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

Thomas Baudin[*] David de la Croix[†]

December 10, 2025

Abstract

Consider the transition out of a stagnant Malthusian system spearheaded by high-human-capital people. Individuals with a stronger preference for offspring quality may be the first to escape the Malthusian regime and adopt modern fertility behaviors (Galor and Moav 2002). To test this hypothesis, we examine the relationship between family size and human capital among academics in Northern Europe in the two centuries before the Industrial Revolution. We measure the human capital of academics using a novel approach based on their publications. We find that scholars with a higher number of publications than the median shifted from having more siblings to having fewer siblings in the first half of the 18th century. Estimating the parameters of an evolutionary growth model by indirect inference, we show how Malthusian constraints initially led the high human capital families to reproduce more, before being endogenously replaced by Beckerian constraints with a tradeoff between child quality and quantity. Our results support an extension of Galor and Moav's (2002) approach, in which the decline of Malthusian constraints is linked to the accumulation of human capital during the 18th century.

Keywords: Unified Growth, Fertility, Scholars, Human Capital, Indirect Inference.

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

In Europe, the waning Malthusian constraints in the 18th and 19th centuries loosened the limits to growth. This unleashed a period of sustained economic development (Galor 2011). In the same historical period, an unprecedented rise in education was made possible by a decline in fertility; taken together, these changes reflect a tradeoff between the quality and quantity of children. Although this narrative is widely accepted, identifying the correct timing of events and the underlying mechanisms remains a challenge. One reason for this is that data on education and human capital from the period before education was organised by the State, towards the end of the 19th century, are quite scarce.

Our study uses an unconventional lens to study the emergence of the quality-quantity tradeoff (hereafter QQ tradeoff) and its relationship with human capital accumulation. We introduce a novel dataset of families of academics from Northern Europe during the early modern period, and reveal a notable shift in the correlation between family size and scholarly success between the 17th and 18th centuries. We find that the most accomplished scholars of the 17th century tended to originate from large families, whereas the opposite pattern was observed during the 18th century. This result implies that, among the parents of academics, the QQ tradeoff emerged in the 18th century.

The positive correlation between quality and quantity of children among the parents of academics in the past likely implies that the families with the highest human capital reproduced more. The same pattern is confirmed when looking at the fertility of academics themselves: the most accomplished scholars had relatively more children until the end of the 17th century. This result is directly relevant in light of Unified Growth Theory. A core prediction of Galor and Moav (2002) is that during the Malthusian era there was a positive association between the quantity and quality of children; consequently, the share of individuals of the "quality type"—those with a stronger preference for educating their children—rose over time, fostering innovation and ultimately triggering industrialization and large-scale investments in education. The model predicts, however, that once the economy escapes the Malthusian constraint, a QQ tradeoff will emerge: less-educated individuals have more children.

The Beckerian tradeoff between quality and quantity of children results from a budget constraint. Since 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 such as the return to education.[1] In one way or another, these elements determine

[1] The literature has stressed the following elements: the returns to education (Galor and Weil 2000; Shiue 2017; Cinnirella and Streb 2017; Madsen and Strulik 2023), preferences for quality vs. quantity (Galor

the shadow price of quantity versus quality and, hence, the choices made by would-be parents. The literature that explains the transition from stagnation to growth and the Rise of the West overwhelmingly relies on mechanisms that are based on changes in this shadow price.

To assess these mechanisms empirically, it is essential to observe both quality and quantity over a sufficiently long period of time. On historical data, the literature has used different proxies for child quality, but in general either a deep time dimension is missing or the data are aggregated rather than individual.[2] In this paper, we construct a measure of individual quality spanning three centuries by assembling a comprehensive database of premodern scholars, and using their publication output as a proxy.[3] The cost of this approach is to limit our scope to academics and their parents: our quality measure covers only academics, which precludes generalizing to the broader population and prevents us from analyzing occupational selection between academia and other sectors.

Our sample includes individuals who were affiliated with institutions of higher education 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 consists of a scholar matched to the institution to which they belonged. To assess individual quality, we cross-referenced these scholars with publication records from hundreds of libraries worldwide, accessible through the Virtual International Authority File (VIAF hereafter). Given the long time span we cover in pre-modern history, we rely on scholars' publications as a proxy for human capital, as this is one of the few dimensions that has persisted over time, while other aspects—such as teaching quality—have been lost. In addition to data on quality, we gathered information on sibship sizes by matching scholars

and Moav 2002), the efficiency of child development and medical technology (Bleakley and Lange 2009; 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), exogenous shocks to fertility (Bhalotra and Clarke 2019), and whether the cost of children is a direct material cost such as food or an opportunity cost such as rearing time (Doepke 2015).

[2] County-level school enrollment in 19th-century Prussia has been used to establish a negative relationship between fertility and schooling (Becker, Cinnirella, and Woessmann 2010). School enrollment data at the department level in 19th century France have been used quite extensively by several authors (Perrin 2013; Murphy 2015). For England, Klemp and Weisdorf (2019) rely on literacy rates and employment in a more prestigious occupation in later life. They show that both measures decline with the number of siblings. China has a centralized exam system, so it is possible to use a dummy variable for passing the civil service entrance exam. 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, as in the study of Hatton and Martin (2010) on children in Britain. In the context of the Ming and Qing dynasties from 1368 to 1911, using genealogies of six lineages, Hu (2024) proxies child quality both by the probability of marriage and the probability of graduation. She documents a positive correlation between the quality of children and their number of siblings but does not identify any reversal of this correlation.

[3] Iaria, Schwarz, and Waldinger (2018) adopt a similar approach for the 20th century.

with genealogical data from major providers such as Geni and Geneanet. Out of the 6,412 scholar-institution pairs in our database (involving 5,335 unique persons), we were able to find genealogical records for 2,926 of them. Overall, our database is a rich and unique resource for investigating the quality and family origins of academic scholars over an extended period and a broad geographical area.

With these unique data, we address the question of whether high quality scholars (i.e., those who publish more) come from large or small families, and whether this has changed over time. Descriptive statistics show that during the 17th century, scholars publishing in the top half of the distribution have slightly more brothers than those publishing in the bottom half of the distribution.[4] They also have more offspring. This "evolutionary" advantage disappears for those active in the 18th century. By the end of our sample period, the pattern is reversed, with well-published scholars having up to 0.4 fewer brothers than those who publish less. This result is confirmed in a rolling regression setup in which we control for various selection and composition biases. The results suggest that, up to the turn of the eighteenth century, families producing well-published scholars enjoyed a reproductive (evolutionary) advantage. This advantage disappears in the 18th century and is replaced by a tradeoff between number of siblings and publications.

To improve our characterization of the change in the historical relationship between fertility and publications of academics, we develop a structural growth model which we estimate using indirect inference (Gourieroux, Monfort, and Renault 1993; Smith 2008). In doing so, we assess whether our theoretical mechanisms can quantitatively reproduce the observed pattern.[5] Indirect inference is a simulation-based method for estimating the parameters of a structural model. The structural parameters are identified by minimizing the distance between the regression coefficients of an auxiliary model using actual data and those using simulated data. We use the rolling regression on actual data as the auxiliary model to capture aspects of the data on which to base the estimation. The structural model is a unified growth model with

[4] The main results are based on males only, as female births are often misrecorded in premodern family trees. We include sisters in a robustness analysis to verify that our findings are not sensitive to their exclusion.

[5] We cannot use the usual identification strategies to disentangle a causal relationship between fertility and education. We do not have enough observations to use the twin-instrument (Bhalotra and Clarke 2019). We do not have precise enough marriage and birth dates to infer exogenous fecundity from the protogenesis interval (i.e., the period between the couple's marriage and their first birth, see Cinnirella, Klemp, and Weisdorf (2017)). Moreover, the dispersion of scholars over a broad geographical area precludes the use of local natural experiments, such as the eradication of the hookworm disease (Bleakley and Lange 2009). As documented in Table 7 of Appendix A, there are only a limited number of recorded professor births within a twenty-year window surrounding major events susceptible of affecting the budget constraint of the parents (battles, fires, or pandemics, etc.). The most significant event is the plague affecting Stockholm in 1710-1711, but only 49 scholars were born in a twenty-year window surrounding this event. As a result, no single micro event can be used as an exogenous force affecting the QQ tradeoff, as none has significantly altered this tradeoff for a large enough group of families.

heterogeneous agents. As in Galor and Moav (2002), heterogeneity affects the preference for quality children. The model implies that before a certain date, households are trapped in a Malthusian regime.[6] 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. Alongside the rise in individual human capital, the Malthusian regime favors individuals with a greater preference for quality, thereby reshaping the distribution of human capital in a manner conducive to economic growth.

The structural estimation shows that the model quantitatively explains the observed pattern in the data without the need for an external shock. The key mechanism is an endogenous switch from a Malthusian constraint to a Beckerian constraint, rooted in human capital accumulation during the Malthusian epoch and its ability to generate some income. Once the initial conditions are fixed, the transition is endogenous. Two additional insights emerge from the estimation. First, the heterogeneity in preferences required to generate the correct differential fertility over time is minimal. Second, errors in measuring fertility, due to the inevitable errors and approximations that genealogists make in constructing their tree, are key to explaining why the regression coefficients of the auxiliary model are usually small. Our coefficients are thus likely to be lower bounds of the true coefficients, given the uneven quality of genealogical data.

Our approach relies on several assumptions. In particular, individuals exercised rational control over their fertility even in premodern 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) versus "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 offspring. We provide an additional argument against the view that the Malthusian epoch is a period of non-rational fertility. If this were the case, one would expect a negative correlation between fertility rates and the development of human capital: if the number of children is random, then education spending should adjust to fluctuations in fertility, implying a negative correlation between the two. In this paper we present evidence contrary to this expectation, showing an exactly opposite relationship. This finding does not reject the notion that children and education were rival goods in Malthusian times; rather, it shows that despite this rivalry, subsistence-related forces were strong enough to prevent the QQ tradeoff from producing a negative correlation between fertility and human capital.

[6] We use the term "Malthusian" in a broad sense to refer specifically to stagnation in consumption levels. In our framework, stagnation does not result from diminishing marginal returns to labor, but instead stems from a minimum consumption constraint.

Second, we model Malthusian forces assuming that consumption is constrained by a subsistence level. Subsistence income is a key concept in the Malthusian model and does not necessarily imply poverty or starvation. Clark (2007) emphasizes 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.*" Thus, a Malthusian regime should not be interpreted as one of mere subsistence; rather, it is characterised by income effects dominating, with fertility increasing in tandem with income.

We contribute to the literature in several ways. First, we inspect one mechanism of Unified Growth Theory. Our evidence that the quantity-quality tradeoff emerged in pre-industrial Europe lends credence to a key tradeoff assumed by the theory. Dating the birth of this tradeoff to the early 18th century invites a broader interpretation of the key trigger, based not only on industrialization but also on human capital accumulation among elite groups. In other words, if 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, this transition may have required an additional impetus for the broader population in the form of an increased return to education. Importantly, we use a new way of measuring human capital with individual-level data.

Our results also help to better characterize 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: focusing on the upper tail reinstates the importance of human capital during the transition from stagnation to growth. We shed new light on the families of members of the upper tail of human capital. We show that there is some heterogeneity within this group. Families of superstars adopted behaviors more compatible with long-run growth earlier and more intensively than families of less productive scholars.

Finally, by analyzing the family pattern of the members of universities and academies, we contribute to the emerging literature on the role of specific institutions of the 17th-18th centuries, such as academies of science, in fostering later development. For example, Koschnick, Hornung, and Cinnirella (2022) examines how economic societies and their members in 18th century Germany facilitated the spatial diffusion of knowledge in the 19th century, while Zanardello (2024) shows that cities with academies having a majority of members in science grew faster 150 years after the birth of such academies. Our results underscore the importance of pre-modern human capital and the academy movement for European development.

2 Data

2.1 Scholars, Institutions, and Publications

We have built a dataset of scholars who were members of 31 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. Focusing the analysis on Protestant countries has an advantage beyond maintaining a shared cultural background: it minimizes potential bias in genealogical data caused by the under-representation of childless clergy, which is characteristic of Catholic countries (Vardè 2024). Unlike Roman Catholic clergy, Protestant pastors and bishops generally marry and have descendants.

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. In Appendix B, Figure 11 shows the distribution of scholars by institution of affiliation.

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)[7] which we match with genealogical data. We define scholars as persons performing 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 specialized in four main fields: theology, law, arts and humanities, and medicine. Their impact on 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*

[7] It is accessible at https://shiny-lidam.sipr.ucl.ac.be/scholars/

of theology provided teachers for the episcopal schools, where the ordinary parish priests were educated." Most academies were created later, in the 17th-18th century, in response to a push to develop new fields of research that were not being taught at universities. The academies ranged 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 for whom we have a genealogy (small red or orange dots). A majority of scholars come from around the Baltic and North Seas. Some come from other European countries, including France, Italy and the Holy Roman Empire,[8] which shows that the academic job market was already integrated at that time (De la Croix et al. 2024).

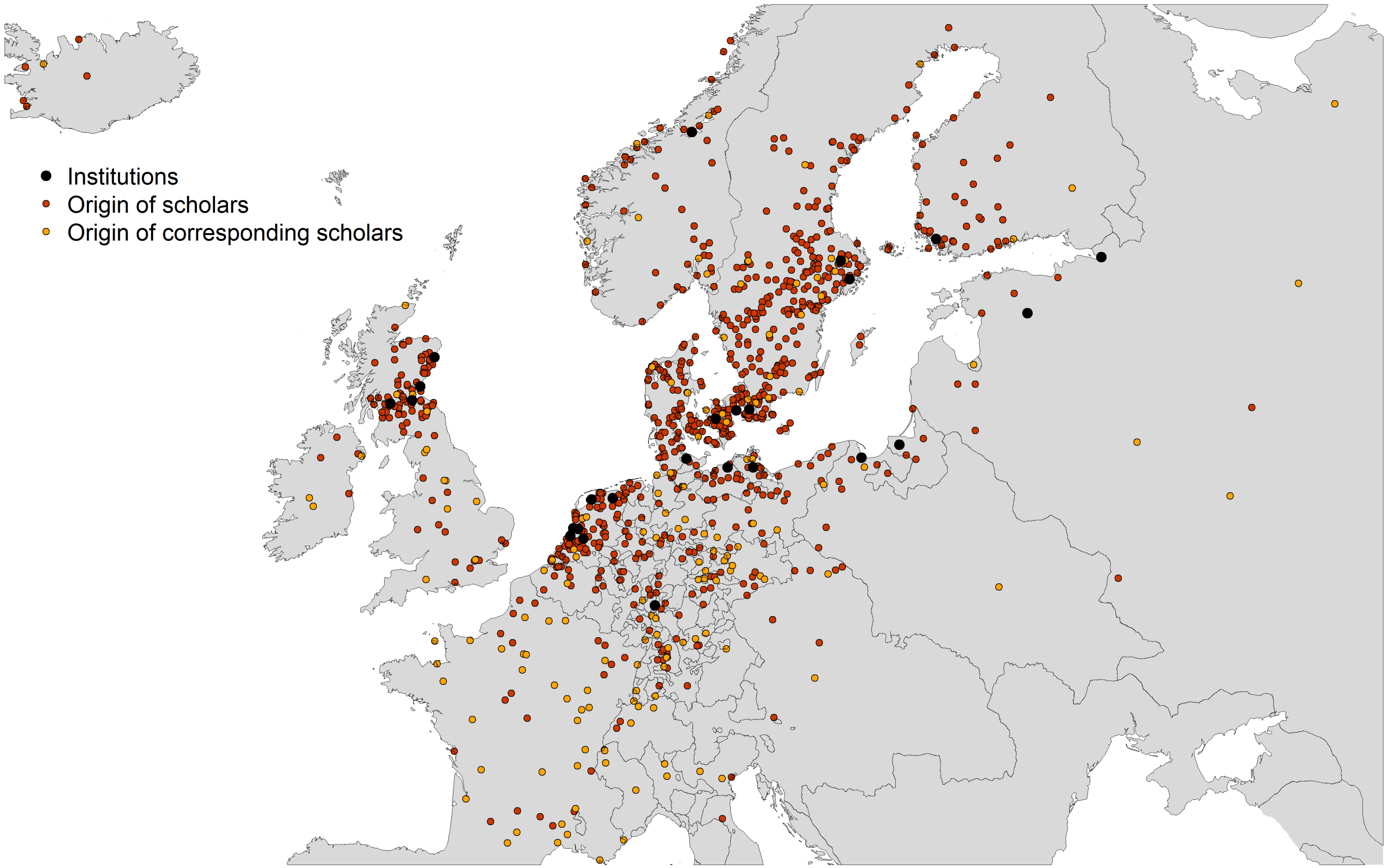
Corresponding members of academies are shown in orange dots. They are located in France, England, Northern Italy, and Russia. Iberia and the countries under Ottoman Rule had no scholars in our sample. Figure 12 in Appendix C shows the same map and includes all members of the institutions we have selected.

Table 1 reports the official creation date of each institution, and a timeline retracing the history of these creations is provided in Figure 13 of Appendix D. 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. All the included institutions are located around the Baltic and North Sea.[9] 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.

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).

[8]In the figure, the countries' borders are those of 1700, as they are drawn in Reed (1999).

[9]Except the "Société patriotique de Hesse-Homburg" which is included because of its special links with Sweden (de Hesse-Hombourg 1777).



Note: origins of scholars is measured by their municipality of birth, institutions are placed according to the historical site of their headquarters.

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)
School of Sorø	Sorø	DNK	1586		Glahn (1975), Olrik (1927)
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)
Royal Norw. Soc. of Sciences and Letters	Trondheim	NOR	1766		Schmidt (1960)
University of Groningen	Groningen	NLD	1612		https://hoogleraren.ub.rug.nl/
Athenaeum Illustre of Amsterdam	Amsterdam	NLD	1632	1877	http://www.albumacademicum.uva.nl/
University of Franeker	Franeker	NLD	1585	1811	Feenstra et al. (2003), Napjus and Lindeboom (1985)
Royal Dutch Society of Sc.	Haarlem	NLD	1752		https://khmw.nl/historische-leden/
University of Leiden	Leiden	NLD	1575		https://hoogleraren.universiteitleiden.nl/
University of Utrecht	Utrecht	NLD	1636		Academia Rheno-Trajectina (1861)
Société patriotique de Hesse-Homburg	Bad-Homburg	DEU	1775	1781	Société patriotique (1777)
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	1936	Schumann (1893)
University of Königsberg	Kaliningrad	RUS	1544		Naragon (2006), Schwinges and Hesse (2019)
Academy of St Petersburg	St-Petersburg	RUS	1724	1917	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)

Note: institutions are placed according to where their headquarters were historically located.

Table 1: Sources used by institution

One key feature of our data is that they include all academic scholars with high human capital 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. 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 hundreds 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 14 in Appendix E shows the histogram of its distribution.[10]

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.

A fundamental mechanism underlying the quantity-quality tradeoff connects parental investment to their children's human capital, which we measure here by academic publications. Essentially, we will assume that the education received in the first period of life is an important determinant of the success of professors in their academic lives as adults. This assumption is fully consistent with the classical theory of human capital (Mincer 1958), where the productivity of individuals in a given sector or occupation depends on their initial education.

Beyond this general argument, examples taken from the biographies in MacTutor[11] illustrate the significant influence of parents of scholars through various channels: direct time investment, school choice, and the role of the family library. For example, Willebrord Snell (1580-1626) was a Dutch mathematician best known for the law of refraction. He was educated by his father, who taught him Latin, Greek and philosophy. He had no further education before entering university. Colin Maclaurin (1698-1746) was a Scottish mathematician who

[10] In Table 8 of Appendix F, we highlight several correlates of scholars' human capital. We document a slightly positive time trend, as well as a positive association between publications and longevity (mean age at death), working in theology, and being a corresponding member of an academy. By contrast, the association is negative for those working in law.

[11] See https://mathshistory.st-andrews.ac.uk/Biographies/

published the first systematic exposition of Newton's methods. He was born in Kilmodan where his father, John Maclaurin, was the minister of the parish. John Maclaurin was more of a scholar than one would expect of a parish minister, having translated the Psalms into Gaelic. His father died when Colin was six weeks old. His mother wanted a good education for Colin and his brother John, so the family moved to Dumbarton where the boys attended school. Finally, Anders Celsius (1701-1744) was a Swedish astronomer best known for the temperature scale he proposed, which is named after him. Anders Celsius grew up with access to a large family library, which survived the fire of 1702. This gave him early access to a copy of the 1687 edition of Isaac Newton's *Principia*.

2.2 Genealogies

To retrieve information about scholars' families, 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 supplement it with information about sibshipsize, number of children, and parents. In recent years, scholars have used the wealth 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 (2024). 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 there is no suitable information in these international sources, we explore smaller -national- databases such as 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 manually reconstruct the completed fertility (total number of children) of their father as well as their own. We add full- and half-siblings indifferently, but construct 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 the fathers of the scholars. Genealogical databases have been very useful for finding places and dates of birth and death when these data are missing from biographical notices.

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

Table 9 of Appendix G presents the number of genealogies found, by institution. For Scandinavian institutions, we are able to match scholars to a genealogy in 57%-81% of cases. Considering 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 32%-50%; for the former German territories, 10%-47%; for Scotland, 34%-48%. Overall, we have a genealogy for 46% of the scholar–institution pairs, i.e., 2,926 linked profiles.

The genealogies may suffer from some biases (Minardi, Corti, and Barban 2023, Stelter and Alburez-Gutierrez 2022). A first source of bias in genealogical data is gender, as women tend to be underrepresented: 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 already documented by Dupâquier (1993): most amateur genealogists have some characteristics that lead them to collect biased information. White amateur genealogists in rich, patriarchal societies have tended to focus on the male branches of their own family trees.

In addition, in pre-modern times, there was a tendency to underreport girls' births, particularly female stillbirths. Finally, but importantly, the data we consider are not necessarily representative of the entire population around the Baltic and North Seas before 1800; rather, they focus on families with 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 leads to a mechanical overrepresentation of men. For example, among the families of professors with 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 arrive at this number by using a natural sex ratio of 1.05 (Chao et al. 2019). This means that each new birth has a 48.8% chance of being a female birth and a 51.2% chance of being a male birth. Extending this logic to other parities, we get the theoretical sex ratios of Table 2. Table 2 shows that our data suffer from misreporting of girl births. This bias is more severe for families with two children and becomes less important as family size increases.[12] A simple way to correct for this gender bias in computing family size is to count only male siblings and use the total number of siblings (male and female) only for robustness.

In Figure 2, we present the distribution of parities (keeping only male children) among the professors for whom we collected information (full sample). We compare this distribution

[12]In a family composed of n siblings with one male academic, the theoretical sex-ratio is equal to $\frac{1+(n-1)*(\frac{1.05}{2.05})}{(n-1)*\frac{1}{2.05}}$, corresponding to the expected number of males divided by the expected number of females.

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	4.09	2.39	2.49	1.69	1.75	-

Table 2: Theoretical versus empirical sex ratio (M/F) as function of parities within families of professors

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 comparative 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 can arise from many issues, the verticality issue is the most important. Many amateur genealogists are interested in discovering their direct ancestors and do not research the siblings of those in their direct line. This implies, especially for ancient data, an overrepresentation of observations with only one parent and no siblings. This bias is nicely discussed in Blanc (2024), who treats it by suppressing observations for which he cannot find an ancestor with at least two children in the pool of 30 ancestors going from parents to grand-grandparents. While Blanc's approach is defensible and works well at the aggregate level, it does not fully overcome the verticality bias in our case. For example, in data like geni.com, lineages are built by merging the inputs of thousands of genealogists. As a result, a person may be connected to a great-great-grandfather who had five children, but still have unrecorded brothers and sisters because the genealogist who coded his profile filled in his own genealogy in a vertical way.

To overcome this, we use a stricter approach than Blanc. First, we excluded all genealogies with a sibshipsize of less than three and where the father's date of birth is not known, in order to exclude data of lower quality. We drop 481 genealogies under this restriction. We also exclude all scholars who have no siblings and who themselves have only one child, a strong indication of verticality bias, dropping an additional 125 genealogies. The right panel of Figure 2 shows the new distribution of parities after implementing our two selection criteria (final sample). We can see a noticeable improvement in the distribution, with a new mode of two instead of one.[13]

Our corrections for fertility measurement ensure that our distribution of parities is closer to that proposed by Wrigley and Schofield (1983). Another point of comparison can be used

[13]Importantly, Figure 15 in Appendix G shows that the share of genealogies excluded by this procedure remains broadly stable over time.

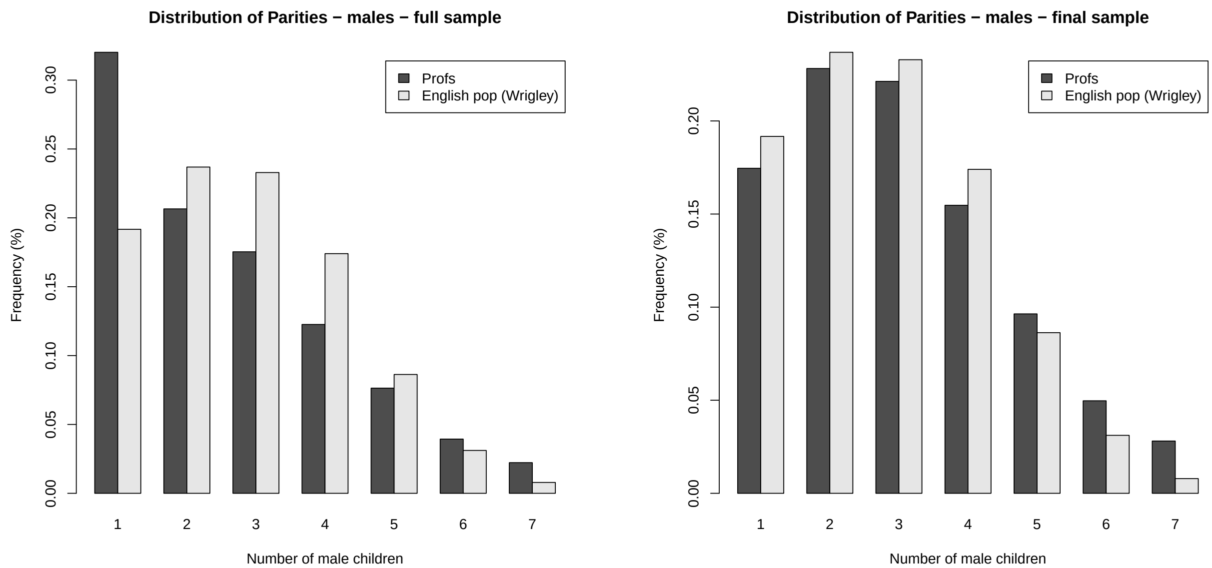


Figure 2: Distribution of parities: all genealogies (left), selected genealogies (right)

to assess whether our 19% share of only son families is reasonable. Galton (1875) looked at a sample of about 200 living scientific members of the Royal Society, and his method was a self-report questionnaire. These scholars were born a little later than our sample, around 1800-1820, but still before any demographic transition in England. Galton reports a proportion of only sons of 20%, very close to our estimate. This reassures us about the quality of our correction.

Another concern is that our sample of scholars with a valid genealogy differs substantially from the sample of those without genealogies in dimensions other than their parents' fertility. An example of where this might occur is if genealogical platforms are more likely to include famous professors than obscure ones. In Tables 10 to 12 of Appendix H we show how the selection changes over time. Academics with a valid genealogy tend to publish more and live longer than their obscure counterparts throughout our observation period. This selection bias is present in all periods and does not change significantly over time, as can be seen in Figure 15 in Appendix G.

As obscure scholars are less likely to have a recorded genealogy, the true variance of publications is higher than what we observe in our sample. Since obscure individuals will be concentrated at lower parities in a Malthusian regime and at higher parities in a Beckerian regime, this induces an attenuation bias: the coefficients estimated in the regressions in the next sections are presumably lower bounds of the true effects. Attenuation holds as long as there are no offsetting biases. For example, if having more brothers increases the probability of having a genealogy, then the correlation between publications and sibship size will be attenuated when the true correlation is positive, but amplified when it is negative. This latter bias is, however, likely to be small. If it were quantitatively important, we would observe implausibly high average fertility rates in our sample.

2.3 Selection into academia

Genealogical websites (through their detailed biographical notice sections) and Wikipedia pages often report the occupations of individuals and their parents. We collect these occupations to better understand the class backgrounds of the scholars. We classify them into three categories according to Van de Putte and Miles (2005): elite, middle class, workers. We observe only very few unskilled workers: they are mostly either skilled or semi-skilled. The middle class includes farmers, local businessmen and non-manual professionals. Table 3 shows the main occupations, with the number of observations in parentheses. As noted by De Candolle (1885), many academics were born into families of pastors and priests. In the following sections, we develop 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 show that our professors do not exclusively come from highly privileged backgrounds such as the nobility, as such groups may not have experienced the full extent of Malthusian constraints. Table 3 provides compelling evidence that the majority of our academics originate from non-elite backgrounds.[14] A notable example is Linnaeus, who came from a modest family and whose father worked as a preacher and built the family home 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 before its Industrial Revolution, as documented by Sandberg (1979). This population had a high level of education despite lacking substantial wealth or privilege. In Figure 3 we illustrate the constancy of the proportion of each social class over time. In particular, a bit less than 50% of scholars come from upper-class backgrounds on average, while a bit more than 40% come from middle-class backgrounds.

Note that, in Figure 3, as in many other figures, we use a moving window that always spans 30 percentile points of the birth-year distribution, instead of using a fixed number of years (e.g., a 10-year window). This means every point is based on roughly the same number of observations (good when data are uneven over time), but the calendar width of the window can vary (narrow where data are dense, wider where sparse).

Among the upper class, some fathers are university professors. De la Croix and Goñi (2024)

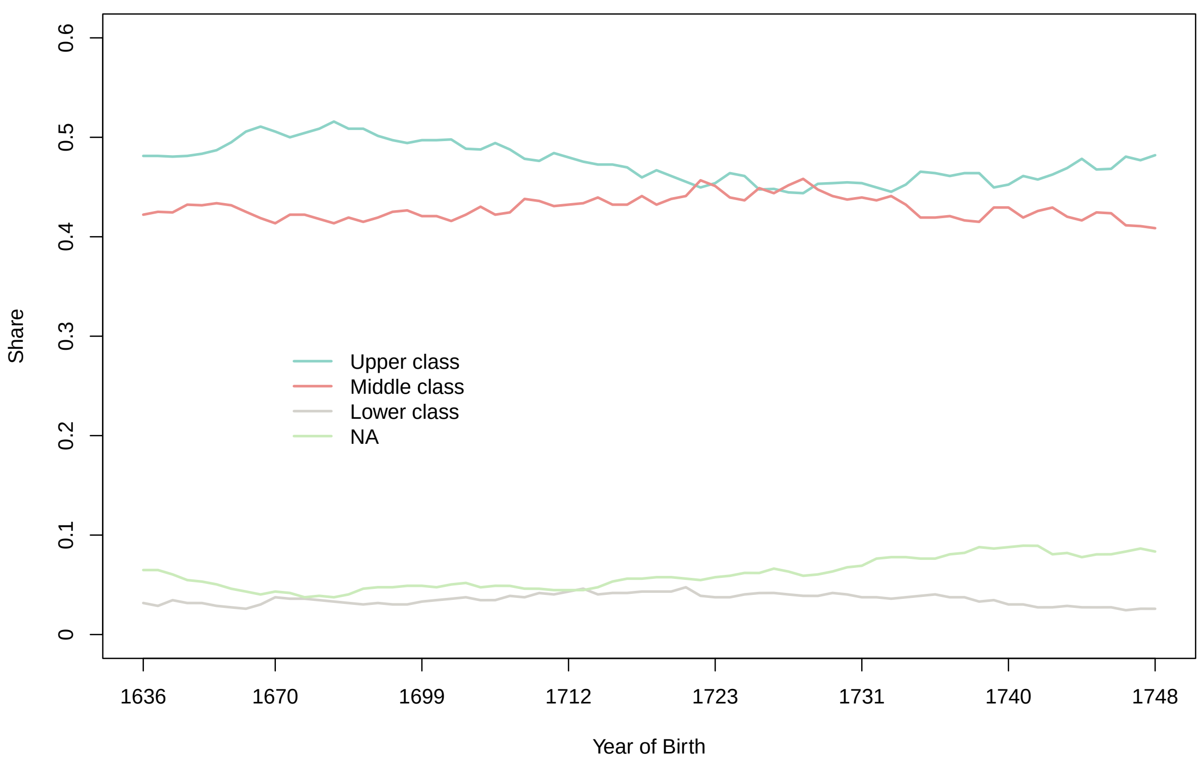
[14] We provide the complete list of occupations and titles and their corresponding class in Appendix I.

analyze the set of father-son pairs in academia (all of Europe, 1088-1800 CE). They find that there was some nepotism (some sons were hired without meeting the human capital requirements to be a professor). However, nepotism declined during the Scientific Revolution and the Enlightenment, reflecting the rise of meritocracy, and was less prevalent in fields experiencing rapid changes in the knowledge frontier.

The environment in which scholars grew up was influenced by their place of birth and shaped their early life experiences. We construct an urban/rural dummy variable that takes a value of 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 came from urban areas, while 47% came from rural areas (see Figure 16 in Appendix J) – the place of birth is unknown for 5% of the sample. These proportions remained roughly constant over time, as shown in Figure 17 in Appendix J.

2.4 Demographic Transition

We turn our attention to the evolution of longevity and fertility over the observation period. In Appendix K Figure 18, we show that longevity increased sharply for professors born in



Note: moving 30–percentile window of birth cohorts. Final sample. Nobs=2314

Figure 3: Parental social class distribution.

Upper class (Elite)	professor (193), councillor (84), minister (74), physician (62), mayor (56), theologian (50), bishop (47), rector (42), doctor (42), general (42), lawyer (36), noble (34), lord (29), governor (27), colonel (25)
Middle class	merchant (203), preacher (164), priest (140), , pastor (134), farmer (74), clergyman (59), officer (42), vicar (34), lieutenant (30), trader (28), master (27), superintendent (25), secretary (23)
Lower class (Workers)	craftsman (19), blacksmith (8), brewer (7), goldsmith (5), fisherman (4), miner (4), builder (4), tailor (3), innkeeper (3), gardener (3), baker (3), grocer (2), tanner (2), saddler (2), carpenter (2), engraver (2)

Table 3: Main occupations of scholars' fathers – occurrence in parenthesis

the 17th century and reached a plateau around 70-72 years along the 18th century. The early increase in longevity we observe is fully consistent with what we know from the literature on academic longevity (Leridon and Mandelbaum 2004, Andreev et al. 2011, Stelter, De la Croix, and Myrskylä 2021), but the magnitude of the increase is stronger than in the literature. Regarding fertility, in Figure 19, the average number of children (sibshipsize of scholars) fluctuates within a narrow interval. That is, we do not observe any fertility transition on average, but this does not mean that fertility does not undergo important transformations through compositional effects, as we show in this paper.

We also recorded the birth order of the scholars. Among the 1853 observations for which a rank can be computed, 970 (52%) are first-born sons. This is in perfect agreement with Galton (1875), who found that 48% of famous English scientists were the first-born sons in their families. If birth order did not matter, we would observe that less than 1/3 of scholars were 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 selection into academia is higher for first-born sons.

Finally, in Figure 4, we divide the scholars into two groups: those who publish more than the median (high quality) and those who publish 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 thus potentially 0.2 more siblings, than their low quality counterparts.[15] In the 18th century,

[15]The fertility differential in favor of the more productive scholars during the first part of our observation period is not significant. This may be due to several factors. First, the further back in time we go, the noisier

high-quality scholars begin to have fewer siblings overall than their low-quality counterparts. The fertility difference between the parents of scholars reaches more than 0.4 boys and thus 0.8 children for births around 1749. To our knowledge, this reversal of the QQ tradeoff is rarely observed in a consistent micro-level dataset. It provides important empirical support for any theory that places the switch in the quality/quantity tradeoff at the center of the European transition to growth.

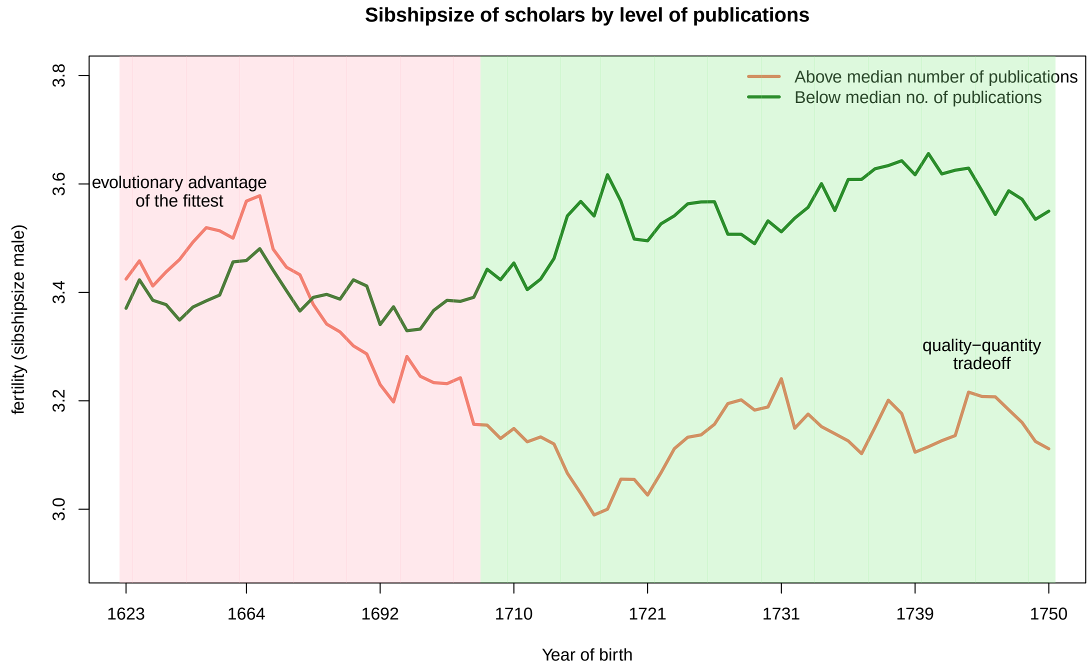
We now extend our analysis to the offspring of the scholars. Maintaining the established distinction between high and low quality scholars, our investigation reveals a noteworthy pattern: in the 17th century, high quality scholars exhibit higher fertility compared to their low quality counterparts, as shown in Figure 5. Strikingly, this trend reverses in the following century, with high-quality scholars born in the 18th century consistently having fewer offspring. The replication of this fertility reversal over two successive generations adds robustness and considerable credibility to our findings.

The positive correlation between quality and quantity within the Malthusian regime, for both parents and academics, occurs despite the mechanisms underlying the quality–quantity trade-off, which push in the opposite direction. The pattern is consistent with an endogenous escape from the Malthusian trap driven by the evolutionary advantage of quality-oriented individuals à la Galor and Moav (2002).

Note that a few papers have documented such a correlation, but without much time depth. Clark and Hamilton (2006) use probate wills from England—specifically male testators' wills that report surviving children and assets— to link wealth/status to net fertility. The core sample covers 1585–1638. An alternative to conducting its own study is to rely on meta-analyses, such as that of Skirbekk (2008). Skirbekk is interested in fertility by social status. He finds that there is a general shift from a positive to a negative or neutral status-fertility relationship. This happens in the 19th century in what are now developed countries. The pattern highlighted by Skirbekk is largely confirmed for the twentieth century by Vogl (2016), who examines the relationship between parents' education and family size using Demographic and Health Surveys covering 48 developing countries.

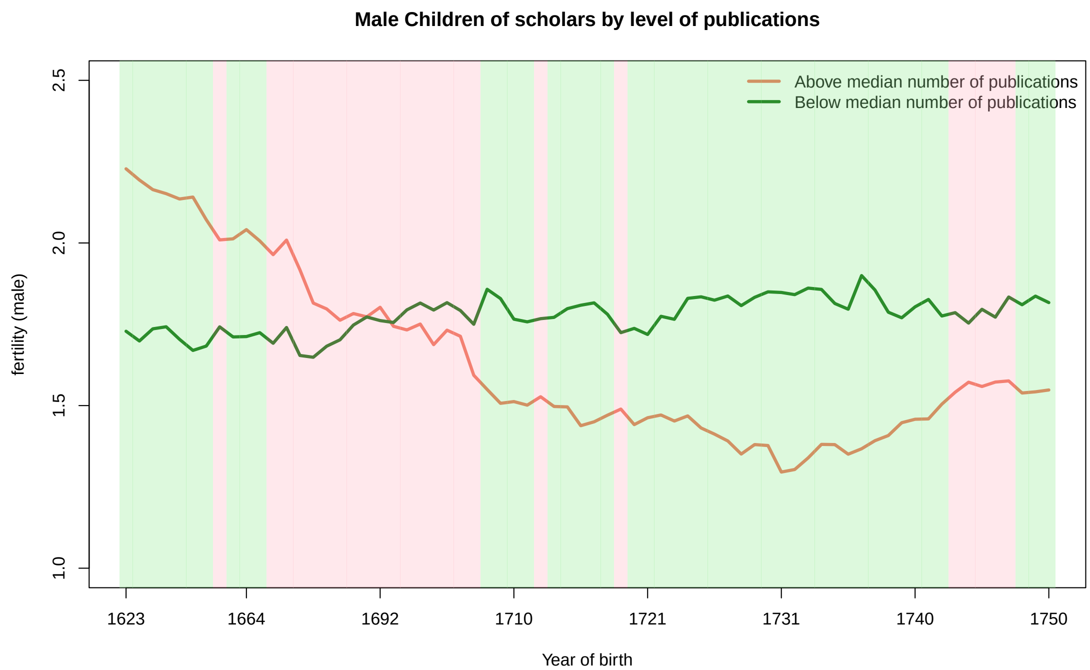
Juxtaposing Figure 5 with Figure 4 reveals a striking difference: scholars, on average, have fewer children than their parents. This discrepancy is substantial, and we attribute it to a key contextual factor: all of our scholars experience adulthood in an urban setting, whereas half of their parents lived in rural areas. The significant disparity in fertility between urban

the fertility measures become. The lack of significance may be due to the fact that we do not control for several important factors, such as field of study and country fixed effects. This second argument is strongly supported by our empirical results in Section 3.



Note: moving 30–percentile window of birth cohorts. Difference is statistically significant at 10% in green areas. Final sample. Nobs=2314

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



Note: moving 30–percentile window of birth cohorts. Difference is statistically significant at 10% in green areas. Final sample. Nobs=2314

Figure 5: Reversal of the QQ tradeoff among scholars.

and rural environments in premodern Scandinavia is well documented (see sources in Baudin and Stelter (2022)).[16]

The reversal of the QQ tradeoff among the parents of scholars has many potential confounding factors and compositional effects. Changes in the relative weight of each institution over time may be important, as may shifts in the weight of alternative disciplines. For example, mathematicians may publish more than theologians, while also coming from smaller and more secular families. If so, a massive entry of mathematicians and other scientists born around 1750 could explain the reversal of the QQ tradeoff, which would have nothing to do with a change in the way parents of professors allocated their resources between quality and quantity of children. The share of the main academic fields over time is shown in Figure 21 of Appendix M. Section 3 is devoted to a detailed analysis of the reversal, controlling for as many factors as we can.

3 The Auxiliary Model

Figure 4 documents a reversal of the fertility differential between high and low-productivity scientists. As it stands, this reversal of the QQ tradeoff could be due to confounding and selection issues that we want to rule out. We divide the range of professors' birthdates into percentiles and run 70 successive regressions with controls, each including the professors born within a specific time interval corresponding to 30% of our total time window. The first regression includes all professors born between 1435 and 1688, the second between 1514 and 1692, and the last between 1736 and 1779. In each iteration, we regress the inverse hyperbolic sine ($asinh$) of the publications of professor i from institution k on the size of his male sibship pool (Sibshipsize_i), controlling for a number of important factors. Let K be the set of institutions and F the set of fields. The field of i in k is d_{ik}. The regression equation is:

$$\begin{aligned}\text{asinh}(\text{Publis}_i) \quad = \quad & \alpha_1 \text{Sibshipsize}_i + \alpha_2 \text{Longevity}_i + \alpha_3 \text{Age Nomination}_{ik} \\ & + \alpha_4 \text{Corresponding}_{ik} + \alpha_5 \text{Urban}_i + \alpha_6 \text{Geni}_i + \alpha_7 \text{Academic Father}_i \\ & + \alpha_8 \text{Father in Upper Class}_i + \alpha_9 \text{Father in Middle Class}_i \\ & + \sum_{j \in K} \delta_{1j}\, \text{I}(k = j) + \sum_{f \in F} \delta_{2f}\, \text{I}(d_{ik} = f) + \delta_3 X_i + \varepsilon_{ik}. \qquad (1)\end{aligned}$$

[16] For a final validation of our fertility data, we conducted an analysis to estimate the intergenerational correlation in fertility. The findings, illustrated in Figure 20 in Appendix L, 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 main objective of this estimation strategy is to obtain the most accurate point estimates of α_1, in order to trace its evolution over time. We then use these estimated values as moments that the theoretical model developed in the next sections must replicate.

We control for longevity (year of death − year of birth) and the age at which the scholar was nominated to the institution k, to capture the fact that the younger a scholar is nominated, the more time he has to develop his publication catalog. We control for corresponding membership in an institution: being a corresponding member is an honorary distinction, so it is likely to select the scholars with the highest quality in terms of publications, coming from distant places where the nature of the QQ tradeoff is different from that prevailing in our region of analysis. We also control for the rural-urban character of the place of birth and whether the source of the genealogy is geni.com or another website. Finally, to prevent nepotism from affecting our results, we control for whether the father was in academia, using the data from De la Croix and Goñi (2024).

By controlling for institution through the fixed effects $\mathrm{I}(k = j)$, we rule out the possibility that the potential reversal of the QQ tradeoff is due to the selection of universities and academies of origin in the sample. This controls, for example, for an increase in the proportion of professors from academies and universities where, for whatever economic or cultural reason, fertility is low and publications are more abundant.[17] For the same reason, we also include in our vector of control variables dummies $\mathrm{I}(d_{ik} = f)$ for the scholar's field. This ensures that we are measuring an association between sibshipsize and publications that is not confounded by the risk that some fields are populated by individuals who are more or less likely to publish and more or less likely to have large families.

The presence of half-siblings indicates unusual parental trajectories, including divorce, widowhood, and remarriage, so we also include in X_i a dummy indicating whether the scholar had half-siblings, to reduce noise in the fertility measure

Some scholars may appear more than once in our database for example because they have belonged to more than one institution and have a high degree of mobility. Duplication can also occur when scholars are corresponding members of academies. For example, Joseph Banks was a British naturalist, born in 1743, who served as president of the Royal Society for more than 40 years and was associated with the academies of Gdansk, Copenhagen, Haarlem, Saint-Petersburg, and Stockholm. To avoid standard errors being artificially deflated by the

[17] Decisions about university affiliation could be endogenous, leading to the suggestion not to include this fixed effect in our main regressions. This is particularly pertinent because the human capital of scholars could influence their geographic distribution in a circular fashion. However, when we do not control for the institution fixed effect, our results remain consistent in both magnitude and significance.

presence of similar observations, we compute robust errors ε_{ik} clustered at the individual level.

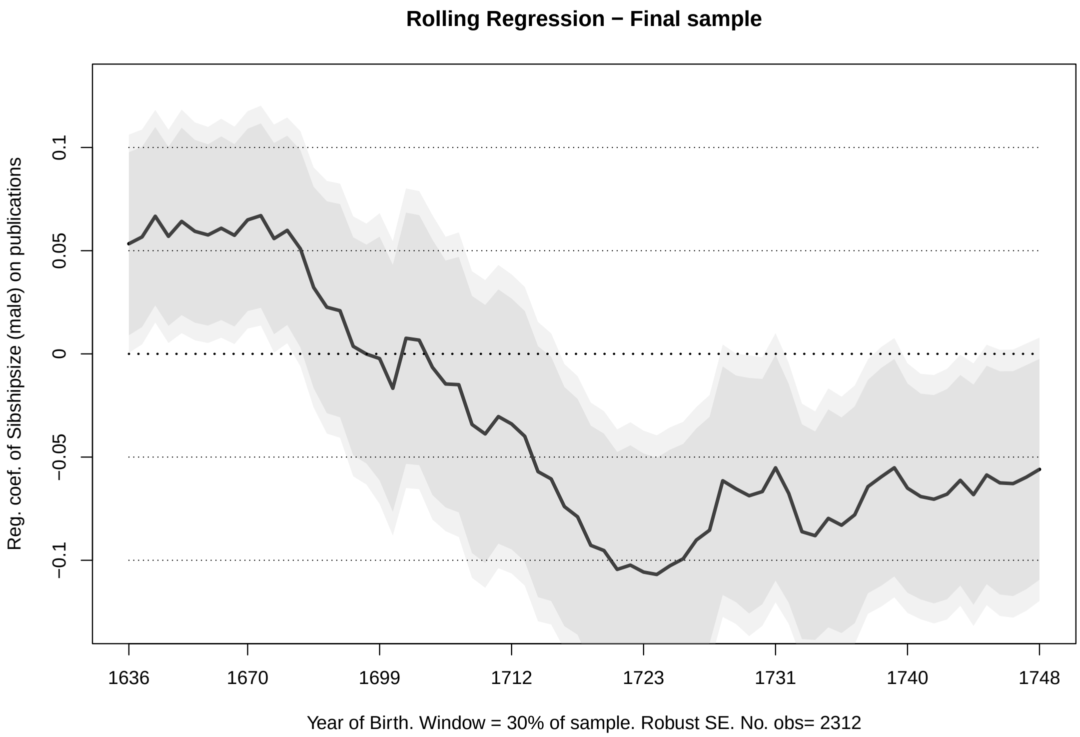
In our main specification, we select observations in the same demanding way as in the previous section, retaining the same final sample. Descriptive statistics are provided in Table 14 of Appendix N. We believe that our rolling time window regression setting constitutes a flexible approach that captures the dynamics of the association between professors' sibship size and their human capital without imposing too demanding a set of constraints.

Figure 6 shows our main results. The black line joins the 70 estimates 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 striking, with an initial period in which scholars from large families tend to publish more than scholars from smaller families, a period of reversal in which the association between fertility and human capital is not significant, followed by the final period in which scholars from smaller families publish more. Notably, the reversal becomes significant (at both the 10 and 5% confidence levels) 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 4 displays the evolution of our main coefficient of interest, pooling the 70 estimates by groups of ten. While the reversal of the QQ tradeoff appears again, it does so in a context where the importance of longevity in explaining publications declines over time. Indeed, the coefficient of association between scholar longevity and their publication metric declines steadily. The share of publication variance that we are able to explain with our main specification evolves between 26 and 33% over our entire time window. Finally, even though they have to be taken with extreme caution since they are computed on only 10 observations/regressions, one can see that the standard deviations of our estimated coefficient are quite small within each time window. We take this as a reassurance that the reversal of the QQ tradeoff is not partly driven by variations in the statistical noise surrounding our estimates.

Figure 6 and Table 4 illustrate a striking symmetry between the intensities of positive association in the early periods and negative association in the late periods. Both show values hovering around 0.05 to 0.06, with a negative peak just below -0.1. Specifically, for professors born between 1597 and 1664, having one more brother is associated with an increase in the number of publications equal to 6.5% (using Bellemare and Wichman (2020) formulas). Conversely, for professors born between 1733 and 1745, this association amounts to -6.2%.

Table 15 in Appendix O shows all the estimated coefficients for 8 instances of the rolling regression. We note that corresponding members of academies and universities tend to publish



Note: moving 30–percentile window of birth cohorts. Point estimates in black, 95% confidence interval in light gray and 90% confidence interval in dark gray.

Figure 6: Rolling regression for main specification, values of $\hat{\alpha_1}$.

more than other scholars, and that controlling for the presence of half-siblings in the sibling pool is important for some periods. Whereas law scholars publish less than others throughout our observation period, scholars in the sciences from the last cohorts—concentrated in the second half of the eighteenth century—tend to publish more than their peers in other fields. This is consistent with the atmosphere of this period. It is characterized by Enlightenment values, where science attracted prestige if not money; where every city wanted its own academy of sciences and arts, and where these academies appointed top scientists as corresponding members. Finally, having a father in academia (c. 12% of the sample) does not show a significant effect in the regression.

In Appendix P, we present a number of important robustness checks. In our main analysis, we restricted our investigations to three types of scholars: (1) those from families with two or more siblings, (2) those with only one sibling and a referenced father's death date, and (3) those with no siblings, a referenced father's death date, and more than one child (to avoid verticality bias). We relax these restrictions to include any scholar with a genealogical link. The overall magnitude of the QQ tradeoff reversal does not change significantly.

	SibShipSize (α_1)		Longevity (α_2)		R2	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
0-30 to 9-39	0.060	0.004	0.032	0.001	0.332	0.007
10-40 to 20-50	0.031	0.026	0.030	0.002	0.329	0.022
20-50 to 29-59	-0.018	0.017	0.025	0.003	0.269	0.024
30-60 to 39-69	-0.081	0.023	0.015	0.016	0.266	0.016
40-70 to 49-79	-0.083	0.018	0.011	0.001	0.273	0.007
50-80 to 59-89	-0.072	0.013	0.011	0.001	0.271	0.005
60-90 to 69-100	-0.064	0.005	0.014	0.001	0.262	0.012

Table 4: Values of α_1, α_2 and R^2 along the rolling regression process

We then examine the extensive and intensive publication margins separately. We find that in Malthusian times, scholars from larger families had a notable advantage, with a higher probability of publishing at least once. In modern times, however, scholars from smaller families have a significant advantage in terms of repeated publication once they have published at least once. This result underscores the importance of including both publication margins in our analysis, as they both contribute to the observed reversal.

We also introduce additional controls, including whether the scholar is the first-born son, the scholar's age at the father's death, and the number of offspring. These controls are potentially correlated with (unobserved) income. Despite some variation in the significance of α_1 for extreme dates, our main results hold after accounting for these controls. This rules out the possibility that the reversal of the QQ tradeoff is solely due to differences in the social environment in which the scholars were raised. Finally, we change the measurement of our outcome variable from VIAF to Worldcat without losing our main results.[18]

To ensure that our results are not driven by a single academic field or a specific geographical area of birth, we re-estimated our main specification, sequentially excluding each academic discipline and each major country of origin. In every case, while minor variations in statistical significance emerged, our core findings remained intact. This robustness confirms that the trend we document is not a historical anomaly of any one country but rather a broader phenomenon that transcends national borders, at least across Northern Europe.

As discussed in previous sections, female children are less well recorded in genealogical data,

[18]WorldCat is the world's largest bibliographic database, maintained by the Online Computer Library Center, aggregating the catalogs of libraries worldwide to help locate resources across institutions (https://www.oclc.org/en/worldcat.html). While the experimental WorldCat Identities project has been suspended, WorldCat Entities remain accessible at https://entities.oclc.org.

which is why our benchmark results focus on males only. Nevertheless, examining sisters may still yield valuable insights. In the left panel of Figure 22, we use the number of sisters as the main regressor. The correlation with their brothers' publications appears negligible during the Malthusian epoch, but becomes negative and significant after the reversal of the QQ tradeoff. This suggests that daughters were effectively costless in the Malthusian regime, but came to play a role in the QQ tradeoff once the economy transitioned to the interior regime. In the right panel of Figure 22, we incorporate sisters of academics into the measurement of sibship size. While the overall reversal pattern persists, it is notably weaker.

Overall, our analysis consistently reveals a significant and systematic change in fertility behavior during the eighteenth century.

4 The Structural model

Spirit of the Model - Our empirical analysis of Northern European academics shows a reversal of the correlation between quality and quantity of children. There is a theoretical growth model 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 who differ only slightly in the importance they attach to their children's education. During the stagnation period, the highly-educated people have a higher income and thus a larger number of children. This gives them an evolutionary advantage and generates differential fertility of the type we observe in our data. At some point, thanks to technological progress, the returns to education increase, pushing the entire population to invest massively in quality. As individuals strive to balance their budgets, investing more in education often leads to a reduction in the number of children. Here the 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 fully supported by the data. First, in their approach, the reversal of differential fertility is closely linked to the Industrial Revolution, when the return to education increases. This is not consistent with our data, as we see a reversal at the beginning of the 18th century, while the Industrial Revolution in Northern Europe took place a century later. Second, Galor and Moav have a strict interpretation of the Malthusian period: income per person oscillated around a constant level, close to subsistence. This view is challenged by recent research that shows some slow growth in the centuries before the Industrial Revolution (Fouquet and Broadberry 2015).

It is precisely because Galor and Moav assumed constant per capita income during the stagnation period that they have to link technical progress to the return to education in order to generate the transition to modern growth. If instead there had been some slow growth in per capita income during the Malthusian epoch, that growth alone would have been sufficient to escape the Malthusian constraints. The escape from Malthusian logic transforms the constraints on households, and the standard quality-quantity tradeoff ultimately prevails.

Our approach is as follows. We start exactly as Galor and Moav do, with two types of people, one with 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 consider both groups to belong to the intellectual elite and neglect the rest of the population. Each household faces two types of constraints: a Malthusian constraint, which requires consumption to be higher than a critical level, and a standard budget constraint, which requires consumption spending and spending on children's education 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, such as 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 minimum level of consumption, the richer people are the quality lovers, and they have more children than the quantity lovers, as in Galor and Moav. Over the Malthusian period, per capita consumption is constant and equal to the critical level for both groups, while education spending increases over time, leading to an increase in human capital and per capita income. Over time, behaviors more compatible to growth become more widespread and a larger share of resources is devoted to education. This view fits very well with both the increase in the number of universities and academies in the 17th and 18th centuries and the rise of the impoverished sophisticated documented by Sandberg (1979).

The main difference between our approach and that of Galor and Moav is that we consider a scenario in which slow economic growth occurs during the Malthusian epoch, leading to a point in time when 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 begin to substitute 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 approach requires that, during the pre-modern period, investing in advanced skills and education could already be converted into substantial economic advantages for the groups at the top of the skill distribution. There is evidence consistent with this view. First, the Dutch Republic in the late seventeenth and early eighteenth centuries combined exceptionally high levels of literacy, numeracy, and commercial/technical skills with the highest real wages and per capita incomes in Europe, suggesting that dense stocks of human capital could underpin high income levels well before mass schooling (de Vries and van der Woude 1997). Second, both states and high-rent sectors had begun explicitly to purchase specialised human capital by 1700–1750. In France, the Crown created specialised engineering corps and, by the 1740s, was training and salarising mathematically competent civil engineers (for roads, bridges, canals) at wages far above those of ordinary labourers, on the explicit premise that such expertise would raise transport efficiency, commerce, and fiscal capacity (Picon 2007). In Britain, the diffusion of early steam technology from the 1710s onward generated rents for technically trained engine specialists in mining, indicating that frontier technology already rewarded scarce technical competence (Kanefsky and Robey 1980). Finally, within elite professions such as the clergy and academia in the early eighteenth century, large income differences existed between ordinary incumbents and those with recognised expertise, patronage, or reputation, even though they nominally held the same occupation (Lindert and Williamson (1982); Best (1964)).

Taken together, these facts imply that human capital at the top of the distribution was not merely a marker of status; it was a pathway to higher earnings and local economic dynamism. Our argument is that families embedded in this environment — including the academic elite we observe — would rationally begin to substitute towards "quality" of children, anticipating that advanced education could be translated into income and status within a single generation.

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 is more consistent with the available data.

Main assumptions - In an overlapping generations setup, we assume that each individual lives two periods: childhood and adulthood. During childhood, the individual is inactive, but receives a portion ϕ of her parents' 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 time $t-1$ is characterized by her level of human capital h_t and a utility function inherited from her parents. Human capital maps into income and there is no occupational choice. If the functional form of the utility function is the same for all individuals, they may differ with respect to 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:

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

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:

$$h_{t+1}^i = \psi e_t^i, \tag{3}$$

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 time 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 minimum consumption constraint such that

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

This constraint introduces a Malthusian dimension to our model because, if binding, it restricts households' fertility decisions and increases the importance of income effects. We assume that individuals have two sources of income: labor income and non-labor income. The wage per efficient unit of labor is normalized to 1, while $a > 0$ represents non-labor income, which may, for example, correspond to home production. The budget constraint of an adult at time t is then:

$$c_t^i + \phi n_t^i h_t^i + e_t^i n_t^i = h_t^i + a. \tag{4}$$

In Equation (4), the cost of quantity n_t^i is a time cost and hence depends on the human capital of the parents h_t^i, while the cost of the quality e_t^i is monetary (goods). Hence the human capital of the parent provides a comparative advantage in quality on account of quantity, as in De la Croix and Doepke (2003) and Moav (2005).

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 minimum consumption constraint can be satisfied. An adult born at time $t-1$ will maximize (2) subject to (3) and (4) and the 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:

$$\bar{h} = (1+\gamma)\bar{c} - a,$$

so that the solutions of the individual maximization program are described in Table 5.

Decisions - When $h_t^i \leq \bar{h}$, the constraint $c_t^i \geq \bar{c}$ is binding and the Malthusian regime prevails. In this situation, provided that non-labor income is not high ($a < \bar{c}$ by Assumption 1), fertility increases with h_t^i. An increase in parental human capital increases the opportunity cost of time spent with children, which should depress fertility, but it also increases total income enough to eventually increase both the quality and quantity of children. In other words, 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 because the opportunity cost effect now dominates the income effect. The opposition between these two effects is illustrated in Figure 7. An increase in $\bar{c}$ increases the range of h_t^i for which the household is trapped in a Malthusian situation.

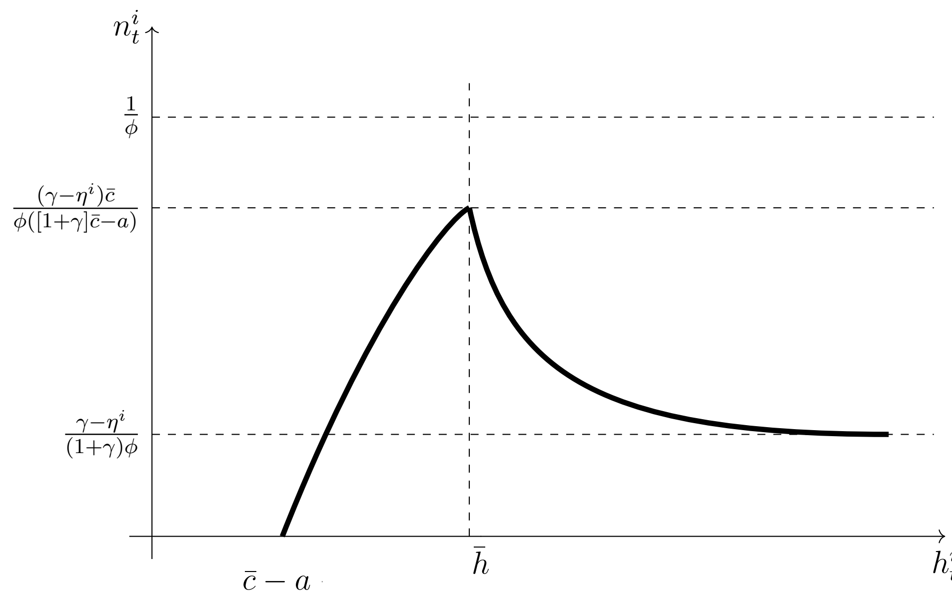


Figure 7: Fertility as a function of parents' human capital

Parental investment in children's education is unaffected by the prevailing regime. This is simpler than the more complex models such as Galor and Weil (2000) and De la Croix and Doepke (2003), but it does not alter the generality of our results and allows us to characterize the accumulation of human capital over time in a simple way:

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

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

h_t^i	$h_t^i \leq \bar{h}$	$h_t^i > \bar{h}$
c_t^i	$\bar{c}$	$\frac{h_t^i + a}{1+\gamma}$
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 5: Individual decisions as a function of own human capital

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 $\psi > \frac{\gamma - \eta^i}{\phi \eta^i} \;\; \forall i$.

From here on, we limit our analysis to situations where Assumption 2 is satisfied. In other words, we limit our analysis to situations in which human capital is strictly increasing for all families. We have found that quality-oriented individuals escape the Malthusian trap earlier than their quantity-oriented counterparts. We now analyze the fertility differentials between these two groups. Proposition 1 summarizes our results.

Proposition 1 *Under assumptions 1 and 2,* $\frac{\partial h_t^i}{\partial \eta^i} > 0 \; \forall i$ *while* $\frac{\partial n_t^i}{\partial \eta^i} < 0 \; \forall t > \bar{t}^i$ *and* $\forall i$. *Furthermore, 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 Q.*

The net impact of η^i on fertility is determined by the opposition of 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 $\bar{t}^i$, in the Malthusian regime ($t < \bar{t}^i$), the accumulation effect dominates the preference effect so that quality-oriented parents have an evolutionary advantage over quantity-oriented parents. Once they enter the Beckerian (interior) regime, they lose this advantage in favor of quantity-oriented parents.[19]

Proposition 1 describes the evolution of the quality-quantity tradeoff at the microeconomic level, but is silent on the aggregate moments, which are the moments we estimate in the previous section.

Proposition 2 *There exist dates $\hat{t}$, $\bar{t}$, and $\breve{t}$ such that $\hat{t} > \bar{t} > \breve{t} > 0$, and:*

$$\forall t \in (\breve{t}, \bar{t}), \frac{\partial n_t^i}{\partial \eta^i} \geq 0 \quad \forall i, \quad \textit{and} \quad \forall t > \hat{t}, \frac{\partial n_t^i}{\partial \eta^i} < 0 \quad \forall i.$$

Proof 2 *See Appendix R*

The difference between Proposition 1 and Proposition 2 is that in the former we define a collection of dates at which families transition from one regime to the other, i.e., each family has a specific transition date, while in the latter we distinguish two specific time periods during which all families adopt the same type of fertility behavior. From $t = \breve{t}$ to $t = \tilde{t}$, all families are in a Malthusian regime where the accumulation effect dominates the taste effect so that $\frac{dn_t^i}{d\eta^i} > 0$. Conversely, when $t > \bar{t}$, all families are in the interior regime such that $\frac{dn_t^i}{d\eta^i} < 0$. Figure 8 illustrates the result.

In the intermediate period $t \in (\tilde{t}, \bar{t})$, the fertility behavior of our families is 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 measuring the association between sibship size and human capital of individuals over a period of time $t \in (\breve{t}, +\infty)$

[19] The astute reader will have noticed that the dominance of the accumulation effect over the preference effect is not necessarily true at any time t within the Malthusian regime. In Appendix Q we show that from $t = 0$ to $t = t_0$ the taste effect dominates. However, this situation is transitory and may correspond to times not covered by our data. Finally, note that the size of this time window is proportional to $\frac{\eta^i}{\gamma}$, which represents the weight of human capital relative to the number of children in the utility function of agent i.

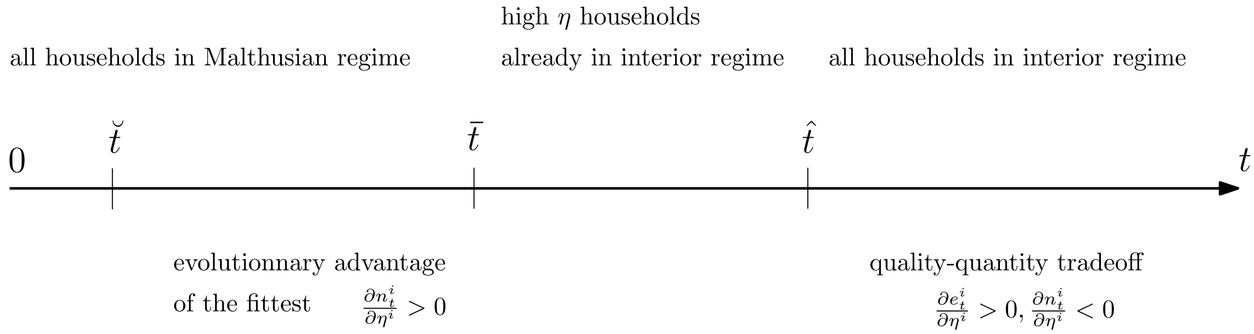


Figure 8: Proposition 2

would identify three distinct periods: a period in which sibship size and human capital are positively associated, followed by the absence of a significant relationship, and then the emergence of a negative association.

5 Quantitative Analysis

In this section, we first estimate the parameters of the model by indirect inference. Then, we simulate the model and discuss its quantitative implications.

5.1 Indirect Inference

Before performing the estimation, we introduce two additional distributional assumptions. First, the preference parameter η is distributed over the interval $\eta_{\min}, \eta_{\max}$ as follows:

$$\eta = \eta_{\min} + \varepsilon^{\eta}(\eta_{\max} - \eta_{\min})$$

where ε^{η} is drawn from a symmetric Beta distribution with shape parameter $\zeta \geq 1$, $\mathcal{B}(\zeta, \zeta)$. The case $\zeta = 1$ corresponds to a uniform distribution. The higher ζ, the lower the variance of ε^{η}. The lower bound $\eta_{\min}$ is given by Assumption 2, i.e. $\eta_{\min} = \frac{\gamma}{1+\phi\psi}$. By definition of its distribution, the upper bound will be $\eta_{\max} = E[\eta] + E[\eta - \eta_{\min}] = 2E[\eta] - \eta_{\min}$. $E[\eta] = \bar{\eta}$ is a parameter to be calibrated.

Second, even after correcting for the underreporting of girls' births and the verticality bias in the genealogical data, our measure of sibship size may still not perfectly reflect reality. To account for residual approximations made by genealogists, we incorporate a measurement error affecting fertility. By explicitly modeling the measurement error, we account for the attenuation bias it introduces in the regression of fertility on publications. We assume that observed fertility follows:

$$n_{obs} = \max(0, n + \varepsilon^n)$$

where ε^n is drawn from a normal distribution $\mathcal{N}(0, \sigma^2)$. The assumption of a zero mean for our misreporting results from the interplay of two opposing phenomena. Blanc (2024) and Charpentier and Gallic (2020) document systematic underreporting of births in family trees, which would suggest a negative mean for our bias. However, our sample selection of genealogies has excluded a high proportion of families with no children or only one child, suggesting a tendency toward systematic over-reporting. The right panel of Figure 2 shows that at the aggregate level these two biases tend to cancel each other out, supporting the white noise hypothesis at the individual level.

We also need to translate the time of the structural model into actual data. We will assume that a period lasts 20 years and focus on the years 1635, 1655, 1675, 1695, 1715, 1735, and 1755.

The full set of parameters to be identified is now:

$$\{h_0, \gamma, \psi, \bar{c}, \phi, a, \bar{\eta}, \zeta, \sigma\}$$

Table 6 summarizes the results of the identification. We proceed in three steps.

Step 1. Two parameters are set ex ante. $h_0 = 1$ (normalization), $\phi = 1/11$. The value of ϕ implies that the maximum number of sons one can possibly have is 11 (i.e. 22 children in total), which is reasonable given the following distribution of parities:

No. boys	1	2	3	4	5	6	7	8	9	10	11	12	13	14	16
No. families	404	528	512	358	223	115	65	34	27	25	16	1	2	2	2

Only seven persons in the dataset had more than 11 boys.

Step 2. Five parameters are set to match exactly five moments that we impose on the model: $\{\gamma, \psi, \bar{c}/a, E[\eta]\}$. Our goal here is to generate reasonable predictions in terms of fertility levels, education levels, and growth. As in De la Croix and Doepke (2003), we impose that for the average family in the long run, i.e. when $h_t \to \infty$, fertility is at its replacement rate ($n = 1$) and education spending is 7.3% of GDP, the value observed in the US today:

$$\frac{\gamma - \bar{\eta}}{\phi(1+\gamma)} = 1 \qquad \frac{\phi\bar{\eta}}{(1-\phi)(\gamma - \bar{\eta})} = 0.073$$

Parameter	value	matched moment	value	fit
Fixed ex ante				
h_0	1	normalization		
ϕ	$^1/_{11}$	Distribution of parities		
Exact identification				
γ	0.187	$\lim_{h_t \to \infty} n_t$ for average family	1	1
$\bar{\eta}$	0.079	$\lim_{h_t \to \infty} \frac{e_t n_t}{a+(1-\phi n_t)h_t}$ for average family	0.073	0.073
ψ	15.452	$\psi\phi E[\eta]/(\gamma - E[\eta])$	1.025	1.025
$\bar{c}/a$	1.207	$\lim_{h_t \to \bar{h}} n_t$	3.316	3.316
a	3.011	regime shift attained after 11 periods		
Indirect inference				
		$\hat{a}_1\|_{1635}$	0.057	0.055
		$\hat{a}_1\|_{1655}$	0.058	0.054
ζ	3.078 (1.201)	$\hat{a}_1\|_{1675}$	0.056	0.028
σ	0.346 (0.066)	$\hat{a}_1\|_{1695}$	0.004	0.002
		$\hat{a}_1\|_{1715}$	-0.061	-0.030
		$\hat{a}_1\|_{1735}$	-0.083	-0.055
		$\hat{a}_1\|_{1755}$	-0.056	-0.101

Note: SE in parentheses from 100 draws of the empirical moments.

Table 6: Estimated parameters

We impose that fertility at the regime shift (when $h = (1+\gamma)\bar{c} - a$) is equal to its average of 3.316 boys. Finally, we impose that the growth factor of human capital for the average person, $\psi\phi\bar{\eta}/(\gamma - \bar{\eta})$, is equal to $1.001256^{20} = 1.025$, reproducing the coefficient of time (birth date) in the third column of Table 8. This gives $\gamma = 0.187$, $\psi = 15.506$, $\bar{c}/a = 1.207$, $\bar{\eta} = 0.079$. The interval from which the preference parameters η are drawn is thus quite narrow.

The value of a will be important for the time at which the regime shift occurs. We want to allow enough time from the initial condition to the regime change, to allow differences in human capital to build up as a function of η's. We set $\bar{h} = 1.3h_0$, which means that at the calibrated growth rate, the regime shift is reached after 11 periods for the average individual. This gives $a = 3.011$.

With the calibrated parameter values we can already calculate $\eta_{\text{min}} = 0.0774$ and $\eta_{\text{max}} = 0.08$.

Step 3. Indirect inference is used to estimate how much heterogeneity is needed, i.e. ζ, and the importance of measurement error in fertility, σ. In practice, for given values of ζ and σ, we simulate the model over a horizon of 20 periods starting in 1455 and with 600 families i (similar to the sample size of the auxiliary model). For each period, we run a regression of simulated fertility in t on simulated human capital in $t+1$:

$$\text{asinh}(h^i_{t+1}) = \kappa_t(\zeta, \sigma) + \beta_t(\zeta, \sigma) n^i_t$$

where $\kappa_t(\zeta, \sigma)$ is a constant and $\beta_t(\zeta, \sigma)$ is the coefficient of interest reflecting the correlation between sibship size and publications. The estimated regression coefficient $\hat{\beta}_t(\zeta, \sigma)$ depends on the chosen parameters (ζ, σ) and is comparable to α_1 of Equation (1). We then minimize W_t, the mean squared error between the OLS coefficient estimated on the simulated data and the one estimated on the observed data:

$$\min_{\zeta,\sigma} W_t = \sum_{t=9}^{15} \left(\hat{\beta}_t(\zeta, \sigma) - \hat{\alpha}_1|_{1455+20t} \right)^2$$

Minimization yields the values $\hat{\zeta}$ and $\hat{\sigma}$ reported in Table 6. The standard errors of these parameters are obtained by drawing 100 values of the moments, assuming they are normally distributed, and re-estimating the model each time.

Figure 9 shows the simulated regression coefficient $\hat{\beta}_t(\hat{\zeta}, \hat{\sigma})$ (solid line) compared to the regression coefficients $\hat{\alpha}_1|_{1455+20t}$. The shaded area represents the 95% confidence intervals around $\hat{\alpha}_1$. The dotted line represents the simulated data under the assumption that there

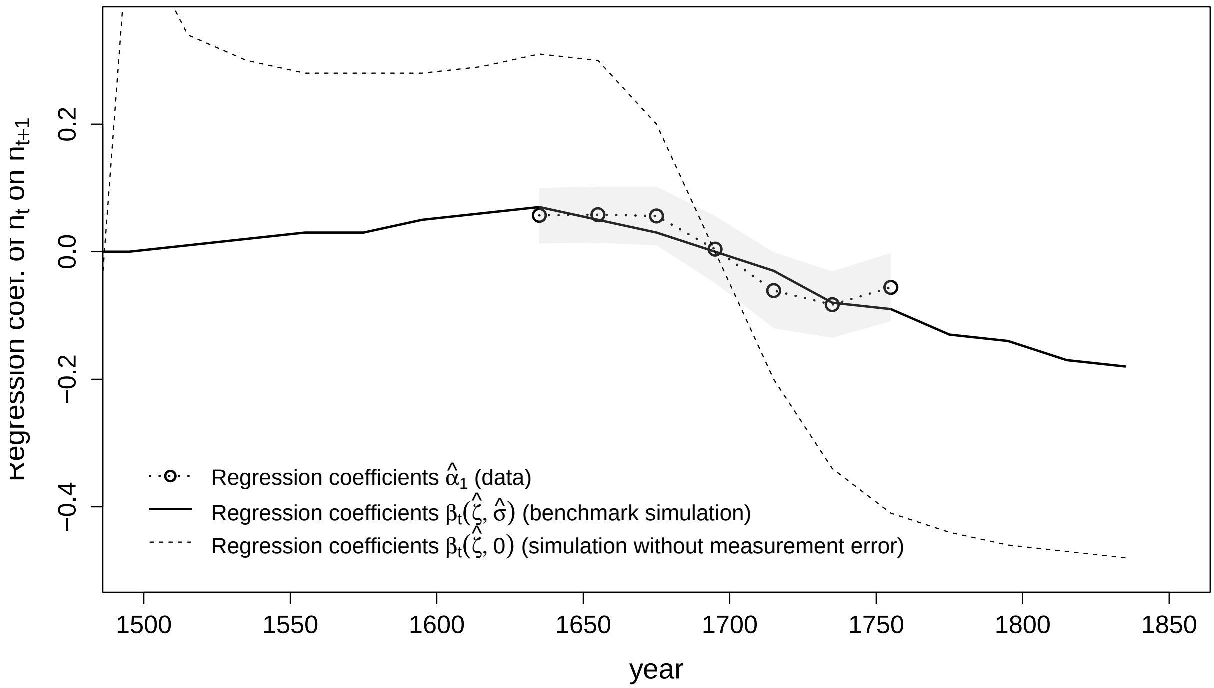


Figure 9: Fit with Parameters Estimated via Indirect Inference

is no measurement error in the genealogies ($\sigma = 0$). The gap between the dotted line and the solid line thus represents the attenuation bias due to measurement error.

The structural estimation demonstrates that the model accurately explains the observed data patterns without the need for external shocks. It also provides two additional insights. First, only minimal heterogeneity in preferences (η) is needed to generate the appropriate differential fertility over time. Second, measurement errors in fertility, caused by the inevitable errors and approximations made by genealogists in constructing family trees, are crucial in explaining why our regression coefficients are generally small. Consequently, our coefficients are likely to be lower bounds on the true values, given the varying quality of genealogical data.

5.2 Simulations

The main mechanism involves an endogenous transition from a Malthusian to a Beckerian constraint, driven by human capital accumulation during the Malthusian era. This transition is endogenous once the initial conditions are set. Introducing a minimum consumption level, $\bar{c}$, serves as a device to generate two distinct regimes: the Malthusian regime and the interior regime. These regimes differ in various ways, but one aspect is particularly relevant to our analysis: the Malthusian regime is not characterized by complete stagnation. While consumption remains constant, investment in human capital steadily rises. This growth is further amplified by a composition effect: quality-oriented individuals enjoy an evolutionary advantage, leading to their increasing prevalence in the population. Consequently, aggregate investment in human capital continues to expand throughout the Malthusian period. We see this very clearly in the following simulation.

Figure 10 depicts the dynamics of three lineages corresponding to the 25th, 50th, and 75th percentiles (Q25, Q50, Q75) of the taste-for-quality distribution η. The top-left panel shows fertility and reproduces the theoretical pattern in Figure 7: among these groups, Q75 exits the corner regime (rising fertility) first and enters the interior regime (declining fertility). The top-right panel tracks each group's population share within the total of 600 households. The initially higher fertility of Q75 raises its population share–an evolutionary advantage emphasized by Galor and Moav (2002)–but this advantage dissipates once its fertility falls; toward the end of the period, the quartile with the lowest taste for quality gains the evolutionary edge.

The middle-left panel reports education spending as a share of income. In the Malthusian regime, this share rises over time with h_t^i:

$$\frac{n_t^i e_t^i}{(1-\phi n_t^i)h_t^i + a} = \frac{1}{1+\frac{\gamma\bar{c}}{\eta\,(h_t^i+a-\bar{c})}}\,.$$

It stays constant in the interior regime. The middle-right panel presents the annualized income growth factor for the three groups. The model predicts very moderate annual income growth in the sixteenth to eighteenth century. This moderate growth persists in the future, showing that our model alone will have limited power to explain the nineteenth and twentieth century growth rates without additional mechanisms.

This is confirmed by computing aggregate variables (for academics). The bottom two panels report aggregate fertility and income, obtained by weighting lineages by their population shares. They show that moderate income growth is feasible in the seventeenth and eighteenth centuries without a demographic transition. The slight decline in aggregate fertility then stalls around 1800, and income growth slows concurrently. Around 1800, composition effects turn adverse: the population share of low-η lineages rises. Thus, as in Galor and Moav (2002) and Collins, Baer, and Weber (2014), differential fertility initially supports an early takeoff but later reverses its effect. This reversal would motivate the introduction of additional mechanisms—such as endogenous technical change, or human capital externalities—to

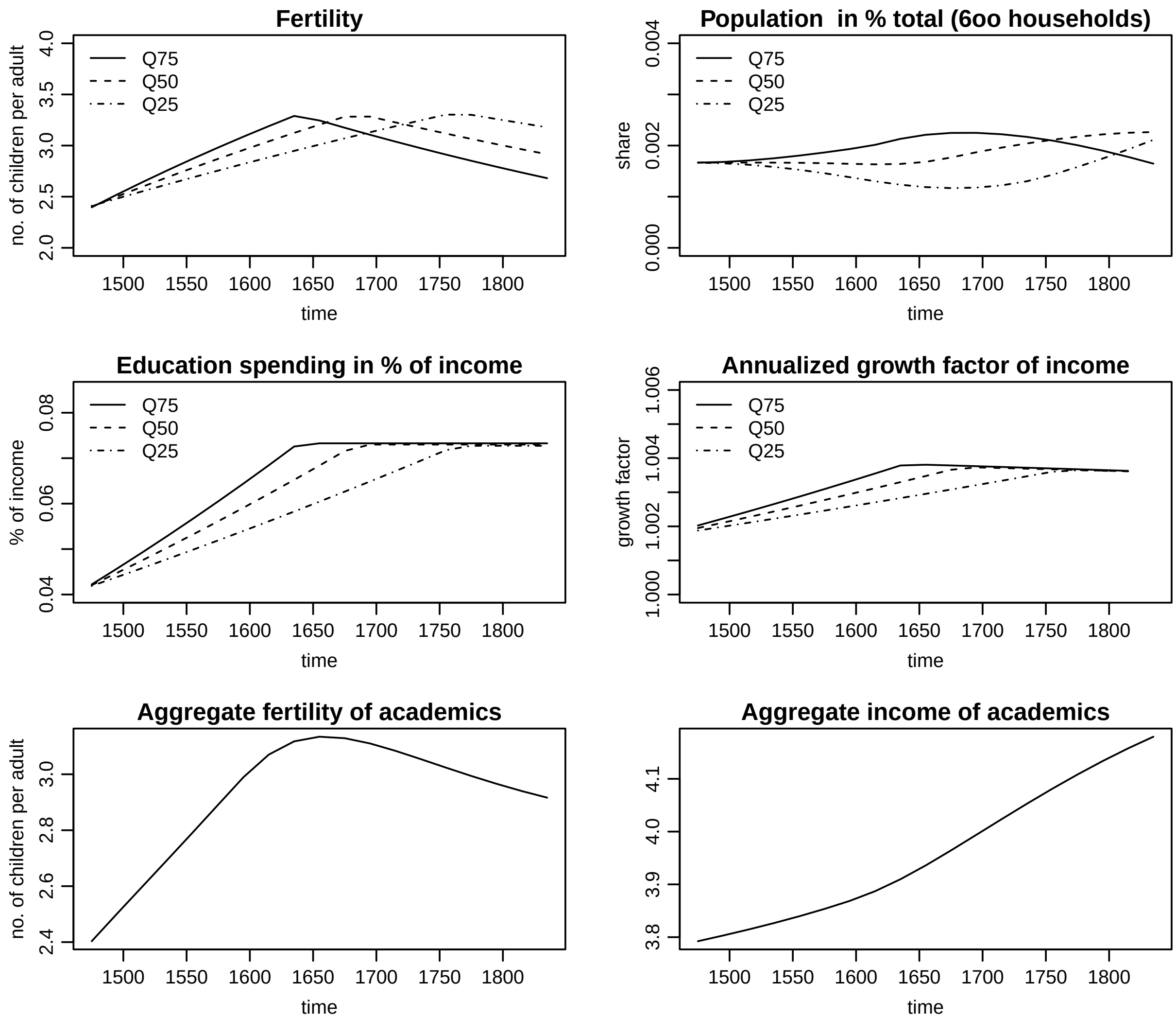


Figure 10: Simulation results for three lineages corresponding to the 25th, 50th, and 75th percentiles (Q25, Q50, Q75) of the taste-for-quality distribution η

account for nineteenth-century growth.

Finally, looking now beyond our model at the whole economy, the accumulation of human capital during Malthusian times is strongly supported by the data. Calculating the share of academic scholars in the total population, we find that Denmark had 0.022 scholars per 1,000 inhabitants in 1536.[20] By 1640, this share had risen to 0.07 per 1,000, reaching 0.223 by 1769. A similar trend is observed in Sweden,[21] where the number of scholars per 1,000 adults increased from 0.033 in 1630 to 0.051 in 1700 and 0.276 in 1760. This growth in the number of scholars was not driven by an expansion of existing institutions but rather by the establishment of new ones over time.

6 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 labour was viewed simply as the wages or salaries paid to workers, not as the investment workers made in their own knowledge and skills. The concept of human capital challenged this view by recognising 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 labour to the value of labour, and from the quantity of labour to the quality of labour.

The paradigm shift also led to the development of new analytical tools and methods for measuring the impact of human capital on economic growth. As these innovations matured, they encountered 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 crucial step toward a more mature understanding of the role of human capital for growth has been to shift the analytical focus away from average levels of literacy and skill and instead to consider the human capital of those at the top 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 to show how members of academic institutions changed their behaviour. We argue that a key mechanism underlying this change was the ability of human capital to generate wealth as early as 1750, which allowed these individuals

[20] Danish population estimates for the years 1536, 1640 and 1660 are obtained from Hilson, Poulsen, and Olesen (2023). We assume that half the population was adult, a standard proportion for the Malthusian epoch. The first population census was in 1769.

[21] Total population from Edvinsson (2015)

to transcend 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 are supportive of the key mechanisms of Unified Growth Theory, namely the reversal of the QQ trade-off over time. This empirical confirmation strengthens the theoretical framework put forward in previous studies. Second, by placing this reversal a century before industrialisation, our findings suggest that mechanisms beyond firms' increased demand for human capital during the Industrial Revolution may have initiated the shift toward behaviours conducive to sustained economic growth.

Acknowledgements

We thank Sascha Becker, Francesco Cinnirella, Matthew Curtis, Matthias Doepke, Oded Galor, Ralf Meizenthal, Stelios Michalopoulos, Joel Mokyr, Stéphane Zuber, and participants at seminars at UNSW (Sydney), ANU (Canberra), Sussex, NYU Abu Dhabi, Paris School of Economics, IESEG Business School (Paris), the EHA conference in Pittsburgh, the Conference on Macroeconomic Implications of Intra-Household Decision Making (Frankfurt), U. of Basel and U. of Brown for comments during seminars. 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.

Academia Rheno-Trajectina. 1861. *Series professorum, qui in academia Rheno-Trajectina publice aut docuerunt aut etiamnunc docent.* Utrecht: De Bruyn, Bibliopolam.

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.

Bain, R Nisbet. 2014. *Scandinavia: A Political History of Denmark, Norway and Sweden from 1513 to 1900.* Cambridge University Press.

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.

Best, Geoffrey. 1964. *Temporal Pillars: Queen Anne's Bounty, the Ecclesiastical Commissioners, and the Church of England.* Cambridge: Cambridge University Press.

Bhalotra, Sonia, and Damian Clarke. 2019. "The Twin Instrument: Fertility and Human Capital Investment." *Journal of the European Economic Association* 18 (6): 3090–3139 (12).

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.

Blanc, Guillaume. 2024. "Demographic Transitions, Rural Flight, and Intergenerational Persistence: Evidence From Crowdsourced Genealogies."

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.

Clark, Gregory, and Gillian Hamilton. 2006. "Survival of the Richest: The Malthusian Mechanism in Pre-Industrial England." *The Journal of Economic History* 66 (3): 707–736.

Collins, Jason, Boris Baer, and Ernst Juerg Weber. 2014. "Economic Growth and Evolution: Parental Preference for Quality and Quantity of Offspring." *Macroeconomic Dynamics* 18 (8): 1773–1796.

Coutts, James. 1909. *A History of the University of Glasgow: From its Foundation in 1451 to 1909.* Glasgow: J. Maclehose and Sons.

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. 2024. "The Academic Market and the Rise of Universities in Medieval and Early Modern Europe (1000-1800)." *Journal of the European Economic Association* 22:1541–1589.

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 Marc Goñi. 2024. "Nepotism vs. Intergenerational Transmission of Human Capital in Academia (1088–1800)." *Journal of Economic Growth* 29:469—-514.

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.

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.

de Hesse-Hombourg, Société Patriotique. 1777. *Programme de la Société Patriotique de Hesse-Hombourg.* Hombourg-ez-Monts: Société Patriotique.

Delen, Första, and Martin Johan Julius Weibull. 1868. *Lunds Universitets Historia, 1668-1868.* Lund: CWK Gleerup.

de Vries, Jan, and Ad van der Woude. 1997. *The First Modern Economy: Success, Failure, and Perseverance of the Dutch Economy, 1500–1815.* Cambridge: Cambridge University Press.

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 historiqu*, pp. 391–395.

Edvinsson, Rodney Benjamin. 2015. "Recalculating Swedish pre-census demographic data: Was there acceleration in early modern population growth?" *Cliometrica* 9 (2): 167–191.

Elgán, Elisabeth, and Irene Scobbie. 2015. *Historical dictionary of Sweden.* Rowman & Littlefield.

Eloy, Nicolas FJ. 1755. *Dictionnaire Historique De La Médecine: Contenant Son Origine, Ses Progrès, ses Révolutíons, 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.

Fulbrook, Mary. 2019. *A concise history of Germany.* Cambridge University Press.

Galor, Oded. 2011. *Unified growth theory.* Princeton University Press.

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.

Glahn, Torben. 1975. *Soraner-Biografier 1747-1800.* Sorø: Soransk Samfund.

Gourieroux, C, A Monfort, and E Renault. 1993. "Indirect Inference." *Journal of Applied Econometrics* 8 (S): S85–118 (Suppl. Dec.).

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.

Heckscher, Eli Filip, and Gunnar Heckscher. 1954. *An economic history of Sweden.* Volume 95. Harvard University Press.

Hilson, Mary, Bjørn Poulsen, and Thorsten Borring Olesen. 2023. "Open Online Course in Danish History: A History of Denmark from the Viking Age to the 21st Century." Edited by Aarhus University, danmarkshistorien.dk.

Hirsch, Theodor. 1837. *Geschichte des academischen Gymnasiums in Danzig, in ihren Hauptzügen dargestellt.* Wedelschen Hofbuchdruckerei.

Hu, Sijie. 2024. "Celebrating legacy: The intergenerational transmission of reproduction and human capital in Ming–Qing Chinese families."

Iaria, Alessandro, Carlo Schwarz, and Fabian Waldinger. 2018. "Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science*." *The Quarterly Journal of Economics* 133 (2): 927–991 (01).

Inno, Karl. 1972. *Tartu University in Estonia during the Swedish rule (1632-1710).* Vaba Eesti.

Israel, Jonathan. 1995. *The Dutch Republic: its rise, greatness, and fall 1477-1806.* Oxford University Press.

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.

Jespersen, Knud J. V. 2004. *A History of Denmark.* Houndmills, Basingstoke, Hampshire and New York: Palgrave Macmillan.

Kanefsky, John, and John Robey. 1980. "Steam Engines in 18th-Century Britain: A Quantitative Assessment." *Technology and Culture* 21 (2): 161–186.

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.

Lindert, Peter H., and Jeffrey G. Williamson. 1982. "Revising England's Social Tables, 1688–1812." *Explorations in Economic History* 19 (4): 385–408.

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.

Meinander, Henrik. 2020. *History of Finland*. Oxford University Press, USA.

Minardi, Saverio, Giulia Corti, and Nicola Barban. 2023. "Historical patterns in the intergenerational transmission of lifespan and longevity: Evidence from the United States, 1700-1900." *SocArXiv. doi*, vol. 10.

Mincer, Jacob. 1958. "Investment in human capital and personal income distribution." *Journal of political economy* 66 (4): 281–302.

Moav, Omer. 2005. "Cheap children and the persistence of poverty." *The Economic Journal* 115 (500): 88–110.

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.

Olrik, HG. 1927. *Oversigt over Forstandere og Lærere ved den kgl. Skole i Sorø 1586–1737*. Sorø: Folkebladets Bogtrykkeri.

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.

Persson, Karl Gunnar, and Paul Sharp. 2015. *An economic history of Europe.* Cambridge University Press.

Picon, Antoine. 2007. "French Engineers and Social Thought, 18th–20th Centuries: An Archeology of Technocratic Ideals." *History and Technology* 23 (3): 197–208.

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.

Schmidt, Olaus Dons. 1960. *Det kongelige norske videnskabers selskab: matrikkel, 1760-1960.* Sentrum.

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.

Singleton, Fred. 1998. *A short history of Finland.* Cambridge University Press.

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 Personalhistorie.

Smart, Robert Noyes. 2004. *Biographical Register of the University of St. Andrews, 1747-1897.* University of St. Andrews Library.

Smith, Anthony A., Jr. 2008. "Indirect Inference." In *The New Palgrave Dictionary of Economics*, edited by Steven N. Durlauf and Lawrence E. Blume. London: Palgrave Macmillan.

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 (10): e2120455119.

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, francois, 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.

Vardè, Antonino Alessio. 2024. "Inheritance Laws and Ecclesiastical Career: a study of Italian Nobility." presented at ULBrussels.

Vogl, Tom. 2016. "Differential fertility, human capital, and development." *The Review of Economic Studies* 83 (1): 365–401.

Volbehr, 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.

Welsh, Frank. 2003. *The four nations: a history of the United Kingdom.* Yale University Press.

Wormald, Jenny. 2005. *Scotland: A history.* Oxford University Press.

Wrigley, Edward., and Roger Schofield. 1983. "English Population History from Family Reconstitution: Summary Results 1600-1799." *Population Studies* 37 (2): 157–184.

Zanardello, Chiara. 2024. "Early Modern Academies and Growth." UCLouvain.

A Main shocks in the 10 most frequent birth places and number of professors' families affected

City	# births	Event	Years	References	# concerned
Stockholm	104	Stockholm Blood Bath	1520		0
		Conquest of Stockholm	1523	Heckscher and Heckscher (1954)	0
		First Great Stockholm Fire	1625	Elgán and Scobbie (2015)	0
		Bubonic Plague	1710-1711	Bain (2014)	49
		Second Great Stockholm Fire	1759		6
Copenhagen	102	Siege of the Count's Feud	1534-1536		0
		Siege of the Second Northern War	1658-1660	Jespersen (2004)	9
		Bubonic Plague	1711-1712		9
		Copenhagen Fires	1728 & 1795	Bain (2014)	28
		1st and 2nd Battles of Copenhagen of the Napoleonic wars	1801 & 1807		0
Edinburgh	50	Burning of Edinburgh	1544		0
		The Lang Siege during the Marian Civil War	1571-1573		0
		Siege of Edinburgh Castle during the Bishops' Wars	1639	Wormald (2005)	0
		Siege of Edinburgh Castle during Oliver Cromwell's invasion of Scotland	1650	Welsh (2003)	0
		The Great Plague	1644-1645		3
		Jacobite Sieges	1689, 1715, 1745		37
Rostock	47	Occupation of Rostock during the 30 Years War	1618-1648		11
		Second occupation of Rostock	1700-1721	Fulbrook (2019)	8
		Great Fire of Rostock	1677		4
Uppsala	40	Conquest of Uppsala during the Swedish War of Liberation	1521	Heckscher and Heckscher (1954)	0
		Uppsala Fire	1702	Elgán and Scobbie (2015)	9
		Bubonic Plague	1710-1711	Bain (2014)	16

City	# births	Event	Years	References	# concerned
Amsterdam	31	First Great Fire of Amsterdam	1421		0
		Second Great Fire of Amsterdam	1452	Israel (1995)	0
		Siege of Amsterdam during the Ostend War	1633-1634	Persson and Sharp (2015)	3
		Bubonic plague	1635-1637		
Leiden	27	Siege of Leiden during the Eighty Years' War	1573-1574	Israel (1995)	1
		Bubonic plague pandemic	1635-1636	Persson and Sharp (2015)	2
Lund	26	Battle of Lund and Danish occupation during the Scanian War	1676-1679	Heckscher and Heckscher (1954), Elgán and Scobbie (2015)	2
		Bubonic plague	1654-1655	Bain (2014)	2
Turku	25	Siege of Abo	1563		0
		Siege of Turku Castle and Abo Blood Bath	1597-1599	Singleton (1998)	1
		Bubonic Plague	1710-1711	Meinander (2020)	5
		Russo-Swedish War	1741-1743		5
Utrecht	25	Bubonic plague	1635-1637	Israel (1995)	4
		Rampjaar	1672	Persson and Sharp (2015)	6

Table 7: Main events in the 10 most frequent birth places and number of professors' families affected. To be affected in their QQ-tradeoff, the event should occur in a time window of -10 years and +10 years around the birth of the professor.

B Repartition of scholars by the institution

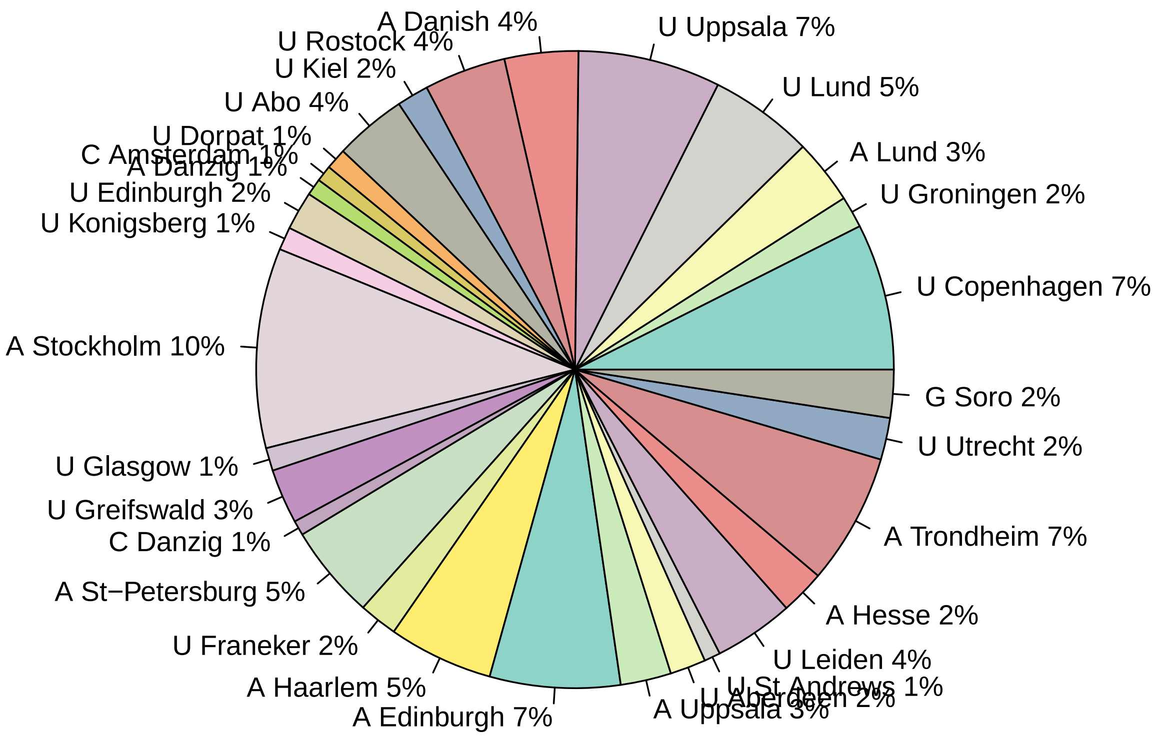


Figure 11: Distribution of scholars by institutions

C Origin of all scholars

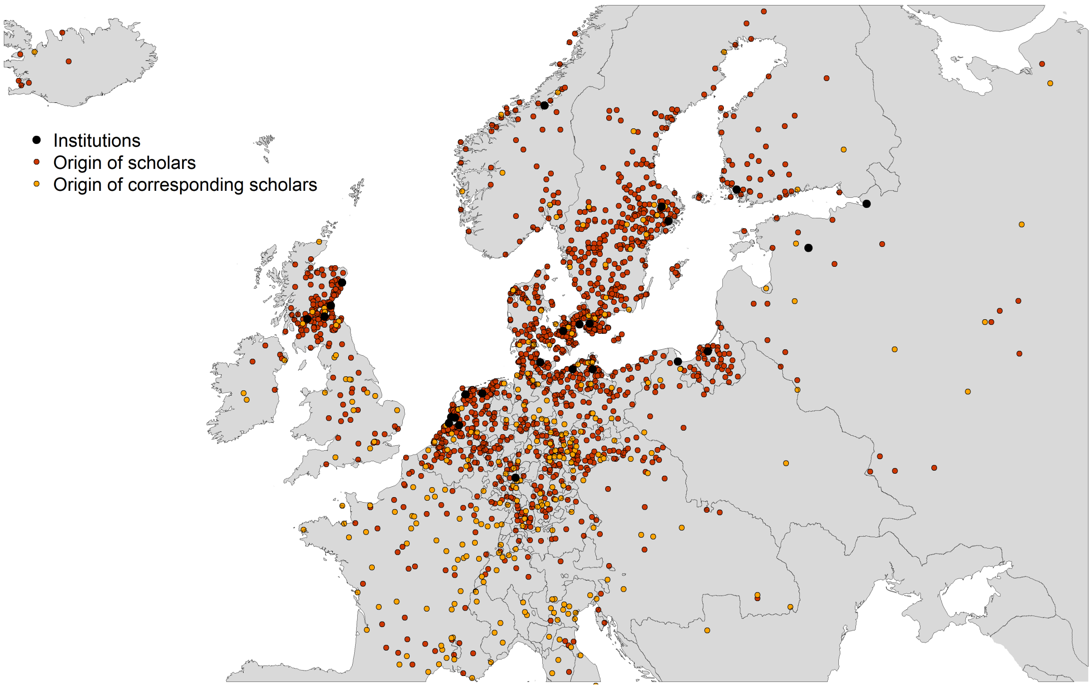


Figure 12: Origin of all scholars, frontiers of 1700

D Timeline of institutions' creation

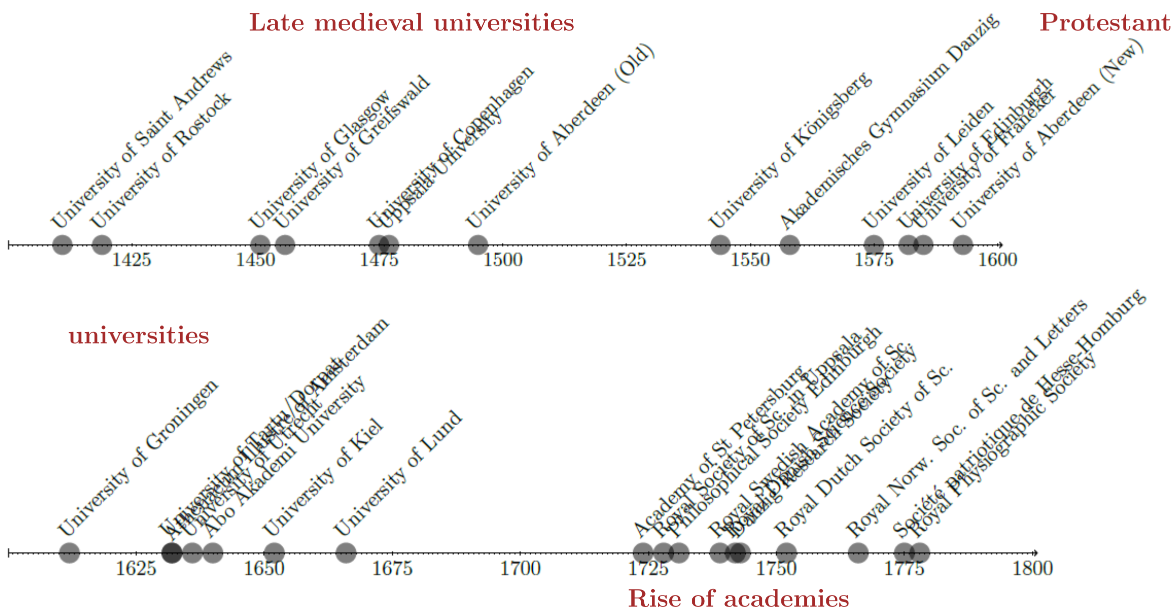
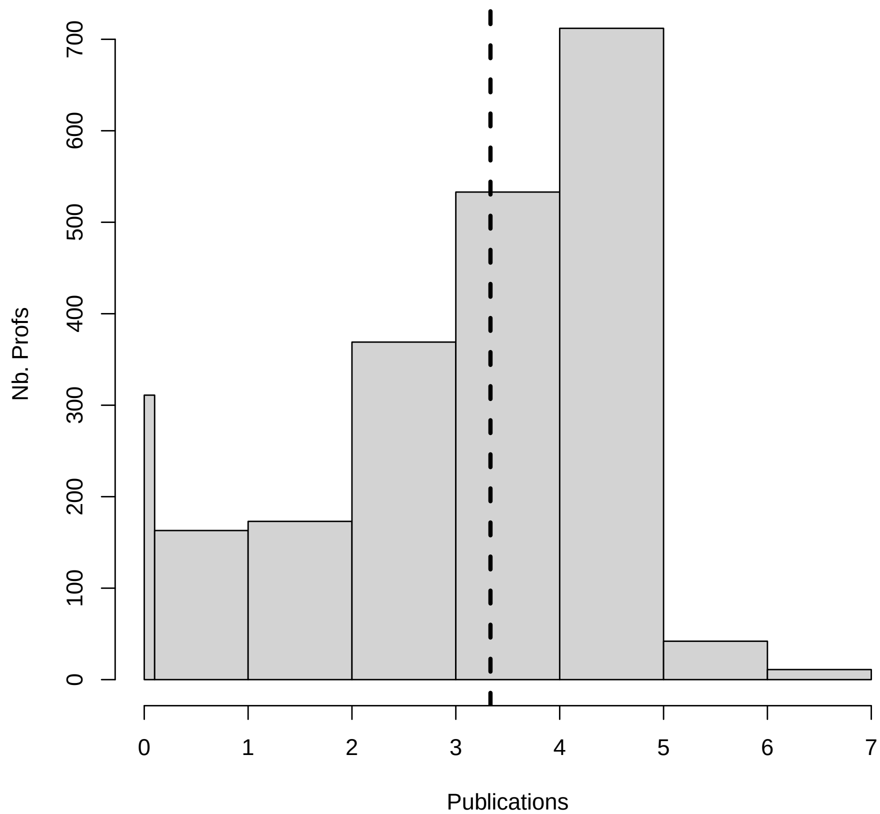


Figure 13: Timeline of institutions' creation

E Histogram of the distribution of number of publications



Note: Publications = asinh(Number of titles in VIAF). Median as dashed line.

Figure 14: Histogram of the distribution of number of publications

F Correlates of publications

We regress the inverse hyperbolic sine of individual number of works published on a time trend based on birth dates:

$$
\begin{aligned}
\operatorname{asinh}(\text{Publis}_i) = \gamma_1 \text{BirthDate}_i + \gamma_2 \text{Longevity}_i + \gamma_3 \text{Age Nomination}_{ik} + \gamma_4 \text{Academic Father}_i \\
+ \gamma_5 \text{With Genealogy}_i + \gamma_6 \text{Corresponding}_{ik} + \sum_{j \in K} \gamma_{7j} \ \mathrm{I}(k = j) \\
+ \sum_{f \in F} \gamma_{8f} \ \mathrm{I}(d_{ik} = f) + \theta X_i + \varepsilon_{ik}. \quad (6)
\end{aligned}
$$

Results are shown in the first column. 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. 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 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.

OLS regression: correlates of publications

birth date (γ_1)	0.0017^{***}	0.0008^{**}	0.0013^{***}
	(0.0004)	(0.0004)	(0.0004)
longevity (γ_2)		0.0179^{***}	0.0184^{***}
		(0.0019)	(0.0017)
age nomin. (γ_3)		0.0086^{***}	-0.0034
		(0.0022)	(0.0021)
academic father. (γ_4)		0.1747^{*}	-0.0094
		(0.0988)	(0.0919)
theology (γ_{7t})			0.3545^{***}
			(0.0640)
law (γ_{7l})			-0.3640^{***}
			(0.0737)
medicine (γ_{7m})			0.0157
			(0.0724)
science (γ_{7s})			0.2613^{***}
			(0.0706)
corresp. member (γ_6)			0.9259^{***}
			(0.0746)
with genealogy (γ_5)			0.5304^{***}
			(0.0498)
Instit. FE. (γ_{8f})	N	N	Y
R^2	0.0051	0.0379	0.2432
Adj. R^2	0.0049	0.0372	0.2377
Num. obs.	5429	5394	5394
N Clusters	4380	4346	4346

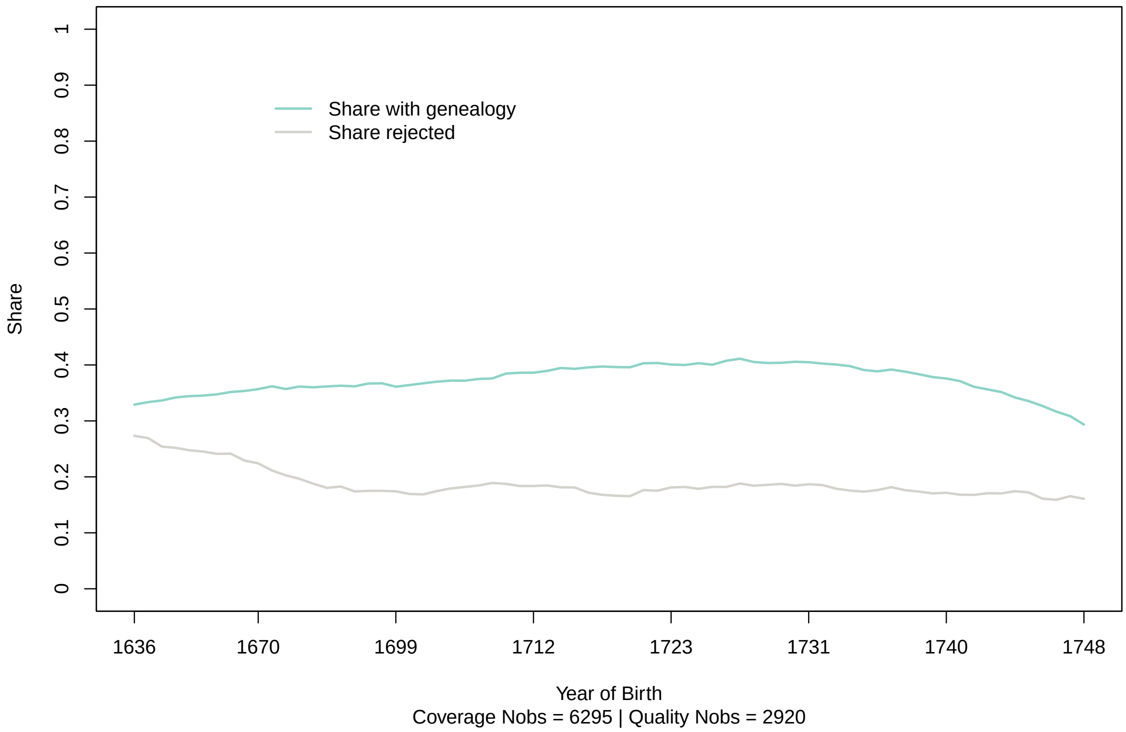
Dependent variable: inverse hyperbolic sine (asinh) of individual number of works

The unit of observation is a scholar-institution pair.

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

Table 8: Correlates of individual publications

G Genealogical coverage



Note: Share with genealogy computes the percentage of academics with a genealogical link among the 6295 recorded in our database. Share rejected computes among the 2920 academics with a valid genealogy, the proportion rejected because they have less than 3 siblings and an unrecorded date of death for the father.

Figure 15: Evolution of the genealogical coverage

Institutions	members	with genealogies	in %
University of Copenhagen	344	217	63
Royal Danish Science Society	155	109	70
Uppsala University	299	211	71
School Sorø	149	72	48
Royal Society of Sciences of Uppsala	98	75	77
Royal Swedish Academy of Sc.	427	296	69
University of Lund	263	154	59
Royal Physiographic Society	146	96	66
Åbo Akademi University	129	107	83
University of Tartu/Dorpat	55	31	56
Royal Norw. Soc. of Sciences and Letters	321	194	60
University of Groningen	103	47	46
Athenaeum Illustre of Amsterdam	74	24	32
University of Franeker	147	57	39
Royal Dutch Society of Sc.	364	155	43
University of Leiden	280	118	42
University of Utrecht	125	62	50
Société patriotique de Hesse-Homburg	143	66	46
University of Greifswald	361	81	22
University of Rostock	325	121	37
University of Kiel	217	47	22
Akademisches Gymnasium Danzig	90	22	24
Danzig Research Society	103	25	24
University of Königsberg	337	34	10
Academy of St Petersburg	304	139	46
University of Edinburgh	160	58	36
University of Glasgow	104	35	34
Academy of Edinburgh	396	193	49
University of Aberdeen (old)	198	34	17
University of Aberdeen (new)	107	21	20
University of Saint Andrews	88	25	28
TOTAL	6412	2936	46

Table 9: Genealogical coverage by institution

H Balance tests

Dependent variable: asinh number of works

quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
(Intercept)	1.57***	2.12***	3.63***	1.72***	1.39***	1.39***	2.72***	2.70***
	(0.20)	(0.18)	(0.33)	(0.36)	(0.18)	(0.18)	(0.17)	(0.17)
has genealogy	0.62***	0.70***	0.58***	0.74***	0.60***	0.62***	0.43***	0.46***
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)
Institut. FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.19	0.19	0.14	0.19	0.17	0.26	0.14	0.14
Num. obs.	1618	1618	1618	1619	1618	1618	1618	1619
N Clusters	1458	1524	1523	1514	1495	1508	1469	1408

Notes: Dependent variable is the inverse hyperbolic sine (asinh) of the number of works. Each column reports an OLS regression for a moving 30–percentile window of birth cohorts (0–30, 10–40, . . . , 70–100). Column 0–30 corresponds to the oldest 30% of academics; column 70–100 corresponds to the youngest 30%. This table shows 8 of 70 regressions. All specifications include institutional fixed effects; robust standard errors clustered at the person level are in parentheses. Significance levels: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

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)	64.66***	61.20***	68.35***	69.78***	64.46***	64.28***	53.84***	54.18***
	(1.73)	(1.84)	(3.14)	(2.58)	(1.58)	(1.58)	(1.89)	(1.90)
has genealogy	2.99***	3.13***	2.88***	4.05***	3.02***	3.53***	2.65***	2.16***
	(0.74)	(0.76)	(0.77)	(0.78)	(0.73)	(0.73)	(0.76)	(0.83)
Institut. FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.08	0.10	0.08	0.08	0.06	0.07	0.08	0.06
Num. obs.	1618	1618	1618	1619	1618	1618	1618	1619
N Clusters	1458	1524	1523	1514	1495	1508	1469	1408

Notes: Dependent variable is the mean age at death. Each column reports an OLS regression for a moving 30–percentile window of birth cohorts (0–30, 10–40, . . . , 70–100). Column 0–30 corresponds to the oldest 30% of academics; column 70–100 corresponds to the youngest 30%. This table shows 8 of 70 regressions. All specifications include institutional fixed effects; robust standard errors clustered at the person level are in parentheses. Significance levels: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

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)	0.14^{***}	0.01	0.59^{***}	0.41^{***}	0.14^{***}	0.14^{***}	0.20^{***}	0.20^{***}
	(0.04)	(0.02)	(0.14)	(0.10)	(0.04)	(0.04)	(0.05)	(0.05)
has genealogy	0.01	0.02	0.01	0.02	0.00	-0.02	-0.07^{***}	-0.07^{***}
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Institut. FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.10	0.09	0.11	0.09	0.16	0.24	0.22	0.17
Num. obs.	1618	1618	1618	1619	1618	1618	1618	1619
N Clusters	1458	1524	1523	1514	1495	1508	1469	1408

Notes: Dependent variable is whether the scholar works in the field of science. Each column reports an OLS regression for a moving 30–percentile window of birth cohorts (0–30, 10–40, . . . , 70–100). Column 0–30 corresponds to the oldest 30% of academics; column 70–100 corresponds to the youngest 30%. This table shows 8 of 70 regressions. All specifications include institutional fixed effects; robust standard errors clustered at the person level are in parentheses. Significance levels: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 12: Balance test: science

I Parental social background

See table next page

Social class	**Titles / occupations used to code the father's social class**
Upper class	archbishop, knighted, marquess, peer, prince, regent, aristocrat, theologian, linguist, king, title, seigneur, professor, councillor, bishop, mayor, doctor, rector, general, governor, lord, chief justice, colonel, captain, advocate, baron, physician, baronet, owner, lawyer, author, dr, prosecutor, lecturer, deputy, magistrate, curator, senator, advisor, king's, medical, scholar, writer, count, advocat, attorney, magister, procurator, commissar, businessman, physicist, marquis, engineer, manager, politician, landlord, artist, orientalist, entrepreneur, industrialist, duke, geographer, conferenceraad, biskop, anatomist, consultant, conventcommissary, noble, lieutenant-colonel, confessor, reichsmarschall, comminister, barrister, commandant, councilor, co-director, commonwealth marsh, prince-episcopal, humanist, magistrat, drtheology, philologist, pedagogue, notaire, banker, exchequer, hetman, geologist, adjutant-general, baroness, lagman, adlad, eversti, oberst, prepositus, logician, chatelain, cardinal, nobleman, prelate, cartographer, councillor-secretary, statesman, counsellor, vice-admiral, assessor, chancellor, attorney-postulant, investor, apothecary, judge, architect, antiquary, consul, sir, earl, knight, colonel, surgeon, laird, principal, historian, mathematician, theologist, admiral, commander, jurist, scholast, notary, dean, marshall, chief marshal, provost, president, chamberlain
Middle class	elder, thesaurier, printer, clergyman, inventeur, poet, trumpeter, painter, violinist, oboist, shipmaster, musician, alderman, administrator, monastic, auditor, reformer, preacher, priest, merchant, pastor, officer, trader, master, superintendent, vicar, secretary, clerk, reverend, minister, lieutenant, teacher, bailiff, baillif, chaplain, pharmacist, inspector, quartermaster, deacon, manufacturer, skipper, schoolmaster, postmaster, foreman, tutor, councilman, chef, shipowner, bookseller, treasurer, bourgeois, organist, civil servant, minister, lector, glazier, shipbuilder, surveyor, weighmaster, explorer, syndic us, commissioner, archivist, protocollist, churchman, controller, turner, librarian, registrar, scribe, greffier, jeweler, wholesaler, burgher, adjutant, marksman, nurse, baccalaureus, papermaker, salesman, cavalry, real estate agent, min. of old kirk, makler, clothier, kornett, adventurer, orfever, reverend, collector, brigadier, armourer, accounter, churchwarden, sheriff, farmer, hundertman, leaseholder, cornet, counsel
Lower class	minstrel, artisan, bierbrouwer, saddler, shoemaker, craftsman, policeman, soldier, cutler, factor, blacksmith, gunsmith, goldsmith, tobacconist, brewer, cossack, piece caster, shear, cartridge, tailor, fisherman, innkeeper, gardener, baker, grocer, tanner, carpenter, assistant, engraver, maker, cutter, receiver, horseman, employee, guard, tinsmith, comrade, weaver, coalminer, fishing, miller, slave, schmidt, coppersmith, janitor, beater, churchman, hunting, miner, shopkeeper, caretaker, shipwright, vigneron, guards, non-commissioned, driver, peasantry, fabrics, burner, military, sweeper, forester, grocery, farm, tractor, lorch, crimond
Unclassifiable	syndict, appellate, citizen, ministry, college, chancery, head, collegium, pharmacies, ordynek, consistory, seeker, writter, musices, council, government, school, trade, administrative, confidant, traveler, buyer, communicator, registar, commercial, advice, secretariat, priech, contractual, senevier, answerer, receiver, piece caster, shear, cartridge, grocery, farm, aids, crimond, hunting, non-commissioned, tractor, fabrics, military, lorch

Note: in case of multiple occupations or titles, the highest one has been kept to assign social class.

Table 13: Mapping between father's title / occupation keywords and assigned social class

J The urban-rural dimension

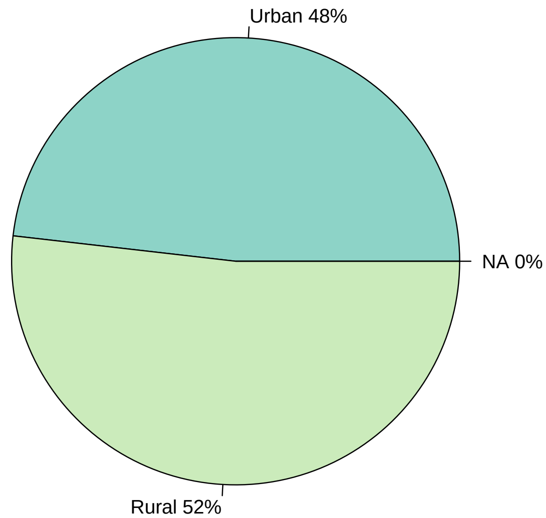


Figure 16: Pie Chart of Parental Origin

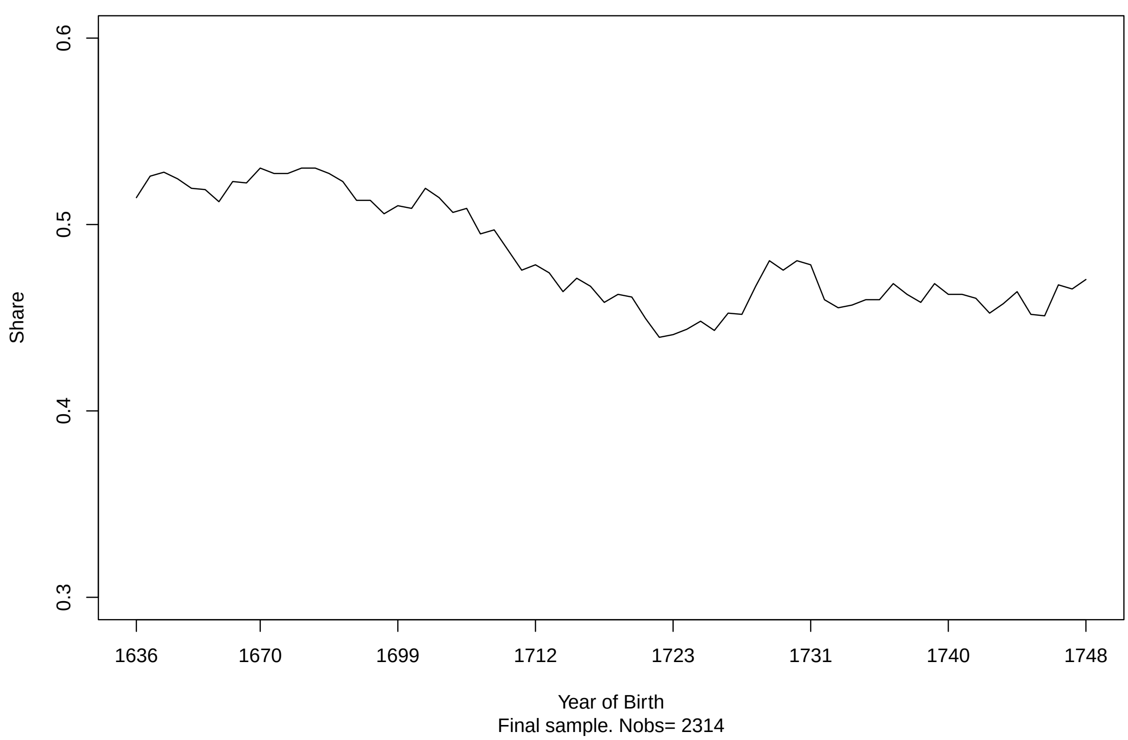


Figure 17: Share of scholars born in cities – moving 30–percentile window of birth cohorts

K The demography

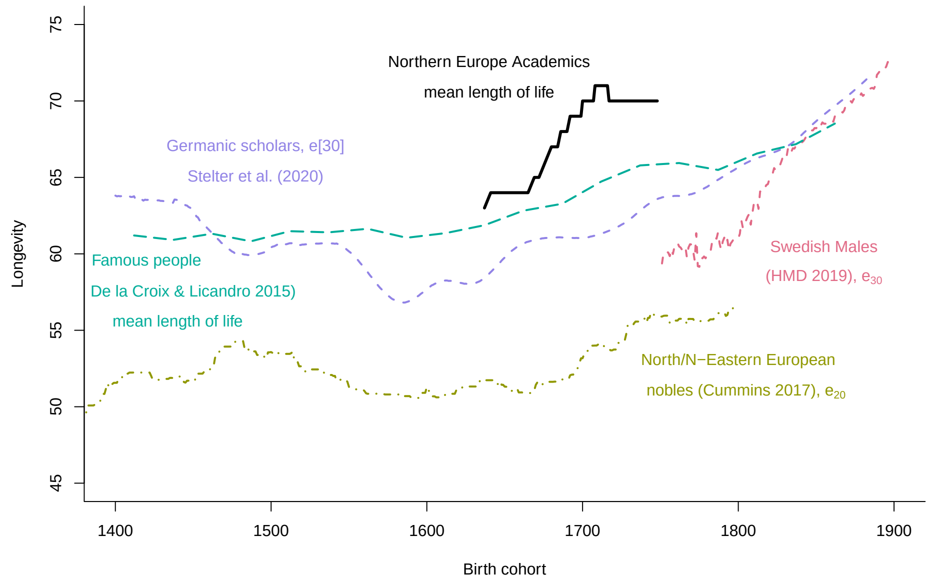


Figure 18: Longevity over time

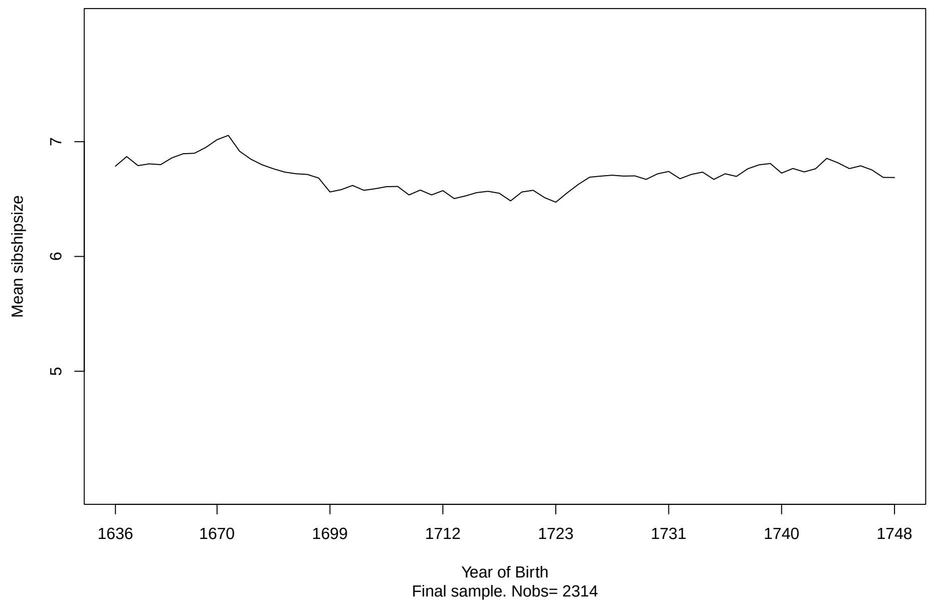


Figure 19: Fertility over time – moving 30–percentile window of birth cohorts

L Inter-generational correlation of fertility

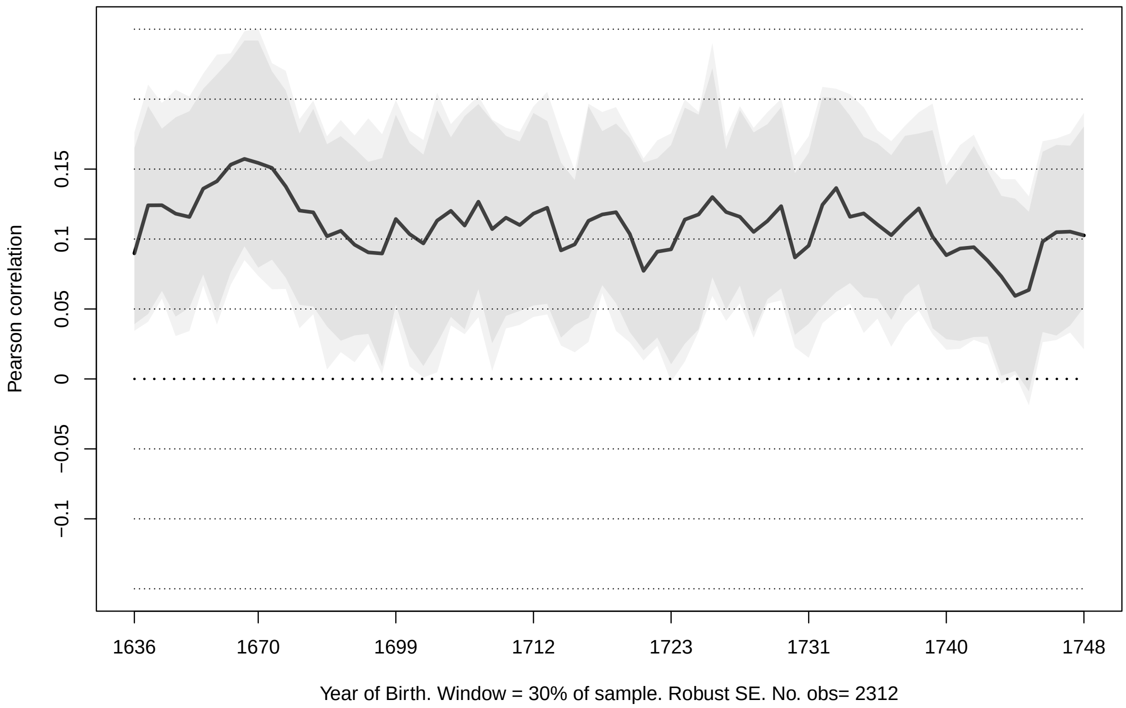


Figure 20: Inter-generational correlation of fertility – moving 30–percentile window of birth cohorts

M Broad academic fields

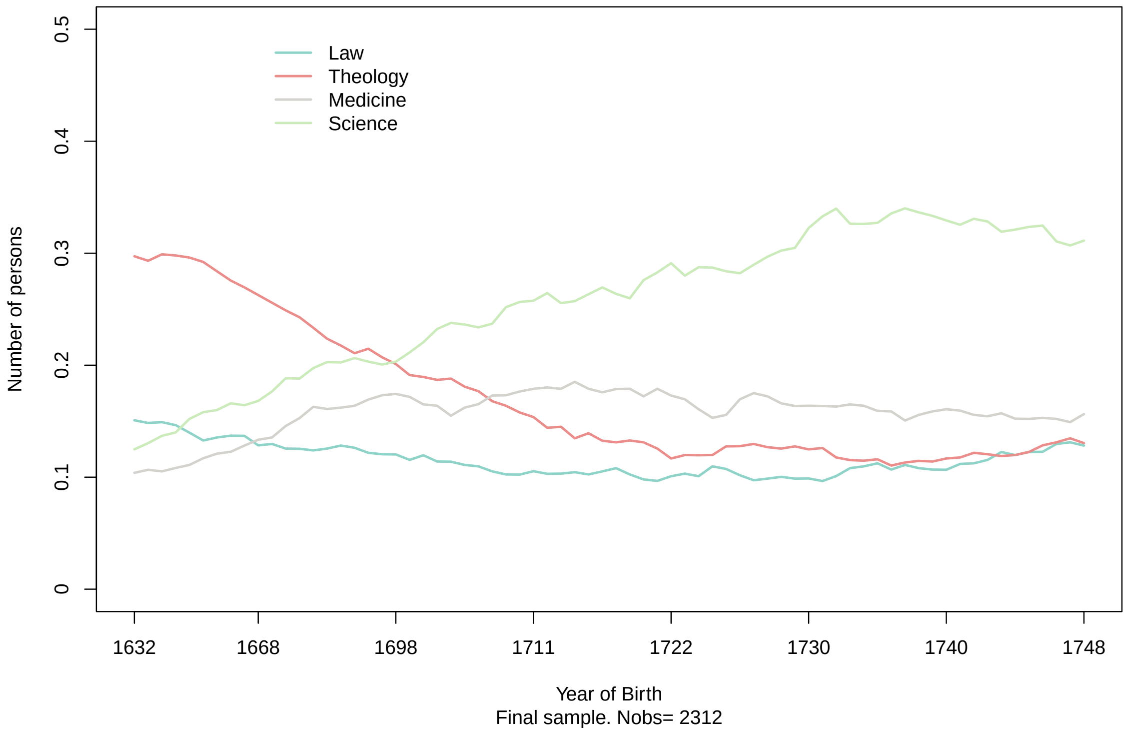


Figure 21: Share of academic fields over time – moving 30–percentile window of birth cohorts

N Descriptive statistics

	Mean	St.Dev.	Min	Max		Mean	St.Dev.	Min	Max
asinh(nworks)	2.922	1.622	0.000	6.669	leiden	0.031	0.174	0	1
longevity	66.139	13.187	20	100	ulund	0.061	0.240	0	1
age nomination	38.356	12.022	13	85	copenhagen	0.080	0.271	0	1
date of birth	1,698	58.749	1,435	1,779	acaddnk	0.042	0.201	0	1
geni	0.860	0.347	0	1	groningen	0.014	0.117	0	1
geneanet	0.050	0.218	0	1	uppsala	0.076	0.265	0	1
field theo	0.184	0.383	0	1	rostock	0.034	0.181	0	1
field law	0.126	0.330	0	1	kiel	0.014	0.117	0	1
field med	0.148	0.342	0	1	abo	0.040	0.196	0	1
field sci	0.240	0.412	0	1	dorpat	0.010	0.099	0	1
SibshipSize	5.390	3.513	1	22	amsterdam	0.006	0.078	0	1
No. sisters	2.050	2.066	0	14	Adanzig	0.009	0.093	0	1
HalfSiblings	0.220	0.414	0	1	edinburgh	0.019	0.138	0	1
No.Descendants	3.349	3.493	0	22	konigsberg	0.007	0.085	0	1
YearDeathFather	1,725	59.914	1,451	1,824	stockholm	0.110	0.313	0	1
RankMale	1.870	1.233	1	10	glasgow	0.009	0.095	0	1
Rank	2.678	2.201	1	15	greifswald	0.020	0.141	0	1
MaleDescendants	1.731	2.022	0	13	Gdanzig	0.006	0.080	0	1
SibshipSizeMale	3.340	2.094	1	16	petersburg	0.042	0.200	0	1
Urban	0.482	0.500	0	1	alund	0.036	0.187	0	1
elder	0.423	0.494	0	1	franeker	0.013	0.115	0	1
soc. class top	0.477	0.500	0	1	haarlem	0.051	0.221	0	1
soc. class mid	0.424	0.494	0	1	aedinburgh	0.068	0.252	0	1
soc. class bot	0.034	0.182	0	1	auppsala	0.030	0.170	0	1
soc. class na	0.065	0.246	0	1	aberdeen	0.018	0.132	0	1
fatherprof	0.128	0.335	0	1	andrews	0.009	0.093	0	1
WeakLink	0.118	0.322	0	1	hamburg	0.026	0.160	0	1
					trondheim	0.076	0.265	0	1
					utrecht	0.019	0.137	0	1
					soro	0.023	0.150	0	1

Note: N=2,312 observations except for YearDeathFather (2,195), Rank (1,576), RankMale (1,864), MaleDescendants (2,240). The variable "aberdeen" includes both Old and New universities.

Table 14: Descriptive statistics

O Detailed Regression results

See table next page

Table 15: Rolling regression results

	Dependent variable: asinh number of works							
quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
(Intercept)	1.363	0.271	0.822	0.714	2.632	3.164**	1.757*	2.844*
	(0.868)	(1.247)	(1.138)	(1.092)	(1.701)	(0.872)	(0.590)	(0.721)
SibshipSizeMale	0.053*	0.065**	−0.002	−0.034	−0.106***	−0.055	−0.065**	−0.056*
	(0.027)	(0.027)	(0.036)	(0.037)	(0.035)	(0.033)	(0.031)	(0.032)
longevity	0.033***	0.031***	0.029***	0.020***	0.012*	0.012**	0.014**	0.016***
	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)
ageentry	−0.013*	−0.010	−0.013**	−0.007	−0.010*	−0.009	−0.008	0.001
	(0.007)	(0.007)	(0.006)	(0.006)	(0.005)	(0.007)	(0.007)	(0.008)
law	−0.575***	−0.699***	−0.614***	−0.453*	−0.527**	−0.662***	−0.585**	−0.517**
	(0.187)	(0.206)	(0.219)	(0.249)	(0.246)	(0.241)	(0.227)	(0.231)
med	0.160	−0.191	−0.108	−0.061	0.249	0.354	0.418*	0.362
	(0.199)	(0.193)	(0.217)	(0.240)	(0.224)	(0.220)	(0.218)	(0.221)
theo	0.617***	0.246	0.148	0.007	−0.031	0.193	0.101	0.310
	(0.148)	(0.159)	(0.182)	(0.217)	(0.232)	(0.239)	(0.240)	(0.219)
sci	−0.166	−0.137	0.128	0.278	0.607***	0.493**	0.631***	0.400**
	(0.239)	(0.227)	(0.243)	(0.218)	(0.206)	(0.211)	(0.202)	(0.193)
WeakLinkTRUE	0.418	0.987***	0.964***	0.979***	1.000***	1.033***	0.929***	0.892***
	(0.446)	(0.293)	(0.240)	(0.187)	(0.142)	(0.143)	(0.151)	(0.173)
HalfSiblingsTRUE	−0.021	−0.177	−0.177	−0.112	0.287	0.281	0.284*	0.090
	(0.156)	(0.152)	(0.182)	(0.205)	(0.211)	(0.189)	(0.163)	(0.162)
Urban	0.189	0.319**	0.231	0.201	0.040	−0.090	−0.086	−0.136

quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
	(0.134)	(0.132)	(0.155)	(0.169)	(0.158)	(0.150)	(0.151)	(0.150)
fatherprof	−0.106	−0.162	−0.167	0.147	0.241	0.085	0.002	−0.070
	(0.211)	(0.217)	(0.244)	(0.261)	(0.300)	(0.294)	(0.285)	(0.244)
geni	0.147	0.249	0.413*	0.298	0.156	−0.040	0.053	0.347*
	(0.166)	(0.188)	(0.210)	(0.222)	(0.185)	(0.195)	(0.213)	(0.200)
soctop	0.118	0.016	0.365	0.227	0.567**	0.228	0.157	0.250
	(0.211)	(0.218)	(0.246)	(0.238)	(0.225)	(0.237)	(0.213)	(0.236)
socmid	0.058	0.085	0.603**	0.451*	0.653***	0.214	0.107	0.128
	(0.211)	(0.220)	(0.250)	(0.243)	(0.232)	(0.243)	(0.217)	(0.238)
copenhagen	−0.996	0.460	−0.172	0.865	−0.422	−0.488	0.495	−1.446
	(0.740)	(1.193)	(1.062)	(1.064)	(1.673)	(0.734)	(0.505)	(0.499)
groningen	0.356	1.560	0.978	1.767	0.783	0.490	1.450*	−0.585
	(0.772)	(1.233)	(1.162)	(1.074)	(1.674)	(0.709)	(0.564)	(0.476)
franeker	0.444	1.536	1.210	1.882	0.737	0.772	1.799*	−0.092
	(0.772)	(1.206)	(1.059)	(1.070)	(1.691)	(0.746)	(0.773)	(0.469)
alund			0.502	0.155	−1.047	−1.246	0.030	−1.695
			(1.210)	(1.051)	(1.679)	(0.734)	(0.506)	(0.487)
ulund	−1.280	−0.189	−0.640	−0.491	−1.943	−2.279*	−0.776	−2.231
	(0.778)	(1.206)	(1.065)	(1.069)	(1.677)	(0.761)	(0.523)	(0.487)
uppsala	−0.673	0.670	0.338	1.044	−0.183	−0.242	1.094	−0.776
	(0.731)	(1.193)	(1.055)	(1.070)	(1.679)	(0.737)	(0.527)	(0.492)
acaddnk	0.092	1.067	0.465	1.028	−0.104	−0.597	0.661	−1.289
	(0.987)	(1.233)	(1.063)	(1.030)	(1.674)	(0.730)	(0.505)	(0.467)
rostock	−0.877	0.067	−0.659	−0.092	−1.546	−1.765	−0.612	−2.512*

quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
	(0.751)	(1.229)	(1.103)	(1.152)	(1.746)	(0.892)	(0.600)	(0.564)
kiel	−0.451	0.777	−0.050	−1.337	−2.035	−2.170*	−0.594	−1.796
	(0.791)	(1.233)	(1.115)	(1.213)	(1.881)	(0.857)	(0.610)	(0.547)
abo	−0.996	−0.025	−0.752	0.044	−0.869	−1.018	0.182	−1.767
	(0.748)	(1.205)	(1.063)	(1.076)	(1.693)	(0.800)	(0.602)	(0.563)
dorpat	−0.394	0.773	0.065					
	(0.769)	(1.211)	(1.065)					
amsterdam	−0.093	1.019	0.691	1.958	0.378		0.668	
	(0.744)	(1.204)	(1.096)	(1.149)	(1.938)		(0.963)	
Adanzig			0.082		−1.092	−1.320	−0.079	−1.963
			(1.066)		(1.679)	(0.771)	(0.580)	(0.720)
Gdanzig	0.630	1.961	1.482	2.303*		−0.267		−1.223*
	(0.761)	(1.209)	(1.089)	(1.079)		(1.259)		(0.457)
edinburgh	−1.093	−0.086	−0.348	0.780	−0.506	−1.397	−0.480	−1.908
	(0.823)	(1.258)	(1.133)	(1.108)	(1.722)	(0.915)	(0.683)	(0.540)
konigsberg	−0.238			1.366	0.307	−0.039	0.648	−0.551
	(0.976)			(1.431)	(1.846)	(1.008)	(0.803)	(0.502)
utrecht	0.419	1.596	0.864	1.236	−0.186	−0.069	1.072	−0.311
	(0.744)	(1.196)	(1.058)	(1.107)	(1.701)	(0.755)	(0.588)	(0.574)
stockholm	−1.439	0.026	−0.233	0.502	−0.720	−0.980	0.231	−1.578
	(0.826)	(1.218)	(1.055)	(1.062)	(1.663)	(0.706)	(0.485)	(0.466)
glasgow	−2.401**	−1.224	−1.360	−0.244	−1.539	−1.883	−0.961	−2.318
	(0.932)	(1.385)	(1.164)	(1.185)	(1.757)	(1.150)	(1.682)	(1.518)
greifswald	−0.628	1.147	0.454	1.380	0.052	−0.604	0.427	−1.311

quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
	(0.777)	(1.212)	(1.074)	(1.121)	(1.708)	(0.787)	(0.791)	(0.692)
petersburg		0.920	0.411	1.179	−0.120	−0.629	0.230	−1.714
		(1.219)	(1.067)	(1.093)	(1.671)	(0.719)	(0.528)	(0.520)
haarlem		0.572	0.202	0.934	−0.225	−0.607	0.522	−1.221
		(1.272)	(1.076)	(1.074)	(1.668)	(0.688)	(0.489)	(0.481)
aedinburgh	−1.423	−0.386	−0.541	0.341	−0.675	−0.853	0.150	−1.863
	(1.231)	(1.275)	(1.079)	(1.084)	(1.676)	(0.715)	(0.506)	(0.466)
auppsala	−0.486	0.685	0.306	0.968	0.001	−0.239	1.205*	−0.800
	(0.734)	(1.222)	(1.065)	(1.086)	(1.670)	(0.726)	(0.522)	(0.573)
aberdeen	−2.360**	−1.316	−1.327	0.155	−1.163	−1.180	−0.288	−3.756**
	(0.798)	(1.274)	(1.173)	(1.259)	(1.803)	(1.154)	(1.166)	(0.483)
andrews	−0.890	−1.350	−2.931**	−1.827	−2.127	−2.345*	−0.865	−2.829*
	(0.895)	(1.449)	(1.117)	(1.127)	(1.676)	(1.053)	(0.625)	(0.662)
leiden	0.560	1.661	0.987	1.844	0.861	0.008	1.267	−0.619
	(0.733)	(1.208)	(1.081)	(1.091)	(1.691)	(0.830)	(0.616)	(0.579)
hamburg		0.149	0.516	0.334	−1.004	−1.532	−0.362	−2.355
		(1.255)	(1.158)	(1.095)	(1.674)	(0.737)	(0.541)	(0.672)
trondheim		0.745	0.147	0.555	−0.769	−1.139	0.040	−1.954
		(1.297)	(1.106)	(1.035)	(1.675)	(0.724)	(0.495)	(0.468)
soro	−2.784***	−2.060	−2.496*	0.348	−0.539	−0.405	−0.535	−2.557*
	(0.805)	(1.224)	(1.116)	(1.204)	(1.747)	(0.823)	(0.674)	(0.598)
R^2	0.324	0.340	0.301	0.240	0.278	0.273	0.276	0.252
Adj. R^2	0.285	0.299	0.255	0.191	0.231	0.226	0.230	0.203

quantiles	0-30	10-40	20-50	30-60	40-70	50-80	60-90	70-100
Num. obs.	693	693	694	693	694	693	694	694
RMSE	1.396	1.377	1.422	1.452	1.392	1.372	1.386	1.435
N Clusters	617	574	507	460	443	439	460	496

Notes: Dependent variable is the inverse hyperbolic sine of the number of works. Each column reports an OLS regression for a moving 30–percentile window of birth cohorts (0–30, 10–40, ..., 70–100). Column 0–30 corresponds to the oldest 30% of academics; column 70–100 corresponds to the youngest 30%. This table shows 8 of 70 regressions. All specifications include institutional fixed effects; robust standard errors clustered at the person level are in parentheses. Significance levels: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

P Robustness checks

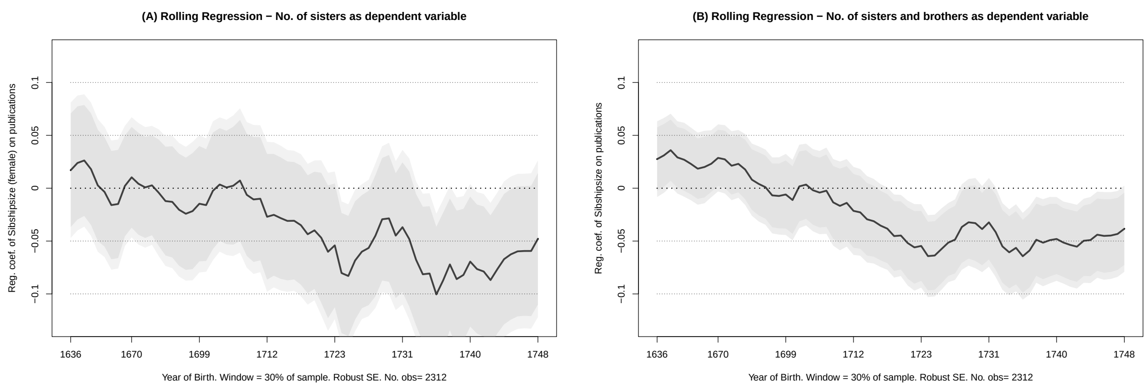


Figure 22: Rolling regression including girls (left: no. of sisters, right: all siblings)

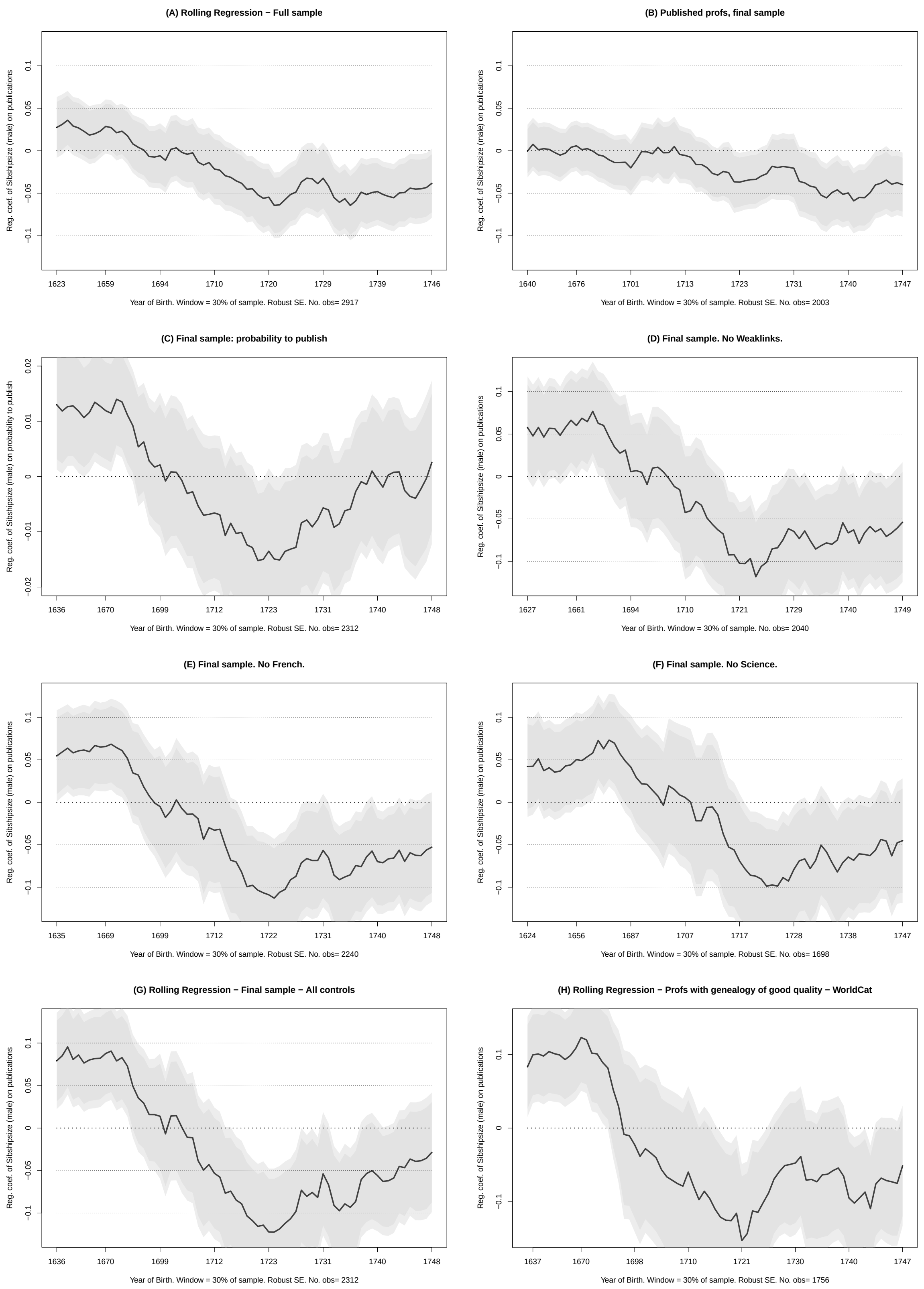


Figure 23: Rolling regression for alternative specification

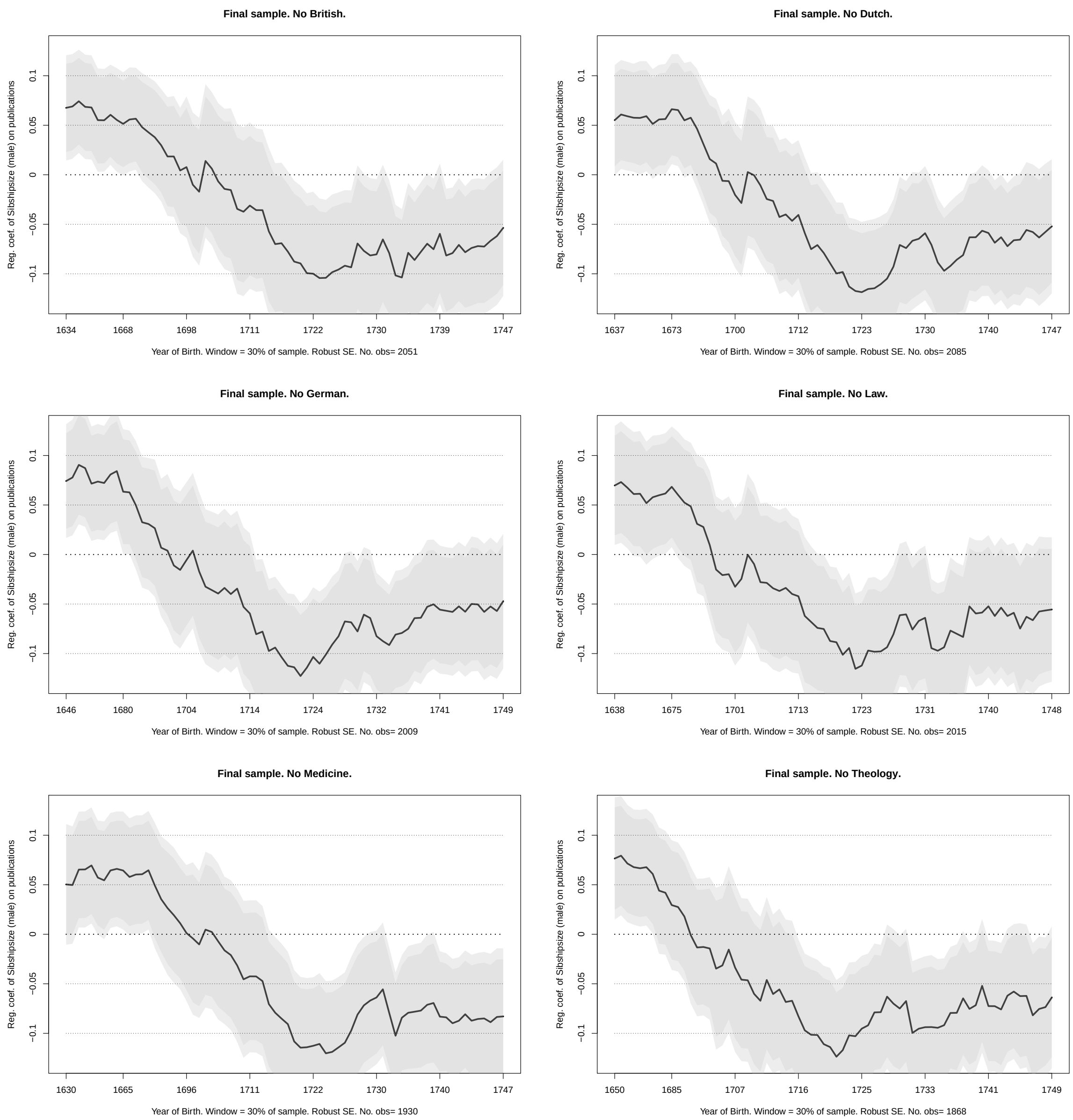


Figure 24: Rolling regression for additional robustness checks

Q 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:

$$\frac{\partial h_t^i}{\partial \eta^i} = t\left(\frac{\psi\phi\eta^i}{\gamma-\eta^i}\right)^{t-1}\frac{\psi\phi\gamma}{(\gamma-\eta^i)^2}h_0 > 0. \tag{7}$$

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 5, we get:

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

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:

$$\frac{dn_t^i}{d\eta^i} = -\frac{1}{\gamma\phi h_t^i}\left[a - \bar{c} + h_t^i + (\gamma-\eta^i)(a-\bar{c})\frac{\frac{\partial h_t^i}{\partial \eta^i}}{h_t^i}\right] \tag{8}$$

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:

$$\frac{dn_t^i}{d\eta^i} \geq 0 \iff h_0\left(\frac{\psi\phi\eta^i}{\gamma-\eta^i}\right)^t + a - \bar{c} + (a-\bar{c})\frac{\gamma}{\eta^i}t \leq 0 \tag{9}$$

Eq. 9 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 $W(.)$ 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 $h_0 > \frac{1}{e}[\frac{(\bar{c}-a)\gamma}{\eta^i(\frac{\psi\phi\eta^i}{\gamma-\eta^i})^{-\frac{\eta^i}{\gamma}}\ln(\frac{\psi\phi\eta^i}{\gamma-\eta^i})}]$, $\Delta \ln \lambda < -\frac{1}{e}$ and Inequation 9 is never satisfied. An easy way to check this is to inspect Equation 9 for $h_0 \to +\infty$. Conversely, if $h_0 < \frac{1}{e}[\frac{(\bar{c}-a)\gamma}{\eta^i(\frac{\psi\phi\eta^i}{\gamma-\eta^i})^{-\frac{\eta^i}{\gamma}}\ln(\frac{\psi\phi\eta^i}{\gamma-\eta^i})}]$, $\frac{dn^i}{d\eta^i} = 0$ admits two solutions:

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

The properties of the Lambert W-function imply that $t_1^i > t_0^i$ and that $\frac{dn^i}{d\eta^i} > 0$ only when $t \in]t_0^i, t_1^i[$. This result can be easily visualized by re-arranging Equation 9 and using logs, which yields to the following condition:

$$\frac{\partial n_t^i}{\partial \eta^i} \geq 0 \quad \Leftrightarrow \quad LHS(\eta^i) \equiv \ln\frac{\bar{c} - a}{h_0} + \ln\left(1 + \frac{\gamma}{\eta^i}t\right) \geq t \ln\frac{\psi\phi\eta^i}{\gamma - \eta^i} \equiv RHS(\eta^i) \tag{10}$$

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 $\bar{h}^i$. 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

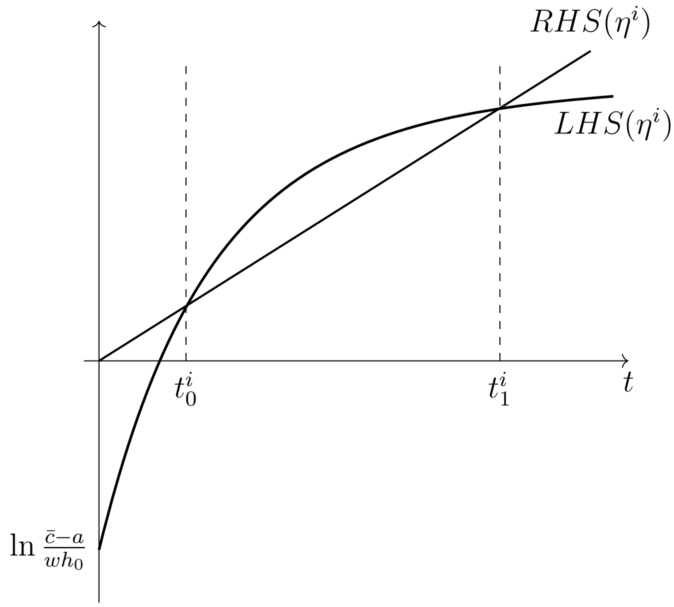


Figure 25: Income expansion path of quality and quantity.

express Equation 8 for $t = \bar{t}^i$ and $h_t = \bar{h}$. Doing this yields to the following condition:

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

Proposition 1 then follows as for any

$$h_0 < \min\left\{ e^{\ln((1+\gamma)\bar{c}-a) - \frac{\bar{c}\eta^i}{\bar{c}-a}\ln\left(\frac{\psi\phi\eta^i}{\gamma-\eta^i}\right)}, \frac{1}{e}\left[\frac{(\bar{c}-a)\gamma}{\eta^i(\frac{\psi\phi\eta^i}{\gamma-\eta^i})^{-\frac{\eta^i}{\gamma}}\ln(\frac{\psi\phi\eta^i}{\gamma-\eta^i})}\right]\right\},$$

t_0^i and t_1^i exists and $\bar{t}^i \in]t_0^i, t_1^i[$. ■

R Proof of Proposition 2

Proposition 2 states that we can identify periods $(\breve{t}, \bar{t})$ 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 25 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 $t_{0,MAX}^i \equiv \breve{t}$ and Proposition 2 directly follows. ■