The Biocultural Origins of Human Capital Formation[*]

Oded Galor   Marc Klemp

Abstract

This research explores the biocultural origins of human capital formation. Using an extensive genealogical record for nearly half a million individuals in Quebec from the 16th to the 18th century, the study establishes that moderate fecundity and thus predisposition towards investment in child quality was conducive for long-run reproductive success, reflecting the negative effect of higher fecundity on the survivability, marriage age and education of each offspring. The finding lends credence to the hypothesis that natural selection has favored individuals with predisposition towards child quality, contributing to human capital formation, the demographic transition and the transition from stagnation to growth.

Keywords Demography, Evolution, Human Capital Formation, Natural Selection, Fecundity, Quantity-Quality Trade-Off, Long-Run Reproductive Success, Economic Growth

JEL Classification Codes J10, N30, O10

1 Introduction

The transition from an epoch of stagnation to an era of sustained economic growth has triggered one of the most significant transformations in the course of human history. While living standards in the world economy stagnated during the millennia preceding the Industrial Revolution, income per capita has experienced an unprecedented twelvefold increase over the past two centuries, profoundly altering the level and distribution of education, health, and wealth across the globe.[1]

Over most of human existence, the process of development was marked by Malthusian stagnation. The Malthusian pressure has governed the evolution of the size of the population, and conceivably, via the forces of natural selection, has shaped the composition of the population as well. Lineages of individuals whose traits were complementary to the economic environment generated higher income, and thus higher reproductive success. The gradual increase in the representation of these growth-enhancing traits in the population presumably has contributed to the process of development and the take-off from stagnation to growth (Galor and Moav, 2002).[2]

In particular, Galor and Moav (2002) have advanced the hypothesis that during the Malthusian epoch, natural selection brought about a gradual increase in the prevalence of traits associated with predisposition towards offspring quality. The effect of this evolutionary process on investment in human capital stimulated technological progress and contributed to the reinforcing interaction between investment in human capital and technological progress that triggered the demographic transition and brought about a state of sustained economic growth.

This research explores the biocultural origins of human capital formation. It presents the first evidence that moderate fecundity and thus predisposition towards investment in child quality was conducive for long-run reproductive success within the human species. It further suggests that individuals with lower levels of fecundity than the median in the population generated an evolutionary advantage in the pre-demographic transition era. These findings lend credence to the hypothesis that during the Malthusian epoch, natural selection favored individuals with lower fecundity and thus larger predisposition towards child quality, contributing to human capital formation, the onset of the demographic transition and the evolution of societies from an epoch of stagnation to sustained economic growth.[3]

---

[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, 2000; Galor and Moav, 2002; Hansen and Prescott, 2002; Lucas, 2002; Galor, 2011).

[2]Evidence suggests that the composition of existing genetic traits may experience rapid evolutionary processes. Voight et al. (2006) detected about 700 regions of the human genome where genes have been reshaped by natural selection within the last 5,000 to 15,000 years. Other notable evidence suggests that lactose tolerance was developed among Europeans and Near Easterners since the domestication of dairy animals in the course of the Neolithic revolution, whereas in regions that were exposed to dairy animals in later stages, a larger proportion of the adult population suffers from lactose intolerance. Furthermore, genetic immunity to malaria provided by the sickle cell trait is prevalent among descendants of Africans whose engagement in agriculture improved the breeding ground for mosquitoes and thereby raised the incidence of malaria, whereas this trait is absent among descendants of nearby populations that have not made the transition to agriculture (Livingstone, 1958; Wiesenfeld, 1967; Durham, 1982).

[3]The interaction between human evolution and the process of development, as was further explored theoretically by Lagerlöf (2007); Dalgaard and Strulik (2011); Galor and Michalopoulos (2012), is applicable to either cultural or genetic intergenerational transmission of entrepreneurial traits (Cavalli-Sforza, 1981; Boyd, 1988; Weibull, 1997;

The influential life-history theory in the field of evolutionary biology suggests that observed fecundity of organisms reflects a trade-off in reproductive success between the quantity and quality of offspring. Central to the theory is the supposition that there exists an optimal level of fecundity beyond which fitness diminishes.[4] A negative association between the quantity and the quality of offspring has been documented in a wide variety of species, ranging from plants to humans. In particular, researchers uncovered an inverse relationship between the number of seeds and their size as well as between the quantity and quality of offspring within and across mammals.[5] Moreover, a trade-off between fertility on the one hand and offspring survival probability and education on the other hand has been documented for pre-industrial human societies.[6]

Nevertheless, as established in the theoretical model, the presence of a static trade-off between the quantity and quality of offspring is merely a necessary but not a sufficient condition for the presence of an adverse effect of fecundity on reproductive success in the long run. In particular, a priori, in an environment in which the carrying capacity was an order of magnitude greater than the size of the founder population, individuals with the highest level of fecundity could have had the largest reproductive success if, for any feasible number of children, an additional child would have contributed directly to the number of offspring in the long-run more than the adverse indirect effect of this additional child on the quality of his siblings and hence their long-run reproductive success. However, if investment in child quality is sufficiently productive in terms of its effect on the potential income and thus the reproductive success of the child (i.e., if the elasticity of fertility of an offspring with respect to fertility of the founding parent (evaluated at the highest feasible level of fertility) is greater than one in absolute value, then an intermediate level of fecundity generates the highest reproductive success and will be therefore favored by natural selection.[7]

This research explores the effect of fecundity (i.e., the capacity to reproduce) on long-run reproductive success within the human species. Using an extensive genealogical record for nearly half a million individuals in Quebec from the 16th to the 18th century, the study examines the effect of fecundity on the number of descendants of early inhabitants of this Canadian province in the subsequent four generations. In particular, in light of the social norm observed in pre-industrial Quebec, in which marriage marked the intention to conceive, the research exploits variation in the random component of the time interval between the date of first marriage and the first birth during this time period to capture the effect of fecundity on fitness. The research establishes that while higher fecundity is associated with a larger number of children, an intermediate level of fecundity is conducive for long-run reproductive success.

Bowles, 1998; Bisin and Verdier, 2000; Doepke and Zilibotti, 2008). The long lasting effects of these historically determined genetic factors on comparative development have been established by Galor and Moav (2007); Spolaore and Wacziarg (2009); Ashraf and Galor (2013); Ashraf et al. (2014).

[4]See Lack et al. (1954); Cody (1966); Roff (1992); Stearns (1992).

[5]See Salisbury et al. (1942); Harper et al. (1970); Roff (2002); Charnov and Ernest (2006); Walker et al. (2008).

[6]See Lee (1993); Hill and Hurtado (1996); Strassmann and Gillespie (2002); Gillespie et al. (2008); Meij et al. (2009); Becker et al. (2010).

[7]Furthermore, few attempts to examine the related phenomenon of the effect of the number of children on fitness are largely inconclusive (Kaplan et al., 1995; Borgerhoff Mulder, 2000).

The research finds that the maximal reproductive success is attained by couples with a moderate level of time to first birth (i.e., those whose first delivery occurs 65 weeks after their marriage, in comparison to a sample median of 53 weeks). In particular, in comparison to highly fertile couples whose first child is born 38 weeks after the marriage, those individuals have on average 0.3 fewer children, but 0.4 more grandchildren, 8.4 additional great-grandchildren, and 15.7 added great-great-grandchildren. In light of the heritability of fecundity, the finding that the optimal level of time to first birth is above the population median suggests that in pre-industrial Quebec, the representation of individuals with lower levels of fecundity, and thus higher pre-disposition towards child quality, has gradually increased in the population.[8]

The research identifies several mechanisms that had contributed to the trade-off associated with higher fecundity and to the observed hump-shaped effect of fecundity on reproductive success in the long run. While individuals with lower fecundity had fewer children, the observed hump-shaped effect of fecundity on long-run reproductive success reflects the beneficial effect of lower fecundity on various measures of the quality of each child. In particular individuals with lower fecundity were more likely to have children that: (i) survived and got married, (ii) married at an earlier age, and (iii) were educated. Thus, despite the positive effect of fecundity on the number of children, the adverse effect of fecundity on child quality and the reproductive success of each child, generated the observed hump shaped relationship between fecundity and long-run reproductive success.[9]

The evidence from pre-industrial Quebec suggests that the forces of natural selection favored individuals characterized by moderate fecundity, increasing the population's predisposition towards investment in child quality. Interestingly, the conditions that were faced by the founder population of Quebec during this high fertility time period may resemble the environment that anatomically modern humans confronted during their migration out of Africa, as they settled new territories where the carrying capacity of the new environment was an order of magnitude greater than the size of the founder population. Thus, the findings suggest that during the high fertility segments of the Malthusian epoch in which evolutionary forces could have made a significant impact on the composition of the population (e.g., during the Neolithic transition and the formation of sedentary agricultural communities), natural selection favored individuals with a larger predisposition towards child quality, contributing to human capital formation, the onset of the demographic transition, and the evolution of societies from an epoch of stagnation to sustained economic growth

2 A Theoretical Model

This section presents a theoretical model that explores the effect of the level of fecundity of members of a founding generation (i.e., individuals who settled new territories where the carrying capacity

---

[8]For the heritability of fecundity, see Christensen et al. (2003); Pettay et al. (2005); Ramlau-Hansen et al. (2008); Kosova et al. (2009). In the present sample relating to Quebec, the time from marriage to first birth is heritable ($h^2 = 0.04$).

[9]As established in Tables 6-8, parental education is positively associated with the child's survivability and education, and negatively with the child's marriage age. Hence, the positive effects of lower parental fecundity on the child quality is further amplified, generating higher reproductive sucess for subsequent generations.

of the new environment was an order of magnitude greater than the size of the founder population) on their long-run reproductive success.

2.1 The Basic Structure of the Model

Consider an overlapping-generations economy originated in a newly acquired (sparsely populated) territory in which the economic environment could support fertility rates at the maximal feasible physiological rate. Individuals live for two periods: childhood and adulthood. In the first period of life, childhood, individuals are passive economic agents, and their consumption needs are provided by their single parent. In the second period of life, adulthood, individuals work and allocate their income between their own consumption and child rearing.

Preferences of a member i of generation t are represented by a utility function, u_t^i, defined over the individual's consumption, c_t^i, the individual's number of children, n_t^i, and the children's level of human capital, h_{t+1}^i (i.e., when they are engaged in production in the subsequent period as adults).[10] Specifically,

$$u_t^i = (1 - \gamma) \ln c_t^i + \gamma [\ln n_t^i + \beta \ln h_{t+1}^i], \tag{1}$$

where $\beta \in (0,1)$ captures the individual's preference for quality, and $\gamma \in (0,1)$ captures the individual's preference for for children relative to consumption.

The income of a member i of generation t, y_t^i, reflects the individual's human capital, h_t^i, and the productivity, w, of each efficiency unit of labor.[11] In particular,

$$y_t^i = w h_t^i. \tag{2}$$

The income of a member i of generation t, y_t^i, is divided between consumption, c_t^i, and expenditure on children. The household's budget constraint is

$$n_t^i (\tau + q_t^i) + c_t^i \leq y_t, \tag{3}$$

where τ is the minimum amount of parental resources required to raise a child, and $\tau + q_t^i$ is the cost of raising a child with quality q_t^i.

The level of human capital of a child of a member i of generation t, h_{t+1}^i, is an increasing and strictly concave function of the parental investment in the education of the child, q_t^i :

$$h_{t+1}^i = h(q_t^i), \tag{4}$$

---

[10]For simplicity, parents derive utility from the expected number of surviving offspring, and the parental cost of child rearing is associated only with surviving children. The introduction of costs associated with non-surviving children, or risk aversion, would not affect the qualitative features of the theory.

[11]In line with the evolution of wages over the sample period, wages are assumed to be constant over time. The introduction of a rising productivity would not alter the qualitative results.

where $h'(q_t^i) > 0$, $h''(q_t^i) < 0, \lim_{q_t \to 0} h'(q_t^i) < \infty$, $\lim_{q_t \to \infty} h'(q_t^i) = 0$, $h''(q_t^i)q_t^i + h'(q_t^i) < 0$, and $h(0) = 1$ (i.e., in the absence of investment in quality, each individual has a one efficiency unit of labor).[12]

2.2 Optimization

A member i of generation t allocates resources between consumption and child rearing in order to maximize utility.

2.2.1 Unconstrained Optimization

In the absence of additional constriants, given the homotheticity of preferences, a member i of generation t allocates a fraction $(1 - \gamma)$ of income to consumption and a fraction γ to to child rearing. Hence,

$$c_t^i = (1 - \gamma)y_t^i;$$

$$(\tau + q_t^i)n_t^i = \gamma y_t^i.$$

(5)

Moreover, applying the Kuhn-Tucker conditions, the level of quality of the children of a member i of generation t, $q_t^i \geq 0$, will be determined such that

$$\beta h'(q_t^i)(\tau + q_t^i) - h(q_t^i) \quad \begin{cases} \leq 0 \quad \text{if} \quad q_t^i = 0 \\ \\ = 0 \quad \text{if} \quad q_t^i > 0. \end{cases}$$

(6)

The maximal number of children, $\left[n_t^i\right]^{\max}$, that a member i of generation t can raise is, therefore,

$$\left[n_t^i\right]^{\max} = \frac{\gamma y_t^i}{\tau} \equiv n^{\max}(y_t^i).$$

(7)

This maximum will be reached only if the individual does not invest in the quality of children (i.e., $q_t^i = 0$).

2.2.2 Constrained Optimization

Given income, y_t^i, each member i of generation t will determine an age of marriage and, therefore, an onset of the reproductive period so as to achieve a desired fertility level, n_t^i, that would not exceed $n^{\max}(y_t^i)$. Suppose, however, that the level of fertility of a member i of generation t is affected by the randomness associated with conception, and that the actual fertility of the individual therefore may deviate from the desired level. In particular, suppose that, in the absence of reliable contraceptive methods or fertility enhancing drugs, nature dictates to each individual the number of children that

---

[12]These conditions are satisfied if, for instance, $h(q_t^i) = (1 + q_t^i)^\alpha$, for $\alpha \in (0, 1)$.

will be born within the feasibility constraint, i.e.,

$$n_t^i \in (0, n^{\max}(y_t^i)], \tag{8}$$

where the number of children is assumed to be strictly positive.[13]

Since preferences are homothetic, the allocation of the individual's resources between consumption and children will not be altered.[14] However, the division of resources between the quantity and quality of each child will be affected by this random assignment of the number of children. A lower number of children will generate higher child quality.

2.3 The Founder Population

Members of the founder population (generation 0) are assumed to be identical in their level of human capital, h_0, and therefore in their income, y_0. Their desired level of fertility, therefore, is identical as well. Nevertheless, due to the randomness associated with conception, the number of children may differ across members of generation 0. In particular, the number of children of a member i of generation 0 is randomly drawn from the interval $(0, n^{\max}(y_0)]$, i.e.,

$$n_0^i \in (0, n^{\max}(y_0)]. \tag{9}$$

Consider a member i of generation 0 that draws n_0^i children. In light of the budget constraint, the optimization of individual i, given by (5), implies that the quality of individual i's children is given by

$$q_0^i = \frac{\gamma y_0}{n_0^i} - \tau, \tag{10}$$

where $\partial q_0^i / \partial n_0^i = -\gamma y_0 / \left[n_0^i \right]^2 < 0$. Thus, the quality of the children of member i of generation 0 is declining in the number of children.

2.4 Subsequent Generations

Members i of generation $t = 1, 2, 3, \ldots, T$ (i.e., the descendants of member i of generation 0) generate an income, y_t^i,

$$y_t^i = wh(q_{t-1}^i), \tag{11}$$

reflecting the parental investment in their human capital, q_{t-1}^i.

---

[13]In light of the possibility that individuals engaged in within-marriage fertility control, the source of exogenous variation in the number of children in the empirical analysis is based on the random component of the variation in the time from the onset of unprotected intercourse as marked by the date of marriage to the first birth. Thus, the source of exogenous variation in the number of children exploited in the empirical analysis is entirely related to the random component of the variation in the timing of first birth and not to subsequent birth intervals.

[14]Given income, the individual determines the marriage age so as to assure that, given the gestation period of a child, the number of children that could be born during the remaining reproductive period could not exceed $n^{\max}(y_t^i)$. The qualitative results will remain unchanged if alternatively individuals are subjected to a non-binding subsistence consumption constraint, $\tilde{c}$, and the the value of $n^{\max}(y_0)$ could be $[y_0 - \tilde{c}]/\tau$, independent of preferences.

Suppose, for simplicity, that a member i of generation t can obtain their desired number of children for all $t \geq 1$.[15] The household's optimization implies, therefore, that parental consumption $c_t^i = (1 - \gamma)y_t^i$, the optimal level of children, n_t^i, and their quality, q_t^i, are obtained by the unconstrained optimization (5) and (6).

Lemma 1. *The level of investment in child quality, q_t^i, and a level of fertility, n_t^i, of members i of generation $t = 1, 2, 3, \ldots, T$ is such that*

$$q_t^i = \begin{cases} 0 & if \quad h'(0) \leq 1/\beta\tau \\ q^* > 0 & if \quad h'(0) > 1/\beta\tau; \end{cases}$$

$$ (12) $$

$$n_t^i = \begin{cases} \frac{\gamma y_t^i}{\tau} & if \quad h'(0) \leq 1/\beta\tau \\ \frac{\gamma y_t^i}{\tau + q^*} & if \quad h'(0) > 1/\beta\tau, \end{cases}$$

where $\frac{\partial q^*}{\partial \beta} > 0$, $\frac{\partial q^*}{\partial \tau} > 0$, $\frac{\partial n_t^i}{\partial \beta} < 0$, $\frac{\partial n_t^i}{\partial \tau} < 0$, and $\frac{\partial n_t^i}{\partial y} > 0$.

Proof. As follows from (6), the household's optimization with respect to $q_t^i \geq$ is given by the implicit function

$$G(q_t^i) \equiv \beta h'(q_t^i)(\tau + q_t^i) - h(q_t^i) \begin{cases} \leq 0 \quad if \quad q_t^i = 0 \\ \\ = 0 \quad if \quad q_t^i > 0, \end{cases}$$

$$ (13) $$

where, as follows from the strict concavity of $h(q_t^i)$.

$$\frac{\partial G(q_t^i)}{\partial q_t^i} = \beta h''(q_t^i)(\tau + q_t^i) - (1 - \beta)h'(q_t) < 0.$$

$$ (14) $$

Moreover, since $h(0) = 1$, $\beta \in (0, 1)$, and $h(q_t^i) - h'(q_t^i)q_t^i > 0$,

$$\lim_{q_t^i \to 0} G(q_t^i) = \tau\beta h'(q_t^i) - 1 \begin{cases} \leq 0 \quad if \quad h'(0) \leq 1/\beta\tau \\ > 0 \quad if \quad h'(0) > 1/\beta\tau; \end{cases}$$

$$ (15) $$

$$\lim_{q_t^i \to \infty} G(q_t^i) < 0.$$

Hence, it follows from the *Intermediate Value Theorem* that, if $h'(0) > 1/\beta\tau$, there exists a unique, interior value of $q_t^i = q^*$ such that $h(q^*) = \beta h'(q^*)(\tau + q^*)$ and a boundary solution $q_t^i = 0$ if $h'(0) \leq 1/\beta\tau$. The level of fertility, n_t^i, is given by (5). □

---

[15]This assumption permits the analysis to focus on the effect of fecundity of the founder generation on long-run reproductive success. Moreover, this is a good approximation for the actual fertility if the offspring generation is sufficiently large, as is the case in our sample where the average fertility rate in nearly 10. In particular, if there is a continuum of children and the realization at the individual level is drawn from a normal distribution with a mean at the desirable level of fertility, the law of large numbers will assure that on average the desired level of fertility is realized in the offspring generation. Clearly, a distribution of fertility may affect the level of average human capital but this is a second-order effect.

The interior unconstrained optimal level of investment in human capital, q^*, reflects the costs and the benefits of this investment and is, therefore, independent of the parental income. Hence, as long as parental income is constant across generations and there are no changes in the cost and the benefits associated with this investment, this level remains constant across generations. In particular, as follow from Lemma 1, the unconstrained level of investment in human capital of a member i of generation t is independent of the individual level of income for all $t \geq 1$.[16]

Hence, the level of human capital, h_t^i, of each member i of generation $t \geq 2$ is equal within and across generations:

$$h_t^i = h(q_{t-1}^i) = h(q^*), \tag{16}$$

and the level of income of each member i of generation $t \geq 2$ is equal within and across generations:

$$y_t^i = wh(q_{t-1}^i) = wh(q^*) = y^*, \quad \text{for all } t \geq 2.$$

Corollary 2.

- *The fertility level of a member i of generation 1 is*

$$n_1^i = \frac{\gamma y_1^i}{\tau + q^*} = \frac{\gamma wh(q_0^i)}{\tau + q^*} \equiv n_1(n_0^i),$$

  where $q_0^i = \frac{\gamma y_0}{n_0^i} - \tau$.

- *The fertility level of member i of generation $t \geq 2$ is*

$$n_t^i = \frac{\gamma y_t^i}{\tau + q^*} = \frac{\gamma y^*}{\tau + q^*} = n^*,$$

  where $y^* = wh(q^*)$.

As established in Appendix A, if parental human capital has a direct positive spillover on the level of the human capital of the child, the level of investment in human capital will increase monotonically and will gradually converge to a steady-state, $\bar{q}$, as long as the complementarity between parental investment in human capital and the spillover exceeds the direct effect of the spillover.

2.5 Long-run Reproductive Success

The long-run reproductive success of a member i of generation 0 is affected by the the initial fertility draw, n_0^i, and the implied level of investment in the quality of each child, q_0^i. In particular, the

---

[16]Since generation 0 faces a constrained optimization, the level of investment in human capital of this generation is affected by income, as established in (10).

number of descendants that individual i will have in generation T, D_T^i, is

$$D_T^i = n_0^i n_1^i n_2^i n_3^i ... n_T^i \equiv \prod_{j=0}^{T} n_j^i, \tag{17}$$

where n_j^i is the number of children that descendants of member of generation i are expected to have in period $j = 1, 2, \ldots, T$. Thus, as follows from Corollary 1,

$$D_T^i = (n^*)^{T-2} n_1^i n_0^i. \tag{18}$$

Nature will give an evolutionary advantage to a member i of generation 0 whose level of fecundity maximizes long-run reproductive success, D_T^i. Hence, the level of fecundity, n_0^i, that would be favored by the forces of natural selection is such that

$$n_0^i = \arg\max D_T^i \tag{19}$$

$$s.t. \quad n_0^i \in \left[0, \left[n_t^i\right]^{\max}\right]. \tag{20}$$

As follows from (18),

$$\frac{\partial D_T^i}{\partial n_0^i} = (n^*)^{T-2} \left[n_1^i + n_0^i \frac{\partial n_1^i}{\partial n_0^i}\right] = (n^*)^{T-2} n_1^i [1 - \eta_{n_1, n_0}], \tag{21}$$

where the elasticity of fertility in period 1 with respect to fertility in period 0, η_{n_1, n_0}, is

$$\eta_{n_1, n_0} \equiv -\frac{\partial n_1^i}{\partial n_0^i} \frac{n_0^i}{n_1^i} = \frac{(\tau + q_0^i) h'\left[q_0^i\right]}{h\left[q_0^i\right]}. \tag{22}$$

Moreover, given the properties of $h\left[q_0^i\right]$ imposed in (4) and the negativity of $\partial q_0^i / \partial n_0^i$ established in (10),

$$\frac{\partial \eta_{n_1, n_0}}{\partial n_0^i} = \frac{\left((\tau + q_0^i) h''\left[q_0^i\right] + h'\left[q_0^i\right]\right) h\left[q_0^i\right] - (\tau + q_0^i)\left(h'\left[q_0^i\right]\right)^2}{\left(h\left[q_0^i\right]\right)^2} \frac{\partial q_0^i}{\partial n_0^i} > 0. \tag{23}$$

Hence, as follows from (21), $\partial^2 D_T^i / \partial \left[n_0^i\right]^2 < 0$, and D_T^i is strictly concave in n_0^i.

The optimal level of fecundity in period 0, n_0^*, is therefore[17]

$$n_0^* \in \left(0, \left[n_0^i\right]^{\max}\right) \iff \frac{\partial D_T^i}{\partial n_0^i}(n_0^*) = 0 \iff \eta_{n_1, n_0}(n_0^*) = 1;$$

$$\tag{24}$$

$$n_0^* = [n_0]^{\max} \iff \frac{\partial D_T^i}{\partial n_0^i}\left([n_0]^{\max}\right) \geq 0 \iff \eta_{n_1, n_0}\left([n_0]^{\max}\right) \leq 1.$$

---

[17]As established in Appendix A, if parental human capital has a direct positive spillover on the level of the human capital of the child, the optimization would depend on the effect of the fecundity of the first generation on all subsequent generation, rather than the next generation only, but the qualitative results would remain intact.

Hence, if over some of the feasible range of fertility in period 0, $\left[0, [n_0]^{\max}\right)$, the elasticity of fertility in period 1 with respect to fertility in period 0 is greater than 1 in absolute value (i.e., if a one percent increase in fertility in period 0 reduces the quality, and therefore the fertility, of the descendant in period 1 by more than one percent), an intermediate level of fecundity $n_0^* \in \left(0, [n_0]^{\max}\right)$ will maximize long-run reproductive success. However, if the elasticity is smaller than 1 for the largest feasible level of fertility, $[n_0]^{\max}$, the level of fecundity that maximizes reproductive success will be the highest feasible one, $[n_0]^{\max}$.

In particular, it follows from (22) and (24) that the highest level of fecundity will maximize long-run reproductive success if and only if $\eta_{n_1, n_0}\left([n_0]^{\max}\right) \leq 1$, which holds if and only if

$$h'[\tilde{q}] \leq \frac{h[\tilde{q}]}{(\tau + \tilde{q})}, \tag{25}$$

where $\tilde{q} = q(([n_0]^{\max})$.

Noting that $q_0^i = 0$ when $n_0^i = [n_0]^{\max}$ and, therefore, $h[\tilde{q}] = h(0) = 1$, it follows that the highest level of fecundity will maximize long-run reproductive success if and only if

$$h'[0] \leq \frac{1}{\tau}. \tag{26}$$

Proposition 3. *The level of fecundity of the founding generation that maximizes reproductive success is*

- *The maximal level,*
$$[n_0]^{\max} = \frac{\gamma y_0}{\tau},$$

if the elasticity of fertility in period 1 with respect to fertility in period 0 evaluated at the highest feasible fertility, $\eta_{n_1, n_0}\left([n_0]^{\max}\right)$, is weakly less than 1 in absolute value, i.e., if

$$h'[0] \leq \frac{1}{\tau}.$$

- *An intermediate level,*
$$n_0^* \in \left(0, \left[n_0^i\right]^{\max}\right),$$

if the elasticity of fertility in period 1 with respect to fertility in period 0 evaluated at the highest feasible fertility, $\eta_{n_1, n_0}\left([n_0]^{\max}\right)$, is greater than 1 in absolute value, i.e., if

$$h'[0] > \frac{1}{\tau}.$$

Proof. The proposition is a corollary of (24) and (26). $\qquad\square$

2.6 Theoretical Insights

The theory establishes that, in contrast to a prevailing perception, the presence of a static trade-off between the quantity and quality of offspring is merely a necessary but not a sufficient condition for the presence of a negative effect of fecundity on reproductive success in the long run. In particular, it demonstrates that individuals with the highest level of fecundity could have had the largest reproductive success if, for any feasible number of children, an additional child would have contributed directly to the number of offspring in the long-run more than the adverse indirect effect of this additional child on the quality of his siblings and hence their long-run reproductive success.

The theory suggests that if investment in child quality is sufficiently productive in terms of its effect on the potential income and thus the reproductive success of the child (i.e., if the elasticity of fertility of an offspring with respect to the fertility of the founding parent, evaluated at the highest feasible level of fertility) is greater than one in absolute value, then an intermediate level of fecundity generates the highest reproductive success and will therefore be favored by natural selection.

3 Empirical Strategy

The analysis focuses on an important heritable determinant of fertility (i.e., the actual number of children born) that could be shaped by natural selection, namely fecundity (i.e., the capacity to reproduce). In order to investigate the hypothesis that an intermediate level of fecundity maximizes long-run reproductive success, the research isolates random variation in the time interval between the date of first marriage and the first birth (TFB) to infer the general effect of TFB, and therefore also fecundity, on long-run reproductive success.

Two major obstacles affect the identification of the effect of fecundity on long-run reproductive success. First, omitted correlates of the quantity of children may also be correlated with their quality, obscuring the effect of the number of children on long-run reproductive success. For instance, if variations in resources across individuals (e.g., income) enable some parents to produce more children as well as higher quality children, failing to account for the effect of resources will obscure the effect of child quantity on long-run reproductive success. In particular, an observed monotonically positive relationship between the number of children and that of grandchildren may be interpreted mistakenly as indicative of a lack of an optimal level of quantity beyond which fitness diminishes.

Second, reverse causality from the quality of children to their aggregate quantity may obscure the presence or the absence of an optimal level of fecundity beyond which fitness diminishes. For instance, the adverse effect of low child quality on the child survival rate may contribute to the total number of children born (via the child replacement channel), generating a downward bias in the correlation between the long-run reproductive success and the quantity of offspring that has no bearing on the presence or the absence of an optimal level of quantity beyond which fitness diminishes.

This research mitigates these major hurdles by focusing on the effect of fecundity, rather than fertility, on reproductive success. Furthermore, it designs an empirical strategy that exploits the inherent uncertainty in the process of human reproduction to identify the effect of fecundity on reproductive success. In particular, in light of the social norm observed in pre-industrial Quebec, in which marriage marked the intention to conceive, the research exploits variation in the random component of the time interval between the date of first marriage and the first birth to capture the effect of fecundity on fitness.

As depicted in Figure 1, a marriage over this period signaled a deliberate attempt to conceive. A sharp spike in birth rates occurs starting in the 35th week after marriage and nearly a third of births occurs within the 36–44 weeks time interval.[18] Furthermore, premarital conception is insignificant, reflecting possibly an adherence to the social and religious norms existing at the time. In particular, only 7.9 percent of the births over this period occurred within 35 weeks of marriage, and the incidence of premature births suggests that even this small fraction overstates the share of babies conceived prior to marriage.[19] Furthermore, the existence of births reported within 35 weeks of marriage reassuringly rules out the possibility that social taboos prevented the recording of pre-nuptially conceived children.

Since fecundity reflects genetic and socio-environmental factors, the time interval between the date of first marriage and the first birth (TFB) is affected by genetic predisposition, socio-environmental conditions, as well as the realization of random elements that affect conception. Accounting for a range of genetic and socio-environmental confounding factors that may affect the time to first birth, reproductive success, and the quality of offspring, the study attempts to isolate the effect of the random variations in TFB across individuals. In particular, genetic, as well as cultural and socio-economic factors that may affect fecundity are accounted for by the inclusion of Maternal Founder fixed effects. Namely, the effect of fecundity on reproductive success is identified based on variations in reproductive success among siblings, capturing the similarities in the genetic predisposition of these genetically linked individuals, as well as their cultural and socio-economic proximity.

Additional confounding variations between siblings are accounted for by the inclusion of dummies for their marriage age, birth year, gender, and literacy. Furthermore, additional control variables include the geographic location at birth and death, birth order, month of birth, month of birth of the firstborn, number of marriages, and immigration and emigration status of the spouse, for each head of lineage.

---

[18]Full term babies are born upon 38 weeks of gestation. Nevertheless, pregnancy is considered at term if the gestation period is within the interval 36-40 weeks. However, since the marriage age may coincide with the ovulation period and may occur at most 4 weeks before it, time to first birth within the interval 36–44 weeks would correspond to babies born at term.

[19]In the sample of all 59,238 mothers, 3.2 percent of births occurred prior to the marriage date, 5.5 percent of the births occurred after two years and 38 weeks of marriage (i.e., two years after first conception), and 1.6 percent of births occurred within one week of marriage (reflecting possibly a tendency of mothers who gave birth before marriage to baptize their firstborn at or shortly after their wedding date).
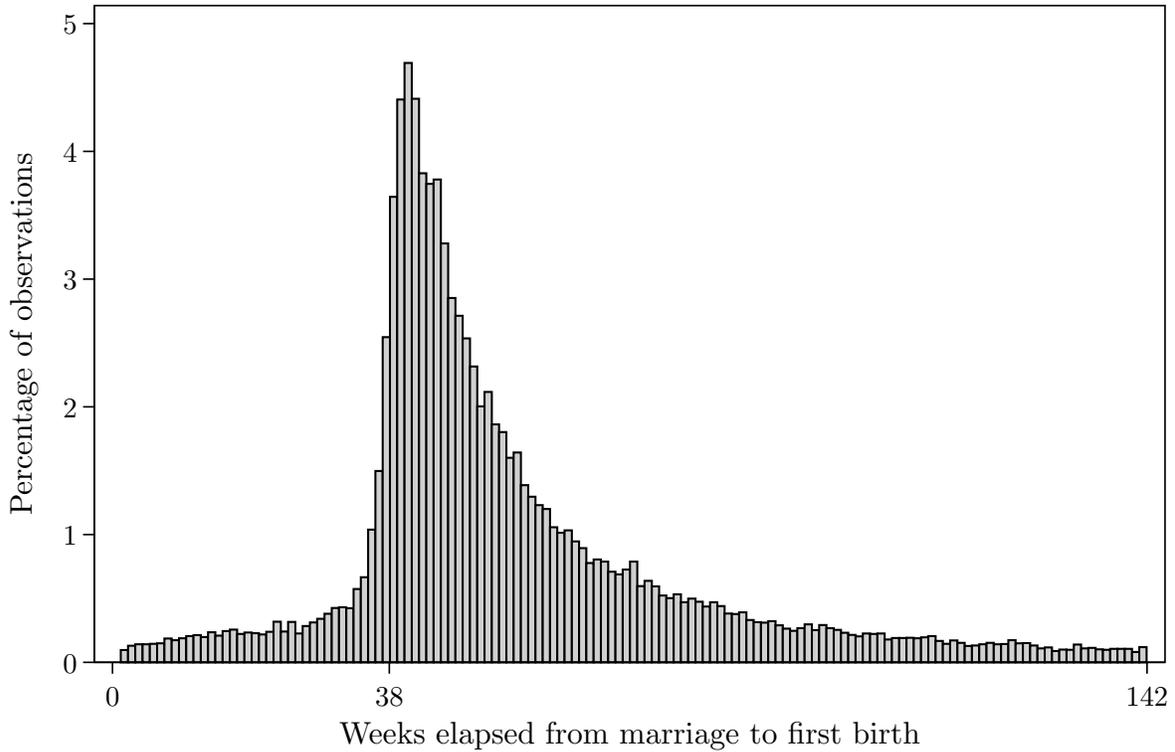
Figure 1: The histogram depicts the durations (in weeks) from first marriages to first births of 53,154 mothers in Quebec between the 16th and the end of the 18th century who gave birth between the 1st and 142nd week of their marriage date.

Finally, to assure the external validity of the empirical analysis for significant evolutionary patterns in fecundity during most of human existence, the research focuses on this evolutionary process among the founder population of Quebec during the pre-industrial, pre-demographic transition era – a high fertility environment that had naturally led to rapid evolutionary changes in the composition of the population. The conditions that were faced by the founder population of Quebec during that time period capture the environment that anatomically modern humans confronted during their migration out of Africa, as they settled new territories where the carrying capacity of the new environment was an order of magnitude greater than the size of the founder population. Thus, the evolution in fecundity in Quebec during this time period is likely to shed light on the evolutionary forces in the high fertility segments of the Malthusian epoch in which natural selection could have made a significant impact on the composition of the population (e.g., the Neolithic transition and the formation of sedentary agricultural communities).

4 Data and Main Variables

This section sets the stage for the empirical examination of the hypothesis that higher fecundity in the pre-demographic transition era was associated with a larger number of children, while an intermediate level of fecundity maximized long-run reproductive success.

4.1 Data

The data is based on the demographic history of Quebec, using the reconstructed genealogy of the entire parish registers of Quebec, covering 471,412 individuals born between 1572 and 1799, spanning the French settlement and colonization of Quebec.[20] Nearly all individuals in the data (i.e., more than 99.96%) were born after the founding of Quebec City in 1608. The data covers all parishes of Quebec, and thus in light of the negligible inter-provincial migration over this period, intra-provincial migration does not prevent the tracking of reproductive success of individuals over several generations. Indeed, more than 94% of these individuals were born and died in Quebec.

The analysis focuses on the reproductive success of individuals who were born in Quebec prior to the end of 1685, died in the province, and had at least one great-great grandchild.[21] The focus on the reproductive success of heads of lineages born prior to the end of 1685 permits tracing of most descendants of these individuals in the subsequent four generations, while accounting for the Maternal Founder fixed effects. Furthermore, since the early period of the data relates to a very small founder population (i.e., 19 births satisfying the sample restrictions in the period 1620–1639, 68 births in 1640–1649, and 287 births in 1650–1659), the focus on reproductive success of heads of lineages born prior to the end of 1685 allows the inclusion of years in which births rates were sufficiently large to permit a precise estimate of the regression coefficients while accounting for a large set of fixed effects. The use of alternative time intervals would not affect the qualitative results. In particular, as established in Tables 3 and 4, the main results are qualitatively unchanged if the analysis focuses on the reproductive success of heads of lineages born: (i) in 1660–1685 or (ii) over the entire sample period.

Furthermore, immigrants among heads of lineages are excluded from the sample for two reasons. First they may differ systematically from natives, reflecting the circumstances that led to their decision to immigrate as well as the effects of immigration on their socio-economic status and thus their reproductive success. Second, reproduction of immigrants prior to their arrival to Quebec is

---

[20]The data is provided by *Le Programme de recherche en démographie historique* at the University of Montreal.

[21]All children of these heads of lineages, virtually all grandchildren, and most of the descendants in the third and the fourth generations are observed. Furthermore, the fraction of extinct lineages from generation to generation in the regression sample is rather constant, implying that the end date of the entire data need not generate concerns regarding the effect of potential censoring. Indeed, as can be derived from the numbers of observations in columns 1–4 in Table C.3, of the 4,240 observations satisfying all the sample restrictions but also includes heads of lineages with no observed great-great-grandchildren, 5.6% of lineages produced no observed grandchild, and of the remaining lineages, 1.7% produced no observed great-grandchild, and of these remaining lineages 3.4% produced no observed great-great-grandchild. Furthermore, as established in Table C.3, the results are robust to the inclusion of extinct lineages. Potential systematic associations between the birth year of the head of the lineage and the number of unobserved descendants in the third and fourth generations are accounted for by the inclusion dummies for the birth year of the heads of lineages.

unknown. Similarly, emigrants are excluded as well since they may possess unique attributes and their subsequent marriages and births outside of Quebec are not observed.

The study focuses on individuals whose time from first marriage to first birth is at least 38 weeks.[22] Moreover, the study follow the convention in the literature (e.g. Milot et al. (2011)) and restricts the sample to individuals whose time to first conception is less than 2 years (i.e., TFB not shorter than 38 weeks and not longer than 2 years and 38 weeks), excluding 10.9% of this subset.[23] This further restriction is designed to mitigate the effect of extreme values of time to first birth which may reflect measurement errors or underlying biological conditions that may directly affect long-run reproductive success. Thus the analysis focuses on the reproductive success of 3,798 heads of lineages in the pre-1685 period (consisting of individuals born in the period 1620–1685) that satisfy the entire sample restrictions.[24]

The sample includes men and women for two reasons. First, due to the occurrence of remarriages, TFB can differ between men and women, whereby the inclusion of both provides a complete picture of the variation in fertility in households due to variation in TFB.[25] Second, by including men and women, the analysis accounts for the differential effects of paternal and maternal characteristics such as birth year, marriage age and literacy. It should be noted that the analysis accounts for the correlation in reproductive success between heads of lineages that produced the same firstborn and thus shared their TFB by clustering the standard errors on the level of the firstborn.[26]

4.2 Descriptive Statistics and Basic Relationships

This subsection explores demographic patterns in Quebec between the 16th and the 18th century that could shed light on the empirical exploration of the effect of fecundity of long-run reproductive success. In particular, it examines: the distribution of the number of children and grandchildren, the distribution of the age at first marriage across individuals, and the relationship between the time to first birth and the number of children.

The high-fertility regime in Quebec over this historical period is evident from Figure 3.[27] The histogram of the fraction of individuals with a given number of children takes on a predominantly mound-shaped form. The density of individuals with a given number of children increases at larger number of children, for less than 10 children, and drops significantly thereafter with each successive

---

[22]For the 8.2 percent of the individuals in the sample of non-migrating heads of lineages whose firstborn's date of birth is unknown, it is estimated to be one week prior to the date of baptism.

[23]The use of alternative cut-offs would not affect the qualitative results. In particular, focusing on individuals whose time from first marriage to first birth is at least 36 weeks, the hump-shaped relationship remains highly significant. Likewise, focusing on individuals whose time from first marriage to first birth is at least 40 weeks, the hump-shaped relationship remains significant. Furthermore, exploiting alternative estimation methods that do not rely on a quadratic specification reveals qualitatively similar patterns in the presence of observations with TFB exceeding 2 years and 38 weeks.

[24]The summary statistics for this sample can be found in Table C.1.

[25]Table C.4 establish that the results are robust to accounting for remarriages.

[26]The results are robust to the use of the sample of men or women only. In particular, Table C.5 establish that the baseline results are robust to the focus on women only.

[27]Figures A.1 and A.2 depicts the data for females and males separately.

birth, with less than 0.5% of individuals (predominantly male) with more than 20 children. In particular, the average fertility of heads of lineages in the main sample, consisting of the early settlers of Quebec, is 9.70 children (Table C.1). However, as a consequence of a high mortality rate among offspring of early settlers, heads of lineages produce on average 4.63 children surviving to age 40 (Table C.1).

Figure 4 depicts the density of individuals with a given number of grandchildren.[28] Interestingly, despite the predominantly mound-shaped distribution of the number of children, the distribution of the number of grandchildren is J-shaped, reflecting the compound nature of long run reproductive success and that the potential role of child quality in generating dispersion in the number of grandchildren. The histograms of the density of individuals with a give number of great-grandchildren and great-great-grandchildren (Figure B.5) display a similar pattern and establish that the range in reproductive success increases dramatically as the number of generations over which it is recorded increases.

The age distribution at first marriage of individuals during this time period is depicted in Figure 2.[29] The mean age at first marriage of 23.85 over this period is substantially lower than the age of 30.1 observed in Canada in the year 2008. Since marriage over this period marked the intention to conceive, this early marriage age combined with a late age at last birth (Figure B.8), contributed to a high-fertility regime in historical Quebec.

Figure B.9 and B.10 depicts the conditional and unconditional relationship between TFB and the number of children nonparametrically, dividing TFB into 15 equal-size groups (15-quintiles) and plotting the mean value of the number of children in each bin.[30] This binned scatter plot provides a nonparametric representation of the conditional expectation function without showing the underlying variance in the individual-level data. The regression coefficient and standard error, corresponding to the overlaid regression line (and reported in the figure caption), is estimated on the micro-data (not the binned averages), with standard errors clustered on the level of the firstborn.

4.3   Main Variables

4.3.1   Dependent and Independent Variables

In the main analysis that explores the effect of the time interval, measured in years, between the first marriage date of the head of a lineage and the birth date of the individual's first child (TFB) on reproductive success, the dependent variable is the number of offspring of each head of lineage in the subsequent four generations (i.e., children, grandchildren, great-grandchildren and great-great-grandchildren). In the additional analysis that examines the mechanism through which TFB affects long-run reproductive success, the dependent variables are the fraction of ever-married children among children observed to have survived to age 40, the average marriage age of children

---

[28]Figures A.3 and A.4 depicts the data for females and males separately.

[29]Figures A.6 and A.7 depicts the distribution for females and males separately.

[30]This flexible nonparametric method is also employed, and further described, by Chetty et al. (2014), investigating relationships found in large datasets.
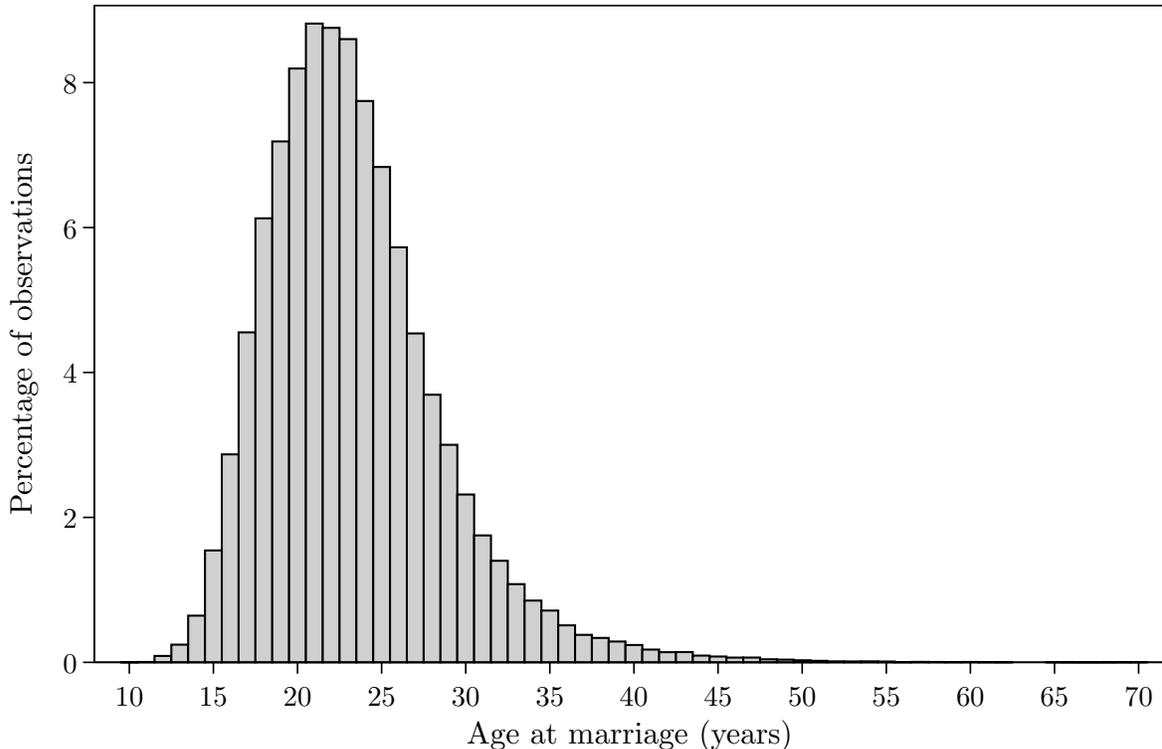
Figure 2: The histogram depicts age at first marriage (in years) of 107,999 non-migrant, married individuals in Quebec between the 16th and the end of the 18th century.

with observed marriage age, and the fraction of literate children among children with observed literacy status (recorded at marriage). The independent variable throughout the analysis is the time interval between the first marriage date of the head of a lineage and the birth date of the individual's first child (TFB).[31]

4.3.2 Maternal Founder Fixed Effects

The effect of fecundity on reproductive success may be affected by variation in genetic predisposition among genetically distinct individuals, as well as variation in cultural and socio-economic background. Hence, similarities in the genetic, cultural, and socio-economic characteristics across siblings within each household are exploited to isolate the effect of random variation in TFB on reproductive success. Accounting for Maternal Founder fixed effects, as well as the confounding

---

[31]In couples where neither spouse remarried, TFB is identical for the husband and the wife. Nevertheless, given that the frequency of remarriage over this period is substantial, reflecting in part a considerable mortality rate, TFB and the number of offspring of each spouse often differ. The correlation in reproductive success between parents sharing the same firstborn, and therefore the same TFB, is accounted for by clustering the standard errors for heads of lineages sharing the same firstborn.

Figure 3: The histogram depicts the number of children of 48,386 non-migrating individuals in Quebec between the 16th and the end of the 18th century with known year of birth and at least one observed grandchild.

factors underlined below, the analysis explores the effect of random variation in TFB on long-run reproductive success within lineages headed by siblings, as opposed to across all heads of lineages.[32]

4.3.3 Control Variables

The analysis accounts for the confounding effects of the marriage age, birth year, literacy, and the maternal identity, for each head of lineage. Furthermore, additional control variables include the geographic location at birth and death, birth order, month of birth, month of birth of the firstborn, number of marriages, and immigration and emigration status of the spouse, for each head of lineage.

The confounding associations between the marriage age of heads of lineages and their affluence, fecundity, and reproduction is accounted for by the inclusion of dummy variables indicating the

---

[32]Accounting for the family characteristics of the spouse of the head of dynasties would restrict the sample for siblings who married siblings from a different household. Accounting for these fixed effects would reduce the sample size by more than 90%, and more importantly, would introduce biases associated with the unique characteristics of these families and their matching technology. Reassuringly however, the Maternal Founder fixed effects do accounts for the characteristics of the spouses of heads of households in the likely scenario of non-random (assortative) mating. Moreover, since the regression analysis is nearly unaffected by the introduction of Maternal Founder fixed effects, the potential effect of random mating does not compromise the analysis and further control for fixed factors in the family of the spouse of the head of lineage are unlikely to alter the results.
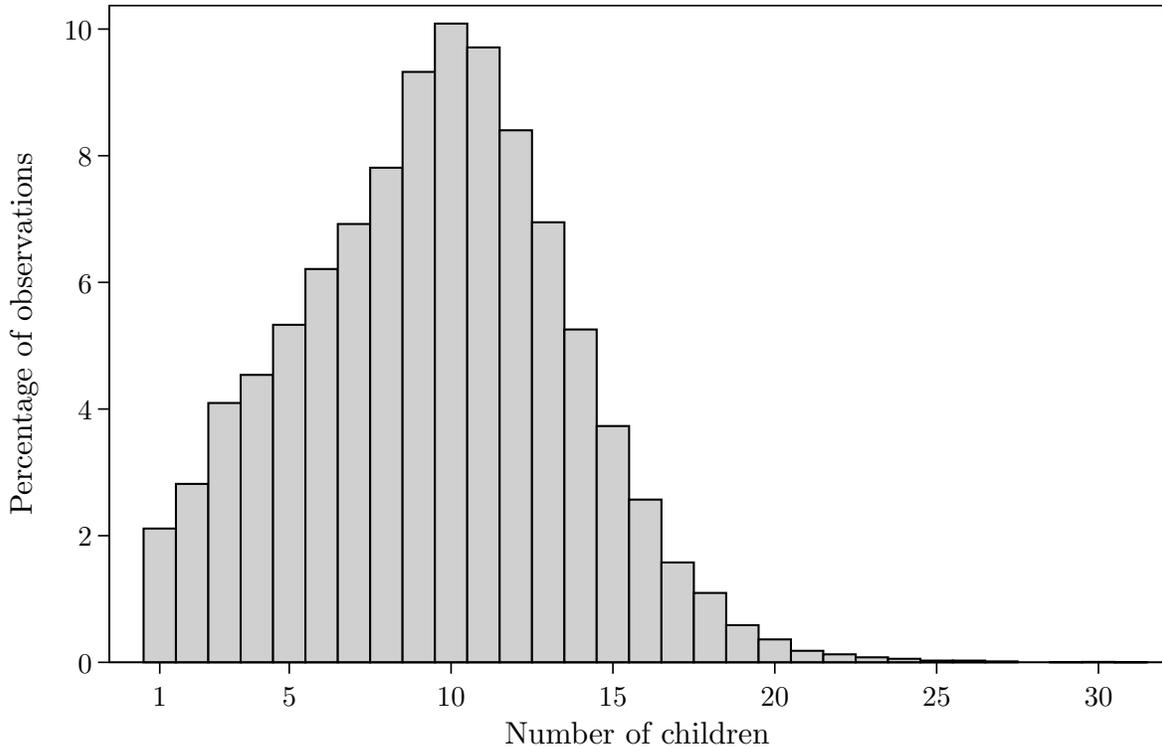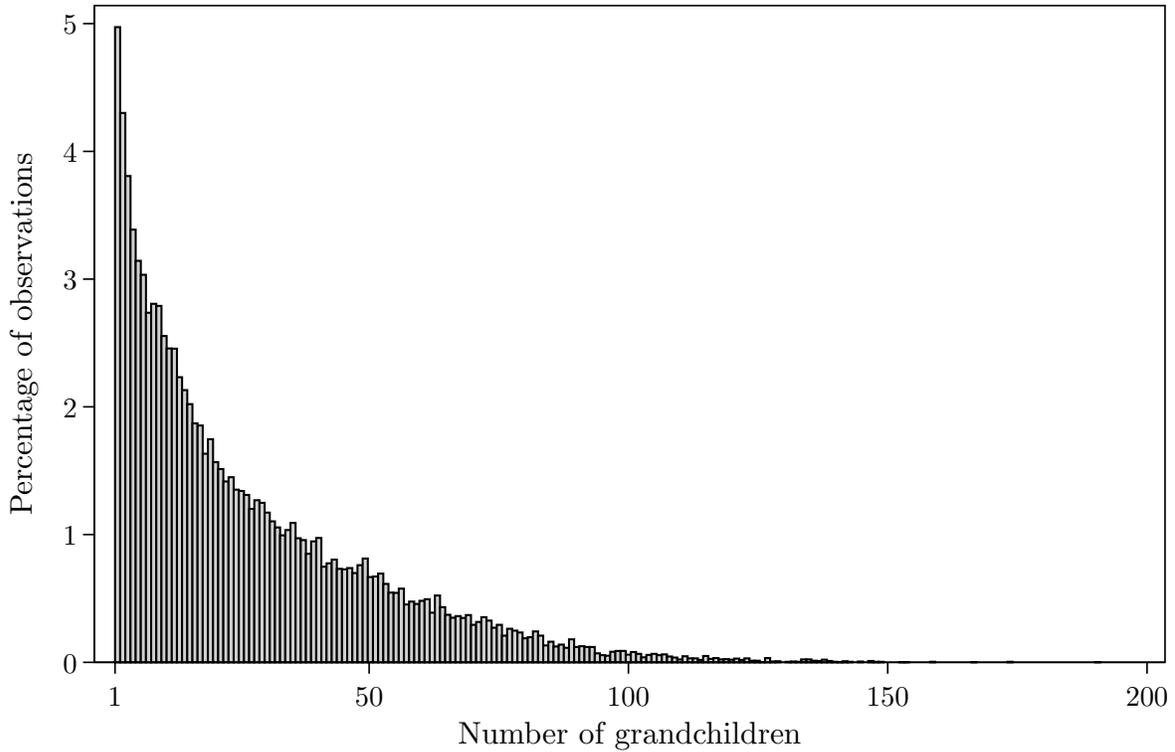
Figure 4: The histogram depicts the number of grandchildren of 48,386 individuals in Quebec between the 16th and the end of the 18th century with known year of birth and at least one observed grandchild.

marriage age of heads of lineages. The marriage age is associated with reproductive success through three channels. First, fecundity is affected by age (Baird et al., 2005). Second, the marriage age affects the length of the reproductive period of the couple. Third, in the pre-demographic transition era that corresponds to our sample, the marriage age was inversely related to the affluence of individuals, and marriage age and its potential association with affluence could have had an independent effect on long-run reproductive success.

The time-path of socioeconomic and demographic factors may differentially affect fecundity and reproductive success across cohorts of heads of lineages. In particular, the affluence, fecundity, and reproductive success of heads of lineages may be affected by the socioeconomic and demographic conditions during their lifetime, as partly captured by their birth year. These confounding factors are accounted for by the inclusion of dummy variables indicating the birth year of heads of lineages.[33]

The human capital attainment that may reflect the socioeconomic status of heads of lineages may affect their TFB and reproductive success. This confounding factor is partly accounted for by

---

[33]In addition, the inclusion of birth year dummies mitigates the potential systematic associations between the birth year and the number of unobserved descendants after three or four generations.

the inclusion of the literacy status of heads of lineages, inferred from the existence of a signature (rather than a mark) on the marriage certificate. Additional confounding variations between heads of lineages are accounted for by the inclusion of dummies capturing gender, geographic location at birth and death, birth order, month of birth, month of birth of the firstborn, number of marriages, and immigration and emigration status of the spouse.

Finally, variations in socioeconomic and physiological factors across heads of lineages may generate variation in the length of the reproductive period that may obscure the effect of TFB on reproduction. In particular, conditional on the marriage age, the age at last delivery determines the length of the reproductive period. Hence, to account for the potential effect of the stoppage age, the analysis is shown to be robust to the inclusion of dummy variables indicating the stoppage age of heads of lineages, in addition to the marriage age, are introduced to account for this confounding factors, permitting the study to capture the effects of TFB on fertility, for a given length or reproductive period.[34]

5   Empirical Analysis

This section examines the proposed hypothesis that higher fecundity in the pre-demographic transition era was associated with a larger number of children, while an intermediate level of fecundity maximized long-run reproductive success. The examination proceeds in two stages. The empirical regularities that emerge from the data are examined initially semi-parametrically, using cubic spline regression models, followed by an examination using OLS regressions models.

5.1   Semi-Parametric Analysis

The proposed hypothesis is confirmed initially using restricted cubic spline regression models. It establishes that while a higher fecundity is associated with a larger number of children, an intermediate level of fecundity maximizes long-run reproductive success.[35] The effect of TFB of heads of lineages on their number of descendants in the subsequent four generations, accounting for the birth year and the marriage and stoppage age of heads of lineages, with 95% confidence intervals based on standard errors clustered on heads of lineages sharing the same firstborn is depicted in Figure 5.

The correlation in reproductive success between parents sharing the same firstborn, and therefore the same TFB, is accounted for by clustering the standard errors for heads of lineages sharing the same firstborn.

In line with the proposed hypothesis, panel A shows an approximately linear negative partial effect of TFB on the number of children, confirming the conventional presumption that *ceteris paribus*, a short time to first birth in the pre-demographic transition era increased the number

---

[34]As depicted in Figure B.8, the stoppage age over this period marked the decline in fecundity and onset of sterility associated with age-related infertility and onset of menopause, with a modal stoppage age of 41.

[35]Figure B.11–B.14 establish that the use of binned scatterplots results in a similar qualitative pattern. Furthermore, the use of multivariate LOWESS also results in a similar qualitative pattern.
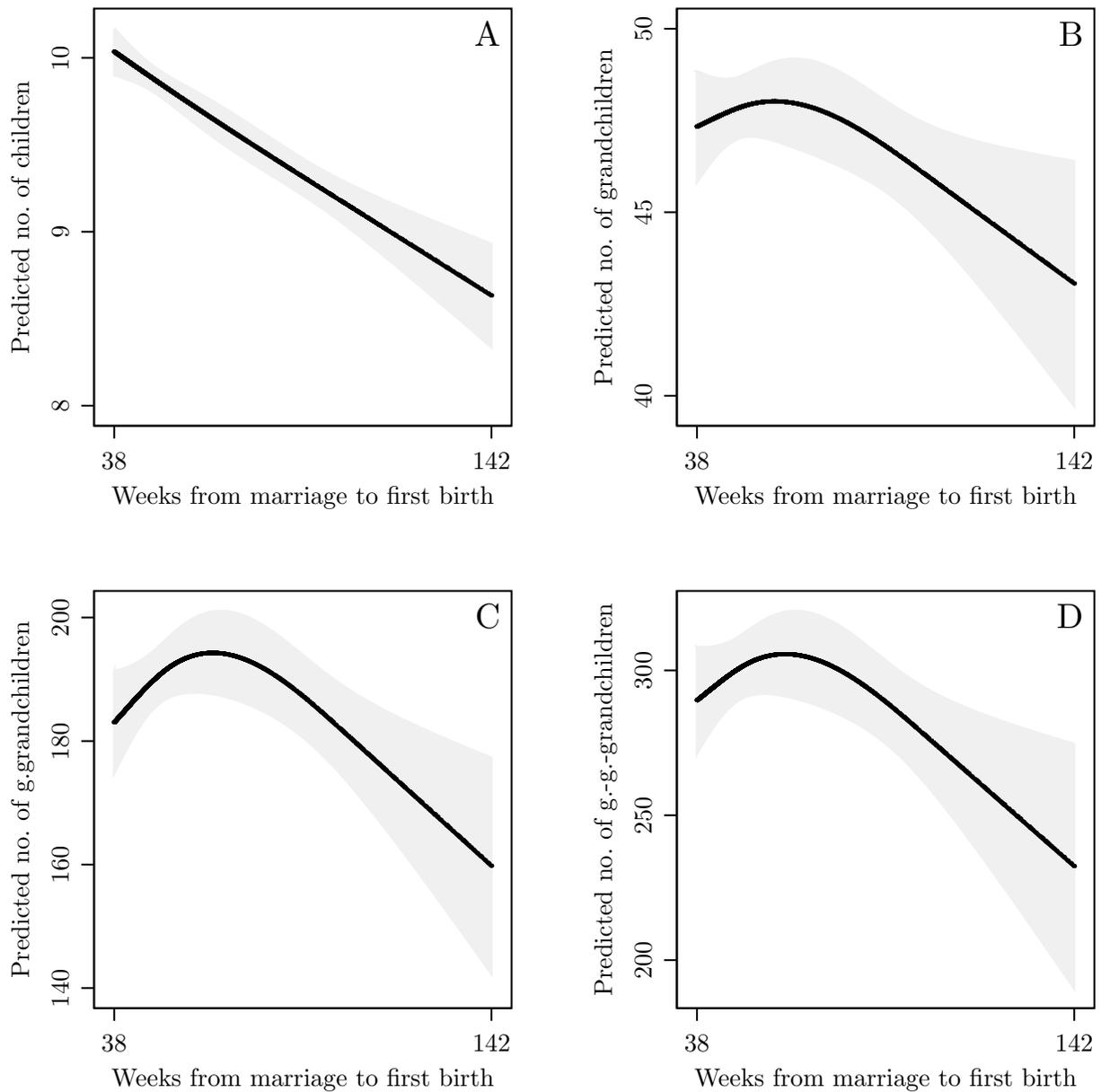
Figure 5: Predicted number of descendants with 95% confidence interval as a function of TFB based on restricted cubic splines with three knots, for 3,798 heads of lineages. Confidence intervals are based on standard errors clustered on heads of lineages sharing the same firstborn. Dummies indicating birth year and marriage and stoppage age are included in the underlying regressions. (A) Number of children. (B) Number of grandchildren. (C) Number of great-grandchildren. (D) Number of great-great-grandchildren.

of children. In contrast, as hypothesized, an intermediate TFB maximizes long-run reproductive success. In particular, panel B depicts a hump-shaped relation between TFB of heads of the lineages and their number of grandchildren. The TFB of heads of lineages that maximizes the number of grandchildren is associated with 48 grandchildren. Panels C and D reveal a similar a hump-shaped relation between TFB of the heads of the lineages and their great-grandchildren and great-great-grandchildren. The TFB of heads of lineages that maximizes the number of great-grandchildren is associated with 194 great-grandchildren, whereas the optimal TFB of heads of lineages for reproductive success in the 4rd generation (62 weeks) is associated with 306 great-great-grandchildren.

Figure 5 shows that, in accordance with the proposed hypothesis, TFB of heads of lineages has a monotonically negative effect on the number of children and a hump-shaped effect on the number of grandchildren, great-grandchildren and great-great-grandchildren. Thus, heads of lineages with an intermediate level of TFB achieved the maximal number of descendants in the long run, despite having a smaller number of children relative to those with lower TFB.

5.2 Econometric Model

The negative relationship between TFB of heads of lineages and the number of children, as well as the hump-shaped relationship between TFB of heads of lineages and long-run reproductive success, is further assessed by estimating a series of regression models.

First, the effect of TFB of the head of lineage on the number of children born to the head of lineage is estimated using the OLS regression model:

$$\ln D_{i,1} = \beta_{0,1} + \beta_{1,1} TFB_i + Z_i \beta_{3,1} + \varepsilon_{i,1},$$

where $D_{i,1}$ is the number of children (i.e., offspring in generation 1) born to head of lineage i; TFB_i is the time from the first marriage to the first birth of the head of lineage i; Z_i is a vector of control variables capturing the characteristics of the head of lineage i; and $\varepsilon_{i,1}$ is an error term clustered at the level of heads of lineages sharing the same firstborn. The coefficient of interest is $\beta_{1,1}$ and it is predicted to be negative, i.e., TFB of heads of lineages negatively affects the number of children.

Second, the effect TFB of heads of lineages and long-run reproductive success is estimated using the OLS regression model:

$$\ln D_{i,t} = \beta_{0,t} + \beta_{1,t} TFB_i + \beta_{2,t} TFB_i^2 + Z_i \beta_{3,t} + \varepsilon_{i,t},$$

where $D_{i,t}$ is the number of descendants that the head of household i, has in the subsequent generations t, $t = 2, 3, 4$; TFB_i is the time from the first marriage to the first birth of the head of lineage i; Z_i is a vector of control variables capturing the characteristics of the head of lineage

i; and $\varepsilon_{i,t}$ is an error term clustered at the level of heads of lineages sharing the same firstborn.[36] The coefficients of interest are $\beta_{1,t}$ and $\beta_{2,t}$. The prediction is that $\beta_{1,t} > 0$ and $\beta_{2,t} < 0$, i.e., TFB has a hump-shaped effect on the number of grandchildren, great-grandchildren and great-great-grandchildren.

5.3 Estimation based on Variation across all Heads of Lineages

The baseline OLS estimates of the effect TFB of the head of lineage on reproductive success are presented in Tables 1 and 2, accounting for the marriage age and the birth year of heads of lineages.[37] The initial estimates in Table 1 are based on variation in TFB across all head of lineages, whereas those in Table 2 accounts for Maternal Founder fixed effects, and thus presents estimates based on variation in TFB within heads of lineages that are originated from the same mother. The correlation in reproductive success between parents sharing the same firstborn, and therefore the same TFB, is accounted for by clustering the standard errors for heads of lineages sharing the same firstborn.

Consistently with the first element of proposed hypothesis, and the pattern depicted in Figure 5, panel A, column 1 of Table 1 establishes a highly significant negative association between TFB (as measured in years) of heads of lineages and the number of children. An increase in the TFB by one year results in a reduction of 0.080 in the log number of children. In particular, an increase in TFB from 38 weeks to 1 year and 38 weeks would result in a reduction of approximately 0.77 children.[38]

The positive association of an intermediate level of TFB and long-run reproductive success is confirmed in columns 2–4, resembling the pattern depicted in Figure 5, panel C–D. Column 2 establishes the formation of a quadratic relationship between TFB of heads of lineages and reproductive success, as measured by the number of grandchildren, with the first-order effect of the quadratic expression being positive and the second-order effect of the quadratic expression being negative. Importantly, the effects are not statistically significant, indicating that a static quantity-quality trade-off is not an exclusive cause of the optimality of moderate fecundity observed in the long run. Indeed, column 3 establishes a significant quadratic relationship between TFB of heads of lineages and the number of great-grandchildren. In particular, the first-order effect of the quadratic expression is positive and significant at the 5% level and the second-order effect of the quadratic expression is negative and significant at the 5% level. Similarly, column 4 establishes a

---

[36]The results are robust to alternative regression methods suitable for count data. In particular, Table 5 and Table C.8 demonstrates that the results are robust to the use of a GLM model with a negative binomial distribution and a logarithmic link function.

[37]Nine of the 3,978 heads of lineages in the baseline sample were charactarized by singular dummies, i.e., they were the only individuals in the sample who were born in a certain year or married at a certain age. Furthermore, an additional five observations were charactarized by singular dummies in the regressions including stoppage age dummies. Thus, the identification of the trade-off is based on variation in the remaining 3,969 observations in the regressions excluding the stoppage age dummies and 3,964 observations in the regressions including the stoppage age dummies.

[38]Throughout the analysis, estimates on the original scale of numbers of descendants reported in the main text are corrected for re-transformation bias in accordance with Duan (1983).

Table 1: The association between the time to first birth (TFB) and the number of descendants for head of lineages born prior to the end of 1685

	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
TFB	-.080***	.137	.509**	.830***	-.081***	.091	.411*	.713**	-.097***	.019	.321	.632**	-.093***	.061	.367*	.664**
	(.024)	(.180)	(.232)	(.311)	(.024)	(.180)	(.230)	(.307)	(.023)	(.177)	(.226)	(.305)	(.011)	(.143)	(.204)	(.300)
TFB^2		-.054	-.174**	-.290***		-.043	-.148**	-.259***		-.026	-.128*	-.241**		-.039	-.141**	-.251***
		(.058)	(.075)	(.099)		(.058)	(.074)	(.098)		(.057)	(.073)	(.097)		(.046)	(.065)	(.095)
Literate					.019	.209***	.382***	.519***	.013	.201***	.373***	.510***	-.059***	.136***	.304***	.462***
					(.027)	(.037)	(.051)	(.070)	(.026)	(.036)	(.050)	(.069)	(.012)	(.030)	(.045)	(.069)
Male									.268***	.385***	.485***	.431***	-.018	.099***	.198***	.276***
									(.025)	(.032)	(.042)	(.053)	(.013)	(.027)	(.040)	(.055)
Stoppage age fixed effects	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Number of observations	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798
Adjusted R^2	.018	.013	.026	.301	.020	.024	.052	.319	.047	.055	.080	.328	.794	.425	.289	.364
Joint sign.-level of TFB & TFB^2	.001	.439	.055	.007	.001	.384	.084	.009	.000	.135	.050	.006	0	.041	.012	.003
Maximizing TFB		1.256	1.46	1.431		1.069	1.391	1.374			.357	1.255	1.309		.785	1.298	1.321
Lower limit of 90% CI		-	1.099	1.170		-	.575	.992			-	-3.120	.759		-	.433	.856
Upper limit of 90% CI		-	1.679	1.588		-	1.634	1.547			-	1.514	1.495		-	1.503	1.494

This table presents the results of a series of OLS regressions of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2 for heads of lineages born prior to the end of 1685. Birth year and marriage age dummies are included as controls. Furthermore, stoppage age dummies are included in columns 13–16. A dummy indicating unknown literacy is included in the regressions underlying column 5–16. Standard errors clustered at the level of the firstborn are reported in parentheses. * p <0.10, ** p <0.05, *** p <0.01.

24

highly significant quadratic relationship between TFB of heads of lineages and the number of great-great-grandchildren. The first and second-order effects are both significant at the one percent level, and jointly highly significant (p=0.007). Moreover, an additional test (not reported in the table) establishes a highly significant hump-shaped relationship under this specification (p=0.009).[39]

Columns 5–8 establish that the results are robust to the inclusion of parental literacy as a control variable. The highly significant negative association between TFB of heads of lineages and the number of children is maintained and the coefficient is rather stable (column 5). Furthermore, the quadratic relationship between TFB of heads of lineages and the number of descendants in the second, third, and fourth generations is stable, although somewhat less significant (column 6–8). Moreover, a test of the joint significance of the two coefficient estimates related to the fourth generation establishes that they are jointly highly significant ($p = 0.009$). Additionally, literacy is positively associated with long-run reproductive success (columns 6–8). As will become apparent in Table 6 and 7, literacy (and its potential association with a quality bias) is positively associated with the fraction of surviving children that got married, negatively associated with their average marriage age, and positively associated with the literacy of offspring, and is thus rewarding in the long run.

Furthermore, columns 9–12 establish that the results are robust to the inclusion of a control for gender. The highly significant negative association between TFB of heads of lineages and the number of children is maintained and the coefficient is rather stable (column 9). Moreover, a quadratic relationship between TFB of heads of lineages and the number of descendants, as well as the significance of this relationship, is stable in the second, third and fourth generations (column 11–12). In addition, a test of the joint significance of the two coefficient estimates related to the fourth generation establishes that they are jointly highly significant ($p = 0.006$).

Finally, columns 13–16 establish that the results are robust to the inclusion of a control for the stoppage age. The highly significant negative association between TFB of heads of lineages and the number of children is maintained and the coefficient is rather stable (column 13). Furthermore, the quadratic relationship between TFB of heads of lineages and the number of descendants in the 3rd and the 4th generations, as well as the significance of this relationship, is rather stable (column 14–16). In addition, a test of the joint significance of the two coefficient estimates related to the fourth generation establishes that they are jointly highly significant ($p = 0.003$).

It should be noted that while columns 3–4, 8–9, 11–12, and 15–16 establishes the presence of a long-run reproductive trade-off, columns 2, 6, 10, and 14 indicate that a static quantity-quality trade-off is not an exclusive cause of the optimality of moderate fecundity observed in the long run.

5.4 Estimation based on Variation within Head of Lineages traced to the same Maternal Founder

The effect of fecundity on reproductive success may be affected by variation in the genetic predisposition among genetically distinct heads of lineages, as well as their variation in cultural and

---

[39]See Lind and Mehlum (2010).

Table 2: The effect of the time to first birth (TFB) on the number of descendants for head of lineages born prior to the end of 1685 – accounting for Maternal Founder fixed effects

	Log number of descendants in:															
	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
TFB	-.052**	.167	.505**	.783***	-.053**	.170	.499**	.788***	-.062***	.140	.463**	.773***	-.077***	.208	.535***	.810***
	(.024)	(.163)	(.205)	(.264)	(.024)	(.163)	(.205)	(.264)	(.024)	(.162)	(.204)	(.264)	(.011)	(.130)	(.181)	(.258)
TFB^2		-.068	-.193***	-.310***		-.070	-.191***	-.313***		-.063	-.183***	-.309***		-.089**	-.210***	-.325***
		(.053)	(.067)	(.087)		(.053)	(.067)	(.087)		(.053)	(.067)	(.087)		(.042)	(.059)	(.084)
Literate					-.008	.060	.145***	.136**	-.006	.063	.148***	.138**	-.027*	.044	.125***	.109*
					(.031)	(.040)	(.051)	(.066)	(.030)	(.040)	(.051)	(.066)	(.014)	(.032)	(.046)	(.066)
Male									.220***	.254***	.299***	.131**	-.028*	.025	.085*	.036
									(.031)	(.039)	(.047)	(.060)	(.015)	(.031)	(.043)	(.063)
Stoppage age fixed effects	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Number of observations	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798
Adjusted R^2	.015	.016	.038	.306	.017	.016	.041	.307	.032	.029	.052	.307	.799	.442	.296	.355
Joint sign.-level of TFB & TFB^2	.031	.196	.002	.000	.026	.184	.003	.000	.010	.130	.002	.000	.002	.000	.000	.000
Maximizing TFB		1.224	1.307	1.261		1.223	1.304	1.260		1.113	1.263	1.249		1.163	1.272	1.247
Lower limit of 90% CI		-	.961	.999		-	.948	1.000		-	.827	.976		-.141	1.012	1.002
Upper limit of 90% CI		-	1.467	1.398		-	1.466	1.397		-	1.435	1.389		1.393	1.403	1.376

This table presents the results of a series of fixed-effects regressions of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2. All regressions account for Maternal Founder fixed effects. Birth year and marriage age dummies are included as controls. Furthermore, stoppage age dummies are included in columns 13–16. A dummy indicating unknown literacy is included in the regressions underlying column 5–16. Standard errors clustered at the level of the firstborn are reported in parentheses. * p <0.10, ** p <0.05, *** p <0.01.

socio-economic background. Thus, the study attempts to further isolate the random variations in TFB across head of lineages by accounting for common characteristics across heads of lineages originated from the same mother. In particular, similarities in the genetic, cultural, and socio-economic characteristics across siblings, as opposed to across the population as a whole, are exploited to isolate the effect of random variation in TFB on reproductive success. Accounting for these Maternal Founder fixed effects, as well as additional confounding factors and clustering of standard errors on the level of the firstborn, the analysis explores the effect of random variation in TFB on long-run reproductive success.

As established in Table 2, the qualitative results established in Table 1 are unaffected, and the statistical precision is increased, by the inclusion of Maternal Founder fixed effects.[40] In particular, consistently with the first element of proposed hypothesis, column 9 establishes a highly significant negative association between TFB (as measured in years) of heads of lineages and the number of children, accounting for the marriage age, birth year, literacy status and gender of the heads of lineages. An increase in the TFB by one year results in a reduction of 0.062 in the log number of children. Hence, an increase in TFB from 38 weeks to 1 year and 38 weeks would result in a reduction of approximately 0.57 children.

The beneficial effects of an intermediate level of TFB on long-run reproductive success is confirmed in columns 10–12. They establish a hump-shaped effect of TFB of heads of lineages on the number of descendants forming in the second generation, and obtaining statistical significance in the third and fourth generations, accounting for the marriage age, birth year, literacy status and gender of the heads of lineages. While the insignificance of the estimates for grandchildren indicate that a static quantity-quality trade-off is not an exclusive cause of the optimality of moderate fecundity observed in the long run, the first and second order terms are highly significant for great-great-grandchildren. In addition, the first-order coefficient and second-order coefficient estimates are jointly highly significant for great-grandchildren and great-great grandchildren. Furthermore, as established in columns 13–16 the findings are robust to the inclusion of control for the stoppage age. In particular, the first and second order terms are highly significant for great-great-grandchildren and the additional test for a hump-shaped relationship for great-great-grandchildren is highly significant ($p = 0.007$).

In addition, while columns 3–4, 8– 9, 11–12, and 15–16 establishes the presence of a long-run reproductive trade-off, columns 2, 6, 10, and 14 indicate that a static quantity-quality trade-off is not an exclusive cause of the optimality of moderate fecundity observed in the long run.

The analysis suggests that the maximal reproductive success is attained by heads of lineages with a moderate TFB (i.e., those whose first delivery occurs 65 weeks after their marriage, in comparison to a sample median of 53 weeks), suggesting that the forces of natural selection may have had a positive effect on the median TFB in the population over this time period. In particular,

---

[40]In the baseline sample of 3,798 heads of households, 437 descended from mothers with no other offspring in the sample. The identification of the trade-off in the fixed-effects regressions is based on variation in 3,353 observations in the regressions excluding the stoppage age dummies and 3,348 observations in the regressions including the stoppage age dummies.

in comparison to highly fertile couples whose first child is born 38 weeks after the marriage, those individuals have on average 0.3 fewer children, but 0.4 more grandchildren, 8.4 additional great-grandchildren, and 15.7 added great-great-grandchildren.

Thus, the regression analysis presented in Table 2 confirms the hypothesis that higher fecundity in the pre-demographic transition era was associated with a larger number of children, while an intermediate level of fecundity maximized long-run reproductive success.

6   Robustness

This section establishes the robustness of the qualitative results to (i) alternative sample periods and sample restrictions, (ii) a wide range of potential confounding factors, accounting for geographic location at birth and death, birth order, month of birth, month of birth of the firstborn, number of marriages, and immigration and emigration status of the spouse, for each head of lineage, (iii) alternative estimation method (GLM), and (iv) estimation of the effect of number of children on long-run reproductive success.

6.1   Alternative Samples

6.1.1   Alternative Sample Periods

While the analysis accounts for cohort effects by the inclusion of birth year dummies, the economic and ecological environment may have changed over the sample period, shifting the costs and benefits of investment in offspring and thus affecting the existence or the nature of a long-run reproductive trade-off. Therefore, it might be a concern whether the baseline regression analysis is representative of alternative sample periods. Reassuringly, the main results are qualitatively unchanged if the analysis focuses on the reproductive success of heads of lineages born over a significantly shorter period, for example the period 1660–1685 (Table 3), as well as over a longer period, for example the entire sample period (Table 4).

6.1.2   Second Births and Marriage Age After Turning 15 Years

Pregnancy-related maternal death may be correlated with TFB and reproductive success due to its association with health-related complications. Since TFB is defined as the duration from marriage to the first birth, maternal death in relation to the first birth can potentially directly affect TFB. Furthermore, as is shown in Figure B.15, depicting the relation between maternal death and the total number of birth produced by mothers, the maximal rate of maternal death associated with the first birth. To account for the potential confounding effect of maternal death on TFB and reproductive success, the sample is restricted to heads of lineages who produced at least two children in the regressions performed in Table 2. As established in Table C.2, the qualitative results are unaffected by restricting the sample to heads of lineages who produced at least two children.

28

Table 3: The effect of the time to first birth (TFB) on the number of descendants for head of lineages born 1660–1685 – accounting for Maternal Founder fixed effects

	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
TFB	-.065**	.237	.544**	.830***	-.066***	.238	.537**	.830***	-.075***	.193	.487**	.802***	-.079***	.201	.498***	.773***
	(.025)	(.175)	(.216)	(.280)	(.025)	(.175)	(.217)	(.282)	(.025)	(.174)	(.216)	(.282)	(.012)	(.139)	(.191)	(.277)
TFB^2		-.096*	-.214***	-.337***		-.097*	-.212***	-.337***		-.086	-.199***	-.330***		-.088*	-.203***	-.323***
		(.057)	(.071)	(.092)		(.057)	(.071)	(.093)		(.057)	(.071)	(.093)		(.046)	(.063)	(.091)
Literate					-.028	.027	.124**	.118*	-.031	.023	.120**	.116*	-.035**	.016	.110**	.087
					(.033)	(.043)	(.055)	(.070)	(.033)	(.042)	(.054)	(.070)	(.015)	(.035)	(.049)	(.070)
Male									.219***	.279***	.315***	.177***	-.035**	.056*	.116***	.115*
									(.032)	(.040)	(.049)	(.063)	(.015)	(.032)	(.044)	(.066)
Stoppage age fixed effects	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Number of observations	3,376	3,376	3,376	3,376	3,376	3,376	3,376	3,376	3,376	3,376	3,376	3,376	3,376	3,376	3,376	3,376
Adjusted R^2	.019	.020	.042	.335	.020	.020	.045	.335	.036	.036	.058	.337	.804	.451	.306	.379
Joint sign.-level of TFB & TFB^2	.010	.066	.001	.000	.009	.066	.001	.000	.003	.042	.000	.000	.005	.000	.000	.000
Maximizing TFB		1.230	1.270	1.232		1.230	1.268	1.231		1.124	1.219	1.214		1.137	1.224	1.196
Lower limit of 90% CI		-11.448	.923	.961		-11.038	.910	.959		-	.761	.919		-1.02	.883	.886
Upper limit of 90% CI		1.505	1.427	1.37		1.504	1.427	1.369		-	1.392	1.357		1.387	1.375	1.343

This table presents the results of a series of fixed-effects OLS regressions of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2 for heads of lineages born in the period 1660–1685. All regressions account for Maternal Founder fixed effects. Birth year and marriage age dummies are included as controls. Furthermore, stoppage age dummies are included in columns 13–16. A dummy indicating unknown literacy is included in the regressions underlying column 5–16. Standard errors clustered at the level of the firstborn are reported in parentheses. * $p <0.10$, ** $p <0.05$, *** $p <0.01$.

Table 4: The effect of the time to first birth (TFB) on the number of descendants for entire sample period – accounting for Maternal Founder fixed effects

	Log number of descendants in:															
	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
TFB	-.064***	.055	.319**	.409**	-.064***	.050	.310**	.400**	-.071***	.031	.293**	.393**	-.090***	.096	.378***	.459**
	(.017)	(.120)	(.141)	(.199)	(.017)	(.120)	(.141)	(.199)	(.017)	(.119)	(.140)	(.199)	(.008)	(.094)	(.125)	(.198)
TFB^2		-.039	-.133***	-.192***		-.038	-.129***	-.189***		-.033	-.126***	-.187***		-.060*	-.158***	-.212***
		(.040)	(.046)	(.065)		(.040)	(.046)	(.065)		(.039)	(.046)	(.065)		(.031)	(.041)	(.065)
Literate					.007	.078**	.133***	.133***	.007	.078***	.133***	.134***	-.014	.057**	.110***	.121**
					(.023)	(.030)	(.037)	(.049)	(.023)	(.030)	(.037)	(.049)	(.011)	(.024)	(.033)	(.048)
Male									.222***	.220***	.191***	.088**	-.013	.005	.019	.019
									(.018)	(.023)	(.027)	(.036)	(.009)	(.019)	(.025)	(.038)
Stoppage age fixed effects	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Number of observations	7,664	7,664	7,664	7,664	7,664	7,664	7,664	7,664	7,664	7,664	7,664	7,664	7,664	7,664	7,664	7,664
Adjusted R^2	.026	.022	.067	.413	.026	.023	.070	.414	.049	.036	.077	.414	.789	.420	.282	.431
Joint sign.-level of TFB & TFB^2	.000	.015	.000	.000	.018	.000	.000	.000	.000	.008	.000	.000	.000	.000	.000	.000
Maximizing TFB		.694	1.200	1.065		.663	1.198	1.060		.457	1.166	1.049		.798	1.196	1.079
Lower limit of 90% CI		-	.741	.466		-	.708	.434		-	.604	.396		-3.176	.924	.610
Upper limit of 90% CI		-	1.376	1.269		-	1.378	1.268		-	1.355	1.261		1.167	1.333	1.264

This table presents the results of a series of fixed-effects OLS regressions of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2 for heads of lineages born in the entire sample period. All regressions account for Maternal Founder fixed effects. Birth year and marriage age dummies are included as controls. Furthermore, stoppage age dummies are included in columns 13–16. A dummy indicating unknown literacy is included in the regressions underlying column 5–16. Standard errors clustered at the level of the firstborn are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Robustness to GLM regression – accounting for Maternal Founder fixed effects

							Number of descendants in:									
	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
TFB	-.041**	.126	.412**	.547**	-.042**	.131	.414**	.557***	-.052***	.101	.387**	.547**	-.073***	.195*	.468***	.595***
	(.020)	(.146)	(.179)	(.215)	(.020)	(.146)	(.178)	(.215)	(.020)	(.145)	(.177)	(.215)	(.010)	(.118)	(.160)	(.216)
TFB²		-.054	-.162***	-.234***		-.056	-.163***	-.238***		-.050	-.158***	-.236***		-.087**	-.191***	-.256***
		(.047)	(.058)	(.068)		(.047)	(.058)	(.069)		(.047)	(.058)	(.068)		(.038)	(.052)	(.069)
Literate					-.006	.059*	.121***	.119**	-.006	.061*	.124***	.121**	-.020	.055*	.118***	.104*
					(.026)	(.036)	(.045)	(.056)	(.025)	(.036)	(.045)	(.056)	(.013)	(.029)	(.040)	(.055)
Male									.248***	.280***	.282***	.127**	-.015	.044	.083**	.072
									(.027)	(.036)	(.044)	(.054)	(.014)	(.029)	(.040)	(.055)
Stoppage age fixed effects	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Number of observations	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798
Joint sign.-level of TFB & TFB²	.045	.207	.001	.000	.037	.189	.001	.000	.009	.097	.000	.000	.000	.000	.000	.000
Maximizing TFB		1.172	1.271	1.169		1.174	1.27	1.169		1.01	1.226	1.155		1.116	1.225	1.159

This table presents the results of a series of GLM regressions, with a negative binomial distribution and a logarithmic link function, of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2. All regressions include dummies for Maternal Founder fixed effects (the results without the Maternal Founder Fixed Effects is presented in Table A.7). Birth year and marriage age dummies are included as controls. Furthermore, stoppage age dummies are included in columns 13–16. A dummy indicating unknown literacy is included in the regressions underlying column 5–16. Standard errors clustered at the level of the firstborn are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Likewise, marriage at low ages, at which some individuals may not have matured their reproductive capabilities, may affect TFB and reproductive success of heads of lineages. To account for the potential confounding effect of a low marriage age on TFB and reproductive success, the sample is restricted to heads of lineages who married after turning 15 years of age in the regressions performed in Table 2. As established in Table C.2, the qualitative results are unaffected by restricting the sample to heads of lineages who did not marry before age 15.

6.1.3 Extinct Lineages

The regression sample is restricted to lineages for which at least one great-great-grandchild is observed. To asses the robustness of the results with respect to including extinct lineages, the restriction of at least one observed great-great-grandchild is relaxed.[41] Furthermore, in an alternative set of regressions, the outcome variable is redefined as $\ln(1 + D_{i,t})$, where the added number 1 ensures that the logarithmic transformation is defined for extinct lineages and all lineages remain in the sample. As established in Table C.3, the qualitative results are unaffected by the inclusion of extinct lineages.

6.2 Additional Attributes of Heads of Lineages

6.2.1 Remarriages

Some head of lineages and their spouses, remarried, possibly multiple times, reflecting in part a considerable mortality rate over this period. The formation of additional unions may affect the reproductive success of heads of lineages via various channels, reflecting possibly the health and socioeconomic circumstances that led to these remarriages, as well as the potential differential treatment of previous and new children in the newly formed household. To account for the effect of remarriages, dummy variables indicating the number of marriages experienced by each head of lineage are included in regressions otherwise similar to those performed in Table 2. As established in Table C.4, the qualitative results are unaffected by accounting for remarriages.[42]

6.2.2 Spousal migration

Immigrants may differ systematically from natives reflecting the circumstances that led to their decision to immigrate as well as the effects of immigration on their socio-economic status and thus their TFB and reproductive success. Thus, the migration status of the first spouse may have affected the TFB and the reproductive success of heads of lineages. In the sample, heads of lineages were neither immigrants nor emigrants. Nevertheless, 23.5% of their first spouses were immigrants and 0.5% were emigrants. To account for the potential effect of spousal migration, dummy variables

---

[41]Since the logarithmic transformation is not defined at zero, extinct lineages drop out of the sample in regressions relating to the generation at which the lineage became extinct.

[42]Excluding remarriages of head of lineages would reduce the sample size considerably and thus would affect the significance of the estimations. Nevertheless, the qualitative results would not be altered and the existence of the hump-shaped relationship would be significant at the 1% level.

indicating the immigration and emigration status of heads of lineages are included in regressions otherwise similar to those performed in Table 2. As established in Table C.4, the qualitative results are unaffected by the migration status of the first spouse of the head of lineage.

6.2.3 Gender

Reflecting an earlier marriage age of women relative to men, the sample of heads of lineages is unbalanced across gender. Although men on average married at a later age than women (i.e., 26.6 for men versus 19.4 for women), their average age at last delivery was higher (i.e., 46.3 for men versus 38.3 for women), and they remarried more often, resulting in a higher number of children per male (i.e., 10 for men versus 9.4 for woman). The effect of gender is directly accounted for as a control in regressions otherwise similar to those performed in Tables 1–5. As an additional robustness check, Table C.5 demonstrate that the results are qualitatively similar in a sample that includes only females.

6.2.4 Birth and Death Parishes

The parishes of birth and death may affect TFB of heads of lineages and their reproductive success due to the influence of cultural and socioeconomic factors in a parish on the resources and preferences of heads of lineages. To account for the effect of these confounding geographical factors, dummy variables indicating the parishes of birth and death of each head of lineages are included in regressions otherwise similar to those performed in Tables 2. As established in Table C.6, the qualitative results are unaffected by accounting for these birth and death parish fixed effects.

6.2.5 Month of Marriage and Month of Birth of Firstborn

The month of marriage may affect TFB and reproductive success of heads of lineages due to influence of climatic conditions on resources, nutrition and human physiology. In addition, the month of birth of the firstborn may affect the resources of heads of lineages and thus their reproductive success. To account for these confounding seasonal factors, dummy variables indicating the month of marriage for each head of lineage and the months of birth of the first born of each head of lineage are included in regressions otherwise similar to those in Table 2. As established in Table C.7, the qualitative results are unaffected by accounting for these seasonal factors.

6.2.6 Birth Order

The birth order of heads of lineages may affect their TFB and reproductive success due to its effect of their nourishment as children, physiology, intergenerational transfers of wealth, and therefore resources as adults. To account for the potential effect of birth order, a dummy variable indicating if the head of lineage is the first birth among individuals sharing the same mother is included in regressions otherwise similar to those performed in Tables 2. As established in Table C.8, the qualitative results are unaffected by accounting for the firstborn status of heads of lineages.

Moreover, the firstborn status has no significant effect on reproductive success. Furthermore, accounting for the entire birth order of each head of lineage does not alter the qualitative results.

6.3 Alternative Estimation Method

The negative relationship between TFB of heads of lineages and the number of children, as well as the hump-shaped relationship between TFB of heads of lineages and long-run reproductive success, is established using quadratic OLS regression models. Table 5 and C.9 demonstrate that these negative relationships are robust to an alternative estimation method, using a generalized linear model (GLM) with a negative binomial distribution and a logarithmic link function.

6.4 Effect of Number of Children on Reproductive Success

While the analysis focuses on an important heritable determinant of fertility that could be shaped by natural selection, namely fecundity, it is possible to exploit the highly significant effect of the isolated random variation in 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 particular, as established by the 2SLS regressions reported in Table C.10, TFB is a strong instrument for the number of children in the presence of stoppage age fixed effects (i.e., the Kleibergen & Paap F-statistic is always above 30). The table establishes a significant hump-shaped effect of the number of children on long-run reproductive success in the extended sample including extinct lineages, i.e. the sample underlying Table C.3.[43] In particular, as established in column 3, the log-transformed number of children has a highly significant positive effect on the number of great-great-grandchildren and the squared log-transformed number of children has a highly significant negative effect on the number of great-great-grandchildren. Columns 6 and 9 establish that this association is robust to accounting for the effects of literacy and gender. Column 9 establishes that the number of children that maximizes the number of great-great-grandchildren is 4.72, well below the regression sample median of 10.

7 Mechanism

The section identifies several mechanisms that had contributed to the trade-off associated with higher fecundity. In particular, it establishes the observed hump-shaped effect of TFB on reproductive success in the long run reflects the positive effect of reduced fertility and thus higher child quality on the reproductive success of each child. While higher TFB reduced the number of chil-

---

[43]The instrumental variable analysis exploits the increased sample size of the sample that includes extinct lineages to obtain a statistically significant hump-shape. In the abscence of these additional observations, the first stage F-statistic would drop to the interval 13–18 and the first-order and second-order effect of children on the number of descendants would be insignificant.

dren, it contributed to the quality of each child, as reflected by health, survivability, education, and earning capacity, and had therefore affected the long-run reproductive success of the lineage.[44]

The findings suggest that child quality enhanced the likelihood that: (i) a child would reach the reproductive age and would have the qualities that would permit a success in the marriage market – preconditions for reproductive success, (ii) a child would have the qualities and the necessary income to be able to marry and start the process of reproduction earlier in life, and (iii) a child would become educated and thus would have higher earning capacity and reproductive success.

In particular, the fraction of children that survived till the end of their reproductive age, and succeeded to get married determined the fraction of offspring that contributed to the number of descendants and thus long-run reproductive success. Moreover, among children that got married, the average marriage age determined their average age of onset of reproductive activity and therefore the long-run reproductive success of the heads of lineages.

These marriage channels are explored in Table 6 and 7. The effect of TFB on the likelihood that a child married is explored in Table 6.[45] As established in column 1, the negative association between the fraction of children surviving to age 40 that got married and the TFB of heads of lineages is highly statistically significant.[46] In particular, accounting for the marriage age and birth year of heads of lineages, as well as gradually for their literacy status as well as gender, there is a significant negative effect of TFB on the fraction of children surviving to age 40 that got married. Thus, it appears that the marriage success channel plays a role in accounting for the established hump-shaped pattern between TFB and long-run reproductive success.[47]

The effect of TFB on the average marriage age of married offspring is explored in Table 7. As established in column 1, the association between the average marriage age among offspring that got married and TFB of heads of lineages is negative and highly significant statistically. In particular,

---

[44]The interaction between demography and human capital formation is at the center of Unified Growth Theory (Cervellati and Sunde, 2005; Boucekkine et al., 2007; Galor, 2011).

[45]Due to the fact that the outcome variable is a fraction, the estimates presented in this table are based on the fractional logit model (Papke and Wooldridge, 1996). The results are robust to the inclusion of marriage age, birth year and stoppage age as continuous variables. In particular, the coefficient on TFB in the first, second, third and fourth column would all increase in size to 0.316 ($p = 0.005$), 0.268 ($p = 0.015$), 0.255 ($p = 0.020$), and 0.256 ($p = 0.020$), respectively. Furthermore, the results are robust to the use of a linear probability model. In particular, the coefficient on TFB in the first, second, third and fourth column would be 0.018 ($p < 0.005$), 0.015 ($p < 0.016$), 0.014 ($p = 0.029$), and 0.013 ($p = 0.035$), respectively. Since the Maternal Founder fixed effects constitute an unbalanced panel, the fractional logit model does not account for these fixed effects (Papke and Wooldridge, 2008). However, accounting for these fixed effects in a fixed-effects linear probability model yields similar results, although less precisely estimated. In particular, the estimate in column 4 would be 0.10 ($p = 0.127$). Under an alternative age criteria of 30 years instead of 40 years the estimate the estimate would be 0.011 ($p = 0.087$)).

[46]This finding is robust to the use of alternative age cut-offs. In particular, the effect of TFB on the fraction of surviving children that got married is statistically significant at the 5% level in all specifications when using 30, 35 or 45 as the surviving cut-off age.

[47]Underlying, in part, this positive effect of TFB on the marriage probability of offspring surviving to age 40 may be a positive association between TFB and the average longevity of offspring. Indeed, OLS regressions on the basis of the entire sample period, i.e., the regression sample underlying Table 4, suggests there is a positive association between TFB and the average longevity of offspring. In particular, measuring average longevity in years, the coefficient on TFB in the first, second, third and fourth column would be 1.063 ($p = 0.038$), 0.963 ($p = 0.060$), 0.980 ($p = 0.056$), and 1.033 ($p = 0.043$), respectively. It should be noted that this result depend on the abscence of control for Maternal Founder fixed effects and the larger sample period.

Table 6: The effect of time to first birth (TFB) on the fraction of children surviving to age 40 that got married

	Fraction of children surviving to age 40 that got married			
	(1)	(2)	(3)	(4)
TFB	.299***	.256**	.235**	.232**
	(.113)	(.112)	(.111)	(.110)
Literate		.770***	.763***	.779***
		(.110)	(.110)	(.109)
Male			.396***	.365***
			(.112)	(.120)
Stoppage age fixed effects	No	No	No	Yes
Number of observations	3,727	3,727	3,727	3,727

This table presents the results of a series of fractional logit regressions of the fraction of children, i.e., individuals in the first generation, surviving to age 40 that got married on time to first birth measured in years, i.e., TFB, for heads of lineages with at least one child surviving to age 40. Birth year and marriage age dummies are included as controls. Furthermore, stoppage age dummies are included in column 4. A dummy indicating unknown literacy is included in the regressions underlying column 2–4. Standard errors clustered at the level of the firstborn are reported in parentheses. * p <0.10, ** p <0.05, *** p <0.01.

Table 7: Time to first birth (TFB) and the average marriage age

	Average marriage age of children			
	(1)	(2)	(3)	(4)
TFB	-.430***	-.400***	-.376**	-.339**
	(.005)	(.008)	(.013)	(.023)
Literate		-.629***	-.621***	-.705***
		(.001)	(.001)	(.000)
Male			-.406***	-.720***
			(.009)	(.000)
Stoppage age fixed effects	No	No	No	Yes
Number of observations	3,796	3,796	3,796	3,796
Adjusted R^2	.006	.010	.011	.036

This table presents the results of a series of OLS regressions of the average marriage age of childre,, i.e., individuals in the first generation, on time to first birth measured in years, i.e., TFB. Birth year and marriage age dummies are included as controls. Furthermore, stoppage age dummies are included in column 4. A dummy indicating unknown literacy is included in the regressions underlying column 2–4. Standard errors clustered at the level of the firstborn are reported in parentheses. * p <0.10, ** p <0.05, *** p <0.01.

Table 8: The effect of time to first birth (TFB) on the fraction of literate children

	Fraction of literate children			
	(1)	(2)	(3)	(4)
TFB	.401***	.351***	.322***	.337***
	(.090)	(.090)	(.091)	(.091)
Literate		1.308***	1.307***	1.305***
		(.094)	(.094)	(.095)
Male			.563***	.407***
			(.090)	(.098)
Stoppage age fixed effects	No	No	No	Yes
Number of observations	3,448	3,448	3,448	3,448

This table presents the results of a series of fractional logit regressions of the share of children, i.e., individuals in the first generation, obtaining literacy on time to first birth measured in years, i.e., TFB, for heads of lineages with at least one surviving child with observed literacy status. Birth year and marriage age dummies are included as controls. Furthermore, stoppage age dummies are included in column 4. A dummy indicating unknown literacy is included in the regressions underlying column 2–4. Standard errors clustered at the level of the firstborn are reported in parentheses. * p <0.10, ** p <0.05, *** p <0.01.

accounting for the marriage age and birth year of heads of lineages, as well as gradually for their literacy status and gender, there is a significant negative relationship between TFB and the average marriage age of children that got married.[48] Thus, it appears that the marriage success channel plays a role in accounting for the established hump-shaped pattern between TFB and long-run reproductive success.

The education channel is investigated in Table 8. As established in column 1, TFB of heads of lineages has a highly significant positive association with the fraction of literate children, accounting for the marriage age and birth year of heads of lineages. As controls are gradually introduced to account for the confounding effects of the literacy status, gender, and stoppage age of heads of lineages, the positive coefficient remains stable and significant at the 1% significance level.[49]

---

[48]In fixed-effects regressions accounting for Maternal Founder fixed effects, the coefficient estimate drops somewhat and the standard error increases resulting in decreased statistical significance. However, the finding remains statistically significant at the 10% significance level in the entire sample period, i.e., the regression sample underlying Table 4. In particular, in the regression corresponding to that underlying column 4, controlling for Maternal Founder fixed effects in the entire sample period, the effect of TFB on the average marriage age of offspring is -0.181 ($p = 0.082$).

[49]As in Table 6, the regressions underlying Table 8 are based on fractional logit models. The results are robust to the inclusion of marriage age, birth year and stoppage age as continuous variables. In particular, the coefficient on TFB in the first, second, third and fourth column would be 0.392 ($p < 0.001$), 0.334 ($p < 0.001$), 0.314 ($p < 0.001$), and 0.311 ($p < 0.001$), respectively. Furthermore, the results are robust to the use of a linear probability model. In particular, the coefficient on TFB in the first, second, third and fourth column would be 0.071 ($p < 0.001$), 0.060 ($p < 0.001$), 0.054 ($p < 0.001$), and 0.055 ($p < 0.001$), respectively. Since the Maternal Founder fixed effects constitute an unbalanced panel, the fractional logit model does not account for these fixed effects (Papke and Wooldridge, 2008). Accounting for the fixed effects in a fixed-effects linear probability model on the baseline sample also yields positive estimates of the coefficient on TFB, although the estimates are insignificant. However, focusing on heads of lineages born prior to the end of 1660, the positive effect of TFB on the average literacy of offspring is statistically significant. In particular, the estimate in column 4 would be 0.137 ($p = 0.002$). Thus, the statistical evidence of the positive

Moreover, literacy of heads of lineages has a highly significant positive effect on the literacy of their children.

8   Concluding Remarks

This research explores the biocultural origins of human capital formation. It presents the first evidence that moderate fecundity and thus predisposition towards investment in child quality was conducive for long-run reproductive success within the human species. Using an extensive genealogical record for nearly half a million individuals in Quebec from the 16th to the 18th century, the study explores the effect of fecundity on the number of descendants of early inhabitants in the subsequent four generations. The research exploits variation in the random component of the time interval between the date of first marriage and the first birth to establish that while higher fecundity is associated with a larger number of children, an intermediate level maximizes long-run reproductive success. Moreover, the observed hump-shaped effect of fecundity on long-run reproductive success reflects the negative effect of higher fecundity on the quality of each child.

The research further indicates that the optimal level of fecundity was below the population median, suggesting that the forces of natural selection favored individuals with a lower level of fecundity and thus higher predisposition towards investment in child quality. The research lends credence to the hypothesis that during the Malthusian epoch, natural selection favored individuals with a larger predisposition towards child quality, contributing to human capital formation, the onset of the demographic transition and the evolution of societies from an epoch of stagnation to sustained economic growth.

---

effect of lower TFB based on variation within siblings sharing the same Maternal Founder on offspring literacy is strongest for the first part of the baseline period.

Online Appendix

Not for publication

A   Intergenerational Transmission of Human Capital

Suppose that parental human capital has a direct positive spillover on the level of the human capital of the child. In particular, the level of human capital of a child of a member i of generation t, h_{t+1}^i, is an increasing strictly concave function of both the parental investment in the education of the child, q_t^i, as well as the parental level of humand capital.

$$h_{t+1}^i = h(q_t^i, h_t^i), \tag{27}$$

where for $j = 1, 2$, $h_j(q_t^i, h_t^i) > 0$, $h_{jj}(q_t^i, h_t^i) < 0$, $\lim_{q_t \to 0} h_1(q_t^i, h_t^i) < \infty$, $\lim_{q_t \to \infty} h_j(q_t^i, h_t^i) = 0$, $h_{11}(q_t^i, h_t^i)q_t^i + h_1(q_t^i, h_t^i) < 0$, and $h(0, 1) = 1$.

Suppose that $h_{12}(q_t^i, h_t^i) > 0$, (i.e., parental human capital increases the productivity of parental investment in the human capital of each child), and suppose that this complementarity exceed the direct spillover from parental human capital to the human capital of the child, i.e.,

$$h_{12}(q_t^i, h_t^i) > h_2(q_t^i, h_t^i). \tag{28}$$

Modifying (13) accordingly, the interior optimal level of investment, q_t^i, is determined by

$$G(q_t^i, h_t^i) \equiv \beta h_1(q_t^i, h_t^i)(\tau + q_t^i) - h(q_t^i, h_t^i) = 0, \tag{29}$$

where $\partial G(q_t^i, h_t^i)/\partial q_t^i = (1 - \beta)h'(q_t^i) - \beta h''(q_t^i)(\tau + q_t^i) < 0$. Hence, it follows from the Implicit Function Theorem that there exist a single-valued function $q_t^i = \xi(h_t^i)$, such that[50]

$$\xi'(h_t^i) = -\frac{\beta h_{12}(q_t^i, h_t^i)(\tau + q_t^i) - h_1(q_t^i, h_t^i)}{\partial G(q_t^i, h_t^i)/\partial q_t^i} > 0. \tag{30}$$

The evolution of human capital within lineage i, is given by the the sequence $\{h_t^i\}_{t=0}^\infty$, such that

$$h_{t+1}^i = h(q_t^i, h_t^i) = h\big(\xi(h_t^i), h_t^i\big) \equiv \phi(h_t^i), \tag{31}$$

where $\phi(0) > 0$ and $\phi(h_t^i)$ is strictly concave in h_t^i. Hence, the level of human capital increases monotonically and it converges to a steady-state equilibrium,

$$\bar{h} = h(\bar{q}, \bar{h}), \tag{32}$$

in which investment in human capital, $\bar{q}$, coupled with parental externality, $\bar{h}$, generate a level of human capital $\bar{h}$.

---

[50]It is further assumed that $\xi''(h_t^i) < 0$ and $\xi'(h_t^i) > -2h_{12}(q_t^i, h_t^i)/h_{22}(q_t^i, h_t^i)$.

The level of fertility of a member i of generation t is, accordingly, along the transition to this steady-state equilibrium, therefore

$$n_t^i = \frac{\gamma y_t^i}{\tau + q_t} = \frac{\gamma w h_t^i}{\tau + \xi'(h_t^i)} = \psi\big(h_0^i(n_0^i)\big), \tag{33}$$

where $h_t^i = \phi(h_t^i) = \phi^{\{T\}}(h_0^i)$

Hence, the level of fecundity of the founding generation, n_0^*, that would maximize reproductive success in period T will be based on the the effect of n_0^* on the level of fertility in each of the subsequent T generations. Moreover, if $h'[0]$ is sufficiently small, the maximal level of fecundity will be favored by the forces of natural selection.

B   Additional Figures

Distribution of Number of Children – Females

Figure B.1: The histogram depicts the number of children of 26,566 non-migrating women in Quebec between the 16th and the end of the 18th century with known year of birth and at least one observed grandchild.

Distribution of Number of Children – Males

Figure B.2: The histogram depicts the number of children of 21,820 non-migrating men in Quebec between the 16th and the end of the 18th century with known year of birth and at least one observed grandchild.

Distribution of Number of Grandchildren – Females

Figure B.3: The histogram depicts the number of children of 26,566 non-migrating women in Quebec between the 16th and the end of the 18th century with known year of birth and at least one observed grandchild.
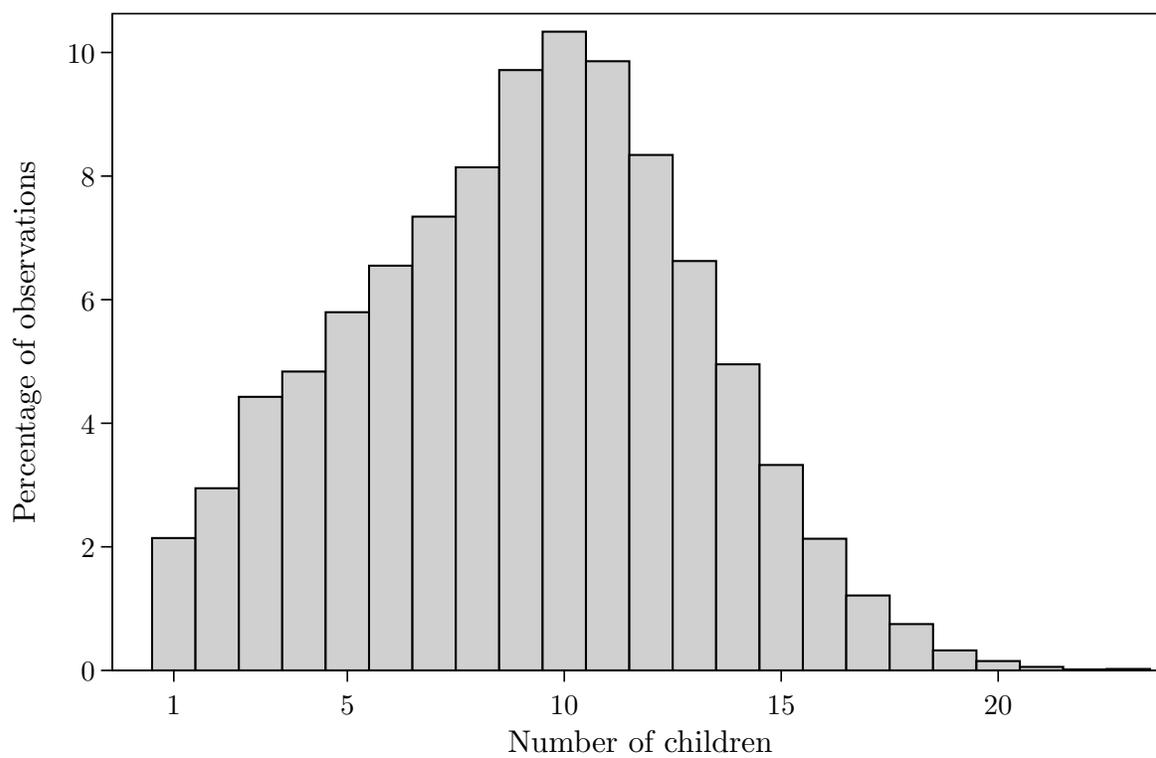
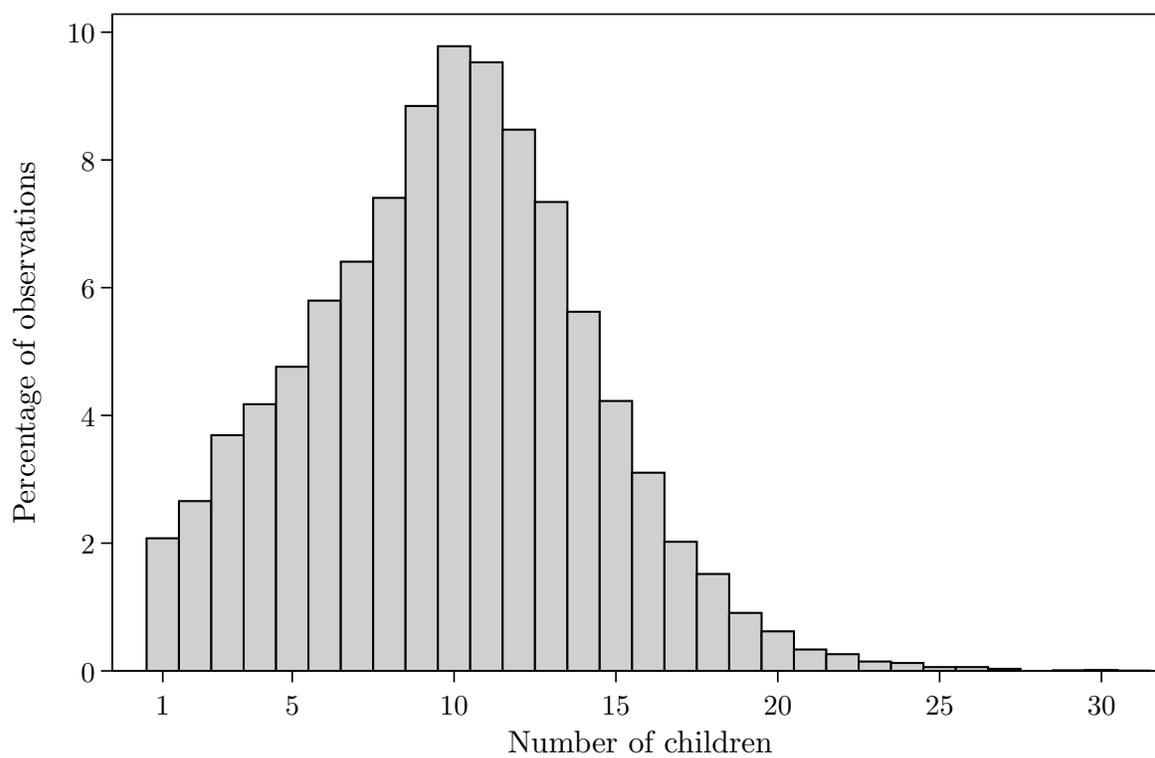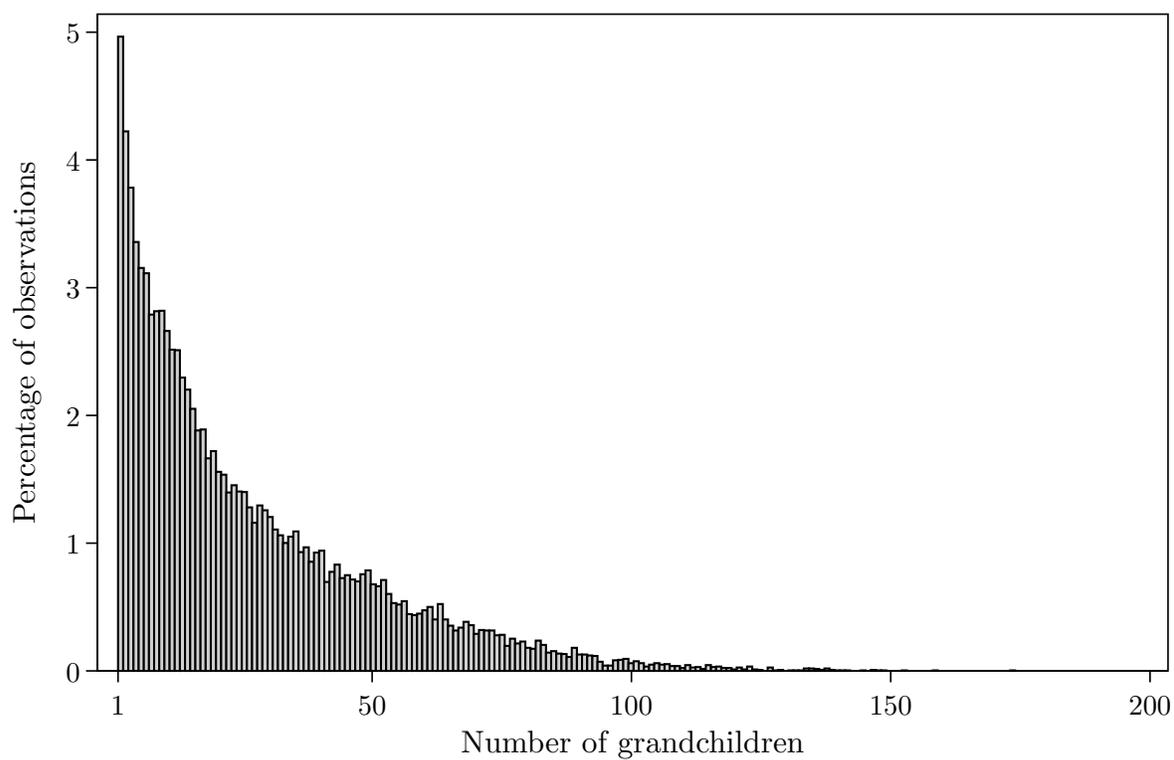**Distribution of Number of Grandchildren – Males**
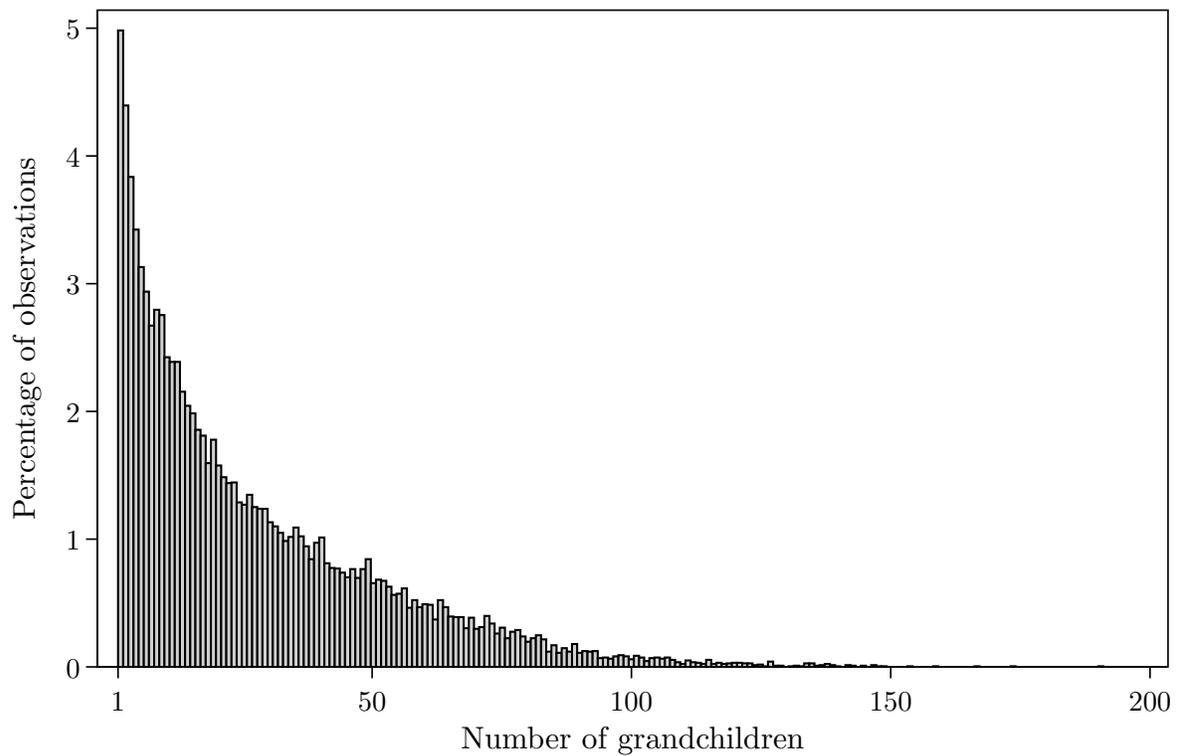
Figure B.4: The histogram depicts the number of children of 21,820 non-migrating men in Quebec between the 16th and the end of the 18th century with known year of birth and at least one observed grandchild.

Distributions of Numbers of Great-Grandchildren and Great-Great-Grandchildren
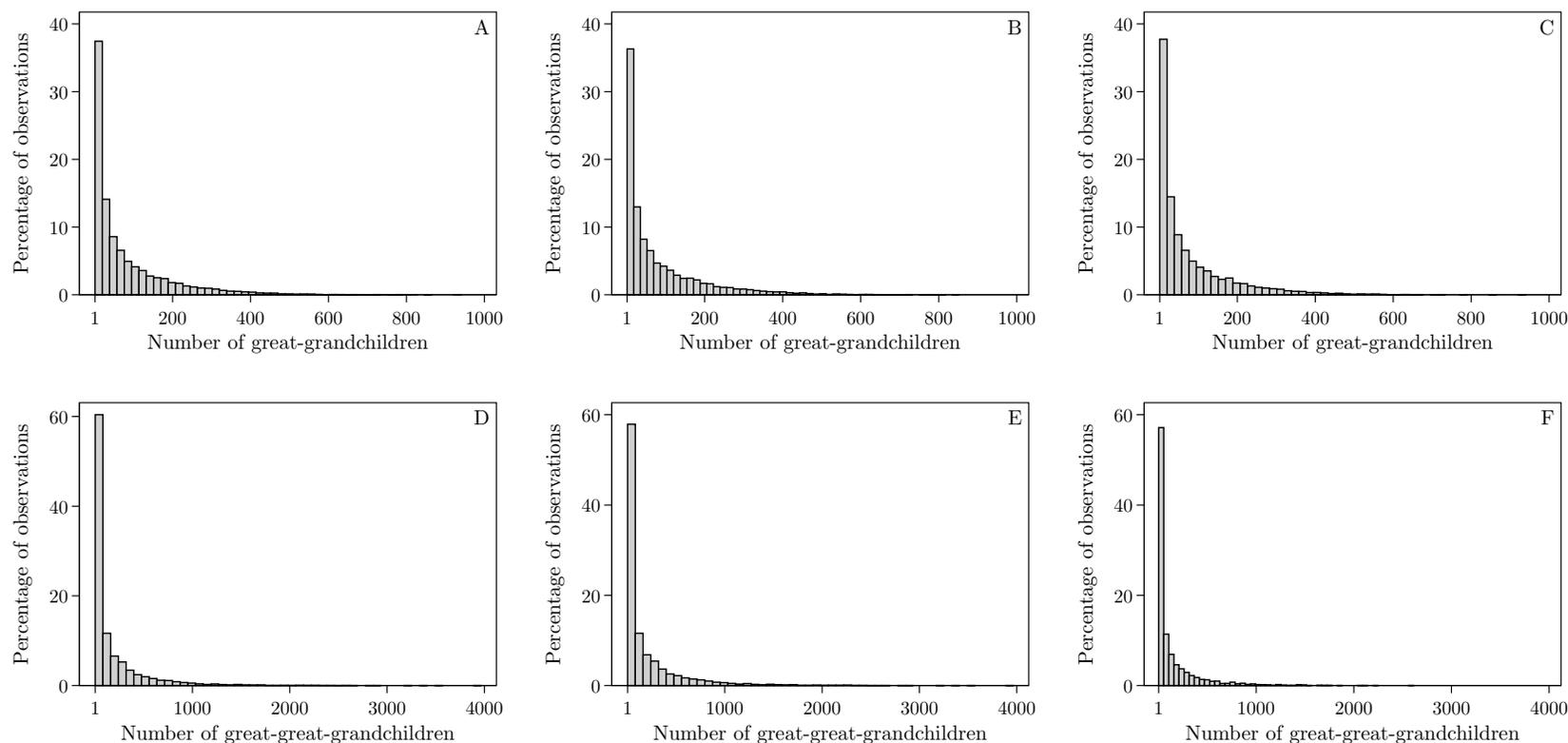
Figure B.5: The histogram depicts the number of descendants of non-migrating individuals in Quebec between the 16th and the end of the 18th century with known year of birth. Panel A–C restricts the sample to individuals with at least one observed great-grandchild. Panel D–F restricts the sample to individuals with at least one observed great-great-grandchild. Panel A depicts the number of great-grandchildren of 23,141 men, panel B depicts the number of great-grandchildren of 12,894 women, panel C depicts the number of great-grandchildren of 10,247 men, panel D depicts the number of great-great-grandchildren of 9,868 individuals, panel E depicts the number of great-great-grandchildren of 5,713 women, and panel D depicts the number of great-great-grandchildren of 4,155 men.

Distribution of Marriage Ages – Females
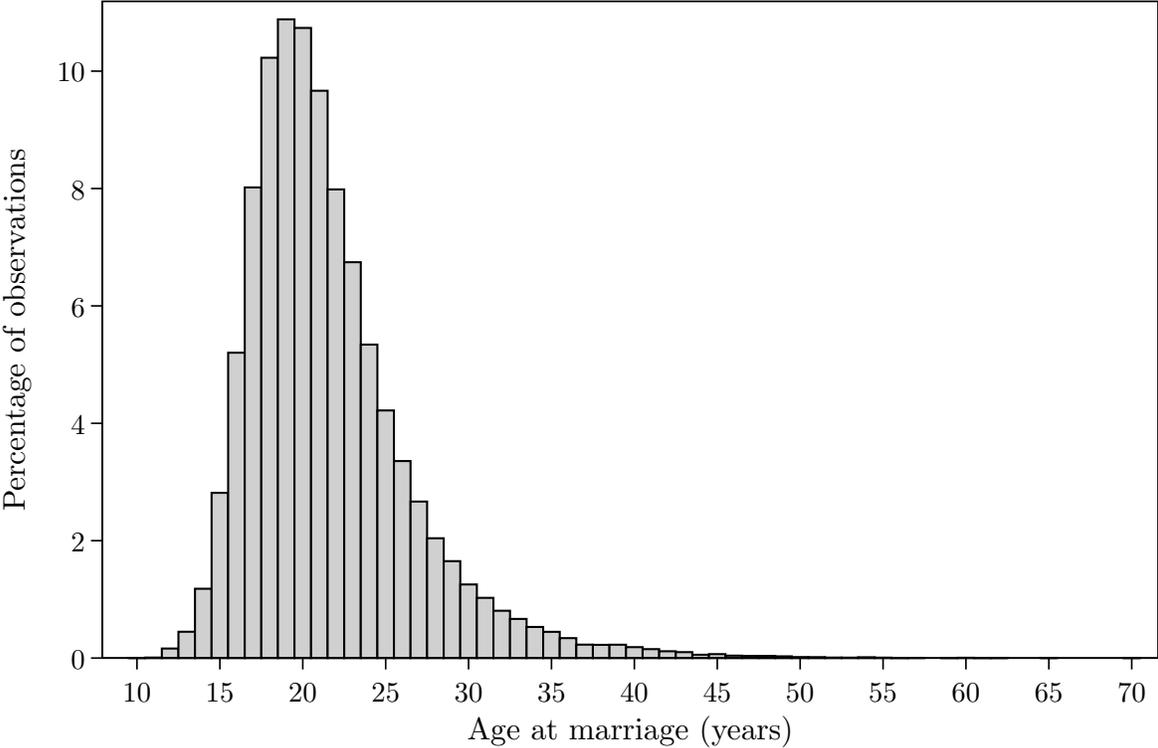
Figure B.6: The histogram depicts age at first marriage (in years) of 58,701 non-migrant, ever-married women in Quebec between the 16th and the end of the 18th century.

Distribution of Marriage Ages – Males
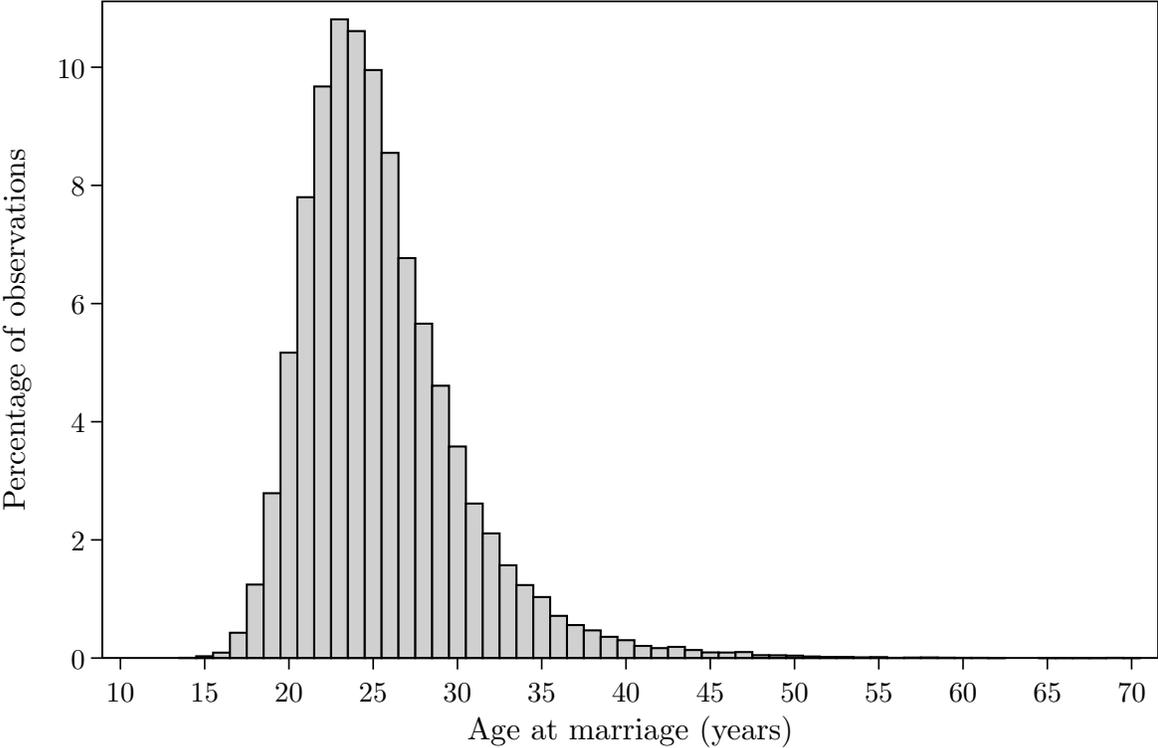
Figure B.7: The histogram depicts age at first marriage (in years) of 49,298 non-migrant, ever-married men in Quebec between the 16th and the end of the 18th century.

Distribution of Stoppage Ages
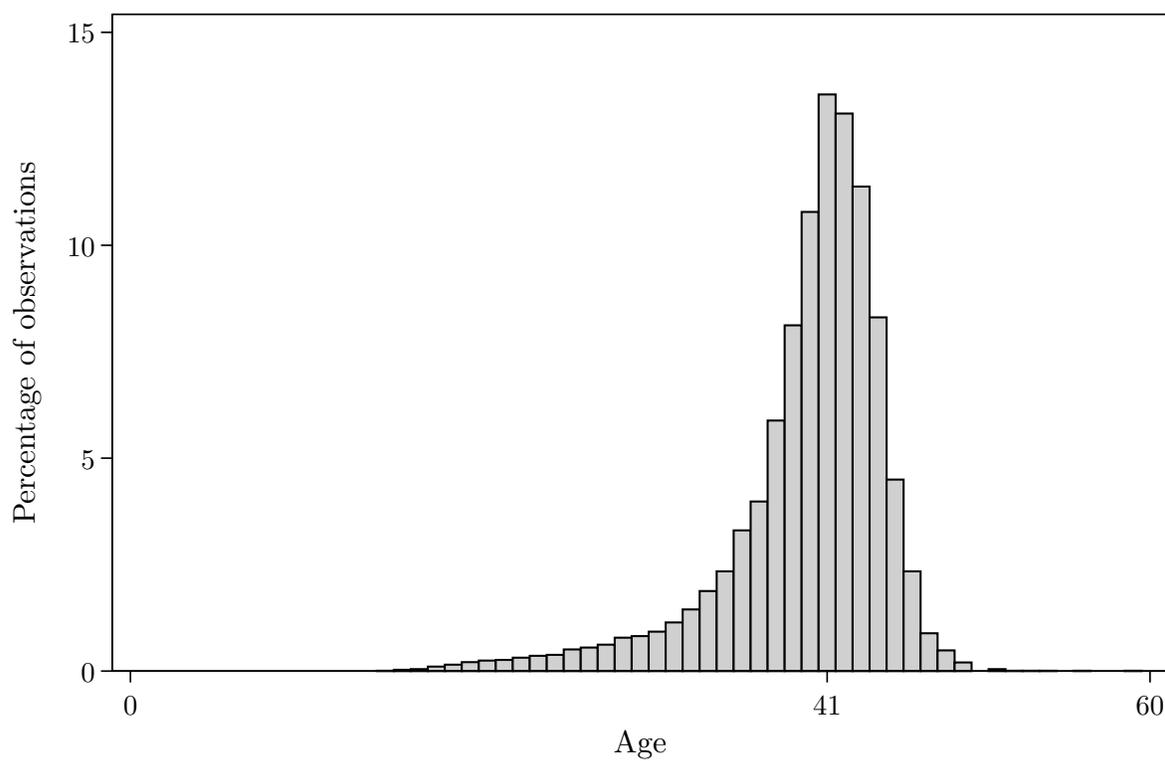
Figure B.8: The histogram depicts the age at last delivery of 13,411 once-married, non-migrant mothers in Quebec born before 1749 (and after 1624) who survived to age 50 and whose husband survived to age 50.

TFB and Number of Children – Conditional Relationship
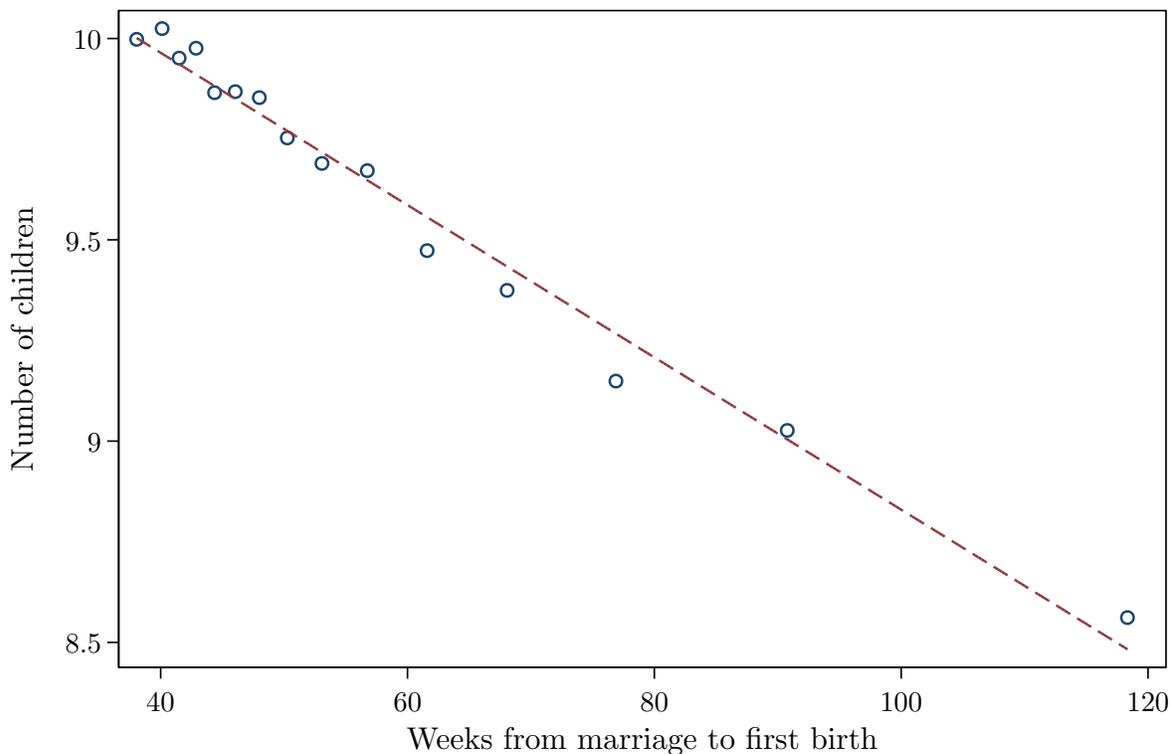

Figure B.9: This figure presents a binned scatterplot of number of children versus TFB for 38,498 non-migrating individuals in Quebec between the 16th and the end of the 18th century with known birth year, at least one observed grandchild and whose first conception occurred within marriage and in less than 2 years (i.e., with a TFB not shorter than 38 weeks and not longer than 2 years and 38 weeks). This binned scatter plot is constructed by first residualizing the the number-of-children variable with respect to marriage age, stoppage age and birth year. Second, TFB is divided into 15 equal-sized groups (15-quintiles) and the means of the number of children residuals within each bin is plotted against the mean value of TFB within each bin. The unconditional relationship is presented in Figure B.9. The line shows the best linear fit estimated on the underlying micro data using OLS. The estimated slope of the best-fit line, is -0.019 (i.e., one additional week from marriage to first birth results in 0.019 fewer children) and highly significant with a standard error clustered at the level of the firstborn of less than 0.001 ($p < 0.001$).

TFB and Number of Children – Unconditional Relationship
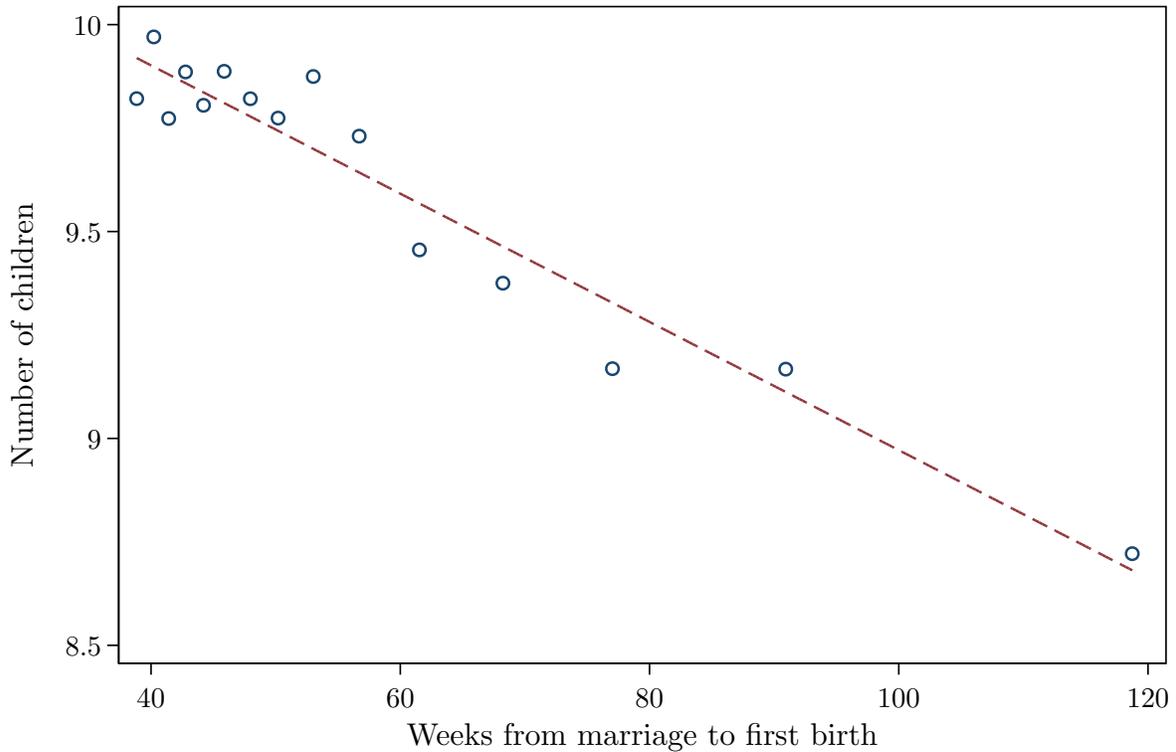
Figure B.10: This figure presents a binned scatterplot of number of children versus TFB for 38,498 non-migrating individuals in Quebec between the 16th and the end of the 18th century with known birth year, at least one observed grandchild and whose first conception occurred within marriage and in less than 2 years (i.e., with a TFB not shorter than 38 weeks and not longer than 2 years and 38 weeks). The binned scatter plot is constructed by dividing TFB into 15 equal-sized groups (15-quintiles) and plotting the means of the number of children within each bin against the mean value of TFB within each bin. The line shows the best linear fit estimated on the underlying micro data using OLS. The estimated slope of the best-fit line, is -0.016 (i.e., one additional week from marriage to first birth results in 0.016 fewer children) and highly significant with a standard error clustered at the level of the firstborn of 0.001 ($p < 0.001$).

Binned Scatterplot: Quartiles

Figure B.11: This figure presents a binned scatterplot of log number of children, grandchildren, great-grandchildren and great-great-grandchildren versus TFB for 3,798 non-immigrating and non-emigrating individuals in Quebec between the 16th and the end of the 18th century with known birth year, at least one observed great-great-grandchild and whose first conception occurred within marriage and in less than 2 years (i.e., with a TFB not shorter than 38 weeks and not longer than 2 years and 38 weeks). These binned scatter plots are constructed by first residualizing the the log-number-of-descendants variable with respect to marriage age, stoppage age and birth year. Second, TFB is divided into four equal-sized groups (quartiles) and the means of the log-number-of-descendants residuals within each bin is plotted against the mean value of TFB within each bin. The points are connected by lines.

Binned Scatterplot: Quintiles

Figure B.12: This figure presents a binned scatterplot of log number of children, grandchildren, great-grandchildren and great-great-grandchildren versus TFB for 3,798 non-immigrating and non-emigrating individuals in Quebec between the 16th and the end of the 18th century with known birth year, at least one observed great-great-grandchild and whose first conception occurred within marriage and in less than 2 years (i.e., with a TFB not shorter than 38 weeks and not longer than 2 years and 38 weeks). These binned scatter plots are constructed by first residualizing the the log-number-of-descendants variable with respect to marriage age, stoppage age and birth year. Second, TFB is divided into five equal-sized groups (quintiles) and the means of the log-number-of-descendants residuals within each bin is plotted against the mean value of TFB within each bin. The points are connected by lines.

Binned Scatterplot: 15-Quantiles
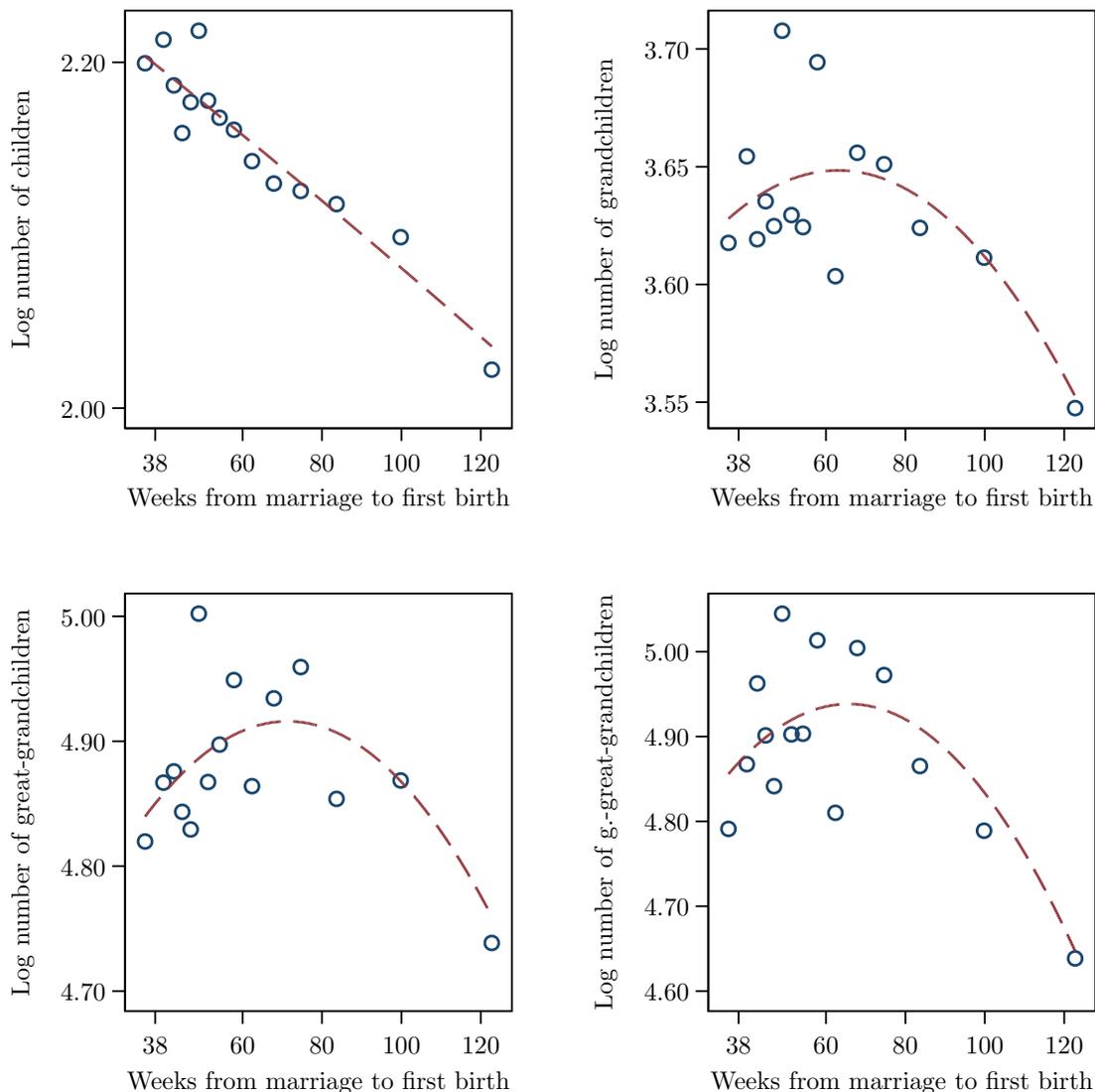
Figure B.13: This figure presents a binned scatterplot of log number of children, grandchildren, great-grandchildren and great-great-grandchildren versus TFB for 3,798 non-immigrating and non-emigrating individuals in Quebec between the 16th and the end of the 18th century with known birth year, at least one observed great-great-grandchild and whose first conception occurred within marriage and in less than 2 years (i.e., with a TFB not shorter than 38 weeks and not longer than 2 years and 38 weeks). These binned scatter plots are constructed by first residualizing the the log-number-of-descendants variable with respect to marriage age, stoppage age and birth year. Second, TFB is divided into 15 equal-sized groups (15-quantiles) and the means of the log-number-of-descendants residuals within each bin is plotted against the mean value of TFB within each bin. The line shows the best quadratic fit estimated on the underlying micro data using OLS.
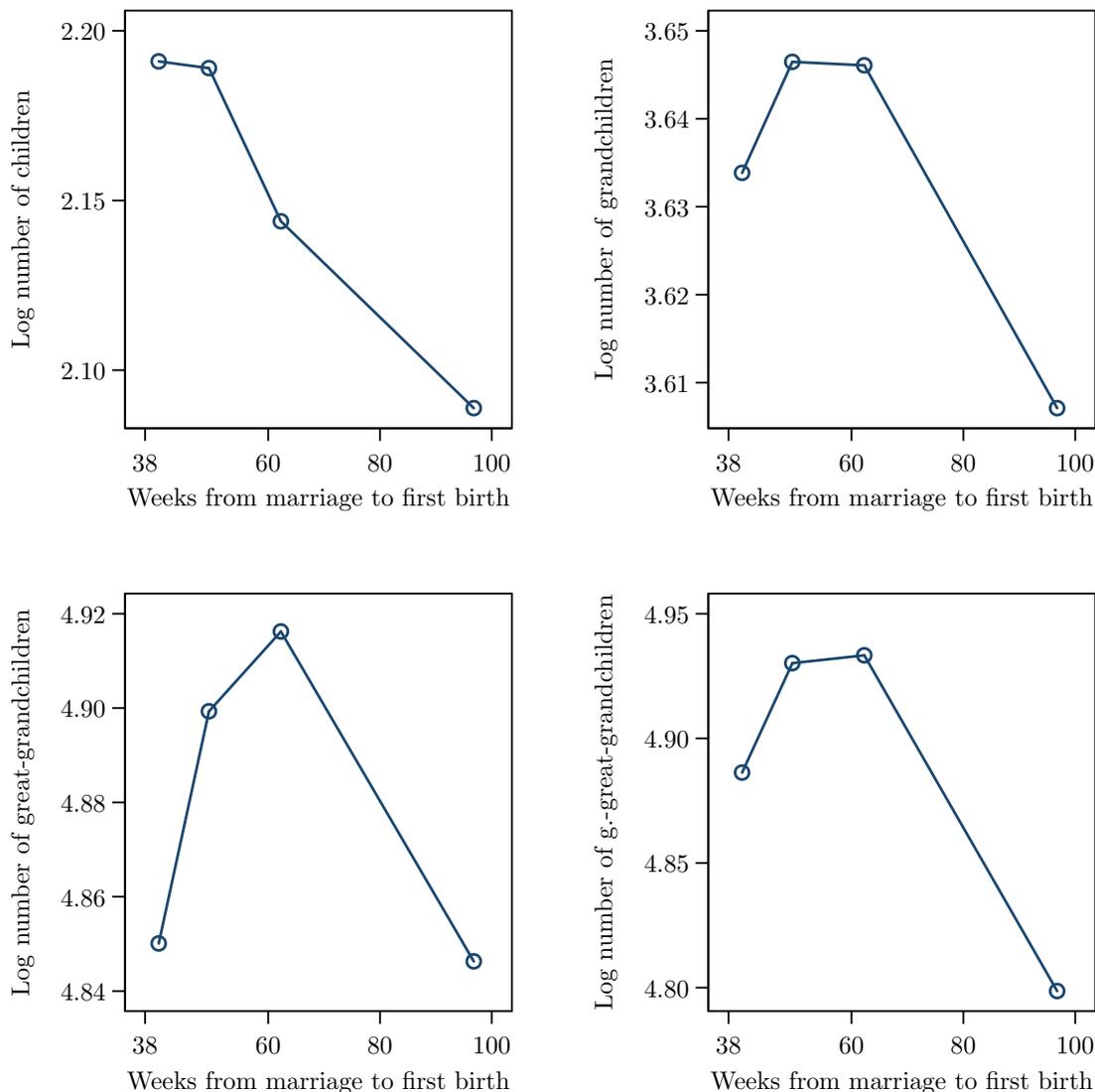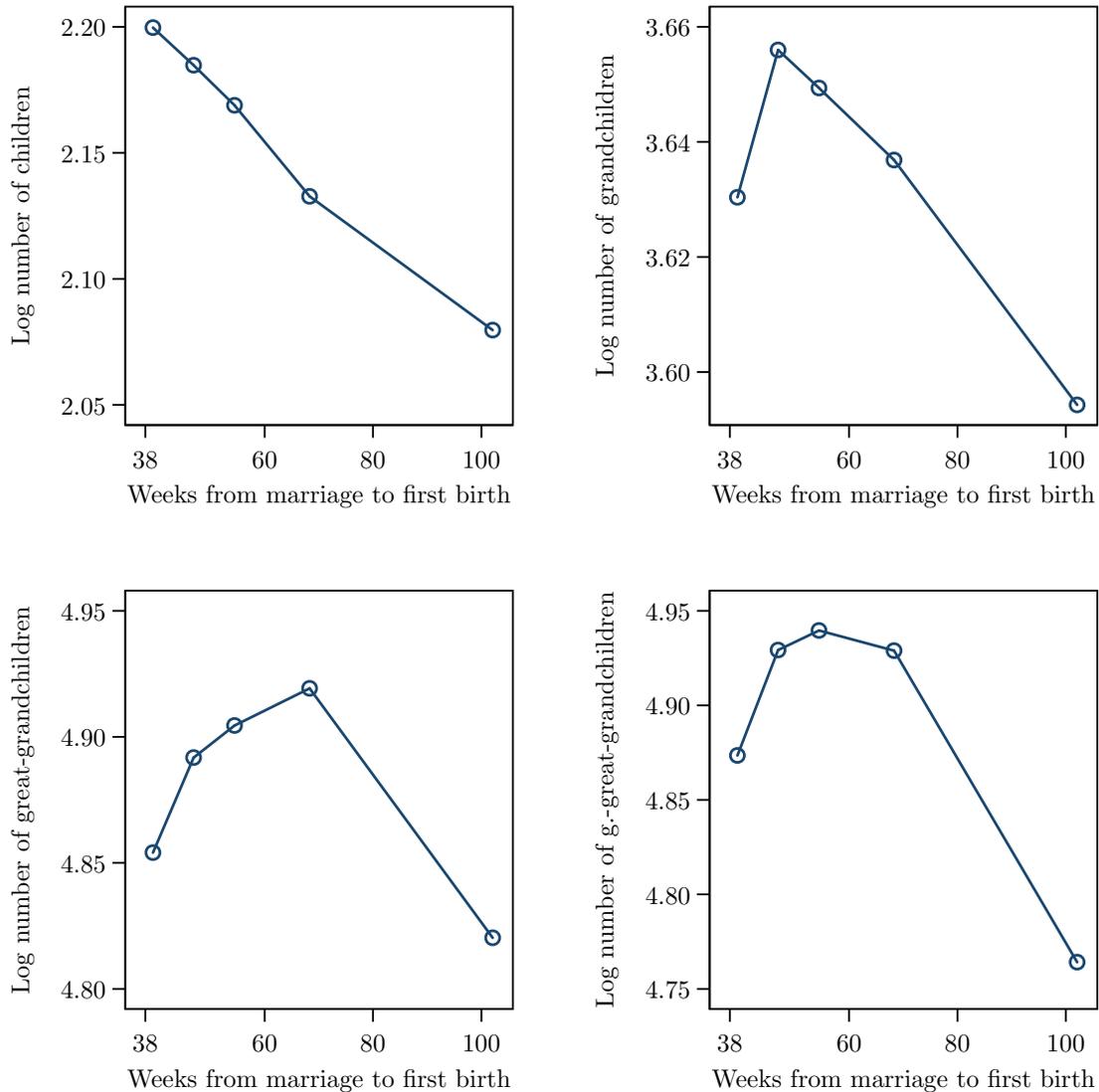
Binned Scatterplot: Vigintiles
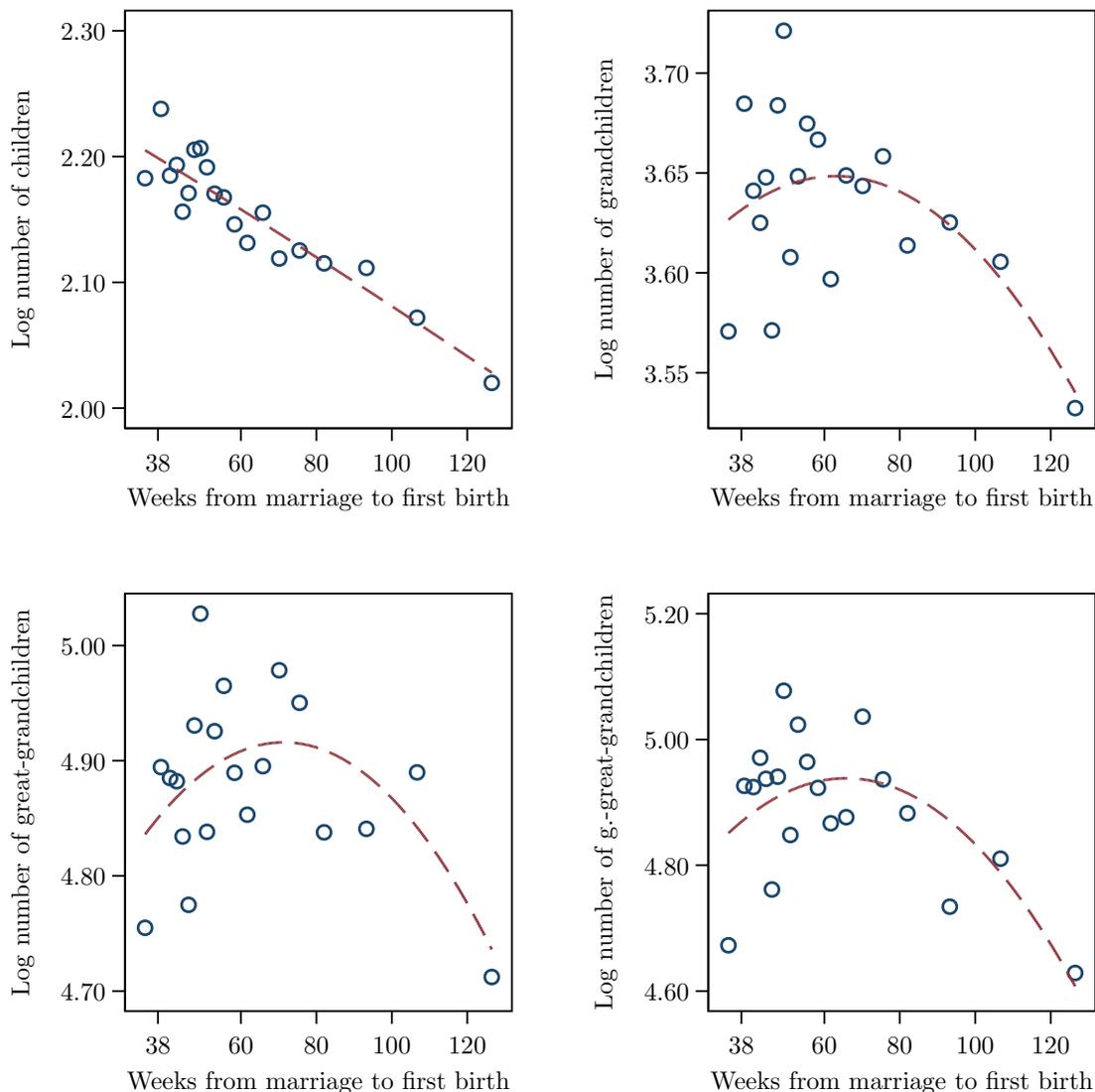
Figure B.14: This figure presents a binned scatterplot of log number of children, grandchildren, great-grandchildren and great-great-grandchildren versus TFB for 3,798 non-immigrating and non-emigrating individuals in Quebec between the 16th and the end of the 18th century with known birth year, at least one observed great-great-grandchild and whose first conception occurred within marriage and in less than 2 years (i.e., with a TFB not shorter than 38 weeks and not longer than 2 years and 38 weeks). These binned scatter plots are constructed by first residualizing the the log-number-of-descendants variable with respect to marriage age, stoppage age and birth year. Second, TFB is divided into 20 equal-sized groups (vigintiles) and the means of the log-number-of-descendants residuals within each bin is plotted against the mean value of TFB within each bin. The line shows the best quadratic fit estimated on the underlying micro data using OLS.

Maternal Death and Number of Children
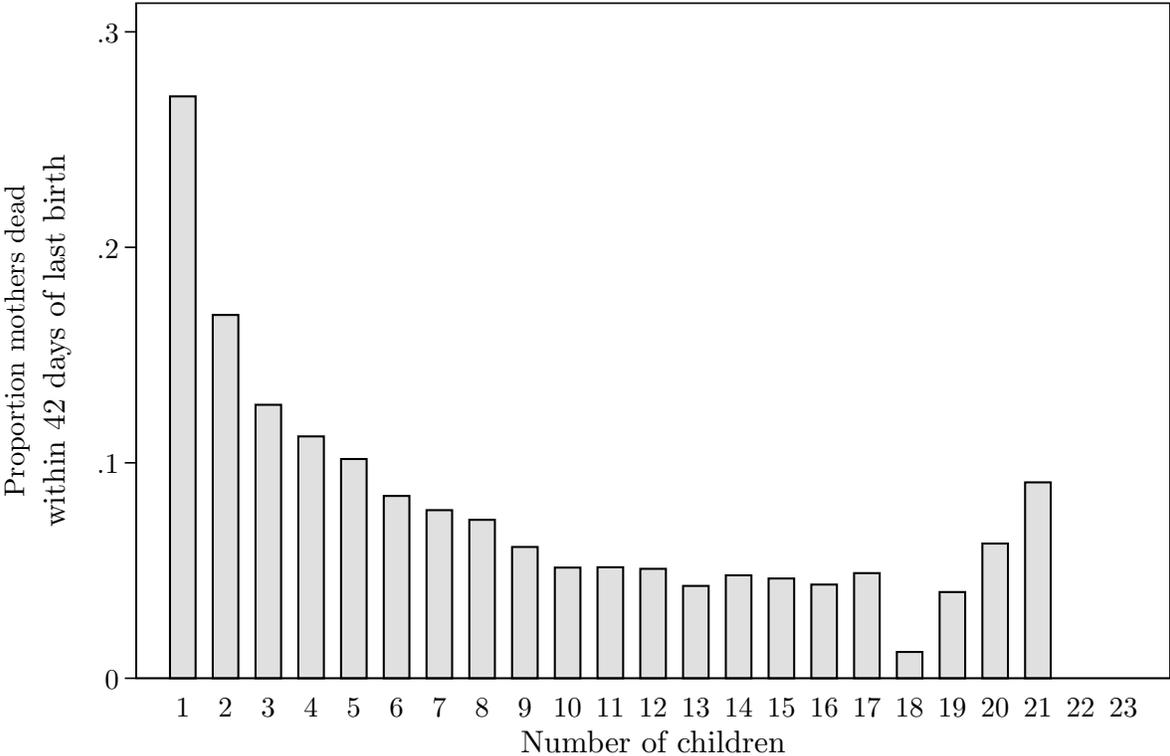

Figure B.15: The graph depicts the proportion of mothers dying within 42 days of their last delivery for 36,371 non-migrant mothers in Quebec.

C   Additional Tables

Summary Statistics for Heads of Lineages

Table C.1: Summary statistics of heads of lineages born prior to the end of 1685

	Mean (1)	Median (2)	S.D. (3)	Count (4)
FEMALES				
Children	9.42	10	3.66	2,058
Surviving children[a]	4.47	4	2.49	2,058
Grandchildren	45.99	43	27.40	2,058
Great-grandchildren	187.65	159	142.74	2,058
Great-great-grandchildren[b]	341.04	206.5	408.07	2,058
Years from marriage to first birth (TFB)	1.23	1.04	0.49	2,058
Literate	0.68	1	0.47	1,192
Fraction of literate children	0.72	1	0.36	1,872
Fraction of surviving children[a]	0.48	0.50	0.21	2,058
Fraction of surviving children with known literacy[a]	0.79	0.75	0.57	2,016
Age at first marriage	19.34	18.7	3.79	2,058
Age at last delivery	38.27	40.3	6.46	2,058
MALES				
Children	10.03	10	4.32	1,740
Surviving children[a]	4.82	5	2.68	1,740
Grandchildren	48.94	45	28.77	1,740
Great-grandchildren	187.53	159	137.10	1,740
Great-great-grandchildren[b]	238.38	136.5	293.17	1,740
Years from marriage to first birth (TFB)	1.16	0.99	0.44	1,740
Literate	0.64	1	0.48	1,030
Fraction of literate children	0.76	1	0.34	1,576
Fraction of surviving children[a]	0.49	0.50	0.21	1,740
Fraction of surviving children with known literacy[a]	0.73	0.67	0.54	1,711
Age at first marriage	26.62	25.9	4.41	1,740
Age at last delivery	46.31	46.9	8.81	1,740
ALL				
Children	9.70	10	3.99	3,798
Surviving children[a]	4.63	4	2.59	3,798
Grandchildren	47.35	44	28.07	3,798
Great-grandchildren	187.59	159	140.17	3,798
Great-great-grandchildren[b]	294.01	171	363.58	3,798
Years from marriage to first birth (TFB)	1.20	1.02	0.47	3,798
Literate	0.66	1	0.47	2,222
Fraction of literate children	0.74	1	0.35	3,448
Fraction of surviving children[a]	0.49	0.50	0.21	3,798
Fraction of surviving children with known literacy[a]	0.76	0.67	0.56	3,727
Age at first marriage	22.67	22.2	5.46	3,798
Age at last delivery	41.95	42.1	8.61	3,798

[c] Survival is recorded at age 40.

[b] The moderate increase in the mean and median number of descendants from the third to the fourth generation (i.e. from great-grandchildren to great-great-grandchildren) reflects the fact that these cohorts are less fully observed. Furthermore, since men produce children at later ages than women, this effect is more pronounced among men.

Robustness: Second Births and Marriage Age > 15 Years

Table C.2: Robustness to additional sample restrictions: more than one child and marriage occurring after turning 15 years – accounting for Maternal Founder fixed effects

| | Log number of descendants in: | | | | | | | |
| | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 4 | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 4 |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TFB	-.053**	.204	.560***	.891***	-.052**	.129	.461**	.779***
	(.021)	(.157)	(.201)	(.263)	(.025)	(.175)	(.218)	(.279)
TFB^2		-.082	-.211***	-.344***		-.057	-.180**	-.310***
		(.051)	(.066)	(.086)		(.058)	(.073)	(.093)
Literate	.025	.090**	.175***	.162**	-.016	.052	.139***	.123*
	(.027)	(.038)	(.050)	(.065)	(.031)	(.041)	(.052)	(.068)
Male	.251***	.272***	.314***	.135**	.225***	.269***	.318***	.153**
	(.029)	(.038)	(.047)	(.061)	(.031)	(.039)	(.048)	(.061)
Number of observations	3,738	3,738	3,738	3,738	3,604	3,604	3,604	3,604
Adjusted R^2	.044	.037	.060	.318	.035	.031	.055	.314
Joint sign.-level of TFB & TFB^2	.011	.104	.001	.000	.038	.294	.009	.000
Maximizing TFB		1.248	1.323	1.296		1.140	1.282	1.256
Lower limit of 90% CI		-	1.059	1.095		-	.798	.973
Upper limit of 90% CI		-	1.464	1.417		-	1.458	1.396

This table presents the results of a series of fixed-effects regressions of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2. All regressions account for Maternal Founder fixed effects. Birth year and marriage age dummies are included as controls. A dummy indicating unknown literacy is also included in the regressions. In columns 1–4, the sample is restricted to heads of lineages who produced at least two children. In columns 5–8, the sample is restricted to heads of lineages who were married after their 15th birthday. Standard errors clustered at the level of the firstborn are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Robustness: Extinct Lineages

Table C.3: Robustness to including extinct lineages – accounting for Maternal Founder fixed effects

	Log number of descendants in:				Log 1+number of descendants in:			
	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TFB	-.046*	.151	.462**	.773***	-.041*	.414*	.822***	1.169***
	(.026)	(.172)	(.208)	(.264)	(.021)	(.215)	(.294)	(.336)
TFB^2		-.071	-.180***	-.309***		-.149**	-.293***	-.427***
		(.057)	(.068)	(.087)		(.068)	(.093)	(.108)
Literate	-.022	.060	.153***	.138**	-.014	.058	.144*	.180**
	(.033)	(.042)	(.053)	(.066)	(.027)	(.056)	(.076)	(.087)
Male	.321***	.321***	.349***	.131**	.276***	.456***	.597***	.478***
	(.033)	(.040)	(.050)	(.060)	(.027)	(.051)	(.069)	(.076)
Number of observations	4,240	4,002	3,933	3,798	4,240	4,240	4,240	4,240
Adjusted R^2	.052	.044	.068	.307	.054	.065	.084	.241
Joint sign.-level of TFB & TFB^2	.073	.083	.003	.000	.053	.046	.002	.000
Maximizing TFB		1.067	1.283	1.249		1.387	1.403	1.368

This table presents the results of a series of fixed-effects regressions of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2. All regressions account for Maternal Founder fixed effects. Birth year and marriage age dummies are included as controls. A dummy indicating unknown literacy is also included in the regressions. In columns 1–4, the restriction of at least one observed great-great-grandchild is relaxed and extinct lineages drop out of the sample in the relevant generations. In columns 5–8, the same extended sample is used, but the outcome is $\ln(1 + D_{i,t})$, where the added number 1 ensures that the logarithmic transformation is defined and all lineages remains included. Standard errors clustered at the level of the firstborn are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Robustness: Number of Marriages and Spousal Migration

Table C.4: Robustness to additional control variables: number of marriages and spousal migration – accounting for Maternal Founder fixed effects

	Log number of descendants in:							
	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TFB	-.059***	.282*	.590***	.828***	-.059**	.117	.423**	.720***
	(.022)	(.157)	(.200)	(.263)	(.024)	(.161)	(.204)	(.264)
TFB^2		-.109**	-.225***	-.328***		-.054	-.168**	-.289***
		(.051)	(.066)	(.087)		(.052)	(.067)	(.087)
Literate	-.013	.059	.143***	.135**	-.007	.067*	.148***	.140**
	(.028)	(.039)	(.050)	(.066)	(.030)	(.040)	(.051)	(.066)
Male	.168***	.204***	.255***	.116*	.198***	.206***	.232***	.037
	(.028)	(.037)	(.046)	(.060)	(.032)	(.039)	(.047)	(.061)
Total number of marriages fixed effects	Yes	Yes	Yes	Yes	No	No	No	No
Total number of marriages of spouse fixed effects	Yes	Yes	Yes	Yes	No	No	No	No
Immigration status of spouse fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Emigration status of spouse fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Number of observations	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798
Adjusted R^2	.188	.116	.097	.312	.035	.035	.062	.315
Joint sign.-level of TFB & TFB^2	.009	.027	.000	.000	.013	.169	.005	.000
Maximizing TFB		1.289	1.311	1.263		1.073	1.260	1.244
Lower limit of 90% CI		.420	1.063	1.024		-	.716	.935
Upper limit of 90% CI		1.509	1.446	1.393		-	1.445	1.392

This table presents the results of a series of fixed-effects regressions of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2. All regressions account for Maternal Founder fixed effects. Birth year and marriage age dummies are included as controls. A dummy indicating unknown literacy is also included in the regressions. In columns 1–4, dummies for the total number of marriages experienced during the lifetime of the heads of lineages, as well as dummies for the total number of marriages experienced by the first spouses of the heads of lineages, are included. In columns 5–8, dummies indicating the immigration and emigration statuses of the head of the first spouses of the heads of lineages are included. Standard errors clustered at the level of the firstborn are reported in parentheses. * p <0.10, ** p <0.05, *** p <0.01.

Robustness: Gender

Table C.5: Robustness to gender distinction – sample restricted to females – accounting for Maternal Founder fixed effects

| | Log number of descendants in: | | | | | | | |
	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TFB	-.079***	.374*	.985***	1.568***	-.080***	.380*	.993***	1.581***
	(.030)	(.214)	(.278)	(.349)	(.030)	(.215)	(.278)	(.349)
TFB^2		-.140**	-.347***	-.560***		-.142**	-.349***	-.564***
		(.069)	(.088)	(.109)		(.069)	(.088)	(.110)
Literate					-.040	.072	.120	.145
					(.043)	(.067)	(.084)	(.107)
Number of observations	2,058	2,058	2,058	2,058	2,058	2,058	2,058	2,058
Adjusted R^2	.089	.064	.091	.267	.089	.064	.092	.267
Joint sign.-level of TFB & TFB^2	.008	.057	.000	.000	.007	.056	.000	.000
Maximizing TFB		1.334	1.418	1.401		1.339	1.422	1.402
Lower limit of 90% CI		.366	1.24	1.263		.428	1.247	1.267
Upper limit of 90% CI		1.58	1.538	1.500		1.583	1.542	1.501

This table presents the results of a series of fixed-effects regressions of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2 for female heads of lineages. All regressions account for Maternal Founder fixed effects. Birth year and marriage age dummies are included as controls. A dummy indicating unknown literacy is also included in the regressions. Standard errors clustered at the level of the firstborn are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Robustness: Location of Birth and Death

Table C.6: Robustness to additional control variables: location of birth and death – accounting for Maternal Founder fixed effects

| | Log number of descendants in: | | | | | | | |
| | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 4 | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 4 |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TFB	-0.062***	0.106	0.416**	0.695***	-0.062***	0.032	0.391*	0.653**
	(0.024)	(0.163)	(0.205)	(0.267)	(0.023)	(0.161)	(0.202)	(0.255)
TFB^2		-0.052	-0.168**	-0.284***		-0.023	-0.154**	-0.262***
		(0.053)	(0.067)	(0.088)		(0.053)	(0.066)	(0.084)
Literate	-0.008	0.062	0.149***	0.144**	-0.015	0.037	0.097*	0.074
	(0.031)	(0.040)	(0.051)	(0.066)	(0.031)	(0.040)	(0.052)	(0.067)
Male	0.214***	0.253***	0.298***	0.134**	0.203***	0.208***	0.226***	0.033
	(0.031)	(0.039)	(0.048)	(0.061)	(0.031)	(0.038)	(0.046)	(0.059)
Constant	3.861***	5.949***	6.288***	6.228***	2.256***	4.130***	5.229***	4.676***
	(0.366)	(0.649)	(0.667)	(0.935)	(0.428)	(0.681)	(0.703)	(0.921)
Birth parish fixed effects	Yes	Yes	Yes	Yes	No	No	No	No
Death parish fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Number of observations	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798
Adjusted R^2	0.034	0.030	0.050	0.308	0.066	0.098	0.130	0.370
Joint sign.-level of TFB & TFB^2	.009	.155	.002	.000	.009	.425	.011	.000
Maximizing TFB		1.024	1.234	1.223		.700	1.264	1.243
Lower limit of 90% CI		-	.646	.875		-	.607	.892
Upper limit of 90% CI		-	1.424	1.379		-	1.461	1.402

This table presents the results of a series of fixed-effects regressions of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2. All regressions account for Maternal Founder fixed effects. Birth year and marriage age dummies are included as controls. A dummy indicating unknown literacy is also included in the regressions. In columns 1–4, dummies indicating the birth (or baptism) parish of the heads of lineages are included. In columns 5–8 dummies indicating the death (or burial) parish of the heads of lineages are included. Standard errors clustered at the level of the firstborn are reported in parentheses. * $p <0.10$, ** $p <0.05$, *** $p <0.01$.

Robustness: Month of Marriage and First Birth

Table C.7: Robustness to additional control variables: month of marriage and first birth –
accounting for Maternal Founder fixed effects

| | Log number of descendants in: | | | | | | | |
| | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 4 | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 4 |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TFB	-0.061**	0.151	0.482**	0.798***	-0.050**	0.181	0.500**	0.787***
	(0.024)	(0.163)	(0.205)	(0.265)	(0.024)	(0.166)	(0.208)	(0.274)
TFB^2		-0.067	-0.190***	-0.318***		-0.073	-0.194***	-0.314***
		(0.053)	(0.067)	(0.087)		(0.054)	(0.068)	(0.089)
Literate	-0.008	0.062	0.146***	0.137**	-0.004	0.065	0.153***	0.140**
	(0.030)	(0.040)	(0.051)	(0.067)	(0.030)	(0.040)	(0.051)	(0.067)
Male	0.224***	0.255***	0.296***	0.123**	0.220***	0.257***	0.298***	0.129**
	(0.031)	(0.039)	(0.048)	(0.060)	(0.031)	(0.039)	(0.047)	(0.060)
Month of marriage fixed effects	Yes	Yes	Yes	Yes	No	No	No	No
Month of birth of firstborn fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Number of observations	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798
Adjusted R^2	0.032	0.028	0.051	0.307	0.032	0.032	0.056	0.308
Joint sign.-level of TFB & TFB^2	.010	.109	.001	.000	.037	.169	.001	.000
Maximizing TFB		1.126	1.267	1.253		1.242	1.286	1.254
Lower limit of 90% CI		-	.865	.994		-	.904	.968
Upper limit of 90% CI		-	1.433	1.389		-	1.452	1.398

This table presents the results of a series of fixed-effects regressions of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2. All regressions account for Maternal Founder fixed effects. Birth year and marriage age dummies are included as controls. A dummy indicating unknown literacy is also included in the regressions. In columns 1–4, dummies indicating the months of marriage of the heads of lineages are included. In columns 5–8, dummies indicating the month of birth of the the firstborn of the heads of lineages are included. Standard errors clustered at the level of the firstborn are reported in parentheses. * p <0.10, ** p <0.05, *** p <0.01.

Robustness: Birth Order

Table C.8: Robustness to additional control variable: birth order – accounting for Maternal Founder fixed effects

| | Log number of descendants in: | | | | | | | |
| | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 4 | Gen. 1 | Gen. 2 | Gen. 3 | Gen. 4 |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TFB	-.062***	.139	.462**	.773***	-.062***	.144	.456**	.752***
	(.024)	(.162)	(.204)	(.264)	(.024)	(.162)	(.204)	(.265)
TFB^2		-.062	-.183***	-.309***		-.063	-.181***	-.303***
		(.053)	(.067)	(.087)		(.053)	(.067)	(.087)
Literate	-.006	.064	.149***	.137**	-.008	.064	.149***	.141**
	(.030)	(.040)	(.051)	(.066)	(.030)	(.040)	(.051)	(.067)
Male	.219***	.250***	.296***	.132**	.220***	.248***	.293***	.131**
	(.031)	(.039)	(.047)	(.060)	(.031)	(.039)	(.047)	(.061)
Firstborn	.016	.064**	.042	-.019				
	(.023)	(.031)	(.038)	(.047)				
Birth order fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Number of observations	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798
Adjusted R^2	.032	.030	.052	.307	.031	.030	.052	.308
Joint sign.-level of TFB & TFB^2	.010	.135	.002	.000	.009	.139	.002	.000
Maximizing TFB		1.117	1.264	1.249		1.133	1.261	1.240
Lower limit of 90% CI		-	.828	.976		-	.807	.949
Upper limit of 90% CI		-	1.436	1.388		-	1.434	1.384

This table presents the results of a series of fixed-effects regressions of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2. All regressions account for Maternal Founder fixed effects. Birth year and marriage age dummies are included as controls. A dummy indicating unknown literacy is also included in the regressions. In columns 1–4, a dummy for the firstborn status of the heads of lineages is included. In columns 5–8, dummies indicating the birth order of the heads of lineages are included. Standard errors clustered at the level of the firstborn are reported in parentheses. * $p <0.10$, ** $p <0.05$, *** $p <0.01$.

Robustness: GLM Regression

Table C.9: Robustness of GLM regression to the exclusion of Maternal Founder fixed effects

	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4	Gen. 1	Gen. 2	Gen. 3	Gen. 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
TFB	-.063***	.120	.371**	.616***	-.063***	.090	.324*	.573**	-.081***	.028	.260	.525**	-.090***	.049	.346**	.573**
	(.018)	(.148)	(.185)	(.230)	(.018)	(.148)	(.186)	(.231)	(.017)	(.145)	(.184)	(.231)	(.010)	(.121)	(.167)	(.227)
TFB^2		-.047	-.133**	-.232***		-.039	-.119*	-.219***		-.026	-.106*	-.209***		-.035	-.137**	-.224***
		(.048)	(.060)	(.074)		(.049)	(.061)	(.074)		(.048)	(.060)	(.074)		(.039)	(.054)	(.073)
Literate					-.002	.143***	.222***	.227***	-.009	.138***	.219***	.225***	-.056***	.098***	.180***	.195***
					(.020)	(.030)	(.038)	(.046)	(.019)	(.030)	(.038)	(.046)	(.011)	(.025)	(.035)	(.047)
Male									.295***	.379***	.402***	.306***	-.004	.094***	.150***	.194***
									(.019)	(.027)	(.035)	(.041)	(.011)	(.023)	(.032)	(.042)
Stoppage age fixed effects	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Number of observations	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798	3,798
Joint sign.-level of TFB & TFB^2	.000	.446	.059	.000	.000	.445	.086	.001	.000	.139	.041	.000	.000	.016	.001	.000
Maximizing TFB		1.273	1.399	1.328		1.159	1.363	1.306		.546	1.229	1.255		.702	1.262	1.277

This table presents the results of a series of GLM regressions, with a negative binomial distribution and a logarithmic link function, of the number of descendants in generation t on time to first birth measured in years, i.e. TFB and TFB^2. Birth year and marriage age dummies are included as controls. Furthermore, stoppage age dummies are included in columns 13–16. A dummy indicating unknown literacy is included in the regressions underlying column 5–16. Standard errors clustered at the level of the firstborn are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Robustness: Number of Children and Long-Run Reproductive Success

Table C.10: Effect of number of children on long-run reproductive success – accounting for Maternal Founder fixed effects

| | Log number of descendants in: | | | | | | | |
| | Gen. 2 | Gen. 3 | Gen. 4 | Gen. 2 | Gen. 3 | Gen. 4 | Gen. 2 | Gen. 3 | Gen. 4 |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log Number of Children	2.334***	3.446***	4.491***	2.339***	3.459***	4.545***	2.384***	3.515***	4.619***
	(.524)	(.842)	(1.078)	(.525)	(.851)	(1.091)	(.520)	(.837)	(1.075)
(Log Number of Children)^2	-.371	-.728**	-1.280***	-.376	-.738**	-1.306***	-.388*	-.749**	-1.324***
	(.229)	(.367)	(.474)	(.230)	(.373)	(.481)	(.228)	(.366)	(.474)
Literate				.068	.151**	.171*	.067	.150**	.170*
				(.047)	(.070)	(.092)	(.047)	(.070)	(.092)
Male							.129***	.266***	.307***
							(.046)	(.071)	(.092)
Number of observations	4,240	4,240	4,240	4,240	4,240	4,240	4,240	4,240	4,240
First stage F (Kleibergen-Paap)	30.701	30.701	30.701	33.094	33.094	33.094	31.891	31.891	31.891
Joint sign. of linear and squared terms	.000	.000	.000	.000	.000	.000	.000	.000	.000
Maximizing number of $\log(1 + D)$	3.142	2.367	1.754	3.113	2.343	1.74	3.069	2.345	1.744
Exp(Maximizing $\log(1 + D)$) $- 1$	22.15	9.665	4.777	21.488	9.412	4.697	20.52	9.433	4.720
Lower limit of 90% CI	1.618	1.770	1.473	1.473	1.756	1.465	2.081	1.767	1.472
Upper limit of 90% CI	6.012	8.556	2.805	2.805	8.445	2.757	57.776	7.475	2.702

This table presents the results of a series of fixed-effects 2SLS regressions of the log number of descendants in generation t, i.e. $\ln(1 + D_{i,t})$, where $D_{i,t}$ is the number of descendants that the head of household i has in generations t, $t = 2, 3, 4$, on the number of children and the number of children squared, i.e., $\ln(1 + D_{i,1})$ and $(\ln(1 + D_{i,1}))^2$, instrumented by variation in TFB. The added number 1 ensures that the logarithmic transformation is defined for extinct lineages and all that lineages remain in the sample. All regressions account for Maternal Founder fixed effects. Birth year, marriage age, and stoppage age dummies are included as controls. Since the second stage of these 2SLS regressions is quadratic in the endogenous regressor, it is necessary to instrument for both the linear and the squared terms in order to identify the parameters. Thus, following (Wooldridge, 2010, pp. 267–268), a zeroth stage is introduced to the analysis, where $D_{i,1}$ is first regressed on TFB and all the second-stage controls, accounting for Maternal Origins fixed effects, to obtain predicted values of the number of children. The predicted number of children from the zeroth stage, $\hat{D}_{1,i}$ is transformed by $\ln(\hat{D}_{1,i})$ as well as $(\ln(\hat{D}_{1,i}))^2$, and these transformed terms are then used as excluded instruments in the second stage. A dummy indicating unknown literacy is included in the regressions underlying columns 4–9. Standard errors clustered at the level of the firstborn are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

References

ASHRAF, Q. AND O. GALOR (2013): "The Out of Africa Hypothesis, Human Genetic Diversity, and Comparative Economic Development," *The American Economic Review*, 103, 1–46.

ASHRAF, Q., O. GALOR, AND M. KLEMP (2014): "The Out of Africa Hypothesis of Comparative Development Reflected by Nighttime Light Intensity," *Brown Working Papers*, 2014.

BAIRD, D. T., J. COLLINS, J. EGOZCUE, L. H. EVERS, L. GIANAROLI, H. LERIDON, A. SUNDE, A. TEMPLETON, A. VAN STEIRTEGHEM, J. COHEN, P. G. CROSIGNANI, P. DEVROEY, K. DIEDRICH, B. C. FAUSER, L. FRASER, A. GLASIER, I. LIEBAERS, G. MAUTONE, G. PENNEY, B. TARLATZIS, AND COAUTHORS (2005): "Fertility and ageing," *Human Reproduction Update*, 11, 261–276.

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.

BISIN, A. AND T. VERDIER (2000): ""Beyond the melting pot": cultural transmission, marriage, and the evolution of ethnic and religious traits," *The Quarterly Journal of Economics*, 115, 955–988.

BORGERHOFF MULDER, M. (2000): "Optimizing offspring: the quantity–quality tradeoff in agropastoral Kipsigis," *Evolution and Human Behavior*, 21, 391–410.

BOUCEKKINE, R., D. DE LA CROIX, AND D. PEETERS (2007): "Early literacy achievements, population density, and the transition to modern growth," *Journal of the European Economic Association*, 5, 183–226.

BOWLES, S. (1998): "Endogenous preferences: The cultural consequences of markets and other economic institutions," *Journal of economic literature*, 36, 75–111.

BOYD, R. (1988): *Culture and the evolutionary process*, University of Chicago Press.

CAVALLI-SFORZA, L. L. L. (1981): *Cultural transmission and evolution: a quantitative approach*, 16, Princeton University Press.

CERVELLATI, M. AND U. SUNDE (2005): "Human Capital, Life Expectancy and the Process of Development," *The American Economic Review*, 95, 1653–1672.

CHARNOV, E. L. AND S. K. M. ERNEST (2006): "The offspring-size/clutch-size trade-off in mammals," *The American Naturalist*, 167, 578–582.

CHETTY, R., J. N. FRIEDMAN, AND J. E. ROCKOFF (2014): "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *American Economic Review*, 104, 2593–2632.

CHRISTENSEN, K., H.-P. KOHLER, O. BASSO, J. OLSEN, J. W. VAUPEL, AND J. L. RODGERS (2003): "The Correlation of Fecundability Among Twins: Evidence of a Genetic Effect on Fertility?" *Epidemiology*, 14, 60–64.

CODY, M. L. (1966): "A general theory of clutch size," *Evolution*, 174–184.

DALGAARD, C.-J. AND H. STRULIK (2011): "The Physiological Foundations of the Wealth of Nations," Discussion Paper dp-480, Leibniz Universität Hannover, Wirtschaftswissenschaftliche Fakultät.

DOEPKE, M. AND F. ZILIBOTTI (2008): "Occupational Choice and the Spirit of Capitalism," *The Quarterly Journal of Economics*, 123, 747–793.

DUAN, N. (1983): "Smearing estimate: a nonparametric retransformation method," *Journal of the American Statistical Association*, 78, 605–610.

DURHAM, W. H. (1982): "Interactions of genetic and cultural evolution: Models and examples," *Human Ecology*, 10, 289–323.

GALOR, O. (2011): *Unified Growth Theory*, Princeton: Princeton University Press.

GALOR, O. AND S. MICHALOPOULOS (2012): "Evolution and the growth process: Natural selection of entrepreneurial traits," *Journal of Economic Theory*, 147, 759–780.

GALOR, O. AND O. MOAV (2002): "Natural selection and the origin of economic growth," *Quarterly Journal of Economics*, 117, 1133–1191.

GALOR, O. AND O. MOAV (2007): "The Neolithic revolution and contemporary variations in life expectancy," *Brown University Department of Economics Working Paper*, 14.

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.

GILLESPIE, D., A. RUSSELL, AND V. LUMMAA (2008): "When fecundity does not equal fitness: evidence of an offspring quantity versus quality trade-off in pre-industrial humans," *Proceedings of the Royal Society B: Biological Sciences*, 275, 713–722.

HANSEN, G. D. AND E. C. PRESCOTT (2002): "Malthus to Solow," *American Economic Review*, 92, 1205–1217.

HARPER, J., P. LOVELL, AND K. MOORE (1970): "The shapes and sizes of seeds," *Annual review of ecology and systematics*, 1, 327–356.

HILL, K. AND A. HURTADO (1996): *Ache life history: The ecology and demography of a foraging people*, Aldine de Gruyter.

KAPLAN, H., J. LANCASTER, S. JOHNSON, AND J. BOCK (1995): "Does observed fertility maximize fitness among New Mexican men?" *Human Nature*, 6, 325–360.

KOSOVA, G., M. ABNEY, AND C. OBER (2009): "Heritability of reproductive fitness traits in a human population," *PNAS*, 107, 1772–1778.

LACK, D. ET AL. (1954): "The natural regulation of animal numbers." *The Natural Regulation of Animal Numbers.*

LAGERLÖF, N.-P. (2007): "Long-run trends in human body mass," *Macroeconomic Dynamics*, 11, 367–387.

LEE, R. D. (1993): "Population dynamics: Equilibrium, disequilibrium, and consequences of fluctuations," in *Handbook of Population and Family Economics*, ed. by M. R. Rosenzweig and O. Stark, Elsevier, vol. 1, chap. 19, 1063–1115, 1 ed.

LIND, J. T. AND H. MEHLUM (2010): "With or Without U? The Appropriate Test for a U-Shaped Relationship," *Oxford Bulletin of Economics and Statistics*, 72, 109–118.

LIVINGSTONE, F. B. (1958): "Anthropological Implications of Sickle Cell Gene Distribution in West Africa," *American Anthropologist*, 60, 533–562.

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

MEIJ, J., D. VAN BODEGOM, J. ZIEM, J. AMANKWA, A. POLDERMAN, T. KIRKWOOD, A. DE CRAEN, B. ZWAAN, AND R. WESTENDORP (2009): "Quality–quantity trade-off of human offspring under adverse environmental conditions," *Journal of evolutionary biology*, 22, 1014–1023.

MILOT, E., F. M. MAYER, D. H. NUSSEY, M. BOISVERT, F. PELLETIER, AND D. RÉALE (2011): "Evidence for evolution in response to natural selection in a contemporary human population," *Proceedings of the National Academy of Sciences*, 108, 17040–17045.

PAPKE, L. E. AND J. M. WOOLDRIDGE (1996): "Econometric Methods for Fractional Response Variables with an Application to 401 (K) Plan Participation Rates," *Journal of Applied Econometrics*, 11, 619–32.

PAPKE, L. E. AND J. M. WOOLDRIDGE (2008): "Panel data methods for fractional response variables with an application to test pass rates," *Journal of Econometrics*, 145, 121–133.

PETTAY, J. E., L. E. B. KRUUK, J. JOKELA, AND V. LUMMAA (2005): "Heritability and genetic constraints of life-history trait evolution in preindustrial humans," *PNAS*, 102, 2838–2843.

RAMLAU-HANSEN, C. H., A. M. THULSTRUP, J. OLSEN, AND J. P. BONDE (2008): "Parental Subfecundity and Risk of Decreased Semen Quality in the Male Offspring: A Follow-up Study," *American Journal of Epidemiology*, 167, 1458–1464.

ROFF, D. (1992): *Evolution of life histories: theory and analysis*, Routledge, Chapman and Hall.

ROFF, D. A. (2002): *Life History Evolution*, Sunderland, MA: Sinauer Associates.

SALISBURY, E. ET AL. (1942): "The reproductive capacity of plants. Studies in quantitative biology." *The reproductive capacity of plants. Studies in quantitative biology.*

SPOLAORE, E. AND R. WACZIARG (2009): "The diffusion of development," *The Quarterly Journal of Economics*, 124, 469–529.

STEARNS, S. (1992): *The evolution of life histories*, vol. 248, Oxford University Press Oxford.

STRASSMANN, B. AND B. GILLESPIE (2002): "Life–history theory, fertility and reproductive success in humans," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269, 553–562.

VOIGHT, B., S. KUDARAVALLI, X. WEN, AND J. PRITCHARD (2006): "A map of recent positive selection in the human genome," *PLoS Biology*, e72.

WALKER, R. M., M. GURVEN, O. BURGER, AND M. J. HAMILTON (2008): "The trade-off between number and size of offspring in humans and other promates," *Proceedings of the Royal Society*, 275, 827–833.

WEIBULL, J. W. (1997): *Evolutionary game theory*, MIT press.

WIESENFELD, S. L. (1967): "Sickle-Cell Trait in Human Biological and Cultural Evolution Development of agriculture causing increased malaria is bound to gene-pool changes causing malaria reduction," *Science*, 157, 1134–1140.

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