

The Emergence of the Child Quantity-Quality Tradeoff - insights from early modern academics

Thomas Baudin*

David de la Croix[†]

December 22, 2023

Abstract

Reflect on the escape from a stagnant or Malthusian system. If this transformation is propelled by human capital, it should be spearheaded by individuals possessing elevated human capital. To explore this hypothesis, we investigate the connection between family size and human capital among academics in Northern Europe in the two centuries leading up to the Industrial Revolution. We gauge scholars' human capital using a novel approach based on their publications. We find that scholars with a high number of publications shifted from having more siblings to having fewer than others during the first half of the 18th century. This shift is consistent with an evolutionary growth model in which the initial Malthusian constraint leads the high human capital families to reproduce more, before being endogenously substituted by a Beckerian constraint with a child quality-quantity tradeoff. Our results support an extension of the Galor and Moav (2002)'s approach, in which the decline of Malthusian constraints is linked to human capital accumulation during the 18th century.

Keywords: Universities, Academies, Fertility, Scholars, Human Capital.

JEL Classification Numbers: N3, J1, O4.

*IESEG School of Management, Univ. Lille, CNRS, UMR 9221 - LEM - Lille Economie Management, F-59000 Lille, France and IRES, UCLouvain, Belgium. E-Mail: t.baudin@ieseg.fr

[†]IRES/LIDAM, UCLouvain, Belgium and CEPR, Paris. E-mail: david.delacroix@uclouvain.be

1 INTRODUCTION

Limits to growth were overcome with the waning of Malthusian constraints during the 18th-19th centuries. This unleashed a period of sustained economic development (Galor 2011). Over the same historical period, an unprecedented rise in education was made possible by a reduction in fertility, reflecting a tradeoff between the quality and the quantity of children. Although this general picture is broadly accepted, identifying the correct timing of events and the underlying mechanisms remains a challenge, partly because data on education and human capital are quite scarce before education became organized by the State towards the end of the 19th century.

Our study focuses on the emergence of the quality-quantity tradeoff and its relationship with human capital accumulation through an unconventional lens. We introduce a novel dataset on families of academics during the early modern period and reveal a notable shift in the correlation between family size and scholarly success between the seventeenth and eighteenth centuries. We find that the most accomplished scholars of the seventeenth century tended to originate from large families, whereas the opposite pattern was observed during the eighteenth century. This result implies that, among the high human capital elite, the quality-quantity tradeoff emerged in the 18th century, that is before the Industrial Revolution.

The Beckerian tradeoff between quality and quantity of children results from a budget constraint. As spending on the quality of each child is a rival good, having many children makes it harder to achieve a high level of quality. The terms of the tradeoff depend on several elements: the return on education (Galor and Weil 2000; Shiue 2017; Cinnirella and Streb 2017; Madsen and Strulik 2023), preferences towards quality vs. quantity (Galor and Moav 2002), the efficiency of child development and medical technology (De la Croix and Licandro 2013), the introduction of education subsidies (Aaronson, Lange, and Mazumder 2014), trade policy (Bignon and García-Peñalosa 2021), urbanization conditions (Baudin and Stelter 2022), and whether the cost of children is a direct material cost such as food or an opportunity cost such as rearing time (Doepke 2015). In one way or another, these elements determine the shadow price of quantity vs quality and, hence, the choices made by would-be parents. The literature devoted to explaining the transition from stagnation to growth and the Rise of the West involves necessarily mechanisms based on changes in this shadow price. To assess these mechanisms empirically, it is essential to observe both quality and quantity over a long enough period of time.

There is a large empirical literature that seeks to measure the importance of the child quality-quantity tradeoff (QQ tradeoff hereafter) using both historical and recent data. The litera-

ture has used different proxy for child quality, but, in general, either a deep time dimension is missing or data are aggregated rather than individual. School enrollment at the county level in 19th century Prussia has been used to establish a negative relation between fertility and schooling (Becker, Cinnirella, and Woessmann 2010). School enrollment data at the department level in 19th century France have been quite extensively used by several authors (Perrin 2013; Murphy 2015). For England, Klemp and Weisdorf (2019) rely on literacy rates and employment in an occupation with greater prestige in later life. They show that both measures decline with the number of siblings. China has a centralized exam system, making it possible to use a dummy variable for passing the entry-level exam of the civil service. This variable has some time depth and is shown to be negatively affected by sibshipsize (Bai, Li, and Lam 2023). Stature is another measure of quality, such as in the study of Hatton and Martin (2010) on children in Britain. Bleakley and Lange (2009) show that school enrollment, regular school attendance, and literacy in the American South increase with the eradication of hookworm disease, an exogenous shifter on the price of child quality vs child quantity. Fighting this disease also resulted in fertility reductions. Overall, these contributions have documented a QQ tradeoff around the epoch of industrialization but none of them have identified a positive correlation between quality and quantity in pre-modern times. This finding is a key feature of the UGT developed by Galor and Moav (2002) and an essential finding of our paper.

To lend credence to the theories of growth based on the child QQ tradeoff, it is important to find in the historical data when the QQ tradeoff started to be relevant, and whether it preceded or followed the take-off to modern growth associated with the demographic transition. Such validations are limited by the availability of data on child quality on a sufficiently long historical period. In their absence, an alternative is to rely on meta-analyses, such that of Skirbekk (2008). Skirbekk is not interested in the intragenerational correlation between children’s human capital and number of siblings, but instead on fertility by social status, i.e. the intergenerational correlation between parents’ education and fertility. But the two dimensions are tightly related, as parents’ education/status correlates with children’s education/status. Skirbekk finds that as fertility declines, there is a general shift from a positive to a negative or neutral status-fertility relation. This happens in the nineteenth century in now-developed countries (the 909 samples used in the meta-analysis include both developed and developing countries). The pattern highlighted by Skirbekk is largely confirmed for the twentieth century by Vogl (2016) who looks at the correlation between education and size of families using Demographic and Health Surveys (DHS, covering 48 developing countries).

In this paper, we establish a comprehensive database of academic scholars, allowing us to as-

sess their quality across several centuries. Our sample includes individuals who were affiliated with higher education institutions in Northern Europe between 1450 and 1800. To construct this sample, we relied on secondary sources that cover scientific academies and universities in the region. Each observation in our sample comprises an academic scholar matched to the institution they belonged to. To evaluate individual quality, we cross-referenced these scholars with publication records from over 10,000 libraries worldwide, which were accessible through VIAF (The Virtual International Authority File). Additionally, we gathered information on sibship sizes by matching scholars with genealogical data from major providers, such as Geni and Geneanet. Out of the 5,178 scholar-institution pairs in our database (involving 4,381 unique persons), we were able to find genealogical records for 2,321 of them (1,912 unique genealogies). Overall, our database is a rich and unique resource for investigating the quality and family origins of academic scholars over an extended period.

With these unique data we can address the question of whether high quality scholars (i.e. those publishing more) come from large or small families, and whether this changed over time. Descriptive statistics show that, during the seventeenth century, scholars publishing in the top half of the distribution have on average at least 0.1 brothers more than those publishing in the bottom half of the distribution. This advantage vanishes for those active in the eighteenth century. At the end of our sample period, the pattern is reversed, with well-published scholars having up to 0.4 fewer brothers than those who are publishing less. This result is confirmed in a rolling regression set-up in which we control for various selection and composition biases. The results suggest that there is an evolutionary advantage in families with well-published scholars until the turn of the 18th century. This advantage disappears in the 18th century, and is replaced by a tradeoff between number of siblings and publications.

To understand the mechanisms behind this pattern, we compare the empirical results to a unified growth model with heterogeneous agents. As in Galor and Moav (2002), heterogeneity affects the preference for quality children. The model shows that, before a certain date, households are trapped in a Malthusian regime. In this regime, there is an evolutionary advantage for those who like quality more. Households gradually escape the Malthusian constraints by accumulating human capital and eventually reach a Beckerian world, where there is a tradeoff between the quality and quantity of children. Our model explains the observed pattern found in the data without the need for any external shock. The key mechanism in our approach is an endogenous switchover from a Malthusian constraint to a Beckerian constraint, which is rooted in human capital accumulation during the Malthusian epoch. Once the initial conditions have been fixed, the transition is endogenous. Our model is a reinterpretation of Galor and Moav (2002) where the regime shift is triggered by the

premodern rise in human capital investments.

We fundamentally assume that individuals exercised rational control over their fertility even in pre-modern societies. Even if the possibility of fertility control within marriage is disputed in the literature (see “Malthus in the Bedroom” (Cinnirella, Klemp, and Weisdorf 2017) against “Randomness in the Bedroom” (Clark and Cummins 2019)), there is a consensus since Wrigley and Schofield (1983) that marriage was the main channel through which individuals controlled their number of descendants. We provide an additional argument against the view that the Malthusian eras is a period characterized by non-rational fertility. If this were the case, one would expect a negative correlation between fertility rates and the development of human capital: when the number of children is random, educational spending would need to adjust, if anything, to fluctuations in fertility, hence implying a negative correlation between the two. Our paper presents evidence contrary to this expectation, showcasing an exact opposite relationship. This finding does not dismiss the concept of children and education being rival goods in Malthusian times; rather, it illustrates that despite this rivalry, subsistence-related forces were potent enough to prevent the QQ-tradeoff producing a negative correlation between fertility and human capital.

Subsistence income is a key notion in the Malthusian model. It is the level of income ensuring that population is constant. Clark (2007) stresses that “*the term subsistence income can lead to the incorrect notion that in a Malthusian economy people are all living on the brink of starvation, like the inmates of some particularly nasty Soviet-era gulag. In fact in almost all Malthusian economies the subsistence income considerably exceeded the income required to allow the population to feed itself from day to day.*” While positing households in a Malthusian regime, it is crucial to clarify that this does not imply their mere survival but rather denotes an income level aligned with a stable population. This assertion is made within the context of a world where fertility rises in tandem with income.

We contribute to the literature in different ways. First, we inspect the mechanisms of Unified Growth Theory. Our evidence that the quantity-quality tradeoff emerged in preindustrial Europe lends credence to a key tradeoff assumed by the theory. Dating the birth of this tradeoff to the early 18th century invites towards a broader interpretation of the key trigger, based not only on industrialization, but also on human capital accumulation among the elite groups. Importantly, we use new way of measuring human capital with individual level data.

While our findings do not negate the relevance of the mechanism proposed by Galor and Moav (2002), it is plausible that the transition was instigated by the accumulation of human capital among the elite. However, it may have required an additional impetus for the broader population, manifesting in the form of an increased return to education.

Our results help to characterize better the behavior of a narrow but important group of people who form the upper tail of the human capital distribution. Squicciarini and Voigtländer (2015) have shown the importance of this group: distinguishing between upper-tail and average skills reinstates the importance of human capital during the transition from stagnation to growth. We shed a new light on the families of the members of the upper tail of human capital. We show that, within this group, there is some heterogeneity. Families of superstars embraced behaviors more compatible with long run growth before and more intensely than families of less productive scholars.

Finally, we contribute to the new literature on the role of specific institutions of the 17-18th centuries, such as academies of sciences, in fostering later development. For example, Koschnick, Hornung, and Cinnirella (2022) look at how economic societies in 18th century Germany facilitated spatial knowledge diffusion in the 19th century, while Zanardello (2023) shows that cities with scientific academies grew faster 150 years after the birth of such academies. Our results stress the importance of pre modern human capital and the academy movement as key roots of Europe’s development.

2 DATA

2.1 SCHOLARS, INSTITUTIONS, AND PUBLICATIONS

We have built a dataset of scholars who were members of 26 universities and scientific academies located around the Baltic Sea and the North Sea, between 1450 and 1800. Our sample of countries includes Denmark, Sweden, Finland, North of Germany, North of Poland, Estonia, Russia, North of the Netherlands, and Scotland. The universities and academies we have selected all share a Protestant background (even St-Petersburg’s academy was initially populated by Protestant scholars coming from Germany and Switzerland). We selected a geographical zone with a relatively homogeneous cultural and religious environment, and a high coverage in the genealogical databases (this is detailed below). Table 1 displays the included institutions. We select academic scholars who were members of these institutions up to 1800, the end of our period of observation to limit our analysis to the pre-industrial era. This also guarantees that our data are not biased by the Humboldt reform of 1810, which is often considered to mark the birth of modern universities in Germany.

The list of scholars is established using secondary sources, often produced by universities and academies themselves. The sample is a subset of the database constructed by De la Croix (2021) (accessible at <https://shiny-lidam.sipr.ucl.ac.be/scholars/>) which we match with genealogical data. We define scholars as persons exerting a research role, a

teaching role, or both, in either a university or an academy. Universities are institutions granting a doctorate degree (Frijhoff 1996). They concentrate on four main fields: theology, law, arts and humanities, and medicine. Their impact on the society is aptly described by Pedersen (1992): “*The faculty of arts gave a basic education to grammar school boys, many of whom would become teachers themselves and contribute to the increase in literacy of the population at large. Others would go on to one of the higher faculties to prepare themselves for other professions. The faculty of medicine produced medical practitioners; the faculty of laws created future administrators with expert knowledge in canon or civil law, and the faculty of theology provided teachers for the episcopal schools, where the ordinary parish priests were educated.*” Academies were usually created later, in the 17th-18th century, responding to a push to develop new fields of research which were not traditionally taught at universities. The academies range from clubs of amateur naturalists or local historians to eminent societies, attracting the best scholars, publishing journals, and building a network of corresponding members.

Figure 1 shows the location of the institutions included in our study (thick black dot) and the birth place of the scholars with a genealogy (small red or orange dots). A majority of scholars comes from around the Baltic and North Seas. Some come from other European countries, including France, Italy and the Holy Roman Empire,¹ showing that the academic job market was already very international at that time (De la Croix et al. 2023).

Corresponding members of academies are shown in lighter (orange) dots. They are located in France, England, Northern Italy, and Russia. Iberia and the countries under Ottoman Rule had no scholars in our sample. Figure 10 in Appendix B shows the same map and includes all members of the institutions we have selected.

Table 1 reports the official creation date of each institution. Several universities were founded before the Reformation, but became Protestant afterwards. The main secondary sources used to build the data on scholars are listed in the last column. These sources of information are complemented with national biographies and other databases such as Taisand (1721) for law, Eloy (1755) for medicine, and Applebaum (2003) for the key actors of the scientific revolution.

From the list of members of universities and academies we remove those who are not clearly scholars, but rather honorary members. These include kings and emperors, military officers (unless they contributed to the development of techniques related to artillery or fortification), diplomats etc.

¹In the figure, the countries’ borders are those of 1700, as they are drawn in Reed (1999).

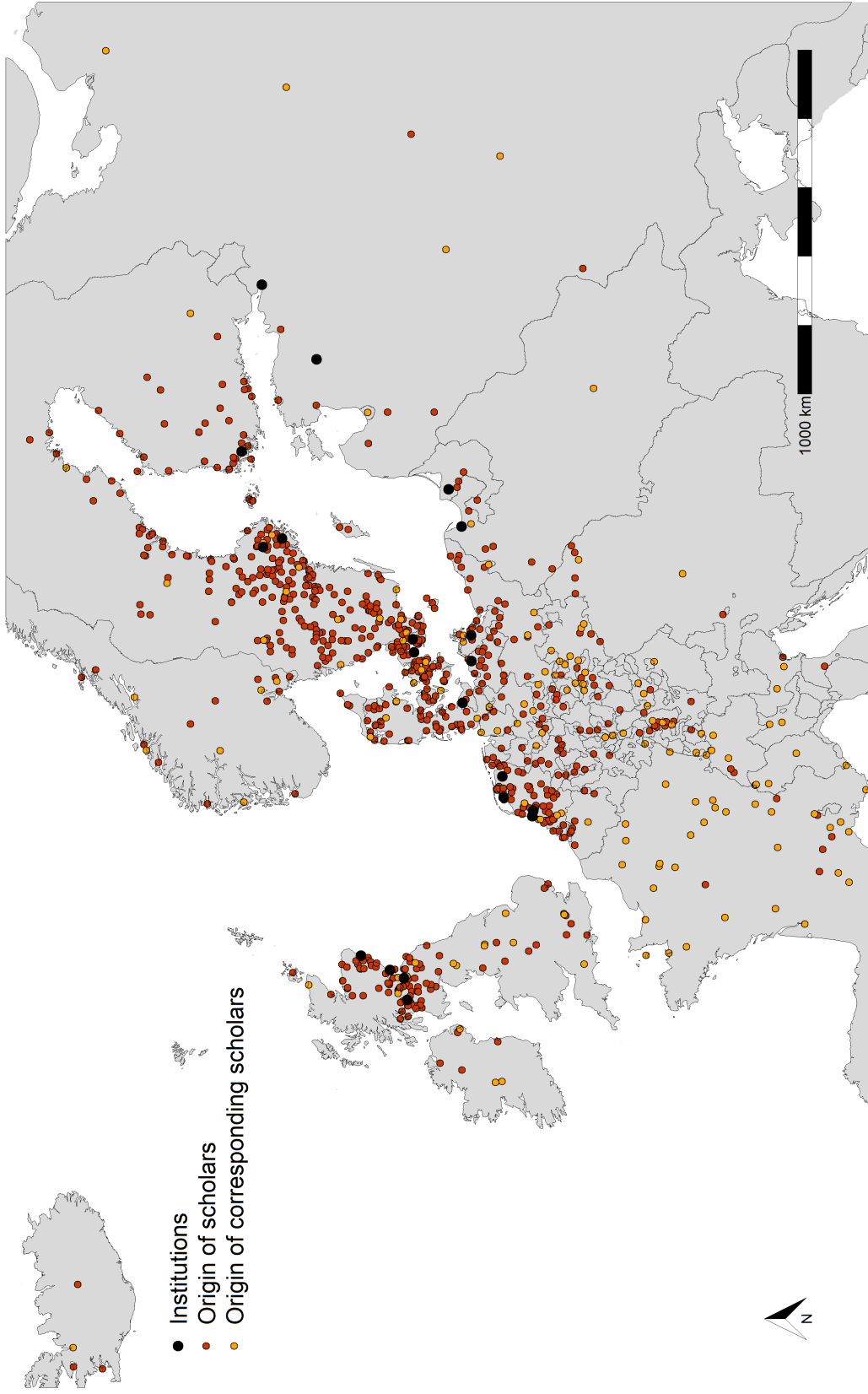


Figure 1: Origin of scholars with genealogies and frontiers of 1700

Institution	City	Country	Dates	Sources
University of Copenhagen	Copenhagen	DNK	1475	Slottved (1978)
Royal Danish Science Society	Copenhagen	DNK	1742	Lomholt (1950)
Uppsala University	Uppsala	SWE	1477	Von Bahr (1945), Astro.uu.se, Jensen (2018)
Royal Society of Sc. in Uppsala	Uppsala	SWE	1728	Karlberg (1977)
Royal Swedish Academy of Sc.	Stockholm	SWE	1739	Dahlgren (1915)
University of Lund	Lund	SWE	1666	Delen and Weibull (1868)
Royal Physiographic Society	Lund	SWE	1778	Gertz (1940)
Åbo Akademi University	Turku	FIN	1640	Klinge et al. (1988)
University of Tartu/Dorpat	Tartu	EST	1632 1710	Inno (1972)
University of Groningen	Groningen	NLD	1612	https://hoogleraren.ub.rug.nl/
Athenaeum Illustre of Amsterdam	Amsterdam	NLD	1632	http://www.albumacademicum.uva.nl/
University of Franeker	Franeker	NLD	1585	Feenstra et al. (2003), Napjus and Lindeboom (1985)
Royal Dutch Society of Sc.	Haarlem	NLD	1752	https://khw.nl/historische-leden/
University of Greifswald	Greifswald	DEU	1456	various encyclopedia
University of Rostock	Rostock	DEU	1419	Krüger (2019)
University of Kiel	Kiel	DEU	1652	Volbehr and Weyl (1956)
Akademisches Gymnasium Danzig	Gdansk	POL	1558	Hirsch (1837)
Danzig Research Society	Gdansk	POL	1743	Schumann (1893)
University of Königsberg	Kaliningrad	RUS	1544	Naragon (2006), Schwinges and Hesse (2019)
Academy of St Petersburg	St-Petersburg	RUS	1724	Shemivot (1873)
University of Edinburgh	Edinburgh	GBR	1582	Grant (1884)
University of Glasgow	Glasgow	GBR	1451	Coutts (1909)
Philosophical Society Edinburgh	Edinburgh	GBR	1731	Emerson (1981), RSE (2006)
University of Aberdeen (Old)	Aberdeen	GBR	1495	Anderson (1893)
University of Aberdeen (New)	Aberdeen	GBR	1593	Anderson (1898)
University of Saint Andrews	St-Andrews	GBR	1411	Smart (2004)

Table 1: Sources used by institution

We also distinguish between members with a strong link to the institutions, including all the professors and ordinary members of academies, and scholars with a weak link. Weak links include corresponding members of academies, who are foreign-based scholars with whom the academicians have regular contact. Occasionally they include some scholars who are linked to a university without having a formal professorship (such as Tycho Brahe, who was connected to the University of Copenhagen without having an official job there).

One key feature of our data we use is that they include nearly all scholars with high human capital (the famous or productive ones) and a large sample of unknown scholars as well (the obscure or less productive). Encoding famous scholars only (for example those in an encyclopedia) would miss a large part of the variance of human capital within institutions before the Industrial Revolution. The use of detailed secondary sources guarantees a satisfactory level of variance in the quality of scholars.

To measure the quality of scholars we consider their visibility in modern-day library catalogues. We use the VIAF search engine, which provides references to the collections of thousands of libraries worldwide. VIAF is an international authority file that links all national authority files through a single platform. For each scholar, we count the total number of titles, including publications by and about the author, and posthumous editions, to capture an element of “citations” and provide a better proxy for their actual human capital. Our measure of quality, labeled “publications”, is actually the inverse hyperbolic sine of the number of titles in VIAF, to accommodate people with no publications. Figure 11 in Appendix B shows the histogram of its distribution. Our measure has two additional advantages. First, the librarians working on VIAF have addressed the issue of author name disambiguation to the best of their abilities. Second, Chaney (2020) has shown that library-led databases like VIAF provide a good approximation of the population of known European authors.

To highlight some correlates of scholars’ human capital, we first regress the inverse hyperbolic sine of individual number of works published on a time trend based on birth dates. Results are shown in the first column of Table 2. The time trend is slightly positive and statistically significant at 1%. In the second column, we include the mean age at death (longevity) and the age at which scholars are first recorded as member of their institution. Longevity is strongly significant, a gain in one year is correlated with a gain of 1.8% in the number of works. It captures part of the correlation with the time trend. The age of entry correlates positively with publications, which is counter-intuitive as the earlier someone enters academia, the more time they have to develop their thinking and academic production. In fact, the age of entry also captures different practices between fields and places that may confound the estimation. This is confirmed in the last column that includes more variables: field dummies,

a dummy for being a corresponding member, a dummy for having a genealogy on genealogical websites, and institutions fixed effects. The correlation with the age of entry then becomes negative and significant. Fields are also important correlates of publications: the scholars working in theology tend to publish more than the reference category, which includes all arts and humanities. Legal scholars tend to publish less. Corresponding members (weak links) publish more than ordinary scholars, as do those with a genealogy. In these regressions, the unit of observation is a scholar-institution pair, and the standard errors are clustered at the individual level.

Dependent variable is asinh(nworks)			
birth date	0.002*** (0.000)	0.001*** (0.000)	0.001* (0.000)
longevity		0.018*** (0.002)	0.019*** (0.002)
age nomination		0.010*** (0.002)	-0.006*** (0.002)
theology			0.344*** (0.069)
law			-0.417*** (0.077)
medicine			-0.074 (0.075)
science			0.105 (0.073)
corresponding member			0.995*** (0.077)
with genealogy			0.609*** (0.053)
Instit. FE.	N	N	Y
Observations	4,471	4,445	4,445
N Clusters	3,698	3,672	3,672
Adjusted R ²	0.009	0.046	0.242

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Robust SE clustered at the individual level.

Table 2: Correlates of individual publications

2.2 GENEALOGIES

To retrieve information about the families of scholars, we first use biographical dictionaries to identify relevant information such as date and place of birth. We rely on crowdsourced genealogical databases to verify this information and complement it with information on sibshpsize, number of children, and parents. Over the past years, scholars have used the richness of public crowdsourced genealogical data to measure fertility, death and migration: see among others Kaplanis et al. (2018), Stelter and Alburez-Gutierrez (2022) and Blanc (2022). We follow this line of research here and use the main online crowdsourced genealogical databases, which are www.geni.com, www.ancestry.com, www.geneanet.org, www.familysearch.org. If no suitable information could be retrieved from these international sources, we explored smaller - national - scale databases like gedbas.genealogy.net for Germany, genealogieonline.nl for The Netherlands and <https://docs.vgd.ru/en/about> (All Russia Family Tree) for Russia.

For each scholar, we reconstruct manually the completed fertility (total number of children) of their father as well as their own fertility. We add full and half-siblings indifferently but build a dummy variable indicating whether the person has half-siblings. We count how many of these children are girls. We collect the year of death and the occupation of scholars' fathers. Genealogical databases have been very useful for finding places and dates of birth and death when that data is missing from biographical notices.

Not surprisingly, some sources have conflicting information. For fertility, we retained the highest provided numbers after correcting for straightforward imputation errors. For instance, if on www.geni.com, a scholar has two siblings but four on www.familysearch.org (and there is no repetition of the same sibling on FamilySearch), we retain the information from FamilySearch and attribute four siblings to the professor. Sometimes, information had to be mixed between sources as each of them provide complementary data. We learned that, for Northern Europe, and especially Scandinavian countries, www.geni.com is the most popular, and therefore the richest and the most reliable source of data. The same can be said for www.geneanet.org for France and www.ancestry.com for England.

Genealogical records in principle include marriage dates, offering a potential means to calculate a proximate determinant of fertility – the bride age at marriage. Marriage dates are also commonly employed in the literature to construct an exogenous measure of fecundity, determined by computing the time interval between marriage and the first birth (Galor and Klemp 2019). Regrettably, the proportion of genealogies within our sample that provide both marriage and first birth dates is insufficient for this analytical purpose. Similarly, dates

Institutions	members	with genealogies	in %
University of Copenhagen	344	216	63
Royal Danish Science Society	155	109	70
Uppsala University	242	175	72
Royal Society of Sciences of Uppsala	98	74	74
Royal Swedish Academy of Sc.	417	287	69
University of Lund	263	154	59
Royal Physiographic Society	142	93	65
Åbo Akademi University	117	94	80
University of Tartu/Dorpat	54	31	57
University of Groningen	104	46	44
Athenaeum Illustre of Amsterdam	73	24	33
University of Franeker	147	56	38
Royal Dutch Society of Sc.	343	135	39
University of Greifswald	261	79	30
University of Rostock	318	121	38
University of Kiel	218	47	22
Akademisches Gymnasium Danzig	90	22	24
Danzig Research Society	102	24	24
University of Königsberg	337	34	10
Academy of St Petersburg	303	137	45
University of Edinburgh	160	58	36
University of Glasgow	103	35	34
Academy of Edinburgh	394	190	48
University of Aberdeen (old)	198	34	17
University of Aberdeen (new)	107	21	20
University of Saint Andrews	87	25	29
TOTAL	5178	2321	45

Table 3: Genealogical coverage by institution

of death are systematically underreported for siblings who presumably died young, preventing us to measure child mortality. Crowdsourced genealogies are often less complete than genealogies based on Parish records, such as those available for Quebec (Galor and Klemp 2019)) or England (De la Croix, Schneider, and Weisdorf 2019), they are nevertheless offered on a much larger geographical scale than these latter.

Table 3 presents the number of genealogies found, by institution. For Scandinavian institutions, we are able to match scholars to a genealogy in 57%-75% of the cases. Keeping in mind that our scholars are active before 1800, this is very high. We do not find such a high level of coverage for the other regions. For the Netherlands, we are at 30%-44%; for former German territories, 9%-44%; for Scotland, 33%-47%. On the whole we have a genealogy for 46% of the scholars-institution pairs, i.e. 2,321 linked profiles.

Genealogical websites (through their detailed biographical notice sections) and Wikipedia pages often report the occupation of the persons and their parents. We collect these occupations to better understand scholars' class backgrounds. We classify them into three categories after Van de Putte and Miles (2005): elite, middle class, workers. We do not observe unskilled workers, so workers are either skilled or semi-skilled. The middle class merges farmers with local business people and non-manual skilled people. Table 4 shows the main occupations, with the number of observations in parentheses. As already noted by De Candolle (1885), many academics were born to families of pastors and priests. In the following sections, we will delve into the interpretation of the reversal of the QQ-tradeoff as a significant indication of the gradual liberation of scholarly families from Malthusian constraints. To substantiate this argument, it is crucial to acknowledge that our professors do not exclusively come from highly privileged backgrounds such as the nobility, as such groups might not have experienced the full extent of Malthusian constraints. Table 4 provides compelling evidence, indicating that a majority of our academics originate from non-elite backgrounds. A notable example is Linnaeus, who emerged from a modest family with a father who worked as a preacher and built the family house with his own hands — a place where Linnaeus began his observations and classifications of living species. This is an example of the “impoverished sophisticated” population in Sweden prior to the Industrial Revolution, as documented by Sandberg (1979). This population boasted high levels of education despite lacking substantial wealth or privilege. In Figure 12, presented in Appendix B, we illustrate the constancy of the proportion of each social class over time. Notably, approximately 40% of scholars hail from an upper-class background, while 33% have a middle-class background.

The environment in which scholars were raised was influenced by their birthplace, shaping their early life experiences. We build a dummy variable, urban/rural, taking a value of

Elite	professor (138), councillor (70), bishop (41), mayor (38), doctor (36), rector (35), general (24), governor (23), lord (21)
Middle class	preacher (157), priest (73), pastor (62), farmer (29), officer (28), trader (27), master (24), superintendent (22), vicar (21), secretary (20)
Workers	tailor (3), innkeeper (3), gardener (3), baker (3), grocer (2), tanner (2), carpenter (2), engraver (2)

Table 4: Main occupations – occurrence in parenthesis

one if the place of birth is a city with at least 2000 inhabitants in 1700 (using data from Buringh (2021) and Bairoch, Batou, and Chevre (1988)). We find that 48% of our scholars originated from urban areas, while 47% came from rural backgrounds (refer to Figure 13 in Appendix B) – the place of birth is not known for 5% of the sample. These shares remained roughly constant over time, as depicted in Figure 14 in Appendix B.

2.3 CORRECTING BIASES IN GENEALOGIES

The genealogies can suffer some biases. A first source of bias with genealogical data is gender, as women tend to be under-represented: see Charpentier and Gallic (2020) or Gavrilov et al. (2002) for a discussion. Some of this under-reporting may be due to the Old-White-Men (OWM) bias as already documented by Dupâquier (1993): most of amateur genealogists have some characteristics pushing them to collect biased information. White amateur genealogists in wealthy, patriarchal societies have tended to focus on the male branches of family trees of white European men.

In addition, there was historically an under-reporting of girls' births at the time of their birth, especially female stillbirths. Finally but importantly, the data we consider are not necessarily representative of the whole population surrounding the Baltic and the North Sea before 1800; indeed, it focuses on families who have at least one university professor in their lineage. Until the beginning of the 20th century, university professors were almost exclusively men (see some exceptions in De la Croix and Vitale (2023)), so looking at these specific families induces a mechanical over-representation of men. For example: among the families of professors having three children, in the absence of gender bias in the reporting of births and applying the law of large numbers, we should end up with a sex ratio of 2.075. We reach this number by using a natural sex ratio of 1.05 (Chao et al. 2019). This implies that

each new birth is 48.8% likely to be a female birth and 51.2% to be a male birth. Extending this logic to other parities, we get the theoretical sex ratios of Table 5. Table 5 shows that our data suffer from a misreporting of girls’ births. This bias is more severe for families of two children and becomes less important as family size increases. A simple way to correct for this gender bias in computing the size of families is to count exclusively the male siblings and use the total number of siblings (male and female) for robustness only.

Number of children	2	3	4	5	6+	$+\infty$
Theoretical sex ratio	3.10	2.07	1.73	1.56	1.31	1.05
Sex ratio in our data	4.18	2.54	2.40	1.76	1.70	-

Table 5: Theoretical versus empirical sex ratio (M/F) as function of parities within families of professors

In Figure 2, we expose the distribution of parities (keeping only male children) among the professors for whom we collected information. We compare this distribution to one we computed using the parish records collected by Wrigley and Schofield (1983) for the English population. Our population of scholars is not strictly comparable to Wrigley and Schofield (1983), in particular because the latter covers all social classes, but it is the best comparison data we can find. We find that the distribution of parities in our data is left-skewed and over-represents parity one, i.e. we have too many single children. Such a bias is well known among scholars using genealogical data; if it may originate from many issues, the verticality issue is the most crucial. Many amateur genealogists are interested in discovering their direct antecedents, and do not research the siblings of those in their direct line. It implies, especially for ancient data, an over-representation of observations having only one parent and no siblings. This bias is nicely discussed in Blanc (2022) who treats it by suppressing observations for which he cannot find one ancestor with at least two children in the pool of 30 ancestors going from parents to grand-grand-grand parents. If the approach of Blanc is defensible, it does not fully overcome the verticality bias in all cases. In the data he uses (Geni/Familinx), lineages are built from the merging of the inputs of thousands of genealogists. As a result, one person can be attached to a grand-grand-grand father having five children but still have unrecorded brothers and sisters because the genealogist who has encoded his/her profile was filling her own genealogy in a vertical way.

To overcome this, we use a stricter approach than Blanc. First, we excluded all genealogies with a sibshpsize less than three and where the date of birth of the father is not known, to exclude data of lower quality. We drop 370 genealogies under this restriction. We further exclude all the scholars having no siblings and having themselves only one child, which is

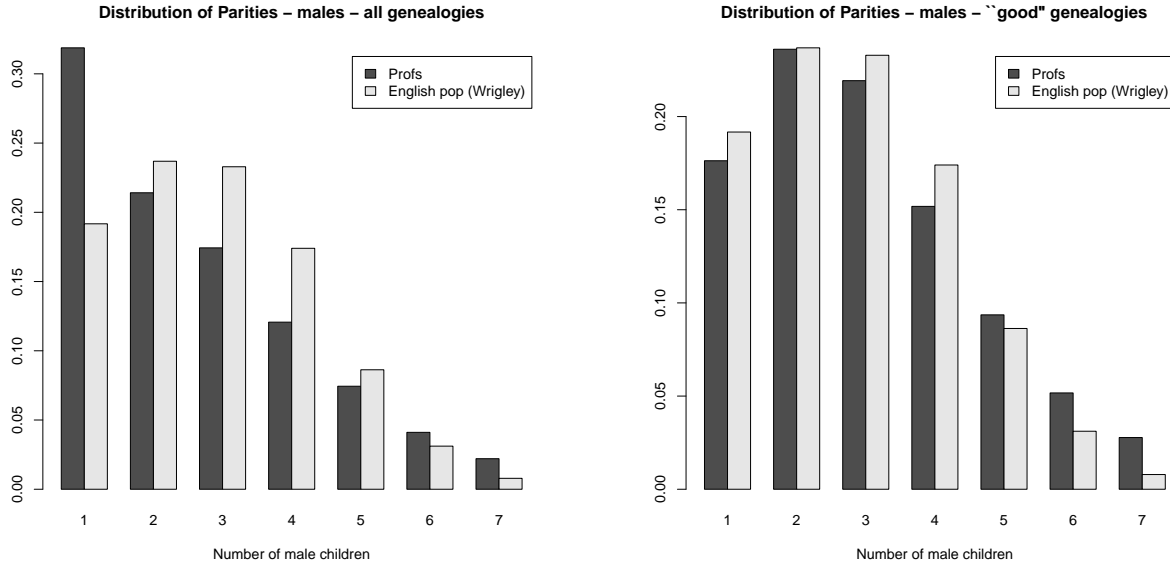


Figure 2: Distribution of parities: all genealogies (left), selected genealogies (right)

strongly indicative of the verticality bias, dropping a further 104 genealogies. The right-panel of Figure 2 shows the new distribution of parities after implementing our two selection criteria. We can see a sensible improvement in the distribution, with a new mode of two instead of one.

Our corrections for the measurement of fertility ensure that our distribution of parities is closer to that proposed by Wrigley and Schofield (1983). Another comparison point can be used to assess whether our percentage of only-sons families of 19% is reasonable. Galton (1875) considered a sample of 200 or so living scientific members of the Royal Society, and his method was a self-report questionnaire. These scholars are born a little later than our sample, around 1800-1820, but still before any demographic transition in England. Galton reports a share of only sons of 20%, very close to our estimate. This reassures us on the quality of our correction.

Another concern is that our sample of scholars with a valid genealogy diverges substantially from the sample of those not having genealogies in dimensions other than the fertility of their parents. An example of where this could arise is if, on genealogical platforms, famous professors are recorded more often than obscure ones. In Tables 10 to 12 of Appendix F, we show how the selection is changing over time. Academics having a valid genealogy tend to publish more and to live longer than their obscure counterparts throughout our period of observation. This selection bias is present in all periods, and does not change significantly over time. It implies that the true variance of publications is higher than the one we measure

in the sample, implying that the estimated coefficients in the regressions in the next sections are presumably the lower bounds of the true estimates.

2.4 DEMOGRAPHIC TRANSITION

We turn our attention to the evolution of longevity and fertility over the period of observation. We take a sub-sample of 30% of the scholars born either side of a given birth year, and measure longevity as the median length of life. We take the median to avoid excessive role of outliers in small samples. We show this estimate (labelled as Northern Europe Academics median length of life) together with other estimates from the literature in the top panel of Figure 3. Longevity increased strongly among professors born in the 17th century and reached a plateau around 70-72 years along the 18th. The early rise in longevity we observe is fully aligned with what we know from the literature on the longevity in academia (Leridon and Mandelbaum 2004, Andreev et al. 2011, Stelter, De la Croix, and Myrskylä 2021) or among the elite (De la Croix and Licandro 2015, Cummins 2017), but the magnitude of the rise is stronger than that in the literature.

Regarding fertility, shown in the bottom panel of Figure 3, the average number of children (sibship size of scholars) fluctuates in a narrow interval. That is, we do not observe any fertility transition on average, but this does not mean that fertility is not going through important transformations via composition effects.

We also recorded the birth order of the scholars. Among the 1452 observations for which a rank can be computed, 767 (51%) are first-born sons. This aligns perfectly with Galton (1875) who found that 48% of famous English scientists were the first-born son in their family. If the birth order did not matter, we would observe fewer than 1/3 of scholars being the first-born (with an average family size of 6 and a sex ratio of 1.34, the average number of males per family is 3.44). This suggests that the probability of selecting into academia is higher for first-born sons (see Black, Devereux, and Salvanes (2005) for an authoritative reference for birth order effects on education in the economic literature. Such effects were recently contested by Clark and Cummins (2023) who find no effect of birth order on various social outcomes in historical England).

Finally, in Figure 4, we divide the scholars in two groups: those publishing more than the median (high quality) and those publishing less than the median (low quality); we then plot the average fertility of these two groups over time. We can see that among scholars born in the 17th century, the high quality scholars tend to have 0.1 more male brothers and so potentially 0.2 more siblings than their low quality counterparts. In the 18th century, high

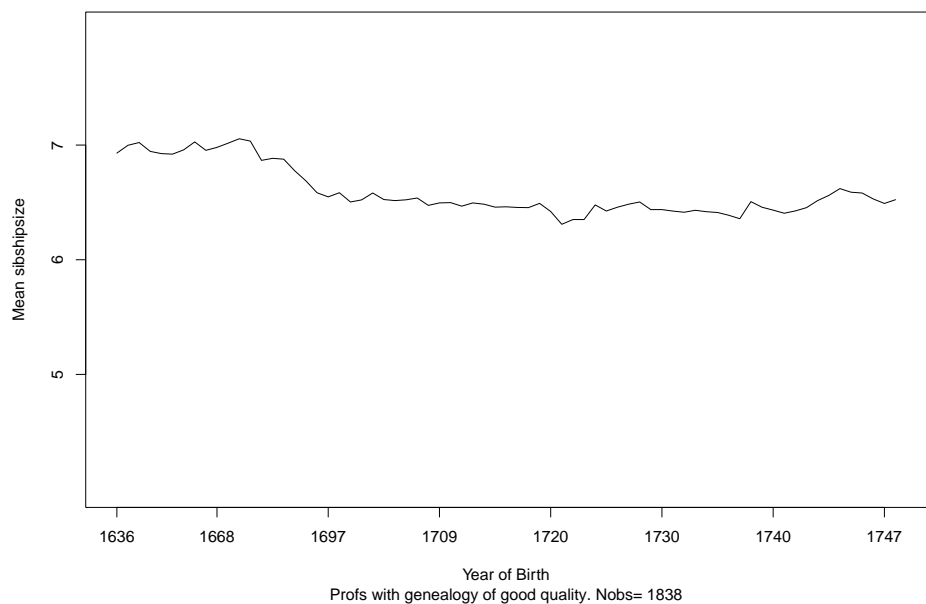
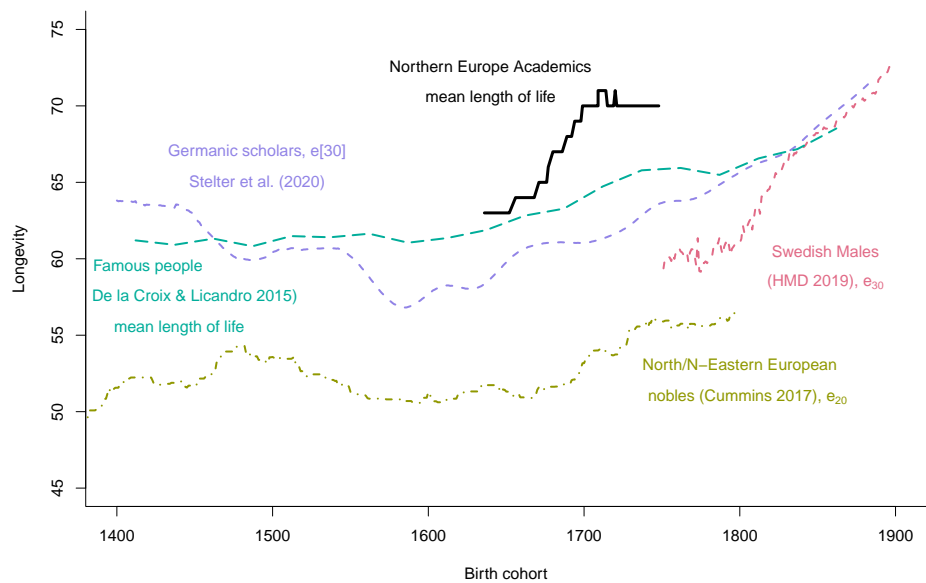


Figure 3: Longevity and fertility over time

quality scholars start to have fewer siblings in total than their low quality counterparts. The fertility differential between scholars' parents reaches more than 0.4 boys and so 0.8 children for births occurring around 1749. To the best of our knowledge, this reversal of the QQ trade-off is rarely observed on a consistent micro-level dataset. It constitutes an important empirical reinforcement of any theory placing the switch in the tradeoff between quality and quantity of children at the heart of the European transition to growth.

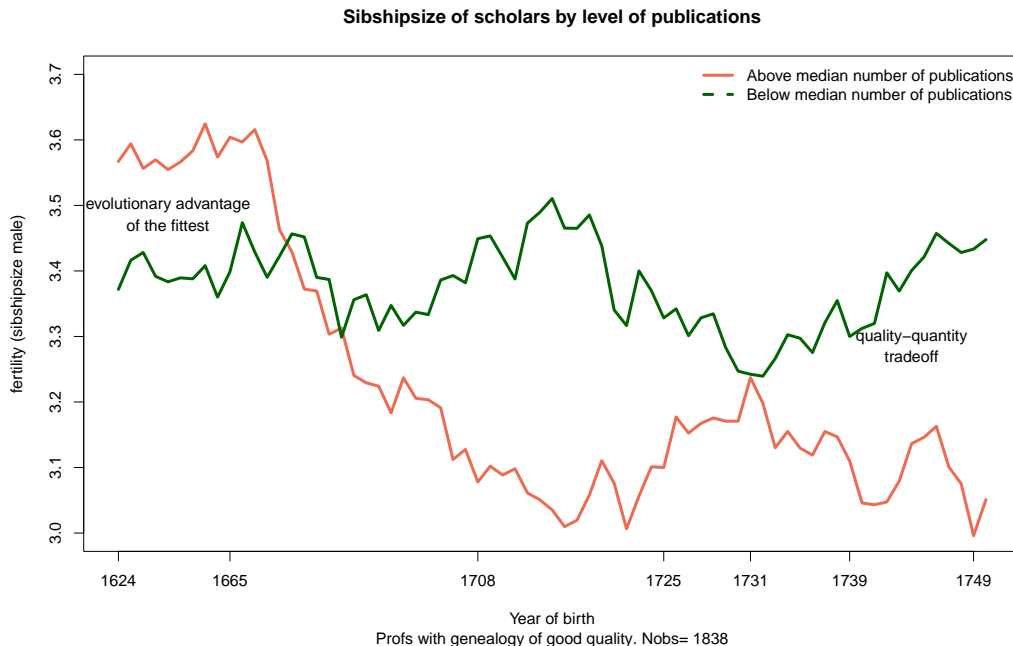


Figure 4: Reversal of the QQ tradeoff among scholars' parents

As a consistency check, we extend our analysis to encompass the offspring of the scholars. Maintaining the established demarcation between high and low-quality scholars, our investigation reveals a noteworthy pattern: in the 17th century, high-quality scholars exhibit a higher fertility compared to their low-quality counterparts, as illustrated in Figure 5. Strikingly, this trend undergoes a reversal in the subsequent century, with high-quality scholars born in the 18th century consistently showing a decreased number of offspring. The replication of this fertility reversal across two consecutive generations adds robustness and significant credibility to our findings.

When juxtaposing Figure 5 with Figure 4, a notable difference emerges: scholars, on average, have fewer children than their parents. This discrepancy is substantial, and we attribute it to a key contextual factor — all our scholars experience adulthood within an urban setting, whereas half of their parents resided in rural areas. The significant disparity in fertility between urban and rural environments during premodern Scandinavia is well-documented

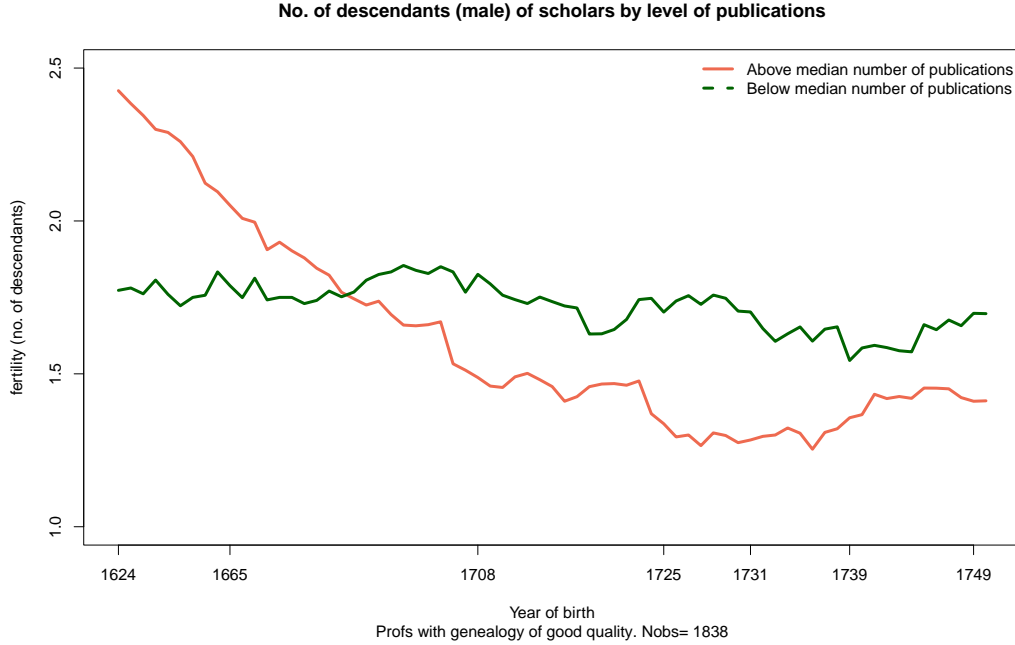


Figure 5: Reversal of the QQ tradeoff among scholars

(refer to the sources in Baudin and Stelter (2022)).

For a final validation of our fertility data, we conducted an analysis to estimate the intergenerational correlation in fertility. The findings, illustrated in Figure 16 within Appendix B, indicate that we cannot dismiss the hypothesis that this correlation remains constant at 0.1 throughout the entire period under consideration. This observation aligns with existing literature on similar correlations in premodern contexts (for example, Pearson Karl and Leslie (1899) finds a correlation of 0.1 for the landed gentry in premodern England).

The reversal of the QQ tradeoff among the parents of scholars has many potential confounding factors and compositional effects. The variations in the relative weight of each institution over time may be important, as could the weight of alternative disciplines. For example, mathematicians may publish more than theologians, while also coming from smaller and secular families. If so, a massive entry of mathematicians and other scientists born around 1750 may explain the reversal of the QQ trade-off, which would have nothing to do with a change in the way parents of professors have allocated their resources between quality and quantity of children. The share of the main academic fields over time is shown in Figure 17. Section 3 will be devoted to an in depth analysis of the reversal, controlling for as many factors as we can.

3 ESTIMATION

Figure 4 documents a reversal of the fertility differential between highly and lowly productive scholars. As it is, this reversal of the QQ tradeoff may be due to confounding factors and selection issues we would like to rule out. We split the range of professors' birthdates in percentiles and run 70 successive regressions with controls, each of them including the professors born within a specific time interval, corresponding to 30% of our entire time window. The first regression then includes all the professors born between 1435 and 1686; the second one includes those born between 1511 and 1689, and the last one includes those born between 1735 and 1777. In each iteration, we regress the inverse hyperbolic sine of the publications of professor i from institution k on the size of his male sibship pool (Sibshipsize_i) controlling for a series of important factors. Denote K the set of institutions and F the set of fields. The field of i at k is d_{ik} . The regression equation is:

$$\begin{aligned} \text{arsinh}(\text{Publis}_i) = & \alpha_1 \text{Sibshipsize}_i + \alpha_2 \text{Longevity}_i + \alpha_3 \text{Age Nomination}_{ik} + \alpha_4 \text{Corresponding}_{ik} \\ & + \alpha_5 \text{Urban}_i + \alpha_6 \text{Geni}_i + \sum_{j \in K} \alpha_{5j} \text{I}(k = j) + \sum_{f \in F} \alpha_{6f} \text{I}(d_{ik} = f) + \gamma X_i + \varepsilon_{ik}. \quad (1) \end{aligned}$$

We control for longevity (death year – birth year) and the age at which the scholar was nominated to institution k , to capture the fact that the younger a scholar is nominated, the more time he has to develop his catalogue of publications. We control for corresponding membership of an institution: being a corresponding member is an honorary distinction, so it likely selects the scholars of the highest quality in terms of publication originating from distant places, where the nature of the QQ-tradeoff is different from that prevailing in our region of analysis. Finally, we also control for the rural-urban character of the place of birth, and for whether the source of the genealogy is geni.com or another website.

By controlling for the institution through the fixed effects $\text{I}(k = j)$, we rule out the possibility that the potential reversal of the QQ-tradeoff is due to the selection of universities and academies of origin into the sample. This controls for, for example, an increase in the proportion of professors coming from academies and universities where fertility is low, for any economic or cultural reasons, and publications more abundant.² In our vector of control variables, we also include for the same reason dummies $\text{I}(d_{ik} = f)$ for scholar's field. This ensures that we measure an association between sibshipsize and publications that is not

²Decisions regarding university affiliation could be endogenous, leading to a suggestion not to include this fixed effect in our main regressions. This is particularly pertinent because the human capital of scholars might influence their geographical allocation in a circular manner. However, when we abstain from controlling for institution fixed effect, our results remain consistent in terms of both magnitude and significance.

confounded by a risk that some fields are populated with individuals who are more or less prone to publish and more or less prone to have large families.

We also include in X_i a dummy determining whether the scholar had half-siblings to reduce the noise in fertility measurement, as the presence of half-siblings indicates unusual parental life courses, including divorce, widowhood and remarriage.³

Some scholars may appear more than once in our database as they may have belonged to more than one institution and they have high mobility. Duplication can also occur where scholars are corresponding members of academies. For example, Joseph Banks was a British naturalist born in 1743, who was the head of the Royal Society for more than 40 years and who collaborated with the academies of Gdansk, Copenhagen, Haarlem, Saint-Petersburg and Stockholm. To avoid standard errors being artificially deflated by the presence of similar observations, we compute robust errors ε_{ik} clustered at the individual level.

In our main specification, we select observation in the same demanding way as in the previous section, keeping only male siblings with “good” genealogies. Descriptive statistics are provided in Appendix A. We believe that our rolling time window regression setting constitutes a flexible approach, capturing the dynamics of the association between the sibship size of professors and their human capital without imposing a too demanding set of constraints.

3.1 MAIN RESULTS

Figure 6 shows our main results (Appendix E provides ten full regression tables for specific years. The black line joins the 70 estimations of α_1 , with the 10% confidence interval in dark gray and the 5% confidence interval in light gray. The reversal of the QQ-tradeoff is salient with a first period where scholars from large families tended to publish more than scholars coming from smaller families, a period of reversal where the association between fertility and human capital is non significant followed by the last period where scholars from smaller families publish more. Remarkably, the reversal becomes significant (both at the 10 and 5% confidence level) among professors born in the 18th century, which also corresponds to the first phase of the reversal identified with our uncontrolled fertility measure in Figure 4.

Table 6 displays the evolution of our main coefficient of interest pooling the 70 estimations by groups of ten. While the reversal of the QQ-tradeoff appears again, it does so in a context where the importance of longevity for explaining publications decreases over time. Indeed the coefficient of association between longevity of scholars and their publication metric

³Results are unaffected if we control for the following additional variables: age of the scholar when his father passed away and fixed effects for the scholar’s country of birth.

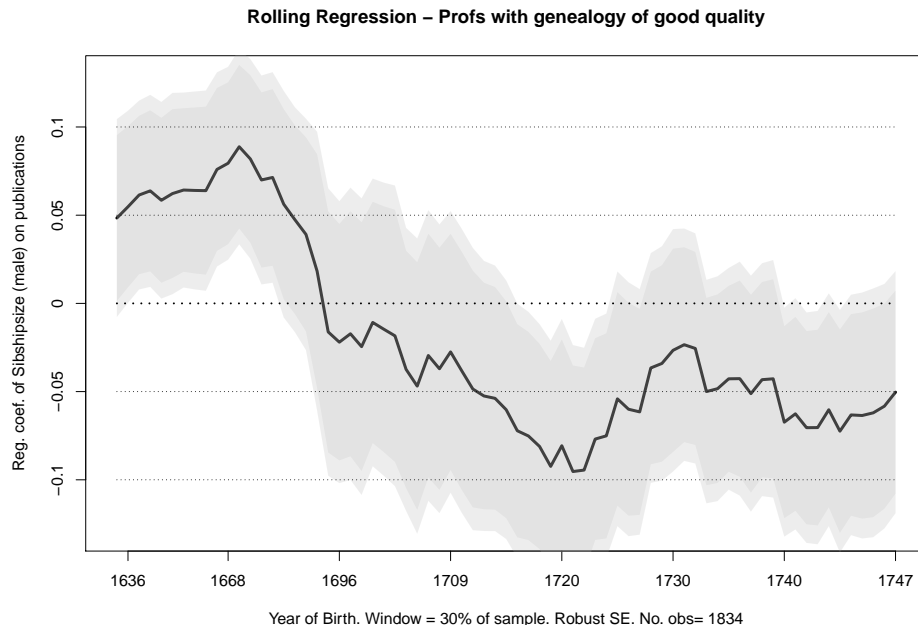


Figure 6: Rolling regression for main specification

decreases continuously. The share of publication variance we are able to explain with our main specification evolves between 26 and 30% on our entire time window. Finally, even if they have to be taken with utmost prudence as computed on only 10 observations/regressions, one can see that the standard deviation of our estimated coefficient is quite small within each time window. We take this as reassurance that the reversal of the QQ-tradeoff is not partly driven by variations in the statistical noise surrounding our estimations.

Figure 6 and Table 6 illustrate a noticeable symmetry between the peak intensities of positive association in the initial periods and negative association in the final periods. Both exhibit values hovering around 0.05 to 0.065, with peaks just below 0.1. Specifically, among professors born between 1611 and 1706, the elasticity of the average scholar's publications with respect to the number of brothers stands at 0.28 (using Bellemare and Wichman (2020) formula). Conversely, for professors born between 1726 and 1753, this elasticity is equal to 0.22. The stability of the fertility rate over time (shown in Figure 3) substantiates the validity of comparing elasticities.

Table 9 in Appendix E shows that corresponding members of academies and universities tend to publish more than other scholars, while controlling for the presence of half-siblings in the pool of siblings is important for some periods. If all along our period of observation, law scholars publish less than others, for the last cohorts concentrated in the second part of

	SibShipSize		Longevity		R2	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
0-30 to 9-39	0.062	0.007	0.033	0.002	0.283	0.007
10-40 to 20-50	0.054	0.033	0.033	0.001	0.304	0.019
20-50 to 29-59	-0.026	0.012	0.028	0.003	0.262	0.007
30-60 to 39-69	-0.060	0.020	0.018	0.017	0.261	0.012
40-70 to 49-79	-0.071	0.019	0.013	0.002	0.275	0.006
50-80 to 59-89	-0.039	0.011	0.013	0.002	0.281	0.010
60-90 to 69-100	-0.064	0.008	0.014	0.001	0.284	0.009

Table 6: Evolution of α_1 , α_2 and R^2 along the rolling regression process

the 18th century, scholars working in the fields of science tend to publish more than others. This is consistent with the atmosphere of this period preceding the industrial revolution. It is characterized by Enlightenment values where science attracts prestige, if not money; where every city wants its own Academy of Sciences and Arts and where these academies appoint top scientists as corresponding members.

3.2 ROBUSTNESS CHECKS

In Figure 6, we limited our investigations to three kinds of scholars only: (1) those from families with two siblings or more, (2) those having only one sibling and a referenced father's date of death and (3) those having no sibling, a referenced father's date of death and more than one child (to avoid the verticality bias). In Panel (A) of Figure 7, we relax these restrictions and include any scholar having a genealogical link. The global magnitude of the QQ-tradeoff reversal is not changed dramatically. Although our sample size increases by around 25%, our coefficients are not more precisely estimated, confirming the need to restrict our sample to rich genealogies.

Our measurement of publication encompasses both an extensive margin (whether an individual has ever published or not) and an intensive margin (the number of publications given that an individual has published). We delve into this distinction in Panels (B) and (C) of Figure 7. Panel (B) focuses on the number of publications conditional on the individual having published, while Panel (C) shifts the dependent variable to the probability of ever publishing a manuscript. Both the extensive and intensive margins of publications play a crucial role in understanding the reversal of the QQ-tradeoff. During Malthusian times, scholars from larger families had a notable advantage, with a higher likelihood of publishing

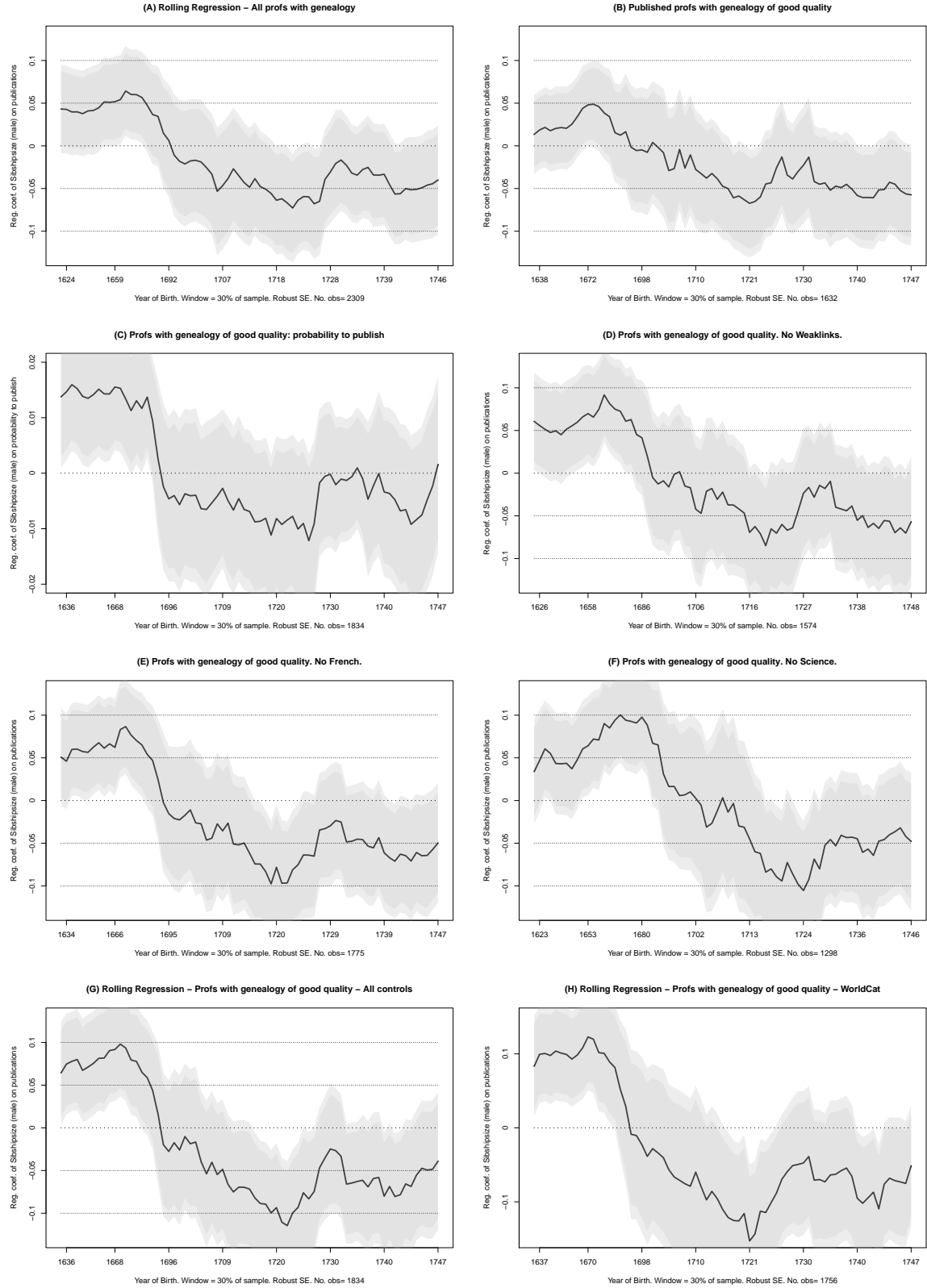


Figure 7: Rolling regression for alternative specification

at least once. In modern times, however, scholars from smaller families exhibit a significant advantage in terms of publishing repeatedly, after they have published at least once. This result reinforces the importance of considering both of these publication margins in our analysis, as they both contribute to the reversal.

In Panels (D) to (F), we present the results of three regressions out of a large series where we try to determine whether our main results are driven by specific sub-groups of our sample.⁴ In Panel (D), we run our main specification, excluding corresponding members of academies. The stability of the α_1 value even after excluding corresponding members indicates that their presence does not significantly affect the emergence of a significant reversal of the QQ-tradeoff. This suggests that the observed reversal is not a statistical artifact arising from an over-representation of highly productive but less fecund scholars residing outside our area of interest.

Panel (E) tests a version of the analysis without French scholars, as France (59 observations) may have initiated its fertility transition earlier than other European countries. It could be the case that French scholars are less fertile than their counterparts while they would tend to be more productive than others, as they are exclusively corresponding members of academies. Our results indicate that this is not the case, neither for French scholars nor for scholars from the Netherlands, the United Kingdom, or Germany. Similarly, in Panel (F), we explore the potential influence of scholars from the field of sciences. Despite the potential for the reduction in sample sizes to affect significance, our main results remain unchanged, indicating that the reversal we document applies across the full range of academic fields covered by our original dataset. Moreover, if we were allowed to arbitrarily remove the physicians from the sample, the results would be more salient (Figure 19 in Appendix).

Panel (G) introduces a series of additional controls: whether scholars are the first-born son or not, four categories of occupation of their father (elite, middle class, workers, NA), their age when their father passed away, and their number of descendants. All these controls are possibly correlated with (unobserved) income. Our main findings hold even after accounting for these controls, which rules out the possibility that the reversal of the QQ-tradeoff solely stems from differences in the social environment in which professors were raised.

Finally, in Panel (H), we alter our outcome variable and measure the quality of scholars using Worldcat. Worldcat furnishes statistics pertaining to the volume of library holdings associated with an individual, offering a comprehensive evaluation of both output quantity and quality. Initially serving as a reliable index of quality, we utilized Worldcat metrics until

⁴The results not presented in this section appear in Appendix G.

its unexpected discontinuation in March 2023 (refer to Curtis and De la Croix (2023) for an in-depth exploration of this metric and its correlation with VIAF statistics). Our findings unveil a consistent pattern of fluctuations, marked by positive correlations pre-1700 and negative correlations post-1700, mirroring the same reversal observed in previous analyses. Notably, the significance of the quality-quantity tradeoff remains robust, comparable to the benchmark. Both measures of quality are subject to noise; the same is true for our fertility measure. But beyond this noise, our analysis consistently reveals a significant and systematic modification of fertility behaviors throughout history, a modification that preceded the Industrial Revolution.

4 MODEL

Spirit of the model - Our empirical analysis of North European academic scholars shows a reversal of the correlation between quality and quantity of children. There is one theoretical model of growth that predicts such a reversal. It is the unified growth model of Galor and Moav (2002). In their approach, there are two types of people which differ only by a small difference in the weight they attach to the education of their children. During the stagnation period, the high-education people have a higher income and hence a larger number of children. This gives them an evolutionary advantage and generates a differential fertility of the type we observe in our data. At some point, thanks to technical progress, the return to education increases, pushing the whole population to invest massively in quality. As individuals strive to balance their budget, investing more in education often leads to a reduction in their number of children. Here differential fertility is reversed, with larger families being less educated.

There are a couple of aspects of Galor and Moav’s model that are not entirely supported by the data. First, in their approach, the reversal of differential fertility is tightly linked to the Industrial Revolution, which is when the return to education increases. This is at odds with our data, as we see a reversal taking place at the beginning of the 18th century, while the Industrial Revolution in Northern Europe takes place one century later. Second, Galor and Moav have a strict interpretation to the Malthusian period: income per person was oscillating around a constant level, close to survival. This view is brought into question by recent research, which shows some slow growth during the centuries preceding the Industrial Revolution (Fouquet and Broadberry 2015).

It is precisely because Galor and Moav assumed constant income per capita during the stagnation period that they need to bind technical progress to the return to education, in order to generate the transition to modern growth. If instead there was some slow growth

in income per person during the Malthusian epoch, this growth alone would have been sufficient to escape the Malthusian constraints. Escaping the Malthusian logic transforms the households' constraints, and the standard quality-quantity trade-off ultimately prevails.

Our approach is as follows. We start exactly as Galor and Moav do with two types of people, one having a slightly higher preference for education than the other. Let us call them quality-lovers and quantity-lovers. To be able to interpret our data, we view both groups as belonging to the intellectual elite, and we neglect the rest of the population. Each household faces two types of constraint: a Malthusian constraint imposing consumption to be higher than a critical level, and a standard budget constraint, imposing consumption spending and education spending on the children to be less than or equal to income. As explained in the introduction, the critical consumption level should not be interpreted as a survival level like the World Bank poverty line of one dollar per day, but as a device to generate the typical income effects found in Malthusian models.

As long as both groups are constrained by the minimal consumption level, the richer people are the quality-lovers, and they have paradoxically more children than the quantity-lovers, as in Galor and Moav. Over the Malthusian period, consumption per capita is constant and equal for both groups to the critical level, while education spending is rising over time, leading to a rise in human capital and in income per capita. As time passes, a greater share of resources is spent on education. This view fits very well both with the rise in the number of universities and academies during the 17-18th centuries and also with the rise of the impoverished-sophisticated documented by Sandberg (1979).

The main difference between our approach and Galor and Moav's is that we consider a scenario where slow economic growth occurs during the Malthusian epoch, leading to a point in time where the Malthusian constraint no longer applies. Quality-lovers are the first to benefit from this enrichment, followed by everyone else. At this point, households face the usual budget constraint and start substituting quantity for quality. As a result, quantity-lovers start having more children than quality-lovers, which means that scholars who publish less come from larger families.

Our model has several appealing features. It is simple and easy to follow, and it generates some income growth during the Malthusian epoch. Moreover, the timing of the reversal of differential fertility is now linked to the expansion of education, rather than the later Industrial Revolution, which better aligns with the available data.

Main assumptions - In an overlapping generations set-up, we assume that each individual lives for two periods: childhood and adulthood. During childhood, the individual is inactive

but receives a part ϕ of her parent's time for childbearing and an education $e_t^i \geq 0$. Each family is mono-parental and reproduction is asexual. In adulthood, a person born at date $t - 1$ is characterized by her level of human capital h_t and a utility function inherited from her parents. If the functional form of the utility function is the same for all individuals, they may differ regarding the weight of the future human capital of their children $\eta^i > 0$. All adults value their level of consumption c_t^i , their number of children n_t^i and the future human capital of the latter h_{t+1}^i , such that:

$$u(c_t^i, n_t^i, h_{t+1}^i) = \ln c_t^i + \gamma \ln n_t^i + \eta^i \ln h_{t+1}^i. \quad (2)$$

η^i is distributed over a set $\mathcal{E} \subset \mathbb{R}^+$. The future human capital of children is produced through an investment into their education (e_t^i) such that:

$$h_{t+1}^i = \psi e_t^i, \quad (3)$$

where $\psi > 0$ is a scaling factor capturing the marginal impact of educational investments on human capital. Equation (3) does not allow for varying returns to education.

At date 0, all families start from the same initial condition $h_0^i = h_0 \forall i$. Following Galor and Weil (2000), we assume that there exists a minimal consumption constraint such that:

$$c_t^i \geq \bar{c}.$$

This condition introduces a Malthusian dimension into our model because when binding, it restricts fertility decisions of households and increases the importance of income effects. We assume that individuals have two sources of income: labor and non-labor income. The wage per efficient unit of labor is normalized to 1 while $a > 0$ represents the non-labor income, which may correspond to home production, for instance. The budget constraint of an adult at date t is then:

$$c_t^i + \phi n_t^i h_t^i + e_t^i n_t^i = h_t^i + a. \quad (4)$$

Assumption 1 $\gamma > \max\{\eta^i\}$, $h_0 > \bar{c} - a > 0$

This assumption ensures that the maximization problem at time 0 is not degenerate and that the minimal consumption constraint can be fulfilled. An adult born at date $t - 1$ will maximize (2) subject to (3) and (4) and usual positivity constraints $c_t^i \geq 0$, $n_t^i \geq 0$ and $e_t^i \geq 0$. Under Assumption 1, we can define a threshold value $\bar{h} = (1 + \gamma)\bar{c} - a$ such that the solutions of the individual maximization program are described in Table 7.

h_t^i	$h_t^i \leq \bar{h}_t$	$h_t^i > \bar{h}_t$
c_t^i	\bar{c}	$\frac{h_t^i}{1+\gamma}$
n_t^i	$\frac{\gamma - \eta^i}{\gamma} \frac{h_t^i + a - \bar{c}}{\phi h_t^i}$	$\frac{\gamma - \eta^i}{1+\gamma} \frac{h_t^i + a}{\phi h_t^i}$
e_t^i	$\frac{\phi \eta^i}{\gamma - \eta^i} h_t^i$	

Table 7: Individual decisions in function of own's human capital

Decisions - When $h_t^i \leq \bar{h}$, the constraint c_t^i is binding and the Malthusian regime prevails. In this situation, provided that non-labor income is not high $a < \bar{c}$ (Assumption 1), fertility increases with h_t^i . An increase in parental human capital increases the opportunity cost of the time spent with children, which should depress fertility, but it also increases total income enough to finally increase both quality and quantity of children. Said differently, the income effect dominates the substitution effect. Once $h_t^i > \bar{h}$, the household enters the interior regime where an increase in labor income reduces fertility as, now, the opportunity cost effect dominates the income effect. The opposition between these two effects is illustrated in Figure 8. An increase in \bar{c} raises the range of h_t^i for which the household is trapped in a Malthusian situation.

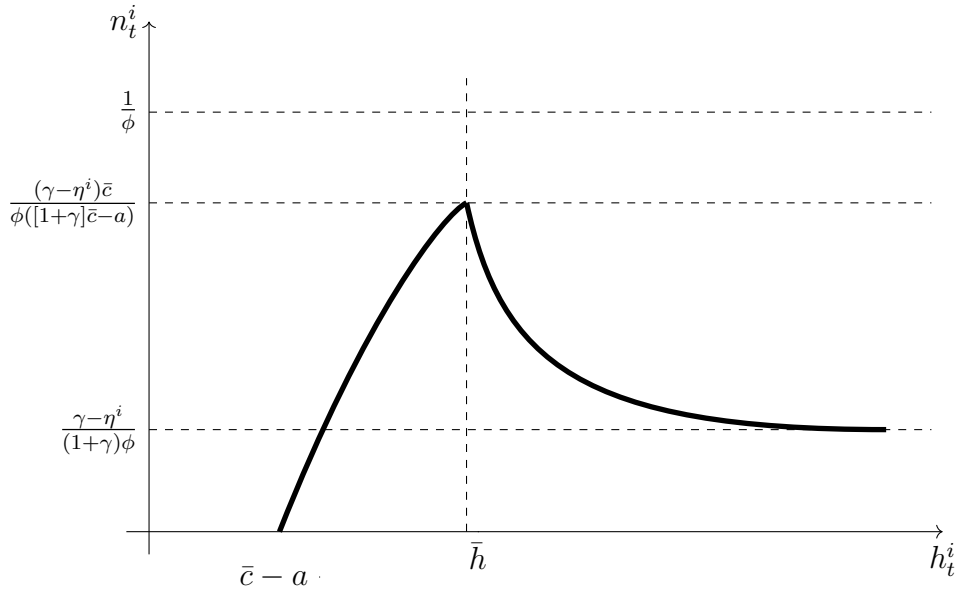


Figure 8: Fertility as a function of parents' human capital

QQ-tradeoff reversal - Parental investment into the education of children is not affected by the prevailing regime. This is simpler than the more complex models like Galor and Weil (2000) and De la Croix and Doepke (2003) but it does not alter the generality of our results and it allows us to characterize the accumulation of human capital over time in a simple way:

$$h_t^i = \left[\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right]^t h_0 \quad (5)$$

For each family i , human capital grows at a constant positive rate if and only if $\psi > \frac{\gamma - \eta^i}{\phi \eta^i}$. Then, members of a dynasty endowed initially with h_0^i will escape the Malthusian regime under the following condition:

$$h_t^i \geq (1 + \gamma)\bar{c} - a \Leftrightarrow t \geq \frac{\ln \frac{(1+\gamma)\bar{c}-a}{h_0}}{\ln \frac{\psi \phi \eta^i}{\gamma - \eta^i}} \equiv \bar{t}^i$$

From this condition, we get that $\frac{d\bar{t}^i}{d\eta^i} < 0$; it means that, for a given h_0 , the quality lovers escape the Malthusian regime sooner than the quantity lovers.

Assumption 2 $\psi > \frac{\gamma - \eta^i}{\phi \eta^i} \quad \forall i$.

From here on we limit our analysis to situations where Assumption 2 is fulfilled. Said differently, we limit our analysis to situations where human capital is strictly growing for all families. We have established that quality-oriented individuals escape the Malthusian trap sooner than their quantity-oriented counterparts. We now analyze fertility differentials between these two groups. Proposition 1 summarizes our results.

Proposition 1 *Under assumptions 1 and 2:*

- $\frac{\partial h_t^i}{\partial \eta^i} > 0 \quad \forall i$,
- $\forall t > \bar{t}^i, \frac{\partial n_t^i}{\partial \eta^i} < 0 \quad \forall i$,
- *there exists a date t_0^i such that:*

$$\bar{t}^i > t_0^i > 0,$$

$$\forall t \in (t_0^i, \bar{t}^i), \frac{\partial n_t^i}{\partial \eta^i} \geq 0 \quad \forall i.$$

Proof 1 *See Appendix C.*

The net impact of η^i on fertility is driven by the opposition between two effects. First, quality-oriented households (high η^i) have a stronger preference for human capital than quantity-oriented households (*preference effect*). Second, they are characterized by a stronger accumulation of human capital (*accumulation effect*). Proposition 1 states that around \bar{t} , in the Malthusian regime ($t < \bar{t}$), the accumulation effect dominates the preference effect such that quality-oriented parents have an evolutionary advantage over quantity-oriented parents. Once they enter into the Beckerian (interior) regime, they lose this advantage in favor of quality-oriented parents.⁵

Proposition 1 describes the evolution of the quality-quantity trade-off at the microeconomic level but remains silent about aggregated moments, which are the moments we estimate in the previous section.

Proposition 2 *There exist dates \hat{t} , \bar{t} , and \check{t} such that:*

$$\begin{aligned}\hat{t} &> \bar{t} > \check{t} > 0, \\ \forall t \in (\check{t}, \bar{t}), \frac{\partial n_t^i}{\partial \eta^i} &\geq 0 \quad \forall i, \\ \forall t > \hat{t}, \frac{\partial n_t^i}{\partial \eta^i} &< 0 \quad \forall i.\end{aligned}$$

Proof 2 *See Appendix D*

The difference between Proposition 1 and Proposition 2 is that with the first one, we define a collection of dates at which families transit from one regime to the other, i.e. each family has a specific date of transition; while in the second one, we differentiate two specific periods of time during which all families adopt the same type of fertility behaviors. From $t = \check{t}$ to $t = \bar{t}$, all families are in a Malthusian regime where the accumulation effect dominates the taste effect such that $\frac{\partial n_t^i}{\partial \eta^i} > 0$. Conversely, when $t > \bar{t}$, all families are in the interior regime such that $\frac{\partial n_t^i}{\partial \eta^i} < 0$. Figure 9 illustrates the result.

⁵The attentive reader has spotted that the dominance of the accumulation effect over the preference effect is not necessarily true at any date t within the Malthusian regime. In Appendix C, we show that from $t = 0$ to $t = t_0$, the taste effect dominates. Nevertheless, this situation is transitory and potentially corresponding to times uncovered by our data. Finally note that the size of this time window is proportional to $\frac{\eta^i}{\gamma}$, representing the weight of human capital relative to quantity of children in the utility function of agent i .

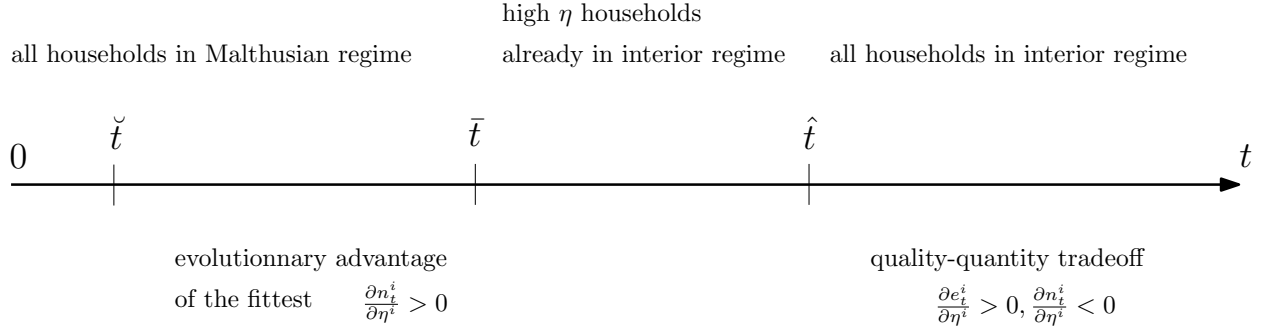


Figure 9: Proposition 2

In the intermediary period $t \in (\check{t}, \bar{t})$, the fertility behaviors of our families are heterogeneous as some of them will be in the Malthusian regime ($\frac{dn_t^i}{d\eta^i} > 0$) while others will be in the interior regime ($\frac{dn_t^i}{d\eta^i} < 0$).

Proposition 2 directly implies that overall, a linear regression model that would measure the association between sibship size and the human capital of individuals over a period $t \in (\check{t}, +\infty)$ would identify three distinct periods: a period in which sibship size and human capital are positively associated, followed by an absence of a significant relationship and then the emergence of a negative association.

5 CONCLUSION

Before the concept of human capital was introduced, growth theory relied mainly on physical capital, such as machinery, buildings, and equipment. The value of labor was viewed simply as the wage or salary paid to workers, and not a worker's investment in their own knowledge and skills. The concept of human capital challenged this view by recognizing that individuals can invest in themselves through education and training, which can increase their productivity and earning potential. This perspective shifted the focus from the cost of labor to the value of labor, and from the quantity of labor to the quality of labor.

The paradigm shift also led to the development of new analytical tools and methods for measuring the impact of human capital on economic growth. In their maturation process, these innovations faced first order difficulties: theoretical models of human capital had to rely on implausibly large externalities to ensure sustained growth, while applied research struggled to find a robust effect of education on growth at the aggregate level.

A critical step towards a more mature understanding of the role of human capital for growth

involved shifting the analytical focus away from the average level of literacy and skill and to look instead at the human capital of those at the upper tail of the distribution, commonly referred to as “upper tail human capital.” This paper adds to this growing body of work by examining the Academy movement of the eighteenth century, demonstrating how members of academic institutions altered their behavior ahead of the Industrial Revolution. We argue that one key mechanism underlying this change was the ability of human capital to generate wealth, as early as 1750, which allowed these individuals to transcend the Malthusian logic and engage in the modern trade-off between the quantity and quality of children.

Our results complement those of Galor and Weil (2000) and Galor and Moav (2002) in two ways. First, they validate empirically one of the key mechanisms of the Unified Growth Theory, the reversal of the QQ-tradeoff over time. This empirical confirmation strengthens the theoretical framework put forth in previous studies. Second, by placing this reversal one century before the Baltic Sea’s Industrial Revolution among an elite group, it shows that mechanisms complementary to firms’ increased demand for human capital, due to the Industrial Revolution, may have triggered the transition to behaviors compatible with sustained economic growth among particular groups.

ACKNOWLEDGEMENTS

We thank Matthew Curtis, Oded Galor, Ralf Meizenthal, Stelios Michalopoulos, Joel Mokyr, Stéphane Zuber, and participants at seminars at UNSW (Sidney), ANU (Canberra), Sussex, NYU Abu Dhabi, Paris School of Economics, EHA conference in Pittsburgh, and U. of Brown for comments during seminar. We thank Anna Maria Gkopi for excellent research assistance. David de la Croix acknowledges the financial support of the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under grant agreement No 883033 “Did elite human capital trigger the rise of the West? Insights from a new database of European scholars”. Thomas Baudin acknowledges the financial support of the French National Research Agency (ANR) under the PRC program under grant agreement ANR-23-CE26-0013 - FaMiGrowth.

REFERENCES

- Aaronson, Daniel, Fabian Lange, and Bhashkar Mazumder. 2014. “Fertility transitions along the extensive and intensive margins.” *American Economic Review* 104 (11): 3701–24.
- Anderson, Peter John. 1893. *Lists of Officers, University and King’s College: Aberdeen, 1495-1860*. Aberdeen: Aberdeen University Press.

- . 1898. *Fasti Academiae Mariscallanae Aberdonensis. Selections from the Records of the Marischal College and University*. New Spalding Club.
- Andreev, Evgeny M, Dmitri Jdanov, Vladimir M Shkolnikov, and David A Leon. 2011. “Long-term trends in the longevity of scientific elites: Evidence from the British and the Russian academies of science.” *Population Studies* 65 (3): 319–334.
- Applebaum, Wilbur. 2003. *Encyclopedia of the scientific revolution: from Copernicus to Newton*. New York: Routledge.
- Astro.uu.se. 2011. “History of astronomy in Uppsala.” <http://www.astro.uu.se/history/>.
- Bai, Yu, Yanjun Li, and Pak Hung Lam. 2023. “Quantity-quality trade-off in Northeast China during the Qing dynasty.” *Journal of Population Economics*, pp. 1–38.
- Bairoch, Paul, Jean Batou, and Pierre Chevre. 1988. *Population des villes européennes de 800 à 1850: banque de données et analyse sommaire des résultats (la)*. Volume 4. Librairie Droz.
- Baudin, Thomas, and Robert Stelter. 2022. “The rural exodus and the rise of Europe.” *Journal of Economic Growth* 27 (3): 365–414.
- Becker, Sascha O, Francesco Cinnirella, and Ludger Woessmann. 2010. “The trade-off between fertility and education: evidence from before the demographic transition.” *Journal of Economic Growth* 15 (3): 177–204.
- Bellemare, Marc F, and Casey J Wichman. 2020. “Elasticities and the inverse hyperbolic sine transformation.” *Oxford Bulletin of Economics and Statistics* 82 (1): 50–61.
- Bignon, Vincent, and Cecilia García-Peñalosa. 2021. “The toll of tariffs: Protectionism, education and fertility in late 19th century france.” CEPR Discussion Paper No. DP16069.
- Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2005. “The more the merrier? The effect of family size and birth order on children’s education.” *The Quarterly Journal of Economics* 120 (2): 669–700.
- Blanc, Guillaume. 2022. “Demographic Change and Development Using Crowdsourced Genealogies.” Mimeo.
- Bleakley, Hoyt, and Fabian Lange. 2009. “Chronic disease burden and the interaction of education, fertility, and growth.” *The review of economics and statistics* 91 (1): 52–65.
- Buringh, Eltjo. 2021. “The population of European cities from 700 to 2000: Social and economic history.” *Research Data Journal for the Humanities and Social Sciences* 6 (1): 1–18.

- Chaney, Eric. 2020. “Modern Library Holdings and Historic City Growth.” University of Oxford.
- Chao, Fengqing, Patrick Gerland, Alex Cook, and Leontine Alkema. 2019. “Systematic assessment of the sex ratio at birth for all countries and estimation of national imbalances and regional reference levels.” *Proceedings of the National Academy of Sciences* 116 (19): 9303–9311.
- Charpentier, Arthur, and Ewen Gallic. 2020. “Can historical demography benefit from the collaborative data of genealogy websites?” *Population* 75 (2): 379–408.
- Cinnirella, Francesco, Marc Klemp, and Jacob Weisdorf. 2017. “Malthus in the bedroom: Birth spacing as birth control in pre-transition England.” *Demography* 54 (2): 413–436.
- Cinnirella, Francesco, and Jochen Streb. 2017. “The role of human capital and innovation in economic development: evidence from post-Malthusian Prussia.” *Journal of economic growth* 22:193–227.
- Clark, Gregory. 2007. *A Farewell to Alms*. Princeton: Princeton University Press.
- Clark, Gregory, and Neil Cummins. 2019. “Randomness in the bedroom: There is no evidence for fertility control in pre-industrial England.” *Demography* 56 (4): 1541–1555.
- . 2023. “Birth Order and Social Outcomes, England, 1680-2022.” paper presented at the Tenth CEPR Economic History Symposium.
- Coutts, James. 1909. *A History of the University of Glasgow: From its Foundation in 1451 to 1909*. Glasgow: J. Maclehose and Sons.
- Cummins, Neil. 2017. “Lifespans of the European elite, 800–1800.” *The Journal of Economic History* 77 (2): 406–439.
- Curtis, Matthew, and David De la Croix. 2023. “Measuring Human Capital: from WorldCat Identities to VIAF.” *Repertorium Eruditorum Totius Europae* 10:17–22.
- Dahlgren, Erik Wilhelm. 1915. *Kungl. Svenska vetenskapsakademien: personförteckningar, 1739-1915*. Almqvist & Wiksells boktr.
- De la Croix, David. 2021. “Scholars and Literati in European Academia before 1800.” *Repertorium eruditorum totius Europae* 5:35–41.
- De la Croix, David, Frédéric Docquier, Alice Fabre, and Robert Stelter. 2023. “The Academic Market and the Rise of Universities in Medieval and Early Modern Europe.” *Journal of the European Economic Association*.
- De la Croix, David, and Matthias Doepke. 2003. “Inequality and Growth: Why Differential Fertility Matters.” *American Economic Review* 93 (4): 1091–1103.

- De la Croix, David, and Omar Licandro. 2013. “The child is father of the man: Implications for the demographic transition.” *The Economic Journal* 123 (567): 236–261.
- . 2015. “The longevity of famous people from Hammurabi to Einstein.” *Journal of Economic Growth* 20:263–303.
- De la Croix, David, Eric B. Schneider, and Jacob Weisdorf. 2019. “Childlessness, celibacy and net fertility in pre-industrial England: the middle-class evolutionary advantage.” *Journal of Economic Growth* 24 (3): 223–256 (Sep).
- De la Croix, David, and Mara Vitale. 2023. “Women in European Academia before 1800 – Religion, Marriage, and Human Capital.” *European Journal of Economic History* 27:506–532.
- De Candolle, Alphonse. 1885. *Histoire des sciences et des savants depuis deux siècles: précédée et suivie d’autres études sur des sujets scientifiques en particulier sur l’hérédité et la sélection dans l’espèce humaine*. Bale.
- Delen, Första, and Martin Johan Julius Weibull. 1868. *Lunds Universitets Historia, 1668-1868*. Lund: CWK Gleerup.
- Doepke, Matthias. 2015. “Gary Becker and the Quantity and Quality of Children.” *Journal of Demographic Economics* 81 (1): 59–66.
- Dupâquier, Jacques. 1993. “Généalogie et démographie historique.” *Annales de démographie historique*, pp. 391–395.
- Eloy, Nicolas FJ. 1755. *Dictionnaire Historique De La Médecine: Contenant Son Origine, Ses Progrès, ses Révolutions, ses Sectes & son Etat chez différens Peuples*. Bassompierre.
- Emerson, Roger L. 1981. “The Philosophical Society of Edinburgh 1748–1768.” *The British Journal for the History of Science* 14 (2): 133–176.
- Feenstra, Robert, Margreet Ahsmann, and Theo Johannes Veen. 2003. *Bibliografie van hoogleraren in de rechten aan de Franeker Universiteit tot 1811*. Amsterdam: Koninklijke Nederlandse Akademie van Wetenschappen.
- Fouquet, Roger, and Stephen Broadberry. 2015. “Seven centuries of European economic growth and decline.” *Journal of Economic Perspectives* 29 (4): 227–244.
- Frijhoff, Willem. 1996. “Patterns.” Chapter 2 of *A History of the University in Europe. Vol. II: Universities in Early Modern Europe (1500–1800)*, edited by Hilde de Ridder-Symoens. Cambridge University Press.
- Galor, Oded. 2011. *Unified growth theory*. Princeton University Press.

- Galor, Oded, and Marc Klemp. 2019. “Human genealogy reveals a selective advantage to moderate fecundity.” *Nature ecology & evolution* 3 (5): 853–857.
- Galor, Oded, and Omer Moav. 2002. “Natural selection and the origin of economic growth.” *The Quarterly Journal of Economics* 117 (4): 1133–1191.
- Galor, Oded, and David Weil. 2000. “Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond.” *American Economic Review* 90 (4): 806–28.
- Galton, Francis. 1875. *English men of science: Their nature and nurture*. D. Appleton.
- Gavrilov, Leonid, Natalia Gavrilova, Jay Olshansky, and Bruce Carnes. 2002. “Genealogical data and the biodemography of human longevity.” *Social Biology* 49 (3-4): 160–173.
- Gertz, Otto. 1940. *Kungl. fysiografiska sällskapet i Lund, 1772-1940: historisk överblick och personförteckningar*. Håkan Ohlssons boktryckeri.
- Grant, Alexander. 1884. *The story of the University of Edinburgh during its first three hundred years*. Longmans, Green, and Company.
- Hatton, Timothy J, and Richard M Martin. 2010. “The effects on stature of poverty, family size, and birth order: British children in the 1930s.” *Oxford Economic Papers* 62 (1): 157–184.
- Hirsch, Theodor. 1837. *Geschichte des akademischen Gymnasiums in Danzig, in ihren Hauptzügen dargestellt*. Wedelschen Hofbuchdruckerei.
- Inno, Karl. 1972. *Tartu University in Estonia during the Swedish rule (1632-1710)*. Vaba Eesti.
- Jensen, Mikkel Munthe. 2018. “From learned cosmopolitanism to scientific internationalism: the patriotic transformation of Nordic academia and academic culture during the long eighteenth century.” Ph.D. diss., European Institute.
- Kaplanis, Joanna, Assaf Gordon, Tal Shor, Omer Weissbrod, Dan Geiger, Mary Wahl, Michael Gershovits, Barak Markus, Mona Sheikh, Melissa Gymrek, et al. 2018. “Quantitative analysis of population-scale family trees with millions of relatives.” *Science* 360 (6385): 171–175.
- Karlberg, Gun. 1977. *Royal Society of sciences of Uppsala, catalogue of publications 1716-1967*. Acta Universitatis Upsaliensis.
- Klemp, Marc, and Jacob Weisdorf. 2019. “Fecundity, fertility and the formation of human capital.” *The Economic Journal* 129 (618): 925–960.

- Klinge, Matti, Jan-Ivar Lindén, Kerstin Smeds, Rainer Knapas, Anto Leikola, and John Strömberg. 1988. *Helsingfors universitet 1640-1990: Kungliga akademien i Åbo 1640-1808*. Otava.
- Koschnick, Julius, Erik Hornung, and Francesco Cinnirella. 2022. “Flow of Ideas: Economic Societies and the Rise of Useful Knowledge.” CESifo Working Paper.
- Krüger, Kersten. 2019. “Catalogus Professorum Rostochiensium.” <http://cpr.uni-rostock.de/>.
- Leridon, Henri, and Jonathan Mandelbaum. 2004. “The demography of a learned society.” *Population* 59 (1): 81–114.
- Lomholt, Asger. 1950. *Det Kongelige Danske Videnskabernes Selskab 1742—1942*. København: Ejnar Munksgaard.
- Madsen, Jakob, and Holger Strulik. 2023. “Testing unified growth theory: Technological progress and the child quantity-quality tradeoff.” *Quantitative Economics* 14 (1): 235–275.
- Murphy, Tommy E. 2015. “Old habits die hard (sometimes).” *Journal of Economic Growth* 20 (2): 177–222.
- Napjus, J. W., and Gerrit Arie Lindeboom. 1985. *De hoogleraren in de geneeskunde aan de Hogeschool en het Athenaeum te Franeker*. Nieuwe Nederlandse bijdragen tot de geschiedenis der geneeskunde en der natuurwetenschappen. Amsterdam: Rodopi.
- Naragon, Steve. 2006. “Kant in the Classroom: Materials to Aid the Study of Kant’s Lectures.” Manchester College. <http://www.manchester.edu/kant/Home/index.htm>.
- Pearson Karl, LA, and BM Leslie. 1899. “Genetic (reproductive) selection: inheritance of fertility in man, and of fecundity in thoroughbred racehorses.” *Philos. Trans. R. Soc. Lond. Ser. A* 192:257–330.
- Pedersen, Olaf. 1992. *Lovers of Learning: A History of the Royal Danish Academy of Sciences and Letters, 1742-1992*. Munksgaard.
- Perrin, Faustine. 2013. “Gender equality and economic growth in the long-run: a cliometric analysis.” Ph.D. diss., Université de Strasbourg.
- Reed, Frank E. 1999. *Centennia: Historical Atlas*. Clockwork Software Incorporated.
- RSE. 2006. “Former Fellows of the Royal Society of Edinburgh.” PDF document.
- Sandberg, Lars G. 1979. “The case of the impoverished sophisticate: human capital and Swedish economic growth before World War I.” *The Journal of Economic History* 39 (1): 225–241.

- Schumann, Eduard. 1893. *Geschichte der Naturforschenden Gesellschaft in Danzig, 1743-1892*. Die Gesellschaft.
- Schwinges, Rainer Christoph, and Christian Hesse. 2019. “Repertorium Academicum Germanicum.” <https://en.rag-online.org/>.
- Shemivot, Vladimir Petrovich. 1873. “General list of members of the Academy of Sciences since its foundation.” St. Petersburg.
- Shiue, Carol H. 2017. “Human capital and fertility in Chinese clans before modern growth.” *Journal of Economic Growth* 22:351–396.
- Skirbekk, Vegard. 2008. “Fertility trends by social status.” *Demographic research* 18:145–180.
- Slottved, Ejvind. 1978. *Lærestole og lærere ved Københavns Universitet 1537-1977*. Copenhagen: Samfundet for dansk Genealogi og Personallhistorie.
- Smart, Robert Noyes. 2004. *Biographical Register of the University of St. Andrews, 1747-1897*. University of St. Andrews Library.
- Squicciarini, Mara P, and Nico Voigtländer. 2015. “Human capital and industrialization: Evidence from the age of enlightenment.” *The Quarterly Journal of Economics* 130 (4): 1825–1883.
- Stelter, Robert, and Diego Alburez-Gutierrez. 2022. “Representativeness is crucial for inferring demographic processes from online genealogies: Evidence from lifespan dynamics.” *Proceedings of the National Academy of Sciences* 119, no. 10.
- Stelter, Robert, David De la Croix, and Mikko Myrskylä. 2021. “Leaders and Laggards in Life Expectancy among European Scholars from the Sixteenth to the Early Twentieth Century.” *Demography* 58:111–135.
- Taisand, Pierre. 1721. *Les vies des plus celebres jurisconsultes de toutes les nations, tant anciens que modernes, savoir, latins ou romains, françois, espagnols, italiens, allemans, anglois, hollandois*. Paris: Sevestre.
- Van de Putte, Bart, and Andrew Miles. 2005. “A social classification scheme for historical occupational data: partner selection and industrialism in Belgium and England, 1800-1918.” *Historical Methods* 38 (2): 61–92.
- Vogl, Tom. 2016. “Differential fertility, human capital, and development.” *The Review of Economic Studies* 83 (1): 365–401.
- Volbehrr, Friedrich, and Richard Weyl. 1956. *Professoren und Dozenten der Christian-Albrechts-Universität zu Kiel: 1665 - 1954*. Kiel: Hirt.

- Von Bahr, Gunnar. 1945. *Medicinska fakulteten i Uppsala*. Stockholm: Almqvist & Wiksell International.
- Wrigley, Edward., and Roger Schofield. 1983. “English Population History from Family Reconstitution: Summary Results 1600-1799.” *Population Studies* 37 (2): 157–184.
- Zanardello, Chiara. 2023. “Early Modern Academies, Universities, and Growth.” UCLouvain.

A DESCRIPTIVE STATISTICS

Table 8: Descriptive statistics

	Mean	St.Dev.	Min	Max		Mean	St.Dev.	Min	Max
nworks	22.973	28.443	0	397	ulund	0.077	0.266	0	1
asinh(nworks)	2.976	1.578	0.000	6.677	copenhagen	0.100	0.301	0	1
longevity	66.329	13.237	20	100	acaddnk	0.053	0.224	0	1
age nomination	38.235	11.867	13	85	groningen	0.017	0.129	0	1
date of birth	1,698	58.320	1,435	1,777	uppsala	0.081	0.273	0	1
geni	0.864	0.343	0	1	rostock	0.043	0.202	0	1
geneanet	0.050	0.217	0	1	kiel	0.017	0.131	0	1
field theo	0.171	0.372	0	1	abo	0.045	0.207	0	1
field law	0.118	0.320	0	1	dorpat	0.013	0.111	0	1
field med	0.162	0.355	0	1	amsterdam	0.008	0.087	0	1
field sci	0.265	0.426	0	1	konigsberg	0.009	0.096	0	1
SibshipSize	5.275	3.446	1	22	stockholm	0.134	0.341	0	1
No. sisters	1.971	2.022	0	14	glasgow	0.011	0.104	0	1
HalfSiblings	0.215	0.411	0	1	greifswald	0.025	0.156	0	1
No.Descendants	3.316	3.515	0	22	Gdanzig	0.008	0.090	0	1
YearDeathFather	1,725	59.425	1,451	1,824	petersburg	0.052	0.223	0	1
RankMale	1.844	1.198	1	10	alund	0.044	0.206	0	1
Rank	2.601	2.134	1	15	franeker	0.016	0.127	0	1
MaleDescendants	1.702	2.026	0	13	haarlem	0.058	0.233	0	1
SibshipSizeMale	3.304	2.070	1	16	auppsala	0.037	0.189	0	1
Urban	0.483	0.500	0	1	aberdeen	0.022	0.146	0	1
elder	0.417	0.493	0	1	andrews	0.010	0.101	0	1
soc. class top	0.397	0.490	0	1	Adanzig	0.010	0.101	0	1
soc. class mid	0.307	0.461	0	1	edinburgh	0.025	0.155	0	1
soc. class bot	0.026	0.160	0	1	aedinburgh	0.085	0.278	0	1
soc. class na	0.269	0.444	0	1					

Note: N=1,834 observations except for YearDeathFather (1,738), Rank (1,199), RankMale (1,449)

B ADDITIONAL FIGURES

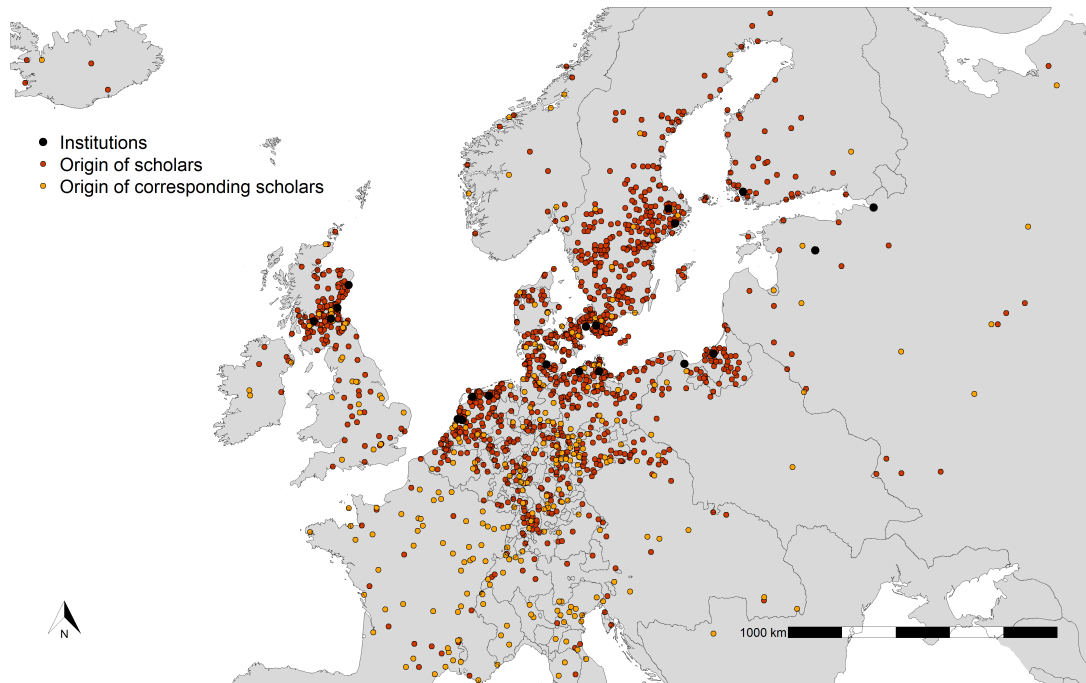
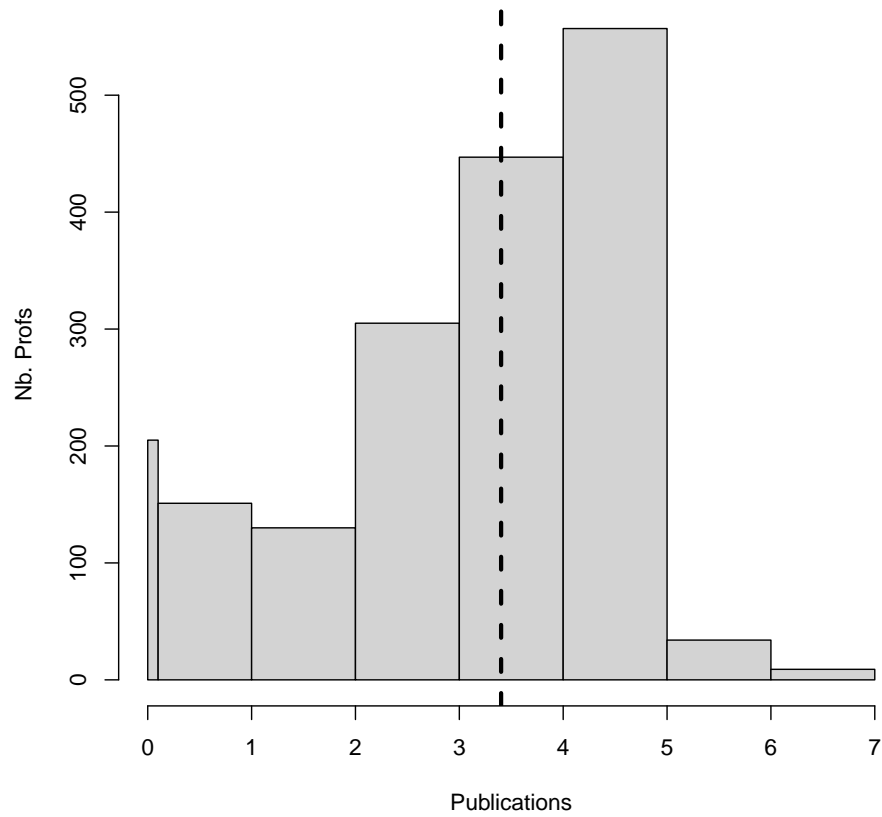


Figure 10: Origin of all scholars, frontiers of 1700



Note: Publications = $\text{asinh}(\text{Number of titles in VIAF})$. Median as dashed line.

Figure 11: Histogram of the distribution of number of publications

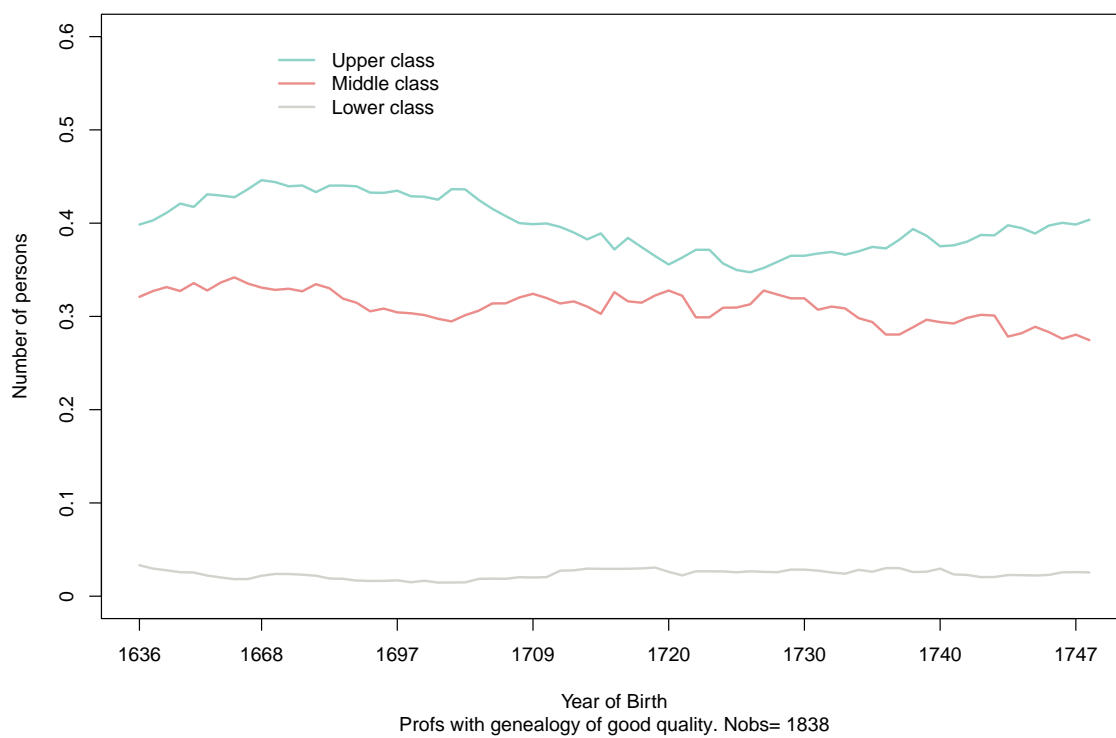


Figure 12: Dynamics of parental social class distribution

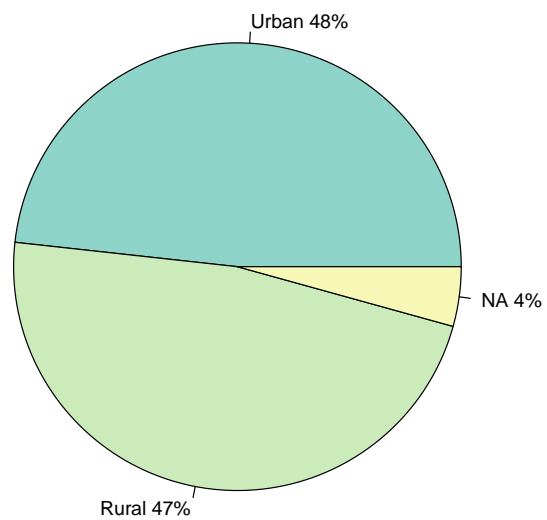


Figure 13: Pie Chart of Parental Origin



Figure 14: Dynamics of the share of scholars born in cities



Figure 15: Dynamics of sex-ratio by birth cohort

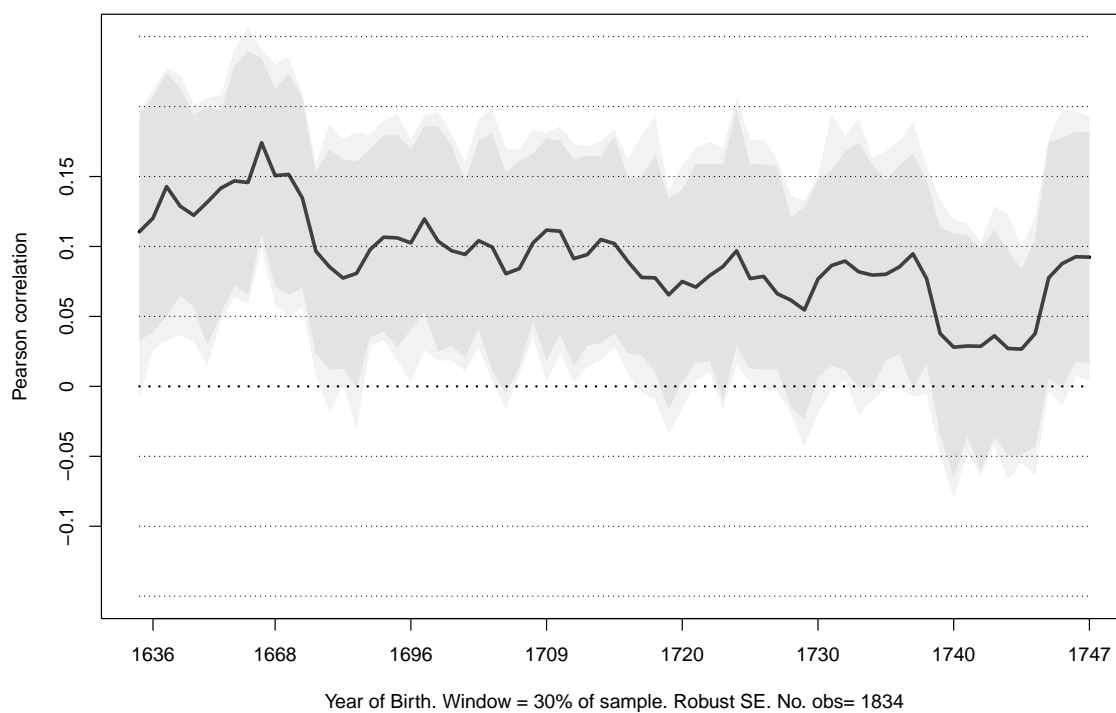


Figure 16: Inter-generational correlation of fertility

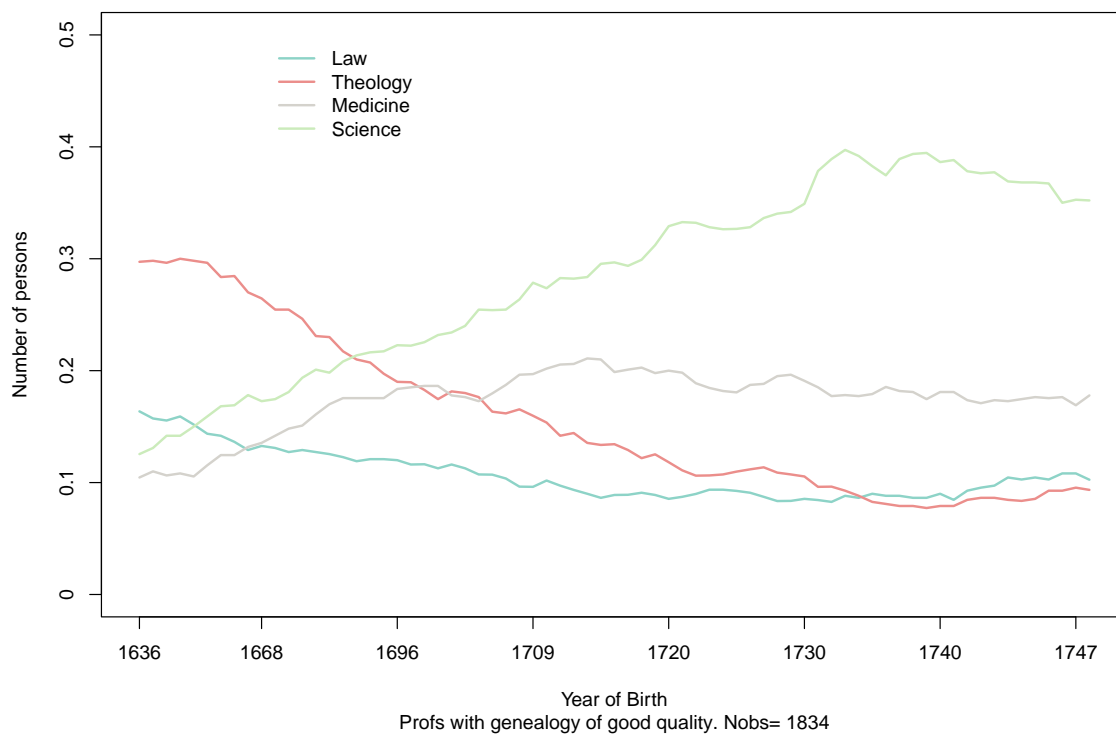


Figure 17: Share of academic fields over time

C PROOF OF PROPOSITION 1

In order to prove Proposition 1, we first analyze the dynamics of human capital, which remains the same in every regime. From Equation 5, we know that:

$$\frac{\partial h_t^i}{\partial \eta^i} = t \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right)^{t-1} \frac{\psi \phi \gamma}{(\gamma - \eta^i)^2} h_0 > 0. \quad (6)$$

It implies that the higher η^i , the higher the level of human capital for a given h_0 .

We now look at fertility differentials in the interior regime where $h_t^i > \bar{h}$. From Table 7, we get:

$$\frac{\partial n_t^i}{\partial \eta^i} = -\frac{1}{\phi(1+\gamma)} \left[\frac{h_t^i + a}{w h_t^i} + \frac{\gamma - \eta^i}{1 + \gamma} \frac{a \frac{\partial h_t^i}{\partial \eta^i}}{(h_t^i)^2} \right] < 0,$$

We then get that within the interior regime where $t > \bar{t}^i$, $\frac{dn_t^i}{d\eta^i} < 0$.

Proposition 1 states that when individuals shift from the Malthusian regime where $c_t^i = \bar{c}$ to the interior regime, the Malthusian regime is characterized by an evolutionary advantage for the quality oriented individuals. In order to prove this result, we first determine under which condition this evolutionary effect may arise. To do so, we first differentiate n_t^i with respect to η^i when $t < \bar{t}^i$; it yields to:

$$\frac{dn_t^i}{d\eta^i} = -\frac{1}{\gamma \phi h_t^i} \left[\bar{c} - a - h_t^i + (\gamma - \eta^i)(c - a) \frac{\frac{\partial h_t^i}{\partial \eta^i}}{h_t^i} \right] \quad (7)$$

From Eq. 5, we get that $\frac{\frac{\partial h_t^i}{\partial \eta^i}}{h_t^i} = \frac{\gamma}{\eta^i(\gamma - \eta^i)} t$ such that:

$$\frac{dn_t^i}{d\eta^i} \geq 0 \iff -h_0 \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right)^t + \bar{c} - a + (\bar{c} - a) \frac{\gamma}{\eta^i} t \leq 0 \quad (8)$$

Eq. 8 is the condition such that, for a given initial endowment of human capital, quality oriented individuals have a higher fertility than quantity oriented individuals. This equation is of the form $a\lambda^x + bx + c = 0$ when it is satisfied at equality. Such kind of equations admit at most two solutions but also potentially none. These solutions are of the form: $x = -\frac{W(\Delta \ln \lambda)}{\ln \lambda} - \frac{b}{c}$, where $W(\cdot)$ is a Lambert W Function with $\Delta = \frac{a}{b} \lambda^{-\frac{c}{b}}$. If $\Delta \ln \lambda > 0$ or $\Delta \ln \lambda = -\frac{1}{e}$, only one solution exists and corresponds to $x = -\frac{W_0(\Delta \ln \lambda)}{\ln \lambda} - \frac{b}{c}$; when

$\Delta \ln \lambda \in]-\frac{1}{e}, 0[$, two solutions exist $x = -\frac{W_0(\Delta \ln \lambda)}{\ln \lambda} - \frac{b}{c}$ and $x = -\frac{W_{-1}(\Delta \ln \lambda)}{\ln \lambda} - \frac{b}{c}$. Finally, when $\Delta < -\frac{1}{e}$, the equation does not admit any real solution.

In the present case, we get that:

$$\Delta \ln \lambda = -\ln \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right) \frac{\eta^i}{\gamma} \frac{h_0}{\bar{c} - a} \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right)^{-\frac{\eta^i}{\gamma}} < 0$$

Then, we may be in two situations: if $h_0 \geq \frac{\frac{\gamma}{\eta^i} \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right)^{\frac{\eta^i}{\gamma}}}{\ln \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right)} (\bar{c} - a)e$, $\Delta \ln \lambda < -\frac{1}{e}$ and Inequation 8 is never satisfied. An easy way to check this is to inspect Equation 8 for $h_0 \rightarrow +\infty$.

Conversely, if $h_0 < \frac{\frac{\gamma}{\eta^i} \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right)^{\frac{\eta^i}{\gamma}}}{\ln \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right)} (\bar{c} - a)e$, $\frac{dn^i}{d\eta^i} = 0$ admits two solutions:

$$\begin{aligned} t_0^i &= -\frac{W_0 \left(-\ln \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right) \frac{\eta^i}{\gamma} \frac{h_0}{\bar{c} - a} \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right)^{-\frac{\eta^i}{\gamma}} \right)}{\ln \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right)} - \frac{\eta^i}{\gamma} \\ \text{and } t_1^i &= -\frac{W_{-1} \left(-\ln \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right) \frac{\eta^i}{\gamma} \frac{h_0}{\bar{c} - a} \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right)^{-\frac{\eta^i}{\gamma}} \right)}{\ln \left(\frac{\psi \phi \eta^i}{\gamma - \eta^i} \right)} - \frac{\eta^i}{\gamma} \end{aligned}$$

The properties of the Lambert W-function imply that $t_1 > t_0$ and that $\frac{dn^i}{d\eta^i} > 0$ only when $t \in]t_0, t_1[$. This result can be easily visualized by re-arranging Equation 8 and using logs, which yields to the following condition:

$$\frac{\partial n_t^i}{\partial \eta^i} \geq 0 \Leftrightarrow LHS(\eta^i) \equiv \ln \frac{\bar{c} - a}{h_0} + \ln \left(1 + \frac{\gamma}{\eta^i} t \right) \geq t \ln \frac{\psi \phi \eta^i}{\gamma - \eta^i} \equiv RHS(\eta^i) \quad (9)$$

From this figure, we can see that the parametric condition $h_0 > \bar{c} - a$ that we imposed along the development of our model, implies that t_0 is always positive.

We have now to remember that an individual i escapes the Malthusian regime at date $t = \bar{t}^i$ corresponding to a level of human capital \bar{h} . It is then crucial to locate \bar{t}^i with respect to t_0^i and t_1^i . Indeed, if for instance $\bar{t}^i < t_0^i$, the evolutionary advantage of the quality oriented individuals would never prevail in the Malthusian regime. In order to locate \bar{t}^i , we need to

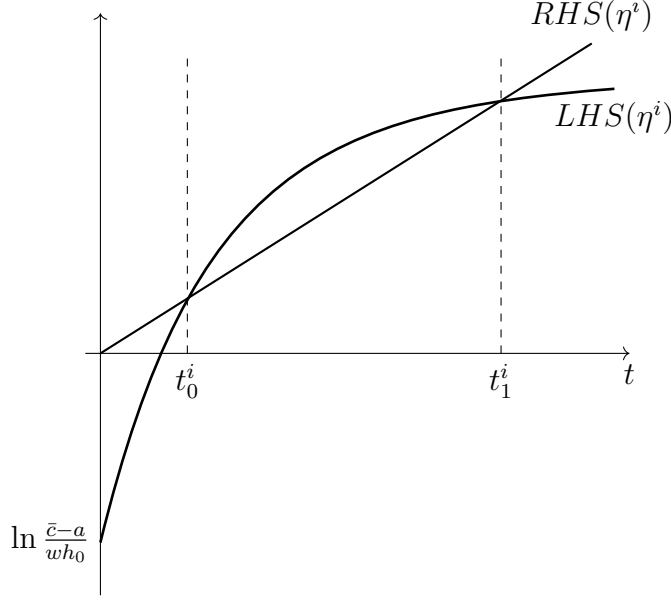


Figure 18: Income expansion path of quality and quantity.

express Equation 7 for $t = \bar{t}^i$ and $h_t = \bar{h}$. Doing this yields to the following condition:

$$\begin{aligned} \frac{dn^i}{d\eta^i} > 0 &\iff -(\bar{h} + a - \bar{c}) + \frac{(\bar{c} - a)\gamma}{\eta^i} \bar{t}^i > 0 \\ &\iff h_0 < e^{\ln((1+\gamma)\bar{c}-a) - \frac{\bar{c}\eta^i}{\bar{c}-a} \ln\left(\frac{\Psi\phi\eta^i}{\gamma-\eta^i}\right)} \end{aligned}$$

Proposition 1 then follows as for any $h_0 < \min\left\{e^{\ln((1+\gamma)\bar{c}-a) - \frac{\bar{c}\eta^i}{\bar{c}-a} \ln\left(\frac{\Psi\phi\eta^i}{\gamma-\eta^i}\right)}, \frac{\frac{\gamma}{\eta^i} \left(\frac{\Psi\phi\eta^i}{\gamma-\eta^i}\right)^{\frac{\eta^i}{\gamma}}}{\ln\left(\frac{\Psi\phi\eta^i}{\gamma-\eta^i}\right)} (\bar{c} - a)e\right\}$, t_0^i and t_1^i exists and $\bar{t}^i \in]t_0^i, t_1^i[$. ■

D PROOF OF PROPOSITION 2

Proposition 2 states that we can identify periods (\check{t}, \bar{t}) and $(\hat{t}, +\infty)$ during which all families are characterized by the same qualitative influence of the preference for quality η^i on their fertility behaviors. In order to prove this statement, we need to inspect more closely the properties of \bar{t}^i and t_0^i . First, we know that both of them depend on η^i , which is distributed on a set \mathcal{E} . Let's denote the minimal and maximal value of η^i on \mathcal{E} respectively η^{MIN} and η^{MAX} . We also know that $\frac{d\bar{t}^i}{d\eta^i} < 0$. It implies that \bar{t}^i is minimum when $\eta^i = \eta^{MAX}$ and maximum when $\eta^i = \eta^{MIN}$. Let's denote these two values respectively \bar{t}^{MIN} and \bar{t}^{MAX} .

We now turn our attention to t_0^i . First, we know that $t_0^i < \bar{t}^i \forall i$. Second, a close inspection

of Figure 18 indicates that when η^i increases, t_0^i increases too. It implies that t_0^i is maximum for $\eta^i = \eta^{MAX}$, let's denote this value $t_{0,MAX}^i$. Consequently, $0 < t_{0,MAX}^i < \bar{t}^{MIN} < \bar{t}^{MAX}$. Let's finally make a notation change such that we denote $\bar{t}^{MAX} \equiv \hat{t}$, $\bar{t}^{MIN} \equiv \bar{t}$ and $t_{0,MAX}^i \equiv \check{t}$ and Proposition 2 directly follows. ■

E DETAILED REGRESSION RESULTS

See table next page

Dependent variable: log number of works

quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
(Intercept)	1.821 (1.111)	0.361 (1.275)	1.854*** (0.498)	1.576 (1.094)	1.141** (0.490)	0.618 (0.966)	3.563** (0.851)	1.537* (0.494)
SibshipSizeMale	0.048* (0.029)	0.079*** (0.028)	-0.022 (0.041)	-0.028 (0.041)	-0.081** (0.037)	-0.027 (0.035)	-0.067** (0.033)	-0.050 (0.035)
longevity	0.035*** (0.005)	0.030*** (0.005)	0.032*** (0.006)	0.021*** (0.007)	0.013** (0.006)	0.015** (0.006)	0.013** (0.006)	0.017*** (0.006)
ageentry	-0.020*** (0.007)	-0.016** (0.007)	-0.018*** (0.007)	-0.012** (0.006)	-0.014*** (0.005)	-0.010 (0.006)	-0.009 (0.007)	0.004 (0.009)
law	-0.623*** (0.198)	-0.670*** (0.221)	-0.673*** (0.254)	-0.544** (0.263)	-0.657** (0.250)	-0.790*** (0.259)	-0.634** (0.262)	-0.529** (0.248)
med	0.188 (0.226)	-0.223 (0.206)	-0.201 (0.230)	-0.220 (0.236)	-0.039 (0.228)	0.089 (0.227)	0.239 (0.228)	0.164 (0.226)
theo	0.558*** (0.165)	0.213 (0.175)	0.168 (0.207)	0.026 (0.228)	-0.048 (0.248)	0.070 (0.251)	0.156 (0.254)	0.329 (0.214)
sci	-0.273 (0.279)	-0.221 (0.245)	-0.002 (0.238)	0.161 (0.206)	0.338* (0.195)	0.261 (0.201)	0.437** (0.204)	0.253 (0.206)
WeakLinkTRUE	0.404 (0.874)	1.211*** (0.325)	1.067*** (0.251)	0.954*** (0.199)	1.145*** (0.143)	1.094*** (0.145)	1.117*** (0.159)	0.902*** (0.178)
HalfSiblingsTRUE	-0.047 (0.170)	-0.267 (0.166)	-0.074 (0.189)	-0.108 (0.200)	0.480** (0.196)	0.312 (0.193)	0.426** (0.168)	0.258 (0.168)
Urban	0.251* (0.125)	0.401*** (0.135)	0.157 (0.125)	0.048 (0.125)	-0.102 (0.125)	-0.212 (0.125)	-0.149 (0.125)	-0.080 (0.125)

quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
geni	(0.143)	(0.142)	(0.167)	(0.170)	(0.156)	(0.154)	(0.149)	(0.150)
	0.117	0.151	0.325	0.133	0.258	0.091	0.220	0.457**
	(0.202)	(0.205)	(0.233)	(0.239)	(0.206)	(0.213)	(0.215)	(0.211)
copenhagen	-1.234	0.685	-0.689**	0.707	1.564***	2.020	-1.218	-0.213
	(0.977)	(1.226)	(0.302)	(1.074)	(0.349)	(0.844)	(0.704)	(0.351)
	0.165	1.800	0.617	1.695	2.935***	3.026**	-0.351	0.553
groningen	(1.021)	(1.267)	(0.582)	(1.086)	(0.351)	(0.875)	(0.766)	(0.355)
	0.139	1.703	0.737*	1.712	2.860***	3.249**		1.119**
	(1.005)	(1.237)	(0.354)	(1.086)	(0.343)	(0.874)		(0.304)
alund			0.597	0.654	1.136***	1.326	-1.813	-0.512
			(0.674)	(0.996)	(0.338)	(0.840)	(0.698)	(0.358)
	-1.372	0.098	-0.999***	-0.551	0.256	0.385	-2.442*	-1.012*
ulund	(1.014)	(1.234)	(0.324)	(1.083)	(0.344)	(0.859)	(0.716)	(0.325)
	-0.420	1.242	-0.033	0.845	2.054***	2.238*	-0.631	0.350
	(0.992)	(1.224)	(0.302)	(1.083)	(0.299)	(0.839)	(0.704)	(0.363)
uppsala	0.091	1.559	0.052	0.877	1.918***	1.971	-1.014	-0.111
	(1.203)	(1.259)	(0.310)	(1.041)	(0.322)	(0.843)	(0.687)	(0.321)
	-1.213	0.179	-1.346**	-0.301	0.473	0.592	-2.272**	-1.291**
acaddnk	(1.003)	(1.265)	(0.508)	(1.174)	(0.614)	(0.962)	(0.757)	(0.400)
	-0.804	0.879	-0.402	-1.542	0.093	0.088	-2.243**	-0.644
	(1.025)	(1.263)	(0.420)	(1.260)	(1.059)	(0.954)	(0.790)	(0.413)
rostock								
kiel								

quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
abo	-1.405 (1.001)	0.074 (1.234)	-1.290*** (0.329)	-0.054 (1.105)	1.206*** (0.425)	1.452 (0.907)	-1.412 (0.789)	-0.554 (0.421)
dorpat	-0.645 (1.006)	0.855 (1.238)	-0.339 (0.301)					
amsterdam	-0.347 (0.996)	1.288 (1.231)	0.389 (0.321)	1.750 (1.141)	2.473** (0.679)	2.788* (1.053)	-0.822 (1.053)	1.208* (0.351)
Adanzig					1.013** (0.353)	1.465 (0.854)	-1.661* (0.722)	-0.293 (0.518)
Gdanzig	0.330 (1.016)	2.094 (1.237)	1.042** (0.378)	2.218 (1.082)	1.768 (1.914)	2.122 (1.313)	-1.716 (0.820)	
edinburgh	-1.246 (1.066)	0.139 (1.297)	-0.823 (0.505)	0.637 (1.110)	1.713*** (0.496)	1.116 (0.963)	-2.221* (0.838)	-0.756 (0.439)
konigsberg	-0.498 (1.192)		-0.550 (1.105)	1.083 (1.404)	2.121** (0.769)	2.289* (1.071)	-1.243 (0.894)	0.491 (0.331)
stockholm	-1.395 (1.074)	0.338 (1.245)	-0.601** (0.279)	0.396 (1.071)	1.438*** (0.251)	1.553 (0.821)	-1.507 (0.686)	-0.438 (0.318)
glasgow	-2.581** (1.129)	-0.887 (1.507)	-1.501* (0.703)	-0.208 (1.318)	0.361 (0.668)	0.467 (1.181)	-2.545 (1.306)	-1.160 (1.432)
greifswald	-1.147 (1.023)	0.947 (1.258)	-0.401 (0.502)	1.125 (1.147)	2.060** (0.460)	2.203* (0.888)	-1.295 (0.924)	-0.140 (0.581)
petersburg		1.082 (1.248)	-0.075 (0.284)	0.962 (1.108)	1.824*** (0.291)	1.826 (0.836)	-1.488 (0.702)	-0.507 (0.416)

quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
haarlem		1.068 (1.288)	-0.006 (0.316)	0.966 (1.089)	1.942*** (0.270)	1.953 (0.826)	-1.225 (0.696)	0.021 (0.312)
aedinburgh	-1.407 (1.386)	-0.403 (1.295)	-1.014*** (0.373)	0.137 (1.094)	1.460*** (0.314)	1.692 (0.838)	-1.521 (0.685)	-0.706 (0.310)
auppsala	-0.400 (0.942)	0.961 (1.250)	-0.072 (0.286)	0.928 (1.087)	2.296*** (0.292)	2.327* (0.839)	-0.416 (0.696)	0.401 (0.439)
aberdeen	-2.504** (1.035)	-0.943 (1.300)	-1.805*** (0.611)	-0.093 (1.299)	0.849 (0.804)	1.490 (1.176)	-1.822 (1.181)	-2.531*** (0.354)
andrews	-1.099 (1.118)	-1.061 (1.479)	-3.288*** (0.492)	-2.033 (1.144)			-2.705** (0.793)	-1.599** (0.578)
Adj. R ²	0.223	0.258	0.217	0.202	0.223	0.235	0.240	0.213
Num. obs.	550	550	551	550	550	551	550	551
N Clusters	500	468	414	385	370	366	375	403

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 9: Rolling regression results

F BALANCE TESTS

Dependent variable: log number of works								
quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
(Intercept)	3.10 (1.35)	3.33*** (0.50)	3.49*** (0.60)	2.40** (0.47)	3.07*** (0.43)	3.29*** (0.29)	3.29*** (0.29)	2.26*** (0.28)
has genealogy	0.44*** (0.10)	0.31*** (0.10)	0.63*** (0.11)	0.86*** (0.10)	0.90*** (0.10)	0.70*** (0.10)	0.58*** (0.10)	0.60*** (0.10)
Institut. FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R ²	0.13	0.13	0.18	0.22	0.23	0.19	0.17	0.18
Num. obs.	1346	1348	1376	1370	1373	1375	1377	1361
N Clusters	1257	1209	1169	1085	1038	1017	1024	1058

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 10: Balance test: publications

Dependent variable: longevity (in years)								
quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
(Intercept)	75.23*** (2.14)	76.16*** (1.48)	75.40*** (1.90)	58.76*** (5.18)	61.26*** (4.02)	67.76*** (4.08)	68.67*** (4.33)	61.46*** (3.92)
has genealogy	2.66*** (0.83)	1.37 (0.88)	1.94** (0.90)	2.24** (0.90)	2.97*** (0.89)	1.73* (0.89)	3.00*** (0.90)	4.18*** (0.94)
Institut. FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R ²	0.06	0.11	0.12	0.11	0.06	0.05	0.04	0.03
Num. obs.	1346	1348	1376	1370	1373	1375	1377	1361
N Clusters	1257	1209	1169	1085	1038	1017	1024	1058

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 11: Balance test: longevity

Dependent variable: work in science (0/1)								
quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
(Intercept)	−0.00	0.38	0.50**	0.26	0.00	0.09	0.17	0.33*
	(0.01)	(0.24)	(0.19)	(0.25)	(0.01)	(0.10)	(0.11)	(0.17)
has genealogy	0.01	−0.02	−0.01	−0.01	−0.00	0.01	0.00	0.01
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Institut. FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R ²	0.04	0.12	0.16	0.15	0.13	0.14	0.13	0.15
Num. obs.	1346	1348	1376	1370	1373	1375	1377	1361
N Clusters	1257	1209	1169	1085	1038	1017	1024	1058

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 12: Balance test: science

G ADDITIONAL ROBUSTNESS CHECKS

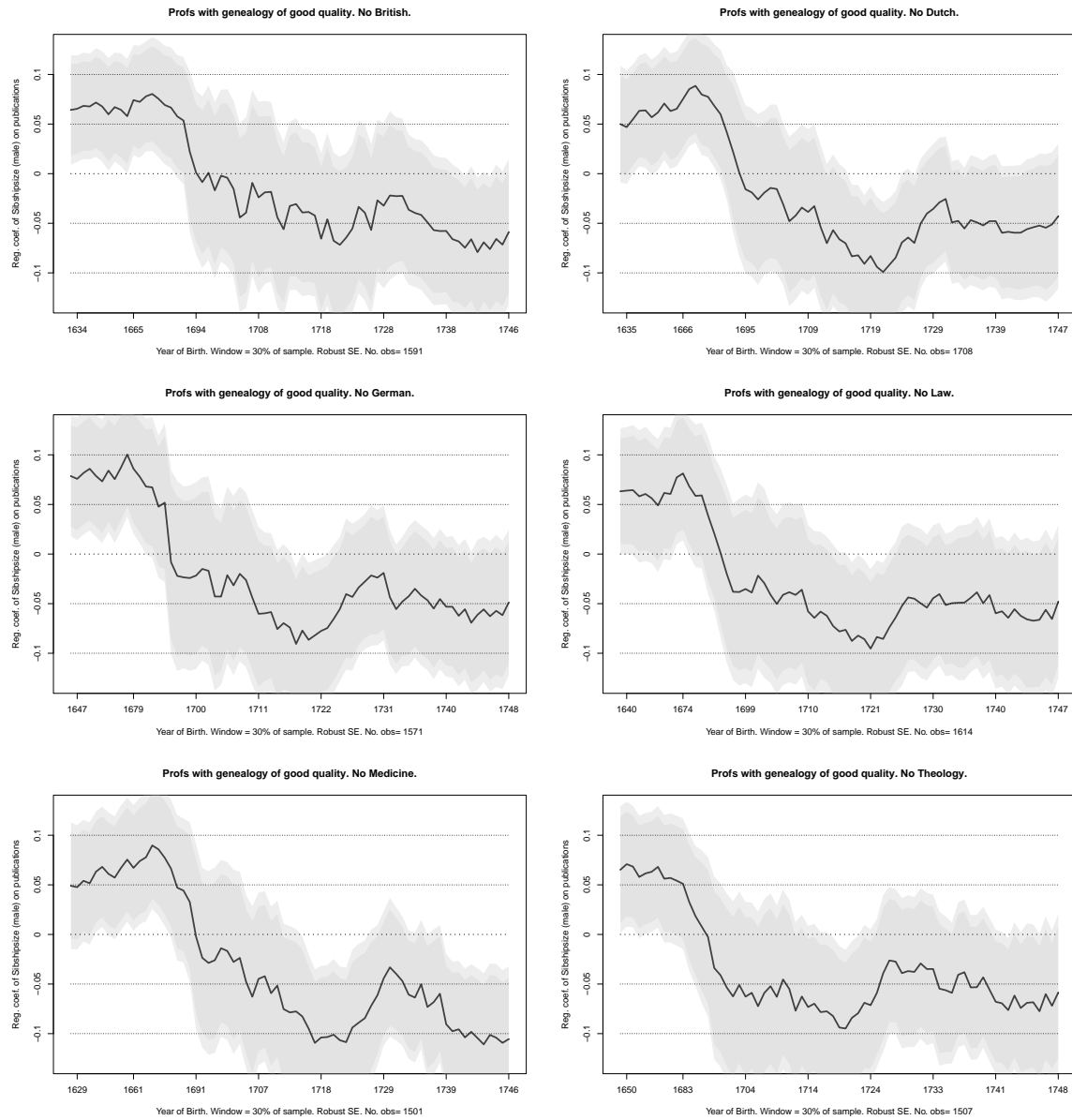


Figure 19: Rolling regression for additional robustness checks