.230	.334	.344	.204	.290	.323
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to accounting for parental longevity, we perform the analysis while including control variables capturing maternal and paternal longevity. In particular, we include dummies indicating the 5-year age interval of death of the mother and dummies indicating the 5-year age interval of death of the father. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the same column in Table 2. The table establishes that the qualitative conclusion is robust to these alternative specifications. In particular, the association between TFB and the three measures of human capital achievements remain significant in all specifications.

Table A.16: Robustness to Accounting for Seasonal Effects

	Literacy			Skilled Occupation			Occupational Wealth		
	Parental Marriage Season	First Birth Season	Parental Marriage Season and First Birth Season	Parental Marriage Season	First Birth Season	Parental Marriage Season and First Birth Season	Parental Marriage Season	First Birth Season	Parental Marriage Season and First Birth Season
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.032***	.033***	.033***	.039***	.038***	.039***	.118***	.114***	.118***
	(.011)	(.011)	(.011)	(.013)	(.013)	(.013)	(.041)	(.041)	(.041)
Marriage Season FE	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
First Birth Season FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Literacy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	.225	.230	.231	.330	.337	.339	.360	.359	.364
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to accounting for seasonal effects, we perform the analysis while including dummy variables capturing the parental marriage season as well as dummy variables indicating the season of the first birth. Seasons are defined based on the quarter of the calendar year. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the full specifications in Table 2, i.e. to columns 3, 6, and 9, except for the inclusion of seasonal dummy variables. The table establishes that the qualitative conclusion is robust to these alternative specifications. In particular, the association between TFB and the three measures of human capital achievements remain significant in all specifications.

Table A.17: Robustness to Temperatures

	Literacy				Skilled Occupation				Occupational Wealth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years to First Birth	.031***	.033***	.031***	.033***	.041***	.042***	.040***	.041***	.122***	.113***	.121***	.134***
	(.011)	(.011)	(.011)	(.011)	(.013)	(.015)	(.013)	(.015)	(.040)	(.043)	(.040)	(.043)
20-Year Average	-.031			-.026	.003			.015	-.218			-.050
Temperature at Birth	(.039)			(.039)	(.051)			(.060)	(.181)			(.213)
20-Year Average		.015		.020		.121***		.103**		.458***		.431***
Temperature at Parent's Marriage		(.041)		(.041)		(.046)		(.050)		(.153)		(.160)
20-Year Average			-.036	-.037			.058	.045			.025	.071
Temperature at Marriage			(.053)	(.053)			(.051)	(.057)			(.197)	(.217)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes			
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes			
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes			
R^2	.222	.220	.222	.219	.333	.352	.341	.351	.342	.341	.347	.345
N. of Observations	1,238	1,224	1,237	1,218	610	502	600	485	639	533	629	514
N. of Families	564	558	563	555	421	364	413	351	432	377	424	363

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to accounting for climatic effects, we perform the analysis while including variables capturing the average yearly surface air temperatures in the 20-year periods surrounding the (i) birth of the first child in the family, (ii) the marriage of the couple, and (iii) the marriage of the individual. The 20-year intervals include the nine years preceding the event, the nine years succeeding the event, and the year of the event itself. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the full specifications in Table 2, i.e. to columns 3, 6, and 9, except for the inclusion of the average temperature variables. The table establishes that the qualitative conclusion is robust to these alternative specifications. In particular, the association between TFB and the three measures of human capital achievements remain highly significant in all specifications.

Table A.18: Robustness to Vital Rates

	Literacy				Skilled Occupation				Occupational Wealth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years to First Birth	.033***	.033***	.032***	.035***	.039***	.040***	.038***	.042***	.112***	.118***	.114***	.125***
	(.011)	(.011)	(.011)	(.012)	(.013)	(.015)	(.013)	(.015)	(.042)	(.044)	(.041)	(.046)
20-Year Average	-.008			-.006	.111*			.144*	.230			.229
Crude Death Rate at Birth	(.056)			(.060)	(.067)			(.079)	(.244)			(.279)
20-Year Average	-.029			-.062	-.016			-.072	.022			-.189
Crude Birth Rate at Birth	(.054)			(.065)	(.065)			(.100)	(.236)			(.339)
20-Year Average		-.045		-.060		.009		.002		-.213		-.223
Crude Death Rate at Parent's Marriage		(.072)		(.076)		(.071)		(.071)		(.280)		(.285)
20-Year Average		.007		.010		.020		.024		-.412**		-.401**
Crude Birth Rate at Parent's Marriage		(.040)		(.039)		(.052)		(.052)		(.184)		(.186)
20-Year Average			-.021	-.020			-.036	-.058			.116	.159
Crude Death Rate at Marriage			(.069)	(.071)			(.070)	(.076)			(.223)	(.241)
20-Year Average			.002	.032			-.006	.034			.095	.209
Crude Birth Rate at Marriage			(.059)	(.074)			(.067)	(.097)			(.232)	(.316)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes			
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes			
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes			
R^2	.225	.222	.225	.223	.332	.341	.329	.349	.356	.350	.355	.353
N. of Observations	1,248	1,229	1,248	1,229	652	516	652	516	686	550	686	550
N. of Families	571	563	571	563	453	376	453	376	468	392	468	392

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to accounting for aggregate birth and death statistics, we perform the analysis while including variables capturing the aggregate yearly birth and death rates in the 20-year periods surrounding the (i) birth of the first child in the family, (ii) the marriage of the couple, and (iii) the marriage of the individual. The 20-year intervals include the nine years preceding the event, the nine years succeeding the event, and the year of the event itself. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the full specifications in Table 2, i.e. to columns 3, 6, and 9, except for the inclusion of the aggregate birth and death variables. The table establishes that the qualitative conclusion is robust to these alternative specifications. In particular, the association between TFB and the three measures of human capital achievements remain highly significant in all specifications.

Table A.19: Robustness to Wages

	Literacy				Skilled Occupation				Occupational Wealth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years to First Birth	.032***	.034***	.032***	.034***	.039***	.041***	.037***	.040***	.116***	.110**	.108***	.107**
	(.011)	(.011)	(.011)	(.011)	(.013)	(.015)	(.013)	(.015)	(.041)	(.044)	(.041)	(.044)
20-Year Average Aggregate	-.050			-.050	-.016			-.016	-.385			-.186
Wage at Birth	(.072)			(.075)	(.070)			(.086)	(.257)			(.325)
20-Year Average Aggregate		.039		.035		-.039		-.036		.300		.316
Wage at Parent's Marriage		(.053)		(.053)		(.073)		(.073)		(.251)		(.251)
20-Year Average Aggregate			.024	.022			-.129	-.047			-.446	-.366
Wage at Marriage			(.094)	(.096)			(.108)	(.119)			(.380)	(.431)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Literacy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	.225	.222	.225	.223	.329	.342	.331	.342	.357	.343	.357	.345
N. of Observations	1,248	1,229	1,248	1,229	652	516	652	516	686	550	686	550
N. of Families	571	563	571	563	453	376	453	376	468	392	468	392

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the robustness of the estimates of the effects of TFB on offspring human capital achievements to accounting for the aggregate economic climate, we perform the analysis while including variables capturing the aggregate average yearly wage rate in the 20-year periods surrounding the (i) birth of the first child in the family, (ii) the marriage of the couple, and (iii) the marriage of the individual. The 20-year intervals include the nine years preceding the event, the nine years succeeding the event, and the year of the event itself. While the coefficients on the control variables are omitted from the table, the specifications underlying each column in the table corresponds to the full specifications in Table 2, i.e. to columns 3, 6, and 9, except for the inclusion of the aggregate birth and death variables. The table establishes that the qualitative conclusion is robust to these alternative specifications. In particular, the association between TFB and the three measures of human capital achievements remain highly significant in all specifications.

Table A.20: Robustness to Heckit Estimation (Instrumental Variable Estimation)

	Literacy		Skilled Occupation		Occupational Wealth	
	(1)	(2)	(3)	(4)	(5)	(6)
Surviving Siblings (> 5 Years)		-.072***		-.079***		-.217***
		(.025)		(.028)		(.081)
Years to First Birth	.044		.006		.014	
	(.033)		(.025)		(.024)	
No Death Date	-.597***		-.182***		-.223***	
	(.079)		(.058)		(.057)	
No Marriage Date	-3.900***		-1.770***		-1.832***	
	(.144)		(.070)		(.069)	
Inverse Mills Ratio		.016		.002		.020
		(.031)		(.030)		(.095)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	Yes	Yes	Yes	Yes	Yes	Yes
Parental Literacy	Yes	Yes	Yes	Yes	Yes	Yes
R^2		.157		.243		.286
F (Kleibergen-Paap)		64.4		56.0		66.4
Anderson-Rubin F stat. p-value		.004		.003		.006
Endogeneity test p-value		.010		.009		.011
N. of Observations	8,619	1,248	8,619	652	8,619	686
N. of Families	1631	571	1631	453	1631	468

Standard errors clustered on the family level are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

To assess the possibility of sample selection bias in the instrumental variables estimation of the effect of the number of siblings on the level of human capital, we perform a series of three-step instrumental-variables Heckit analyses (Wooldridge, 2010, Procedure 19.2). In particular, we exploit the fact that a missing marriage and/or death date significantly predict the availability of human capital information. A missing death date indicates that the individual migrated to another parish before death, and a missing marriage date indicates that the individual either did not get married or that the individual migrated before marriage. To the extend that these events do not independently affect human capital achievements, they act as useful sample selection predictors.

In the first stage we extend the sample to also include observations where literacy and skill status are unknown, thus expanding the sample to 8,647 individuals representing a total of 1,639 families.[45] Next, we estimate the probability of observing human capital with a probit model, using dummies for missing marriage or death dates as instruments in addition to TFB (and covariates). We have 6,037 observations with missing marriage dates; 4,405 observations with missing death dates; and 2,976 cases where both dates are missing. Based on the predicted probabilities, we calculate the inverse Mills ratio, proceeding to estimate Equation (2) by 2SLS including the inverse

[45] The results from this stage establishes that there is no relationship between TFB and the chances of observing the children's human capital.

Mills ratio as a control variable. We conduct the procedure for both outcome variables (i.e. literacy and skill status). If the inverse Mills ratio is statistically significant in the first or second stage, then it means our estimations possibly suffer from a sample selection bias.

Table A.20 shows that the dummies for missing marriage and death dates are both highly significant, confirming their significance in predicting a missing literacy or skill status. The inverse Mills ratio turns out to be highly insignificant in both stages of both regressions, verifying the absence of a sample selection bias. Indeed, the coefficient on the years to first birth remains highly significant for all three measures of human capital.