

# FECUNDITY, FERTILITY AND THE FORMATION OF HUMAN CAPITAL

## Online Appendix

*Marc Klemp and Jacob Weisdorf*

### A Additional Text

#### A.1 Variables

##### A.1.1 Independent Variables

**Time to first birth (i.e., protogenesic interval).** The main independent variable is the time, measured in years, from the parental marriage date to the birth date of the first sibling.

**Number of surviving siblings.** On the premise that surviving children consume more parental resources than those dying during childhood, the independent variable of interest here is family size measured as the number of siblings surviving to age five.<sup>1</sup>

##### A.1.2 Control Variables on the Individual Level

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

**Birth order.** Many studies have linked birth order with human capital achievements, both in historical and in present times (see e.g., Ejrnæs and Pörtner, 2004; Black et al., 2005; Klemp et al., 2013). Due to the mechanical association between the number of siblings and birth order, and due to the fact that we are interested in the effect of the protogenesic interval on human capital independently of birth order effects, we include dummies for each birth order to account for birth order fixed effects in all the main regression analyses.<sup>2</sup>

**Non-Sunday baptism.** Although the Prayer Books of the English Church prescribed that baptisms take place on Sundays, many families did not submit to this rule. Non-Sunday baptism

---

<sup>1</sup>Table C4 in the appendix establishes that the results are robust to the inclusion of all births in the calculation of the family size rather than children surviving to age five.

<sup>2</sup>Appendix Table C10 establishes that the findings are robust to excluding these dummies, as well as to controlling for birth order effects in other ways. Furthermore, that table also establishes that controlling for birth order is the most conservative approach.

services were possible for an additional fee, which means that non-Sunday baptisms might positively reflect family income.<sup>3</sup> Meanwhile, the occurrence of a non-Sunday baptism could also indicate a perceived higher risk of infant death and hence an immediate baptism. The occurrence of a non-Sunday baptism is therefore included in the regression analysis to account for these possibly confounding factors.

### A.1.3 Control Variables on the Family and Parish Level

By controlling for measures of parental human capital attainments that are similar to those of the offspring, the analysis accounts for relevant educational heterogeneity between families.<sup>4</sup>

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

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

**Paternal occupational wealth.** Fathers with occupations that are associated with higher wealth can simultaneously afford larger families and devote more resources to their offspring. Furthermore, fathers holding those occupations may have an increased propensity and ability to direct their children towards similar professions. Since occupational wealth and skills are both based on occupational titles, paternal occupational wealth is divided into two main categories, with labourers and husbandmen making up the poorest segments of the English society.<sup>5</sup> Thus, a dummy variable indicating paternal occupational wealth is included in the regression analysis to account for the possibly confounding factors associated with paternal wealth, along with a dummy variable indicating unobserved occupational wealth.<sup>6</sup>

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

---

<sup>3</sup>The fact that non-Sunday baptisms were often requested by affluent families is supported by the positive associations between a non-Sunday baptism and the number of surviving siblings as well as their level of human capital, as established in the regression analyses below.

<sup>4</sup>The parental human capital control variables enter as dummy variables. This allows the inclusion of families with unobserved parental human capital, captured by a dummy variable indicating missing information.

<sup>5</sup>Table C9 in the appendix establishes that the results are robust to the inclusion of dummies for all seven paternal wealth groups.

<sup>6</sup>Given the low female labour participation rate, information on maternal occupational wealth is omitted.

indicating the time period of marriage of the parents (for 20-year intervals) is included in the regression analysis to account for marriage time-period fixed effects.<sup>7</sup>

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

**Local occupational structures.** The sampled parishes range from market towns to remote rural villages and have been organised by Schofield (2005) in four groups: “agriculture”, “industry”, “retail and handicraft” and “other” (a mix). The local occupational structure may affect the return to different types of human capital investments and may furthermore be correlated with genetic differences determining fecundity (see e.g. Juul et al., 1999, for related evidence for modern populations). Thus, dummies capturing the four different types of occupational structure are introduced to account for these confounding effects.<sup>8</sup>

## A.2 Socio-economic Elite Classification

As mentioned in Section 3.3, we classify families in which paternal literacy or skills are unobserved as not being part of the socio-economic elite. This is consistent with the following cross-generational correlations observed in the union of the samples in Table 3 ( $N = 1,517$ ). There is a positive and significant correlation (i.e., a Pearson correlation coefficient of 65% with  $p < 0.001$ ) between a dummy variable indicating if an individual is both literate and skilled and a similar paternal dummy variable in the subset of observations in which both paternal literacy and skills are known, indicating that children of literate and skilled fathers are more likely to become literate and skilled themselves. Furthermore, there is a negative and significant correlation (i.e., a coefficient of -29% with  $p < 0.001$ ) between the dummy variable for the individual and a dummy variable indicating unknown paternal literacy or skills.

Table C14 in the appendix examines the robustness of the results to controlling for unknown information on paternal literacy or occupational skills as well as to accounting for long protogenetic intervals.

---

<sup>7</sup>The results are robust to accounting for marriage year (i.e., 1-year interval) fixed effects (see Table C12 in the appendix).

<sup>8</sup>Table C11 in the appendix demonstrates that the results are robust to controlling for parish-level fixed effects.

## B Additional Figures



Figure B1: Map of England with the locations of the parishes.

## C Additional Tables

Table C1: Example Family

		PARISH LEVEL INFORMATION		FAMILY LEVEL INFORMATION					
		Parish	Occupational Type	Years to First Birth					
		Odiham	Mixed	0.92					
INDIVIDUAL LEVEL INFORMATION									
Family Member	Name	Birth Date	Death Date	Age at Death	Occupation	Skilled Profession	Occupational Wealth	Literate	Marriage Age
Mother	Hannah Sury	21 July 1740	10 Nov. 1816	76.3	-	-	-	No	21.2
Father	Edward Neville	14 May 1733	3 Nov. 1816	83.5	Labourer	No	1	No	28.4
Daughter	Ann	8 Oct. 1762	-	-	-	-	-	No	-
Son	John	17 Apr. 1765	13 Oct. 1850	85.5	Sawyer	Yes	2	No	-
Son	Edward	3 Mar. 1767	8 May 1852	85.2	Baker	Yes	4	Yes	-
Son	James	3 May 1769	14 Apr.1849	79.9	Labourer	No	1	No	-
Son	Thomas	6 Mar. 1771	20 Mar. 1771	0.0	-	-	-	-	-
Son	Daved	28 Mar. 1773	13 May 1858	85.1	-	-	1	No	-
Son	Thomas	23 Apr. 1775	21 Dec. 1855	80.7	Sawyer	Yes	2	No	-
Son	Francis	8 June 1777	9 May 1780	2.9	-	-	-	-	-
Daughter	Hannah	5 Dec. 1779	-	-	-	-	-	-	-

“Years to First Birth” is the length of the protogenetic interval, i.e., the length of time from the marriage to the first birth, measured in years. “-” indicates missing information.

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

Table C2: Robustness to Cutoff at 38 Weeks (Instead of 40 Weeks)

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

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C3: Summary Statistics

	Mean	S.D.	Count	P10	P90
Surviving Siblings (> 5 Years)	4.83	2.51	1517	2	8
Literate	0.56	0.50	1248	0	1
Skilled	0.68	0.47	652	0	1
Years to First Birth	1.58	1.18	1517	0.81	2.99
Male	0.53	0.50	1517	0	1
Non-Sunday Baptism	0.53	0.50	1485	0	1
Skilled Father	0.69	0.46	925	0	1
Skilled Mother	0.63	0.49	35	0	1
Literate Father	0.60	0.49	969	0	1
Literate Mother	0.32	0.47	942	0	1
Longevity of Father (Years)	72.37	9.79	1517	58.96	84.08
Age of Marriage of Mother (Years)	25.08	4.68	1517	19.84	31.04
Agricultural Location	0.25	0.43	1517	0	1
Industrial Location	0.24	0.43	1517	0	1
Retail Location	0.16	0.36	1517	0	1
Birth Year	1771.22	38.19	1517	1733	1807
N. of Observations	1,517				

Table C4: Robustness to Use of all Births Instead of Surviving Births (Instrumental Variable Regression)

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

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C5: Robustness to Birth Year

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

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C6: Robustness to Inclusion of Paupers and Gentry in the Analysis of Skills

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

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C7: Robustness to Paternal Marriage Age

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.031** (.012)	.032** (.012)	.032*** (.011)	.040*** (.014)	.039*** (.013)	.039*** (.013)	.120*** (.043)	.122*** (.042)	.113*** (.042)
Paternal Marriage Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
$R^2$	.170	.181	.233	.286	.325	.334	.315	.338	.359
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, paternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

The paternal marriage age may have an effect on the protogenetic interval and offspring human capital achievement. Thus, dummy variables indicating the paternal marriage age (on the 5-year level) are included in the regression analysis to account for this potentially confounding factor. While the coefficients on these additional control variables are omitted from the table, the specifications underlying each column in the table correspond to the columns in Table 2, except for the additional inclusion of paternal marriage age 5-year group dummy variables. The table establishes that the qualitative conclusion is robust to these alternative specifications. In particular, the association between the protogenetic interval and the three measures of human capital achievements remains significant in all specifications.

Table C8: Robustness to Paternal Marriage Age (Instrumental Variable Regression)

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Surviving Siblings (> 5 Years)	-.067** (.026)	-.069*** (.027)	-.071*** (.025)	-.083*** (.031)	-.081*** (.029)	-.079*** (.028)	-.231*** (.087)	-.239*** (.084)	-.216*** (.082)
Paternal Marriage Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
<i>F</i> (Kleibergen-Paap)	66.5	63.6	60.7	56.8	53.0	56.2	63.4	61.2	65.7
Anderson-Rubin <i>F</i> stat. <i>p</i> -value	.010	.010	.005	.004	.003	.003	.005	.004	.007
Endogeneity test <i>p</i> -value	.008	.010	.010	.008	.008	.009	.005	.005	.011
Plausibly Exogenous <i>p</i> -value	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Clusters	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, paternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

The paternal marriage age may have an effect on the protogenetic interval and offspring human capital achievement. Thus, dummy variables indicating the paternal marriage age (on the 5-year level) are included in the instrumental variable regression analysis to account for this potentially confounding factor. While the coefficients on these additional control variables are omitted from the table, the specifications underlying each column in the table correspond to the columns in Table 6, except for the additional inclusion of paternal marriage age 5-year group dummy variables. The table establishes that the qualitative conclusion is robust to these alternative specifications. In particular, the effect of the number of siblings on the three measures of human capital achievements remains significant in all specifications.

Table C9: Robustness to Accounting for Paternal Occupational Wealth

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

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C10: Robustness to Alternative Birth Order Specifications

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

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C11: Robustness to Parish Fixed Effects

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.029** (.012)	.029** (.012)	.030*** (.011)	.029** (.014)	.028** (.013)	.028** (.013)	.096** (.044)	.097** (.042)	.092** (.042)
Parish Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
Signif. of Parish FE	< .1%	< .1%	< .1%	< .1%	< .1%	< .1%	< .1%	< .1%	< .1%
$R^2$	.185	.194	.250	.332	.362	.371	.356	.377	.398
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C12: Robustness to Fixed Effects on the Yearly Level

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.030** (.013)	.032** (.012)	.029*** (.011)	.046** (.019)	.048*** (.018)	.046** (.018)	.139*** (.051)	.148*** (.051)	.139*** (.049)
Parental Marriage Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Maternal Marriage Age on the Yearly Level FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
$R^2$	.292	.303	.355	.509	.545	.550	.545	.563	.582
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C13: Robustness to Logit Regression

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

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C14: Robustness to Missing Information on Parental Literacy and Skills and Long Protogenetic Intervals

	Literacy					Skilled Occupation					Occupational Wealth				
	Full Sample			No Long PI's		Full Sample			No Long PI's		Full Sample			No Long PI's	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Years to First Birth	.026			.092***	.150**	.101***			.114***	.170**	.133*			.237**	.505**
	(.026)			(.029)	(.059)	(.030)			(.030)	(.066)	(.075)			(.097)	(.199)
Years to First Birth × Literate and Skilled Father	-.017			-.113**	-.169**	-.110***			-.127**	-.182**	-.177			-.383*	-.649**
	(.031)			(.050)	(.072)	(.036)			(.064)	(.086)	(.112)			(.224)	(.280)
Years to First Birth (Winsorized)		.068***	.072*				.082***	.149***				.233***	.301**		
		(.020)	(.044)				(.023)	(.048)				(.071)	(.140)		
Years to First Birth (Winsorized) × Literate and Skilled Father		-.069*	-.073				-.110***	-.177***				-.372**	-.447**		
		(.037)	(.053)				(.042)	(.060)				(.156)	(.196)		
Literate and Skilled Father	-.070	.082	.007	.138	.132	.225*	.229**	.310**	.263**	.336**	.150	.728*	.516	.792*	.829*
	(.098)	(.096)	(.113)	(.106)	(.125)	(.116)	(.103)	(.130)	(.118)	(.145)	(.358)	(.384)	(.422)	(.437)	(.471)
Skilled Father	.134*	.097	.133*	.105	.138*	.194***	.193***	.198***	.173**	.174**	.222	.159	.237	.131	.202
	(.072)	(.070)	(.071)	(.074)	(.074)	(.071)	(.071)	(.071)	(.074)	(.074)	(.161)	(.164)	(.158)	(.173)	(.164)
Literate Father	.222***	.197***	.224***	.188***	.215***	.026	.020	.027	.022	.022	.531**	.372*	.532**	.377*	.532**
	(.058)	(.057)	(.058)	(.059)	(.060)	(.081)	(.076)	(.081)	(.078)	(.083)	(.225)	(.222)	(.224)	(.228)	(.232)
Missing Information	Yes	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes	No	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Literacy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R <sup>2</sup>	.229	.229	.233	.219	.223	.336	.337	.340	.339	.340	.365	.361	.368	.357	.366
N. of Observations	1,248	1,248	1,248	1,180	1,180	652	652	652	616	616	686	686	686	647	647
N. of Clusters	571	571	571	525	525	453	453	453	422	422	468	468	468	436	436

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

To assess the robustness of the results presented in Table 3, we perform a number of robustness checks. The specifications underlying columns 1, 6, and 11 are similar to those of columns 3, 6, and 9 of Table 3, respectively, except for the inclusion of a dummy variable indicating missing information for either paternal literacy or paternal skills as well as its interaction with the protogenesic interval (both omitted from the table). The specifications underlying columns 2, 7, and 12 exclude this dummy variable and its interaction while the protogenesic interval measure is winsorized at three years and 40 weeks (corresponding to a time to conception of around three years). The specifications underlying columns 3, 8, and 13 include the dummy variable and its interaction while the protogenesic interval measure is winsorized at three years and 40 weeks. The specifications underlying columns 4, 9, and 14 exclude the dummy variable and its interaction while observations with a protogenesic interval longer than three years and 40 weeks are excluded. The specifications underlying columns 5, 10, and 15 include the dummy variable while observations with a protogenesic interval longer than three years and 40 weeks are excluded.

The table establishes that the qualitative conclusions in Table 3 are robust to controlling for missing information for either paternal literacy or paternal skills when excluding long protogenesic intervals or winsorizing the protogenesic interval measure. Furthermore, when using the baseline protogenesic interval measure in the full samples, the qualitative conclusions of Table 3 remain unchanged, although the level of statistical significance tends to fall and differences between the elite and the non-elite only obtains statistical significance when using occupational skill as the outcome variable.

Table C15: Robustness to Accounting for Seasonal Effects

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

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C16: Robustness to Wages

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

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C17: Robustness to Temperatures

	Literacy				Skilled Occupation				Occupational Wealth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years to First Birth	.031*** (.011)	.033*** (.011)	.031*** (.011)	.033*** (.011)	.041*** (.013)	.042*** (.015)	.040*** (.013)	.041*** (.015)	.122*** (.040)	.113*** (.043)	.121*** (.040)	.134*** (.043)
20-Year Average Temperature at Birth	-.031 (.039)			-.026 (.039)	.003 (.051)			.015 (.060)	-.218 (.181)			-.050 (.213)
20-Year Average Temperature at Marriage		.015 (.041)		.020 (.041)		.121*** (.046)		.103** (.050)		.458*** (.153)		.431*** (.160)
20-Year Average Temperature at Parent's Marriage			-.036 (.053)	-.037 (.053)			.058 (.051)	.045 (.057)			.025 (.197)	.071 (.217)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Literacy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$R^2$	.222	.220	.222	.219	.333	.352	.341	.351	.342	.341	.347	.345
N. of Observations	1,238	1,224	1,237	1,218	610	502	600	485	639	533	629	514
N. of Families	564	558	563	555	421	364	413	351	432	377	424	363

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C18: Robustness to Vital Rates

	Literacy				Skilled Occupation				Occupational Wealth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Years to First Birth	.033*** (.011)	.033*** (.011)	.032*** (.011)	.035*** (.012)	.039*** (.013)	.040*** (.015)	.038*** (.013)	.042*** (.015)	.112*** (.042)	.118*** (.044)	.114*** (.041)	.125*** (.046)
20-Year Average	-.008 (.056)				-.006 (.067)				.144* (.244)			
Crude Death Rate at Birth												
20-Year Average	-.029 (.054)				-.062 (.065)				-.072 (.236)			
Crude Birth Rate at Birth												
20-Year Average	-.045 (.072)				-.060 (.076)				.002 (.280)			
Crude Death Rate at Marriage												
20-Year Average	.007 (.040)				.010 (.039)				.024 (.184)			
Crude Birth Rate at Marriage												
20-Year Average					-.021 (.069)				-.036 (.076)			
Crude Death Rate at Parent's Marriage												
20-Year Average					.002 (.059)				.034 (.097)			
Crude Birth Rate at Parent's Marriage												
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Literacy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$R^2$	.225	.222	.225	.223	.332	.341	.329	.349	.356	.350	.355	.353
N. of Observations	1,248	1,229	1,248	1,229	652	516	652	516	686	550	686	550
N. of Families	571	563	571	563	453	376	453	376	468	392	468	392

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C19: Robustness to Alternative Marriage and Stoppage Age Specifications

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

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C20: Robustness to Gender Division

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

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C21: Robustness to Parental Longevity

	Literacy			Skilled Occupation			Occupational Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Years to First Birth	.034*** (.014)	.034*** (.013)	.034*** (.012)	.049*** (.015)	.047*** (.013)	.048*** (.013)	.120*** (.051)	.120*** (.048)	.114*** (.046)
Paternal and Maternal Longevity Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Parental Literacy	No	No	Yes	No	No	Yes	No	No	Yes
Signif. of Maternal Longevity FE	.109	.121	.087	< .001	< .001	< .001	.219	.385	.487
Signif. of Paternal Longevity FE	< .001	< .001	< .001	< .001	< .001	< .001	< .001	< .001	< .001
$R^2$	.188	.200	.248	.336	.366	.375	.351	.365	.383
N. of Observations	1,248	1,248	1,248	652	652	652	686	686	686
N. of Families	571	571	571	453	453	453	468	468	468

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C22: Robustness to Heckit Estimation

	Literacy		Skilled Occupation		Occupational Wealth	
	Non-Missing	Value	Non-Missing	Value	Non-Missing	Value
	(1)	(2)	(3)	(4)	(5)	(6)
Years to First Birth		.032*** (.011)		.038*** (.013)		.113*** (.041)
Missing Death Date	-.593*** (.079)		-.181*** (.058)		-.222*** (.057)	
Missing Marriage Date	-3.897*** (.146)		-1.770*** (.070)		-1.832*** (.069)	
Inverse Mills Ratio		.020 (.033)		-.004 (.029)		.014 (.094)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	Yes	Yes	Yes	Yes	Yes	Yes
Parental Literacy	Yes	Yes	Yes	Yes	Yes	Yes
$R^2$		.225		.329		.355
N. of Observations	8,619	1,248	8,619	652	8,619	686
N. of Families	1,631	571	1,631	453	1,631	468

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

Table C22 shows that the dummies for missing marriage and death dates are both highly significant, confirming their significance in predicting a missing literacy or skill status. The inverse Mills ratio turns out to be highly insignificant in both stages of both regressions, verifying the absence of a sample selection bias. Indeed, the coefficient on the years to first birth remains highly significant for all three measures of human capital.

Table C23: The Association Between Inter-Birth Interval Length, the Protogenetic Interval Length, and Observed Socio-Economic Variables

	(1)	(2)	(3)	(4)	(5)	(6)
Years to First Birth	-.065*** (.010)	-.065*** (.010)	-.065*** (.010)	-.066*** (.010)	-.065*** (.010)	-.065*** (.010)
Age of Marriage of Mother (Years)	-.013 (.020)					-.007 (.020)
Age of Marriage of Mother (Years) Squared	.000 (.000)					.000 (.000)
Poor Father		-.115*** (.034)				-.026 (.040)
Skilled Father			.136*** (.040)			.078* (.047)
Skilled Mother			-.011 (.211)			-.092 (.210)
Literate Father				.126*** (.044)		.113** (.046)
Literate Mother				.117** (.051)		.081 (.052)
Retail Location					.126*** (.037)	.106*** (.038)
Industrial Location					-.024 (.032)	.020 (.035)
Agricultural Location					-.042 (.034)	.013 (.038)
N. of Observations	10,930	10,930	10,930	10,930	10,930	10,930
N. of Families	2,586	2,586	2,586	2,586	2,586	2,586

This table presents the results of a series of Cox Proportional Hazard regressions analyses of the time to a subsequent birth on various observable parental and locational characteristics. The estimates are log hazard ratios and larger coefficients correspond to a higher risk of birth (i.e., shorter inter-birth interval lengths). All regressions account for parental marriage time-period fixed effects and include dummies indicating unknown information. The coefficient on a constant term is omitted from the table. Standard errors clustered on the family level are reported in parentheses. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

Table C24: The Association Between Inter-Birth Interval Length, the Protogenetic Interval Length, and Observed Socio-Economic Variables — Controlling for the Maternal Age at the Beginning of Birth Intervals

	(1)	(2)	(3)	(4)	(5)	(6)
Years to First Birth	-.055*** (.010)	-.060*** (.010)	-.060*** (.010)	-.061*** (.010)	-.060*** (.010)	-.055*** (.010)
Maternal Age at Beginning of Birth Interval	-.024*** (.002)	-.016*** (.002)	-.016*** (.002)	-.016*** (.002)	-.017*** (.002)	-.024*** (.002)
Age of Marriage of Mother (Years)	.009 (.021)					.015 (.021)
Age of Marriage of Mother (Years) Squared	.000 (.000)					.000 (.000)
Poor Father		-.109*** (.034)				-.025 (.041)
Skilled Father			.127*** (.040)			.074 (.048)
Skilled Mother			.003 (.209)			-.088 (.204)
Literate Father				.117*** (.045)		.113** (.047)
Literate Mother				.115** (.052)		.078 (.053)
Retail Location					.123*** (.037)	.095** (.039)
Industrial Location					-.035 (.033)	.023 (.036)
Agricultural Location					-.048 (.034)	.013 (.039)
N. of Observations	10,930	10,930	10,930	10,930	10,930	10,930
N. of Clusters	2,586	2,586	2,586	2,586	2,586	2,586

This table presents the results of a series of Cox Proportional Hazard regressions analyses of the time to a subsequent birth on various observable parental and locational characteristics, including the maternal age at the beginning of the intervals. The estimates are log hazard ratios and larger coefficients correspond to a higher risk of birth (i.e., shorter inter-birth interval lengths). All regressions account for parental marriage time-period fixed effects and include dummies indicating unknown information. The coefficient on a constant term is omitted from the table. Standard errors clustered on the family level are reported in parentheses. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

Table C25: The Effect of the Length of the Protogenetic (PI) Interval on Completed Fertility — Family Level Regressions

	Number of Surviving Offspring (> 5 Years)			
	Literacy Sample	Skills Sample	Wealth Sample	Total Sample
	(1)	(2)	(3)	(4)
Years to First Birth	-.477*** (.052)	-.488*** (.065)	-.512*** (.064)	-.478*** (.047)
Poor Father	.529* (.304)	.550* (.303)	.563* (.293)	.427* (.253)
Skilled Father	-.054 (.325)	-.146 (.306)	-.091 (.300)	-.106 (.275)
Literate Father	.621** (.277)	.818** (.332)	.801** (.325)	.561** (.266)
Literate Mother	-.231 (.280)	-.329 (.350)	-.342 (.347)	-.209 (.269)
Agricultural Location	.142 (.274)	.455 (.309)	.404 (.307)	.147 (.245)
Industrial Location	.049 (.303)	.577* (.343)	.480 (.340)	-.008 (.260)
Retail Location	-.371 (.307)	-.251 (.254)	-.389 (.249)	-.387* (.216)
Skilled Mother	3.187* (1.753)	3.226* (1.769)	3.136* (1.772)	2.705* (1.475)
$R^2$	.340	.382	.378	.347
N. of Observations	571	453	468	729

This table presents the results of a series of OLS regression analyses of the number of surviving offspring on the time to first birth (i.e., protogenetic interval) in the family. All regressions account for couple marriage time-period fixed effects, wife marriage age interval fixed effects, and dummies indicating unknown information. The coefficient on a constant term is omitted from the table. Standard errors clustered on the family level are reported in parentheses. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

Table C26: Robustness to Heckit Estimation (Instrumental Variable Regression)

	Literacy		Skilled Occupation		Occupational Wealth	
	Non-Missing	Value	Non-Missing	Value	Non-Missing	Value
	(1)	(2)	(3)	(4)	(5)	(6)
Surviving Siblings (> 5 Years)		-.072*** (.025)		-.079*** (.028)		-.217*** (.081)
Years to First Birth	.044 (.033)		.006 (.025)		.014 (.024)	
Missing Death Date	-.597*** (.079)		-.182*** (.058)		-.223*** (.057)	
Missing Marriage Date	-3.900*** (.144)		-1.770*** (.070)		-1.832*** (.069)	
Inverse Mills Ratio		.016 (.031)		.002 (.030)		.020 (.095)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Parental Skills	Yes	Yes	Yes	Yes	Yes	Yes
Parental Literacy	Yes	Yes	Yes	Yes	Yes	Yes
<i>F</i> (Kleibergen-Paap)		64.4		56.0		66.4
Anderson-Rubin <i>F</i> stat. <i>p</i> -value		.004		.003		.006
Endogeneity test <i>p</i> -value		.010		.009		.011
Plausibly Exogenous <i>p</i> -value		< 1%		< 1%		< 1%
N. of Observations	8,619	1,248	8,619	652	8,619	686
N. of Families	1,631	571	1,631	453	1,631	468

Standard errors clustered on the family level are reported in parentheses. All regressions account for parental marriage time-period fixed effects, maternal marriage age interval fixed effects, birth order fixed effects, and dummies indicating unknown information. \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level.

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

In the first stage we extend the sample to also include observations where literacy and skill status are unknown, thus expanding the sample to 8,647 individuals representing a total of 1,639

families.<sup>9</sup> Next, we estimate the probability of observing human capital with a probit model, using dummies for missing marriage or death dates as instruments in addition to the protogenetic interval (and covariates). We have 6,037 observations with missing marriage dates; 4,405 observations with missing death dates; and 2,976 cases where both dates are missing. Based on the predicted probabilities, we calculate the inverse Mills ratio, proceeding to estimate Equation (2) by 2SLS including the inverse Mills ratio as a control variable. We conduct the procedure for both outcome variables (i.e., literacy and skill status). If the inverse Mills ratio is statistically significant in the first or second stage, then it means our estimations possibly suffer from a sample selection bias.

Table C26 shows that the dummies for missing marriage and death dates are both highly significant, confirming their significance in predicting a missing literacy or skill status. The inverse Mills ratio turns out to be highly insignificant in both stages of both regressions, verifying the absence of a sample selection bias. Indeed, the coefficient on the years to first birth remains highly significant for all three measures of human capital.

---

<sup>9</sup>The result from this stage establishes that there is no significant relationship between the protogenetic interval and the chances of observing the children's human capital.

## Appendix References

- BLACK, S. E., P. J. DEVEREUX, AND K. G. SALVANES (2005): “The More the Merrier? The Effect of Family Size and Birth Order on Children’s Education,” *Quarterly Journal of Economics*, 2, 669–700.
- EJRNÆS, M. AND C. C. PÖRTNER (2004): “Birth order and the intrahousehold allocation of time and education,” *Review of Economics and Statistics*, 86, 1008–1019.
- JUUL, S., W. KARMAUS, J. OLSEN, E. I. S. S. GROUP, ET AL. (1999): “Regional differences in waiting time to pregnancy: pregnancy-based surveys from Denmark, France, Germany, Italy and Sweden,” *Human Reproduction*, 14, 1250–1254.
- KLEMP, M., C. MINNS, P. WALLIS, AND J. WEISDORF (2013): “Picking winners? The effect of birth order and migration on parental human capital investments in pre-modern England,” *European Review of Economic History*, 17, 210–232.
- SCHOFIELD, R. S. (2005): ““Monday’s child is fair of face”: favoured days for baptism, marriage and burial in pre-industrial England,” *Continuity and Change*, 20, 93–109.
- WOOLDRIDGE, J. M. (2010): *Econometric Analysis of Cross Section and Panel Data*, London: MIT Press, second ed.