Seeds of Knowledge: Premodern Scholarship, Academic Fields, and European Growth

Matthew Curtis[a] and David de la Croix[b]

[a]Syddansk Universitet, Denmark
[b]IRES/LIDAM, UCLouvain, Belgium & CEPR, Paris

March 21, 2024

Abstract

Human capital is an engine of modern economic growth. Using a novel database of premodern European academics (1000–1800), we find that scholarship also fostered growth in the past. Combining secondary sources on the history of academia with data from worldwide library catalogs, we measure both the quantity and productivity of scholars. We find that a 10% increase in scholarly output was associated with 1.14 % higher income per capita in the region of the scholars' birth in 1900. We use machine learning to group scholars into academic fields. Income per capita was positively associated with the share of scholars in science (including mathematics, physics, and astronomy), botany, and one type of theology (centered around the Bible), but negatively with the share in law. Only the share of science and botany matter once we look within countries. Exogenous variations in local scholarship caused by forced migration suggests endogeneity biases are small. Finally, we propose a mechanism consistent with the empirical evidence: scholars inspire their fellow natives to accumulate human capital.

JEL codes: N33, O47, I23

1 Introduction

The accumulation of knowledge is a crucial factor in economic development. It helps explain the West's prosperity and disparities in income among countries (Mokyr 2002, 2016; Galor 2022). Europe has had a strong tradition of scholarship, reaching back to the earliest Medieval Universities. However, the link between historical scholarship and economic development is unclear. First, it is hard to quantify academic knowledge, as it covers diverse subjects and can be embedded in various forms. Second, it is unclear what types of formal scholarship would eventually lead to practical applications in the economy.

In this study, we propose a novel method to quantify the knowledge that developed in premodern academia. We find evidence of a positive relationship between regional GDP per capita in 1900 and the birthplaces of academic scholars (university professors and members of academies of sciences and arts) during the period of 1000–1800. Next, to unravel which types of knowledge was more conducive to economic development, we group scholars in different clusters using a machine learning algorithm, with each cluster representing one academic field. We find a particularly strong association between growth and the field related to mathematics and physical sciences and the field related to botany and life sciences. Then, to mitigate endogeneity concerns, we develop an instrumental variable strategy based on three distinct historical episodes of forced migration. This strategy suggests that endogeneity is not a major concern. Finally, we investigate the underlying mechanisms linking past scholarship to regional income, hypothesizing that scholars serve as catalysts for inspiration among successive generations, fostering the pursuit and application of knowledge.

Our dataset contains sixty thousand scholars compiled from five hundred secondary sources on the members of universities and academies. To measure the productivity of these scholars, we count every work and edition attributed to them in WorldCat. This approach is complementary to that of De Courson, Thouzeau, and Baumard (2023), who use Wikipedia as both the index of individuals and the measure of output. The important difference is in which individuals are assigned a measure of productivity. Our sample is both more exclusive, in that it only considers members of academia, and more inclusive, in that we are not selecting based on retroactive notability. Our approach is also complementary to that of Johnson, Thomas, and Taylor (2023), who use texts as a measure of the local adoption of printing presses. We focus on the locations of authors, not publishers, and thus measure the human capital of scholars instead of the physical capital of printing presses.

While our scholars database ends in 1800, estimates of income are too sparse before 1900 to compare all regions where scholars were born. Thus we focus on outcomes in 1900, using estimates of GDP per capita for the contemporary NUTS2 regions from the Rosés-Wolf database on regional GDP (Rosés and Wolf 2021). One advantage of this approach is we are looking at gains from both the First and Second Industrial Revolutions. While during the First Industrial Revolution there was a major role for the human capital and idiosyncrasies of craftsmen and inventors, the Second Industrial Revolution saw a more direct pipeline between scientific knowledge, applied innovations, and an educated workforce (Cinnirella and Streb 2017).

Beyond measuring the quantity of academic knowledge in general, we also consider the types of knowledge produced. It has been argued that specific types of knowledge were important for economic growth. For example, scientific knowledge pushed the envelope of propositional knowledge, leading to future economic applications (Wootton 2015; Danna 2022). Academic knowledge contributed to building better political and economic institutions as far back as the Middle Ages (Mitterauer 2010). Theologians promoted nuclear family structures (Henrich 2020; Schulz 2022), held beliefs compatible with the spirit of capitalism (Weber 1930), and encouraged education to read the Bible (Becker and Woessmann 2009). Lawyers developed Roman and civil law encouraging trade (Cantoni and Yuchtman 2014), and physicians laid the ground for advances in botany (Hill 1915). Our fields of study are based on a list of subjects associated with the works by or about the author from the WorldCat Identities database. The subjects are based on the FAST subject terminology schema developed by OCLC (the organization that develops WorldCat) and the Library of Congress. Using these subjects, we use an unsupervised machine learning algorithm — k-means clustering — to assign each author to a cluster. This approach is similar in spirit to Grajzl and Murrell (2023), Almelhem et al. (2023), and Koschnick (2023). Their algorithms classify texts into topics. In our work, we classify scholars into fields based on their associated topics. Our paper also shares an interest in how types of knowledge matter for economic growth. Our approach is complementary, as we focus on the production of knowledge in academia and look at impacts over European regions.

To mitigate endogeneity concerns arising from omitted variables, this study identifies refugee scholars originating from the expulsion of the Huguenots from France, the flight of British Catholics post-Reformation, and the escape of Byzantine scholars following the fall of Constantinople. By instrumenting the aggregated output of scholars from a region with an indicator reflecting the presence of a refugee scholar, our analysis indicates that the potential for endogeneity bias is minimal. These migrants belonged to the intellectual elite, and we also know that forced migrants shift their investment towards human capital because it is portable (Becker et al. 2020). In each case, scholars appeared to have resettled based on physical proximity, not factors strongly correlated with economic growth. As we show below in an event study, refugees did not select their destination based on previous regional levels of scholarship. This instrumental variable approach suggests endogeneity is not a major concern.

Finally, we delve into the underlying mechanisms that establish a connection between past scholarship and regional income. The discovery of a correlation at the local level is intriguing, as one might expect the impact of figures like Galileo to transcend their birthplace and influence the entirety of Europe. Our hypothesis posits that scholars serve as catalysts for inspiration among successive generations, fostering the pursuit and application of knowledge. Unlike other historical studies which focused on how teachers inspire students (Borowiecki 2022; Koschnick 2023), our paper considers inspiration taking place at the birth location. Our approach gains support from three compelling observations. Firstly, regional development is better explained when scholars are aggregated by birthplace rather than their activity location. Secondly, scholars who remain in proximity to their origins exhibit a more pronounced association with development compared to those who relocate or meet an untimely demise. Thirdly, areas with a greater number of scholars born had higher general population numeracy in the late 19th century, even after controlling for GDP per capita.

Additionally, we bolster our argument with numerous anecdotal instances, drawing on monuments and street names that underscore the existence of a role model effect.

2 Scholars, Universities and Academies

Medieval universities concentrated on four main fields: theology, law, arts and humanities, and medicine. Their impact on society is well described by Pedersen (Pedersen 1992): "The faculty of arts gave a basic education to grammar school boys, many of whom would become teachers themselves and contribute to the increase in literacy of the population at large. Others would go on to one of the higher faculties to prepare themselves for other professions. The faculty of medicine produced medical practitioners; the faculty of laws created future administrators with expert knowledge in canon or civil law, and the faculty of theology provided teachers for the episcopal schools, were the ordinary parish priests were educated." Academies were usually created later, in the 17th-18th century, responding to a need of developing new fields of research which were not traditionally taught at universities. The academies range from clubs of amateur naturalists or local historians to eminent societies, gathering the best scholars, publishing journals, and building a network of corresponding members, called the Republic of Letters (McClellan 1985; Mokyr 2016).

The full database of scholars we built and used contains information on 60,145 scholars who were appointed to universities or were nominated to academies over the period 1000–1800. The data were harvested manually from 535 different sources. We took the list of universities from Frijhoff (1996) and the list of academies from McClellan (1985), and added to this the language academies, the most important Italian Renaissance academies from British Library (2021), and several other higher education institutions which conferred academic degrees.

To assemble the list of scholars from each academy and university we use secondary sources, i.e. books on the history of institutions and their members based on primary sources. For academies, this task is usually straightforward, as comprehensive lists of members are often available (our data on academies have already been used in Blasutto and De la Croix (2023) for Italian academies and in De la Croix and Goñi (2020) for father-son pairs in academies and universities). These lists encompass various membership statuses, including ordinary members, corresponding members, and honorary members. Corresponding members are individuals who do not attend academy meetings but contribute to its work from a distance. Honorary members typically include local authorities such as bishops, wealthy merchants, and governors, who provide support and protection to the academy. To ensure that our results are not influenced by publications concerning these sometimes prominent figures, we exclude anyone with honorary status or individuals who are clearly not scholars or intellectuals (such as Napoleon, who was elected to the Académie des Sciences in 1797).

For universities, our goal is to include scholars who have participated in teaching in some capacity. This encompasses various positions, from royal chairs in France to fellowships in England. More detailed information on the criteria for including university scholars in our database can be found in De la Croix et al. (2023), while additional global statistics are provided in De la Croix (2021) and in various issues of the Repertorium Eruditorum Totius Europae.

The resulting database is accessible at https://shiny-lidam.sipr.ucl.ac.be/scholars/. Notably, the gender distribution heavily favors male scholars, with only 108 women among 58,995 scholars (De la Croix and Vitale 2023).

To assign a measure of productivity to each scholar, we use the Worldcat search engine which provides references to the collections of thousands of libraries around the world. To measure the quality of each author, we count the number of publications by the author. This measure thus cover both output of the scholar and impact, as it includes new editions, translations etc. Worldcat provides a good approximation of the population of known European authors, for example, Chaney (2020) compares the Universal Short Title Catalogue (St. Andrews 2019) to the references in the Virtual International Authority File (VIAF), on which WorldCat is based. Chaney successfully locates 81% of USTC authors in the VIAF. Hnence scholars with missing Worldcat publications were likely unproductive.

Finally, we use secondary sources to document each scholar's academic field. For university professors, this corresponds to their area of instruction, while for academicians, it is derived from the descriptions provided in the sources. We categorize scholars into the following broad fields: lawyers, physicians, theologians, scientists, applied scientists, and arts and humanities scholars. These classifications align with the traditional higher faculties of early universities, with the addition of the arts faculty, where the prominence of scientists grew over time. While these collected fields provide valuable insights on average, they are not without imperfections. For instance, for some contexts it may have made sense to include more nuanced distinctions, such as differentiating between canon law and civil law within the field of law. Additionally, some scholars' teaching roles may not fully represent their expertise; for example, Gassendi, who taught theology at Aix-en-Provence but was renowned for his contributions to astronomy.

In the case of academicians, we assigned the field of "law to members of courts, such as the regional French Parliaments, which may not always accurately reflect their actual skills. A similar issue arises with individuals associated with the Church, where we may have hastily attributed the field of "theology." This critique extends to Protestant countries as well; for instance, many fellows in Oxford hold a D.D. (Doctor of Divinity) but may have pursued teaching and research in other fields, as evidenced by Gunther's work on "Early Science in Oxford - Oxford Colleges and their Men of Sciences" Gunther 1937.

3 Identifying Academic Fields

For each scholar with a WorldCat Identities page,[1] we collected the tag cloud of their "Associated Subjects" (excluding the persons who are honorary members). We then drop subjects associated with fewer than 30 scholars or that are about a specific country (e.g. "French history"). This leaves us with 1,360 subjects and 16,149 scholars with at least one subject.

WorldCat assigns each subject a relative importance. We quantify the importance of a subject from 1–5. Thus, for each scholar i and subject j, we have weights $\gamma_{ij} \in \{0, 1, 2, 3, 4, 5\}$. We then construct a data matrix Γ of dimensions $1,360 \times 16,149$ containing every γ_{ij}. Each row is an academic, each column a subject.

We use the k-means algorithm which treats each row of Γ as the coordinates point in a $1,360$-dimensional space. It partitions the data into k clusters, minimizing the total within-cluster sum of squared deviations (TWCSS). This is the sum of squared deviation of each

1. Sadly, the 2 million pages of the WorldCat Identities project were suddenly retired in March, 2023. This is bad news for those interested in measuring human capital from publications data. For the future however, we found a viable alternative using statistics drawn from the VIAF platform. See Curtis and De la Croix (2023) for more details.

point from the centroid of its cluster.

k-means must be estimated using numerical methods as there is no closed-form solution. We use the default R package which implements the Hartigan-Wong algorithm (Hartigan and Wong 1979). This algorithm starts with random guesses for the centroids of each cluster and then iteratively improves the centroids until a certain convergence threshold is reached. As the improvements converge to a local optimum, not a global optimum, we repeat the estimation 500 times, picking the replication with the lowest TWCSS.

The choice of k can be made using various criteria. We minimize the Bayesian information criterion (BIC): $TWSS_k + \log(I)Jk$, where $I = 16,149$ and $J = 1,360$. This is minimized at $k = 10$. More details are in Appendix A.5. Ten clusters are thus the most informative yet parsimonious way to describe academic fields.

Table 1 presents the ten clusters. The first column contains a description we chose to represent the various subjects included in the cluster. Column 2 gives the total number of published scholars in each cluster. One cluster is much bigger than the others; it appears to contain both humanists, classicists and scholars who were unrelated to any other cluster. The smallest cluster is Botany, with 543 persons.

To better grasp the nature of each cluster, we show in Column 3 the names of the scholars belonging to the cluster who published the most. Column 4 gives the median number of publications of scholars in each cluster. Theology 2 leads and Classics lags. Column 5 shows the date of activity of the earliest scholar in each cluster. It shows that all ten clusters started before 1200, thus having deep roots in the Middle Ages. The last column shows the median year of activity in the cluster. Law is the cluster with the earliest median date, while Politics is the cluster with the most recent median date.

The clusters are further explored in the Appendix. The most important topics and scholars by cluster are described in Appendix A.1. Most clusters are strongly associated with a few key terms, however the Classics cluster is not. Classics contains authors who write on many diverse topics, perhaps related to the Humanistic Revolution. Appendix A.2 plots the shares of scholars by cluster over time. In Appendix A.3, we provide ten graphics with names of published scholars over time by cluster, allowing to see through whom each field has medieval roots.

Theology is the only field to have two clusters (see the maps in Appendix A.4). The division between Theology 1 and Theology 2 is related to the Catholic-Protestant divide, but is not a simple denominational split. In Theology 1, we find some leading figures of Catholicism such as Aquinas (professor at University of Paris 1252–72 and Naples 1272–4), Bossuet (member of Académie Française 1671–1704), and Robert Bellarmin (professor at the Gregorian University in Rome 1576–1593) but also some unorthodox catholics such as Pascal

Table 1: Clusters of WorldCat Topics

Cluster / Field	N. Scholars	Top 3 Names	Median N. Publ.	Earliest Year	Median Year
Theology 1	1581	Aquinas, Bossuet, Pascal	143	975	1615
Theology 2	940	Luther, Melanchthon, Wesley	315	1039	1671
Politics	990	Swift, Machiavelli, Corneille	184	1043	1756
Law	727	Stryk, Bentham, Bohmer	156	1090	1593
Science	661	Newton, Euler, Galilei	177	1116	1714
Classics	7317	Schiller, Erasmus, Pope	54	970	1712
Philosophy	653	Rousseau, Kant, Diderot	258	980	1700
Botany	543	Linnaeus, Bernardin, Trew	189	1176	1753
Culture	1086	Arouet, Humboldt, Homman	211	1140	1749
Medicine	1651	Haller, Hohenheim, Gessner	125	1025	1698

Note: Clusters estimated by k-means clustering. Top 3 Names are the top three scholars assigned to a cluster based on their number of publications.

(member the Mersenne academy of c. 1639, close to Jansenism, a controversial Catholic movement with similarities to Calvinism) and some important Protestant figures such as Gilbert Burnet (professor of Divinity at the University of Glasgow 1669–74, and member of the Royal Society). Theology 2 is led by the main figures of Protestantism, such as Luther (professor at University of Wittenberg 1508–46), Melanchthon (professor at University of Tübingen 1512–18 and Wittenberg 1518–60), John Wesley (fellow of Lincoln College at University of Oxford 1725–7), and Jean Calvin (professor at the University of Geneva 1541–64). But it also includes medieval (Catholic) theologians such as Hugues de Saint-Victor (University of Paris 1133–41). Looking at the subjects with the highest frequency in both clusters, we find "Catholic Church" and "Clergy" in Theology 1, and "Bible" in Theology 2.

Scientific fields are split in three clusters. The cluster Sciences includes the subjects "Mathematics", "Astronomy", "Geometry", "Physics", led by Newton (professor at University of Cambridge 1661–1696, member of several academies), Euler (professor at University of St Petersburg 1727–41, member of several academies), and Galilei (professor at University of Pisa 1589–92 and Padua 1592–1610). The cluster Botany includes the subjects "Plants" and "Natural History", and is led by Linnaeus (professor at University of Uppsala 1742–78, and member of many academies). The cluster Medicine includes subjects "Human anatomy" and "Surgery". Together with the clusters on Politics, Law, and Philosophy, the clustering procedures seems to lead to a very coherent set of academic fields. Only Classics and Culture have vague boundaries. We are thus confident interpreting these clusters as academic fields.

4 Academic knowledge and regional development

We now analyze whether academic knowledge is associated with historical development at the subnational level. This allows us to determine if scholarship matters both at a local and a national level. We interpret a higher GDP per capita in 1900 as evidence of economic growth. Before 1800, GDP per capita was restricted by the Malthusian trap, albeit with some geographic and temporal variation. Moreover, we control for log ruggedness from Nunn and Puga (2012), the log post-1500 caloric suitability index of land from Galor and Özak (2015, 2016) and Galor, Özak, and Sarid (2017), and the log area of the region in km^2. With these controls, and given the low initial levels of development, we interpret a higher GDP in 1900 as evidence of stronger 19th century economic growth.

Figure 1 shows the geographical area we cover with the NUTS2 regions. The map's background color for each region reflects its GDP per capita in 1900, with darker shades indicating higher levels. Color dots indicate the place of birth of scholars belonging to two example fields. Red dots correspond with scholars belonging to the field of Law, blue dots with scholars belong to the field of Science.

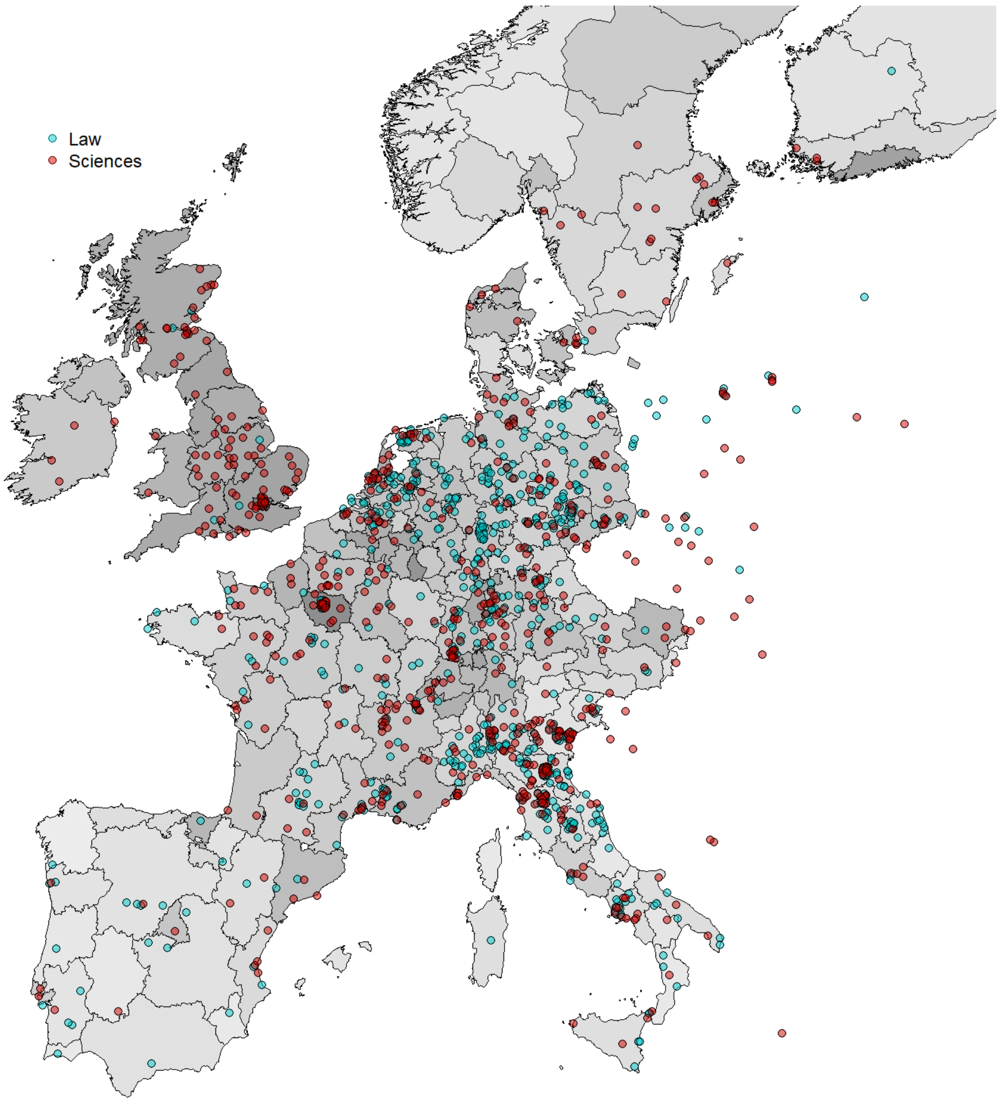
When summing over scholars at the regional level, we weight each scholar by a function of the number of his publications. The number of publications, which includes multiple editions and translations, ranges from 1 to 111,660 (Martin Luther). Given these values, it is not reasonable to weight scholars by their number of publications. It would imply that Luther worth a hundred thousand obscure theologians (those with only one publication). If, instead of the number of publications, we take its square root, Luther would be worth 334 obscure theologians. If we take the fourth root of the number of publications, Luther would be worth 18 obscure theologians. Galileo would be worth 9 mathematicians with one publication. We adopt this last formula, which gives a weight from 1 to 18 to each scholar. (In Appendix B.1, we show that this choice of weighting does not drive our results.)

We estimate the following regression model:

$$y_{r,s} = \alpha_0 + \alpha_1 \log(n_{r,s}) + \sum_{c=1}^{10} \beta_c share^c_{r,s} + \beta X_{r,s} + \phi_s + \epsilon_{r,s} \tag{1}$$

where $n_{r,s}$ is the weighted sum of published scholars born in r from 1000–1800, weighted by their number of publications raised to the power of 0.25; c is one of the ten fields identified

Figure 1: Map of Birthplace of Law and Science Scholars



Note: Every scholar is assigned a field and a birth NUTS 2 region. GDP per capita from the Rosés-Wolf database on regional GDP Rosés and Wolf 2021.

by the K-means algorithm; $share^c_{r,s}$ is the share of $n_{r,s}$ that belong to field c; $X_{r,s}$ is a vector of controls, ϕ_s is a country fixed effect, and $\varepsilon_{r,s}$ is an error term.

In our main set of regressions, scholars are allocated to their region of birth. We estimate a second set in which scholars are allocated to their region of activity (see Appendix B.3). We find that the first set gives stronger results. This suggests the presence of mechanisms beyond the mere effect of universities and academies on the region in which they are located. Moreover, in Appendix B.2, we show that these results are robust to the inclusion of controls for the location of such academic institutions.

Table 2 presents the results. As shown in the first line, we find an overall association between $\log(n_{r,s})$ before 1800 (the weighted sum of published scholars) and GDP per capita in 1900. A ten percent increase in the weighted sum of scholars born between 1000 and 1800 in region i is associated with a 1.14 percent increase in GDP per capita in 1900, all else equal. The estimate is reduced in magnitude but still significant after adding country fixed effects (second line). This shows that human capital in the past is associated with future growth. Below, we argue that this is likely a causal effect. Regardless of the exact mechanism, our findings lend credence to theoretical frameworks in which human capital plays a role in development.

When we additionally look at the shares of the different fields (Figure 2), we find that the fields Theology 2, Science, and Botany have a positive association with growth. The field Law has a negative association. In that regression, the reference category is the share of scholar in Classics. All else equal, a ten percentage point increase in the number of scholars that are in the fields of Law, Science, and Botany (at the expense of Classics) are associated with a -7.8 % , 11.43 %, and 7.93 % percent change in GDP per capita in 1900.

We also still estimate the impact of the total number of scholars (third line of Figure 2), with a 10 percent increase in the weighted sum of scholars born between 1000 and 1800 in region i being associated with on average a 0.97 percent increase in GDP per capita in 1900, *ceteris paribus*. Therefore, the coefficients for the shares are estimating the *additional* impact from specialization in a field compared to the others. Scholarship, regardless of field, is associated with higher GDP per capita.

Law appear to vary substantially across countries (using contemporary boundaries). Law is rare in common law Britain (Figure 1). As common law are commonly studied as determinants of growth, the associations we find might be related to more broad factors relating to legal systems (Porta, Lopez-de-Silanes, and Shleifer 2008). To control for any country-specific characteristics, we add country fixed effects. The association disappears for Law. For Science and Botany, the changes associated with a 10 percentage point increase are8.1 % and 6.9 %. This suggests that Law related to growth through some mechanism occurring

Table 2: Regional GDP per capita and academic fields, 1000–1800

	log GDP per capita, 1900				2015
	(1)	(2)	(3)	(4)	(5)
$log(n_{r,s})$	0.12***	0.10***	0.10***	0.10***	0.10***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.01)
Share Theology 1			0.04	-0.08	-0.06
			(0.25)	(0.21)	(0.20)
Share Theology 2			0.87	-0.12	-0.26
			(0.64)	(0.46)	(0.23)
Share Politics			0.30	0.02	-0.18
			(0.38)	(0.13)	(0.21)
Share Law			-0.81*	-0.29	-0.00
			(0.41)	(0.23)	(0.23)
Share Science			1.08**	0.78*	0.34
			(0.39)	(0.37)	(0.24)
Share Philosophy			-0.06	0.03	0.14
			(0.13)	(0.16)	(0.17)
Share Botany			0.76**	0.67**	0.50***
			(0.28)	(0.23)	(0.13)
Share Culture			-0.07	-0.33	-0.30
			(0.36)	(0.29)	(0.18)
Share Medicine			-0.21	-0.57	-0.12
			(0.40)	(0.39)	(0.27)
N	172	172	172	172	165
Adj. R Squared	0.40	0.68	0.48	0.70	0.72
Country FEs		X		X	X

Note: 95% confidence intervals displayed. Standard errors clustered by country in parentheses. The unit of observation is a NUTS2 region. Every scholar is assigned a field, a weight equal the fourth root of their number of publications, and to a birth region. GDP per capita from the Rosés-Wolf database on regional GDP (Rosés and Wolf 2021) and from Eurostat. For comparison, regions with GDP in 2015 but not in 1900 are omitted. $\log(n_{r,s})$ is the log of the weighted total of scholars. Shares are the share of the total scholars who are assigned to a given field. GDP per capita from the Rosés-Wolf database on regional GDP (Rosés and Wolf 2021). All regressions include as controls log ruggedness from Nunn and Puga (2012), log post-1500 calorie suitability index from Galor and Özak (2016), and the log of the region's area.

at the national level.

In the last column we also look at GDP per capita in 2015 (from Eurostat). The positive overall association remains. A one percent increase in the weighted sum of scholars born between 1000 and 1800 in region i is associated with a 9.66 % percent increase in GDP per capita in 2015. Having a high concentration of scholars gave regions an initial advantage, and while the advantage is smaller in 2015, the other regions have not fully converged. On the other hand, the initial advantage from having a high proportion of scientists and botanists appears to have partially converged by 2015. The share of scientists is no longer significant, and the magnitude of the association with botanists has decreased. Perhaps scientists were particularly important for the early adoption of the technologies of the Industrial Revolution, leading to an initial but temporary edge.

5 Endogeneity

Many economic growth models view GDP per capita and human capital as two endogenous variables evolving together along their dynamical path (Azariadis and Drazen 1990; Lucas Jr 1988). If we interpret each region as a closed economy starting from specific initial conditions, these models imply a positive correlation across locations between regional GDP per capita and regional human capital, which is the result we find in the analysis above. More generally, we can imagine that both upper tail human capital and GDP per capita in 1900 depend on common unobserved variables. An example is provided in Sandra de Pleijt and Wallis (2023). They argue that upper tail human capital depended on high school foundations in the region after the death of wealthy donors. Such schools may also affect GDP directly through enhancing basic skills.

To evaluate the extent of a possible endogeneity bias, we rely on the following strategy. Shocks to a region's stock of human capital were frequent in history. We exploit three of them to build an instrumental variable for the total output of scholars born within a NUTS2 region. The instrument gives similar results to OLS, suggesting that endogeneity bias is not a major concern.

Our instrument is based on forced migration linked to the following episodes: the fall of the Eastern Roman Empire, the Reformation in England, and the revocation of the Edict of Nantes. Combining these three migration waves gives us a set of plausibly exogenous variation in local scholarship that spans a large area of Europe.

Many of the Byzantine Greek scholars in European academia were refugees from the collapse of the Eastern Roman Empire. We consider all scholars born in modern Turkey, Greece, Cyprus, Albania between 1389 (the battle of Kosovo) and 1699 as potential refugees.

As shown in Figure 2, most of them settled in the closest cities, in Italy. There is a literature stressing their importance in bringing books and knowledge from the Greek Antiquity, fostering the Italian Renaissance (Harris 1995; Link 2023).

Several British Catholic scholars departed to the continent in the decades after the Church of England split from Rome in 1534. We consider all scholars born in Britain between 1558 (the coronation of Queen Elizabeth I) and 1699 who worked or died in a Catholic region potential refugees.[2] As shown in Figure 3, they mostly settled in France, Belgium, and Italy.

Many Huguenots scholars left France in the 17th century in response to religious suppression, most notably the Siege of La Rochelle and the revocation of the Edict of Nantes (1685). We consider all scholars born in France between 1572 (the Bartholomew's Day massacre) and 1699 who worked or died in a region where Protestantism was at least tolerated, if not the state religion, potential refugees.[3] Figure 4 shows their migrations.

In Table 3 we use an indicator variable that is one if at least one of these scholars was active in the region as an instrument for the total weighted output of scholars born in that region after 1700. We pool the three migration waves as each affected a different region; in the appendix, we estimate each separately. As shown by the first stage regression, the instrument is relevant.[4] This in itself supports our proposed inspiration mechanism. Refugee scholars are a positive shock to regional academic activity that appears to increase the amount of local scholarship.

For the instrument to be valid, the exclusion restriction requires that the refugee scholars affect GDP per capita in 1900 only through the academic output of the region. One potential violation would be if the refugees choose their destination region for a reason correlated with economic growth. Looking at the map (figures 2, 3,4), we suspect that scholars likely chose their destination based on physical distance, which would not be particularly associated with economic growth.

If the association between refugees and growth was spurious due to migrants selecting a location, we would expect the field of the migrant to be irrelevant. Forced migration is too rare of an event to provide substantial variation in the shares of research in different fields. However, if we pool the two clusters shown in Table 2 to be significantly associated with growth, Science and Biology, we find that the presence of a refugee from that cluster is associated with higher GDP conditioning on the presence of at least one refugee (3, Column 4).

We can also look at the effect of the refugees on scholarship in an event study. This allows

2. Modern France, Belgium, Italy, Spain, Portugal, Austria, and Czechia.

3. Modern Germany, the Netherlands, Great Britain, Sweden, Denmark, and Switzerland.

4. The first stage F-statistic is 54.82. Pooling the migration waves also increases the strength of the instrument.

Figure 2: Byzantine Academic Refugees, 1389–1699

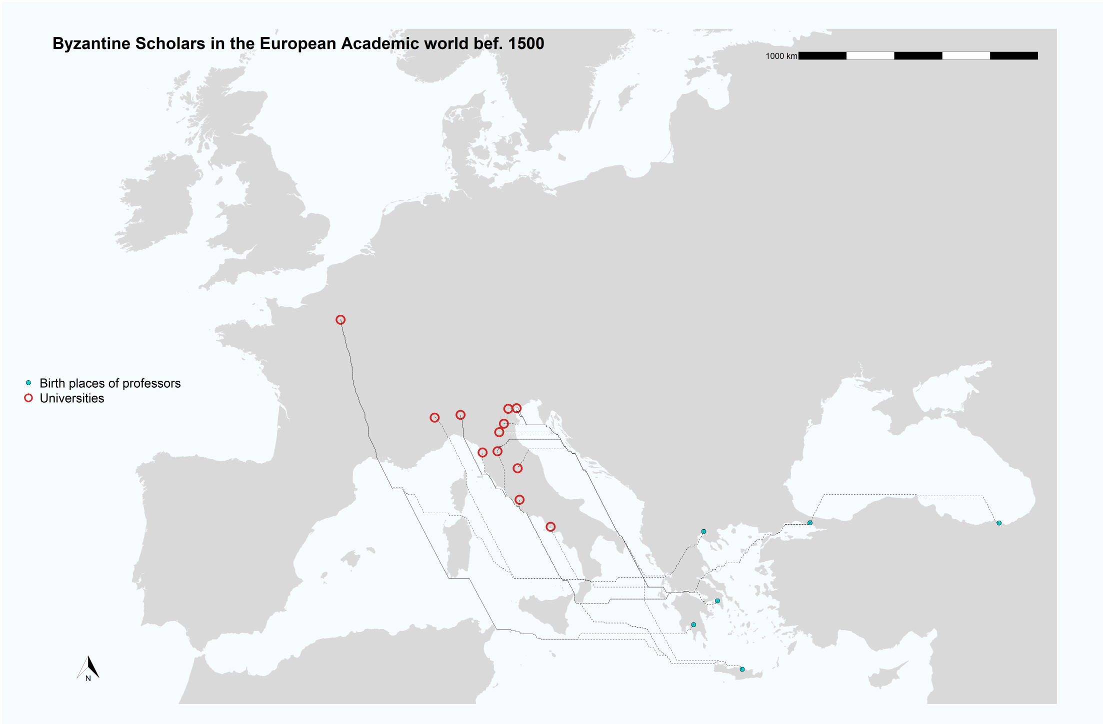


Figure 3: British Catholic Academic Refugees, 1389–1699

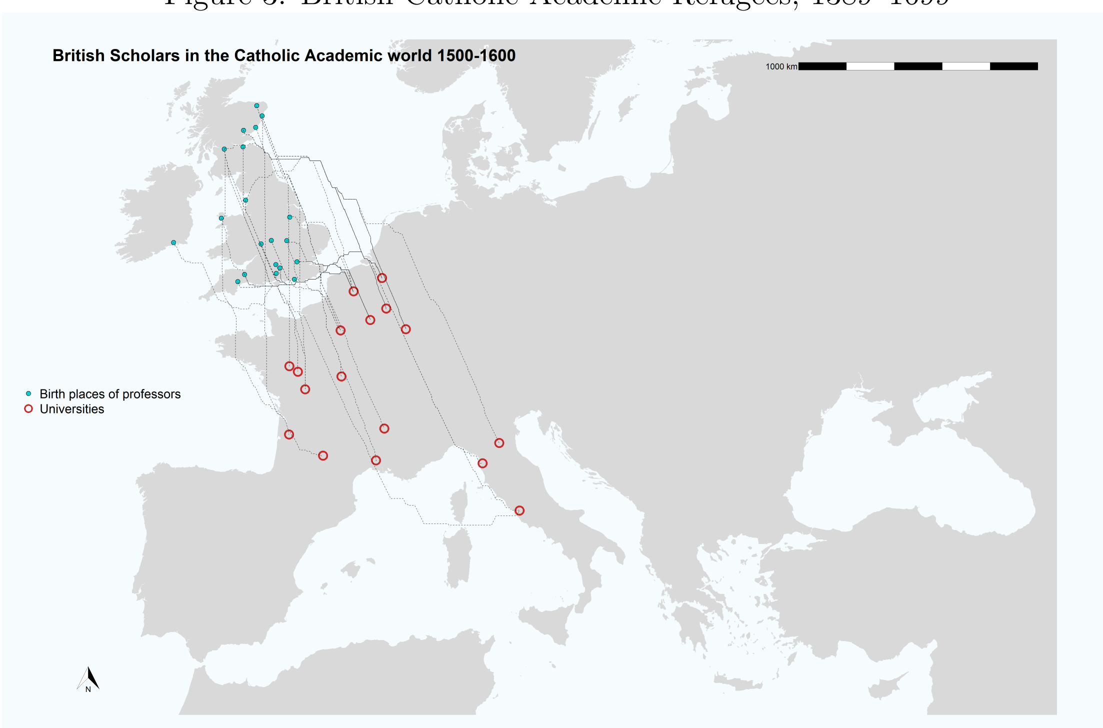


Figure 4: Huguenot Academic Refugees, 1572–1699

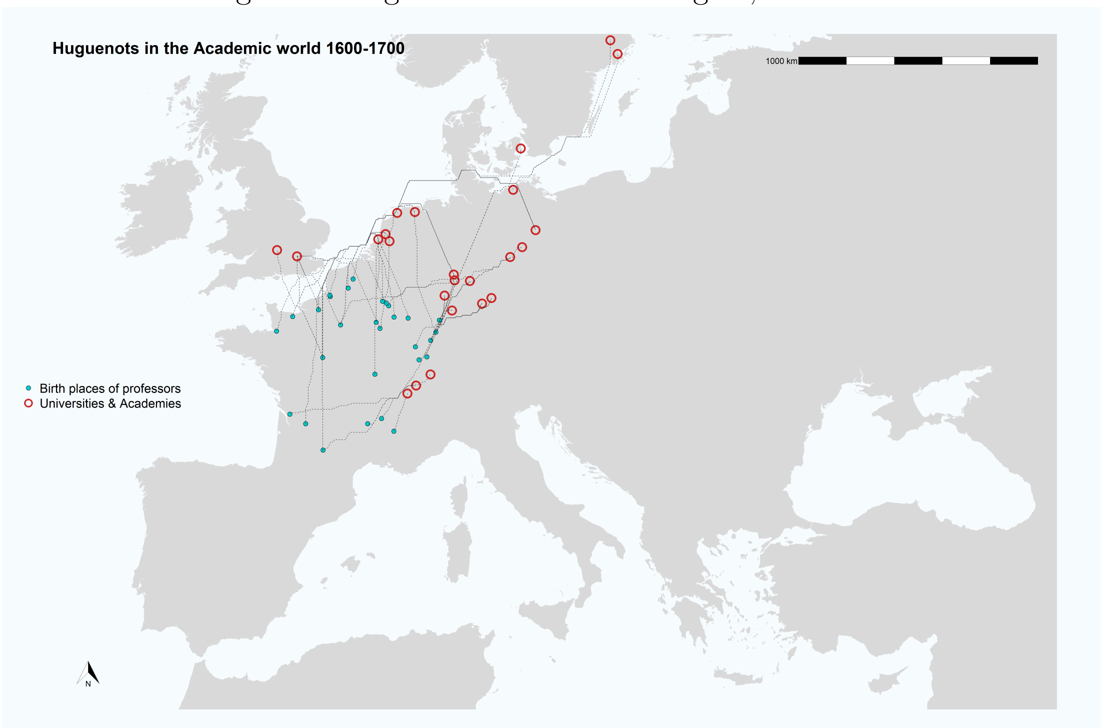


Table 3: Refugee scholars as an instrument for total scholarly output

	log GDP per capita, 1900 OLS	IV	log total scholarship 1st-stage	log GDP per capita, 1900 OLS
$log(n_{r,s}+1)$	0.10***	0.11***		
	(0.03)	(0.02)		
1+ refugee botanist / scientist				0.62**
				(0.28)
1+ refugee scholars			1.12***	0.94***
			(0.18)	(0.17)
N	172	172	172	172
Adj. R Squared	0.68	0.67	0.61	0.62
Country FEs	X	X	X	X
1st stage F-stat.		54.82		

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors clustered by country in parentheses. The unit of observation is a NUTS2 region. Every scholar is assigned a weight equal the fourth root of their number of publications and to a birth region. GDP per capita from the Rosés-Wolf database on regional GDP (Rosés and Wolf 2021). $\log(n_{r,s,i})$ is the weighted total of scholars born in the region 1700–1800. 1+ refugee scholars is an indicator that is one if a refugee scholar was active or died in the region; see text for details. All regressions include as controls log ruggedness from Nunn and Puga (2012), the log post-1500 calorie suitability index from Galor and Özak (2016), and the log of the region's area.

us to rule out pre-trends, demonstrating that scholars did not choose their destination based on academic output. It also allows us to illustrate the dynamics of the effect of a migrant.

For region r in period t, let $D^h_{r,t}$ be an indicator that is one if the region is "treated" h periods in the future with exposure to a refugee scholar. As we do not always know the exact date of arrival, we consider any cohort of scholars dying in a period after the refugee was born as treated. Then we estimate:

$$\log(n_{r,t}) = \sum_{h=-K}^{K} \beta^h D^h_{r,t} + \alpha_1 D^{>K}_{r,t} + \alpha_2 D^{<-K}_{r,t} + \phi_s + \rho_t + \varepsilon_{r,t} \tag{2}$$

where $n_{r,t}$ is the weighted sum of published scholars born in r who died in period t, weighted by their number of publications raised to the power of 0.25; β^h are the event study coefficients; $D^{>K}_{r,t}$ and $D^{<K}_{r,t}$ capture treatment outside the window of interest; ϕ_s is a region fixed effect; ρ_t is a time fixed effect; and $\varepsilon_{sr,t}$ is an error term (clustered by period). As we use quarter-centuries for periods, look at refugees born through 1699, and look at scholars born through 1800, we set K equal to 4.

We also estimate a stacked event study to avoid issues with negative weights in two-way fixed effects (Cengiz et al. 2019; Dube et al. 2023).[5] For each potential period of treatment c, we take a subset of the data where either the region is treated for the first time in period c or has not yet been treated. In other words, the sample includes observations where:

$$D_{it} = \begin{cases} \text{newly treated} & \Delta D^c_{r,t} = 1 \\ \text{clean control} & D^{<c}_{r,t} = 0 \end{cases} \tag{3}$$

Then, we "stack" all the subsets into one dataset and estimate:

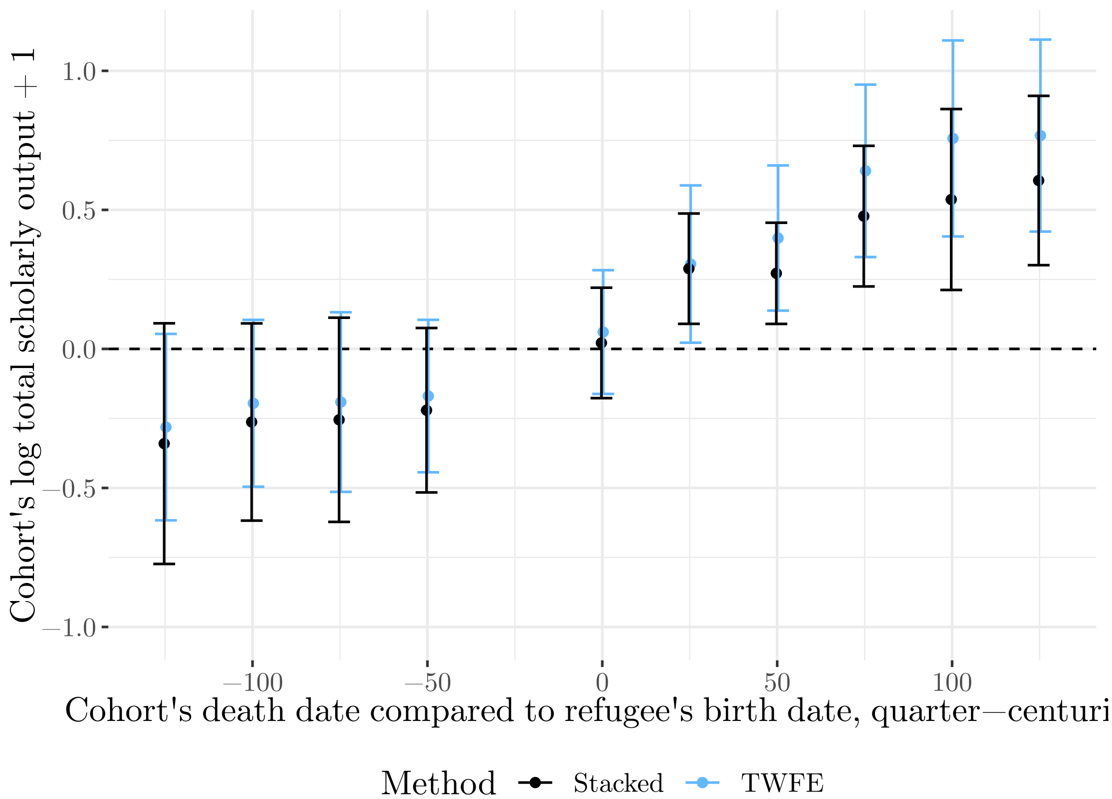
$$\log(n_{r,t,c}) = \sum_{h=-K}^{K} \beta^h D^h_{r,t,c} + \alpha_1 D^{>K}_{r,t,c} + \alpha_2 D^{<-K}_{r,t,c} + \phi_{s\times c} + \rho_{t\times c} + \varepsilon_{r,t,c} \tag{4}$$

This is the equivalent of the previous regression except the fixed effects are now subset specific. The results of both regressions are shown in figure 5. There appears to be no obvious pre-trends before the arrival of a refugee scholar. Afterwards, the level of academic output is elevated for each death cohort for at least 100 years.

We interpret the event study as evidence that the refugees cause locals to produce more scholarship. The lack of pre-trends is reassuring for our instrumental variable strategy, as it suggests the refugees are not selecting destinations based on their scholarly output. In turn,

5. If the effects are heterogenous, then the stacked difference-in-difference estimate is a weighted combination of the average treatment effect on the treated. However, by using only "clean" controls it avoids any issues with negative weights.

Figure 5: Refugee scholars and academic output, event study



we interpret the IV results as evidence that historical scholarship did have a causal effect on later economic development.

6 Mechanisms

We believe that the mechanism that most likely explains our results is that scholars inspire fellow natives to accumulate human capital. This proposition finds support through three compelling observations. Firstly, regional development is better understood when scholars are aggregated by birthplace rather than their activity location. Secondly, scholars who stay closer to their roots demonstrate a stronger correlation with development compared to those who relocate or face an untimely demise. Thirdly, regions boasting a higher number of scholars born within them exhibited elevated levels of general population numeracy in the late 19th century, even after adjusting for GDP per capita.

In our main set of regressions, scholars are allocated to their region of birth. One might wonder about the effects of academic knowledge at the location where it was produced. This reduces the amount of geographic variation as institutions tend to be more centralized than scholar birthplaces (106 regions hosted a university or an academy, while 172 witnessed the birth of some scholars). Moreover, assigning scholars to locations based on activity is challenging, as many academies had corresponding members. To address this, we do two adjustments. First, we drop all corresponding members. Second, for scholars who were faculty at a university, we ignore their membership of any academies. Table 11 shows the results. Compared to Table 2 in the main text, the coefficients on the weighted number of scholars is about half the magnitude and the adjusted R squared is lower. There is no statistically significant difference between the different subfields. We interpret this as evidence that the effect of scholars is much stronger in their region of birth, consistent with role-model effects.

Table 4 tests a key part of this inspiration mechanism: that growth is related to the connection between a scholar and his region of birth. The output of scholars weakly attached to their home region had little effect on growth. Because we control for more strongly attached scholars, we are indirectly controlling for any omitted variables that increased the demand for or supply of scholars. We thus argue the lack of effect suggests that it is scholars influencing development in their home region, not vice versa. Moreover, the measures of weak attachment that we use are particularly relevant for our inspiration mechanism. It is hard to picture a Catholic region building monuments to a Protestant emigrant, or a scholar who died young encouraging the next generation.

The first group are scholars who died in a foreign country. This is a proxy for scholars who emigrated, and thus had a weaker connection to their home region. The second group are those who died after the Peace of Augsburg (1555) in a country with a different state religion than their home country (omitting Germany due to its religious heterogeneity). This proxies for what we term émigrés, that is scholars who migrate due to a religious or political conflict in their home regions. The most prolific of these émigrés was René Descartes, who was born in France and died in Sweden. He did not convert to Protestantism, but was placed on the *Index Librorum Prohibitorum* in 1663. Other notable examples are Helen Maria Williams, an English Girondin revolutionary, Alban Butler, an English Catholic priest, and Jacques Abbadie, a French Anglican minister. The third group has a more inclusive definition of migrant, consisting of any scholars who died in a different location from their birth. The final group are scholars who died before age 40. These scholars had less time to build a local reputation even if they had produced scholarly works.

The regressions are of the form:

$$y_{r,s} = \alpha_0 + \alpha_1 \log(n_{r,s,i} + 1) + \log(n_{r,s,j} + 1) + \beta X_{r,s} + \phi_s + \varepsilon_{r,s} \tag{5}$$

Notation is the same as in Equation (1). $n_{r,s,i}$ is the sum of published scholars in the group i of interest, born in r from 1000–1800, weighted by their number of publications raised to the power of 0.25; is $n_{r,s,j}$ the same but for scholars not in the group of interest. We also run the same regressions for regions of death.

As shown in Table 4 Lines 1–10, the output of those scholars weakly attached to their birthplace appear to have little to no association with growth after controlling for the output of the rest. Moreover, the coefficients for the output of scholars who were not weakly attached are very similar to the coefficient of the baseline regression. This suggests that the associations in Table 2 are driven by scholars with a close association with their birthplaces.

A contrasting result is found in the last two lines of Figure 4, which splits scholars born before and after 1600. This crude periodization attempts to split the sample roughly before and after the Scientific Revolution. A scholar born in 1600 could read Bacon's *Novum Organum* at age 20 and Galilei's *Dialogue Concerning the Two Chief World System* at age 38. While the coefficient for the earlier scholars is slightly weaker, the difference is marginal. In other words, the output of pre-Scientific Revolution scholars seem as important as that of post-Scientific Revolution scholars.

These results suggest that there really was a mechanism tying scholars to growth in their place of birth. One possibility is that successful scholars encouraged others from the same region to accumulate human capital. Early Modern Europe's "Republic of Letters" was a small elite network, but provided notable scholars with both prestige and financial patronage

Table 4: Analysis by strength of attachment to birth region

	log GDP per capita, 1900					
	(1)	(2)	(3)	(4)	(5)	(6)
$log(n_{r,s}+1)$	0.11***					
	(0.03)					
×died in other country		0.01				
		(0.03)				
×died in same country		0.09**				
		(0.04)				
×émigré			-0.01			
			(0.03)			
×not émigré			0.12***			
			(0.03)			
×died in other place				0.04		
				(0.04)		
×died in same place				0.05**		
				(0.02)		
×died before age 40					-0.00	
					(0.02)	
×died after age 40					0.10**	
					(0.04)	
×born before 1600						0.04**
						(0.02)
×born after 1600						0.06**
						(0.03)
N	172	172	172	172	172	172
Adj. R Squared	0.68	0.69	0.68	0.68	0.67	0.68
Country FEs	X	X	X	X	X	X

Note: 95% confidence intervals displayed. Standard errors clustered by country in parentheses. The unit of observation is a NUTS2 region. Every scholar is assigned a weight equal the fourth root of their number of publications and to a birth region. GDP per capita from the Rosés-Wolf database on regional GDP (Rosés and Wolf 2021). $\log(n_{r,s,i}+1)$ is the logarithm of one plus the weighted total of scholars belonging to a subset of scholars. All regressions include as controls log ruggedness from Nunn and Puga (2012), log post-1500 calorie suitability index from Galor and Özak (2016), and the log of the region's area.

and was relatively open to new talent (Mokyr 2016). Examples of scholars born nearby may have been illustrative of the potential returns to human capital.

A key mediating variable in the inspiration story is human capital. The available data for regional human capital for Europe in 1900 are more limited than those for income. Nevertheless, Table 5 estimates the same regressions as Table 2 for a measure of human capital: numeracy. To measure numeracy, we use the ABCC index from Baten and Hippe (2018). This index is defined as $125(1-s)$, where s is the share of reported ages between 23 and 72 which end in 0 or 5. It measures a very rudimentary level of human capital: do people know their own age. We find that, even controlling for GDP in 1900, areas with higher scholarship had higher *lower-tail* human capital, even after controlling for economic development.

Table 5: Human capital and academic fields, 1000–1800

	ABCC		ABCC, imputed	
	(1)	(2)	(3)	(4)
$log(n_{r,s})$	2.68**	3.21**	1.89*	1.81
	(1.03)	(1.09)	(0.86)	(1.06)
N	62	62	102	102
Adj. R Squared	0.57	0.57	0.70	0.70
Country FEs	X	X	X	X
Census year FE	X	X	X	X
Controlling for GDP		X		X

Note: 95% confidence intervals displayed. Standard errors clustered by census year in parentheses. x-axis normalized to be standard deviations. The unit of observation is a NUTS2 region. Every scholar is assigned a weight equal the fourth root of their number of publications, and to a birth region. Numeracy is the ABCC index from Baten and Hippe (2018): $125(1-s)$, where s is the share of reported ages ending in 0 or 5. Imputed numeracy assigns a value of 1 to Germany and Scandinavia in 1900. GDP per capita from the Rosés-Wolf database on regional GDP (Rosés and Wolf 2021). All regressions include as controls fixed effects for the year for which the ABCC index was computed, fixed effects for country, log ruggedness from Nunn and Puga (2012), log post-1500 calorie suitability index from Galor and Özak (2016), the log of the region's area, and the log of the population in 1900. Columns (2) and (4) also control for the GDP per capita in 1900.

In addition to the suggestive evidence provided above, the inspiration channel is backed by a large number of anecdotes. We report here a few salient ones.

This inspiration could be through social networks around the place of birth. Leonhard

Euler was born in 1707 in Basel as the son of Paul Euler, a Reformed pastor. Paul had befriended Jacob and Johann Bernoulli (1655 and 1667, Basel). Johann later convinced him to let his son Euler study mathematics instead of theology. Both Bernoullis, notable mathematicians in their own right (they are the 82nd and 83rd most prolific members of the field Science), thus directly contributed to young Euler becoming the most second most productive member of our field Science (behind only Isaac Newton).

This inspiration could also be an indirect effect on future generations of academics. On February 5th, 1835, the Lincoln Mechanics' Institute received a bust of Isaac Newton (born in 1642 in Lincolnshire) from a wealthy benefactor. To celebrate, the 19 year old son of the society's curator (and local shoemaker) gave a lecture on the "Life and Discoveries of Newton." (Burris 2022). The young man, George Boole, born in 1815 in Lincolnshire, would become the founder of modern algebraic logic. The story is interesting because it shows how there was interest by the local broader community in science, inspired by Newton, which in turn led to the birth of additional talents in the field.

Inspiration could be embedded in local culture. Pierre de Fermat (1605–1665), one of the greatest French Mathematicians, member of the Academy of Castres, was born in a small village, Beamont-de-Lomagne. His working life was spent in Toulouse at the Parliament (a court). Today, Beamont-de-Lomagne has a statue of him, a street named after him, a tourism office located in the house where he was born, and a yearly *fête des maths* in his honor. Every year kids learn to like mathematics at this festival.

Råshult is the name of trolley sold by IKEA, but it is also a village in Småland, Sweden, notable as the birthplace of the "father of modern taxonomy," Carl Linnaeus (1707–1778). Råshult has a monument to him, a reconstruction of the cottage where he was born, and garden based on his famous *Adonis Stenbrohultensis*, in which he first used taxonomy to classify every plant in his father's garden.

Even the medieval scholar Pierre Abelard (1079–1142), is honored in his hometown, the tiny Breton village of Le Pallet, with both a street name and a statue. His intellectual influence, philosophical writings, and his tragic romance with Héloïse (resulting in a son named Astrolabe and the castration of Abelard by an angry uncle) have left a lasting impact over several centuries.

Finally, we can also think to the inspiration channel in terms of theory. Embedding inspiration into a growth model can follow two paths. One is to make preferences time dependent, introducing habits or aspirations into utility (Ryder Jr and Heal 1973; De la Croix and Michel 1999). The other is to make the depreciation rate of human capital endogenous, depending on the level of inspiration. High levels of inspiration in a region would slow down the depreciation of human capital. The intuition is that future generations will be more likely to maintain the knowledge of past generations if they are inspired by it. In Appendix B.5 we show that, in such a model, two regions with identical total factor productivity and stock of capital can differ in output: the one with the highest stock of inspiration will produce more. In the long-run, however, this effect of inspiration does not hold anymore, which is what we found in Table 2.

7 Conclusion

We find a strong relationship between economic growth and premodern European scholarship. Our findings support the view that upper tail human capital was important for growth (Squicciarini and Voigtländer 2015). Moreover, we find that certain fields of scholarship had a stronger influence on growth than others.

Perhaps it is not surprising that we find that Science and Botany were particularly important. Fundamental scientific research paved the way for future applied technologies. For example, engineering has been critical to the development of infrastructure and technology throughout history (Maloney and Valencia Caicedo 2022). Engineering was not part of curricula in the period we consider (1000–1800), but is strongly grounded in mathematics and physics, two important components of our field Science. Medical research and advancements have been crucial to improving public health, curing diseases, and extending lifespan, in particular in the nineteenth century. Modern medicine is based on natural sciences such as botany, which appears as a strong correlate of growth as well.

The negative role of Law is also interesting. It appears only when comparing across, not within, countries. It could be that the share of law among academic scholars reflects the local legal system. Indeed, in common law countries, legal education and training are often not solely confined to universities, and there is more emphasis on practical training through apprenticeships, clerkships, and other forms of legal practice. Civil law countries have more lawyers in academia, and there is a large literature showing that these countries tend to perform less well than common law countries (Porta, Lopez-de-Silanes, and Shleifer 2008). The association disappears after controlling for country fixed effects, again suggesting this is mostly a process occurring at the national level.

One theory for the rise of the West argues that universities and academics played a central role. However, there have been no quantitative studies of historical academia and growth for Europe as a whole. This paper develops a methodology to measure academic productivity using a large novel database of scholars 1000–1800. We find that the output of academics predicts 19th century economic growth, providing Europe-wide evidence that the sciences paved the way for the Industrial Revolution. Moreover, approaches to theology and legal systems also mattered for economic development. Despite the fact that ideas and written knowledge were highly mobile, the birthplaces of scholars in particular appear to have higher growth, suggesting scholars played a role-model for the next generations.

Acknowledgements

This project has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program under grant agreement No 883033 "Did elite human capital trigger the rise of the West? Insights from a new database of European scholars." We thank Joel Mokyr, Julius Koschnick, and participants to seminar presentations in Aix-Marseilles, Odense, Louvain, and CEPR Paris Conference for their comments on an earlier draft.

References

Almelhem, Ali, Murat Iyigun, Austin Kennedy, and Jared Rubin. 2023. Enlightenment Ideals and Belief in Science in the Run-up to the Industrial Revolution: A Textual Analysis.

Azariadis, Costas, and Allan Drazen. 1990. Threshold externalities in economic development. *The Quarterly Journal of Economics* 105 (2): 501–526.

Baten, Joerg, and Ralph Hippe. 2018. Geography, land inequality and regional numeracy in Europe in historical perspective. *Journal of Economic Growth* 23, no. 1 (March): 79–109.

Becker, Sascha O, and Ludger Woessmann. 2009. Was Weber wrong? A Human Capital Theory of Protestant Economic History. *The Quarterly Journal of Economics* 124 (2): 531–596. https://doi.org/10.1162/qjec.2009.124.2.531.

Becker, Sascha O., Irena Grosfeld, Pauline Grosjean, Nico Voigtländer, and Ekaterina Zhuravskaya. 2020. Forced migration and human capital: evidence from post-wwii population transfers. *American Economic Review* 110 (5): 1430–63. https://doi.org/10.1257/aer.20181518.

Blasutto, Fabio, and David De la Croix. 2023. Catholic censorship and the demise of knowledge production in early modern Italy. *The Economic Journal* 133 (656): 2899–2924. https://doi.org/10.1093/ej/uead053.

Borowiecki, Karol Jan. 2022. Good reverberations? teacher influence in music composition since 1450. *Journal of Political Economy* 130 (4): 991–1090.

British Library. 2021. *Database of Italian academies.* https://www.bl.uk/catalogues/ItalianAcademies/.

Burris, Stanley. 2022. George boole. In *The stanford encyclopedia of philosophy,* edited by Edward N. Zalta. Stanford University.

Cantoni, Davide, and Noam Yuchtman. 2014. Medieval universities, legal institutions, and the commercial revolution. *The Quarterly Journal of Economics* 129 (2): 823–887. https://doi.org/10.1093/qje/qju007.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. 2019. The Effect of Minimum Wages on Low-Wage Jobs. *The Quarterly Journal of Economics* 134 (3): 1405–1454.

Chaney, Eric. 2020. Modern library holdings and historic city growth. University of Oxford.

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. https://doi.org/10.1007/s10887-017-9141-3.

Curtis, Matthew, and David De la Croix. 2023. Measuring human capital: from WorldCat identities to VIAF. *Repertorium Eruditorum Totius Europae* 10:17–22. https://doi.org/10.14428/rete.v10i0/hc.

Danna, Raffaele. 2022. Elaboration and diffusion of useful knowledge in the long run: the case of european practical arithmetic (13th-16th centuries). *Rivista di storia economica* 38 (1): 57–84.

De Courson, Benoît, Valentin Thouzeau, and Nicolas Baumard. 2023. Quantifying the Scientific Revolution. *Evolutionary Human Sciences,* 1–37. Accessed May 25, 2023. https://doi.org/10.1017/ehs.2023.6.

De la Croix, David. 2021. Scholars and literati in european academia before 1800. *Repertorium Eruditorum Totius Europae* 5:35–41. https://doi.org/10.14428/rete.v5i0/global21.

De la Croix, David, Frédéric Docquier, Alice Fabre, and Robert Stelter. 2023. The academic market and the rise of universities in medieval and early modern Europe (1000-1800). *Journal of the European Economic Association,* https://doi.org/10.1093/jeea/jvad061.

De la Croix, David, and Marc Goñi. 2020. Nepotism vs. intergenerational transmission of human capital in academia (1088–1800). CEPR Discussion Paper No. 15159. https://doi.org/10.2139/ssrn.3807314.

De la Croix, David, and Philippe Michel. 1999. Optimal growth when tastes are inherited. *Journal of Economic Dynamics and Control* 23 (4): 519–537.

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 (4): 506–532. https://doi.org/10.1093/ereh/heac023.

Dube, Arindrajit, Daniele Girardi, Òscar Jordà, and Alan M. Taylor. 2023. *A Local Projections Approach to Difference-in-Differences Event Studies.* Working Paper Series 2023-12. Federal Reserve Bank of San Francisco, April. https://ideas.repec.org/p/fip/fedfwp/96259.html.

Frijhoff, Willem. 1996. Patterns. Chap. 2 in *A History of the University in Europe. vol. ii: Universities in Early Modern Europe (1500-1800),* edited by Hilde de Ridder-Symoens. Cambridge University Press.

Galor, Oded. 2022. *The journey of humanity: the origins of wealth and inequality.* Penguin.

Galor, Oded, and Ömer Özak. 2015. Land Productivity and Economic Development: Caloric Suitability vs. Agricultural Suitability. *SSRN Electronic Journal,* accessed May 22, 2023. https://doi.org/10.2139/ssrn.2625180.

———. 2016. The Agricultural Origins of Time Preference. *American Economic Review* 106 (10): 3064–3103. Accessed May 22, 2023. https://doi.org/10.1257/aer.20150020.

Galor, Oded, Ömer Özak, and Assaf Sarid. 2017. Geographical Origins and Economic Consequences of Language Structures. Available at SSRN 2820889. https://doi.org/10.2139/ssrn.2820889.

Grajzl, Peter, and Peter Murrell. 2023. A Macroscope of English Print Culture, 1530-1700, Applied to the Coevolution of Ideas on Religion, Science, and Institutions. https://doi.org/10.2139/ssrn.4336537.

Gunther, Robert Theodore. 1937. *Early science in oxford - oxford colleges and their men of sciences.* Oxford: Oxford for the author.

Harris, Jonathan. 1995. *Greek emigres in the west 1400-1520.* Porphyrogenitus.

Hartigan, J. A., and M. A. Wong. 1979. A k-means clustering algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 28 (1): 100–108. https://doi.org/10.2307/2346830.

Henrich, Joseph. 2020. *The WEIRDest people in the world: how the west became psychologically peculiar and particularly prosperous.* Penguin UK.

Hill, Arthur W. 1915. The history and functions of botanic gardens. *Annals of the Missouri Botanical Garden* 2 (1/2): 185–240. Accessed January 25, 2023. https://doi.org/10.5962/bhl.title.31145.

Johnson, Noel D, Andrew J Thomas, and Alexander N Taylor. 2023. The Impact of the Black Death on the Adoption of the Printing Press.

Koschnick, Julius. 2023. Breaking tradition: teacher-student effects at english universities during the scientific revolution. London School of Economics.

Link, Andreas. 2023. The fall of constantinople and the rise of the west. *Available at SSRN 4372477.*

Lucas Jr, Robert E. 1988. On the mechanics of economic development. *Journal of monetary economics* 22 (1): 3–42.

Maloney, William F, and Felipe Valencia Caicedo. 2022. Engineering growth. *Journal of the European Economic Association* 20 (4): 1554–1594. https://doi.org/10.1093/jeea/jvac014.

McClellan, James E. 1985. *Science reorganized: scientific societies in the eighteenth century.* Columbia University Press. https://doi.org/10.2307/2739060.

Mitterauer, Michael. 2010. *Why Europe?: the medieval origins of its special path.* University of Chicago Press. https://doi.org/10.7208/chicago/9780226532387.001.0001.

Mokyr, Joel. 2002. *The gifts of Athena: historical origins of the knowledge economy.* Princeton: Princeton University Press. https://doi.org/10.1515/9781400829439.

———. 2016. *A culture of growth.* Princeton: Princeton University Press. https://doi.org/10.2307/j.ctt1wf4dft.

Nunn, Nathan, and Diego Puga. 2012. Ruggedness: the blessing of bad geography in Africa. *Review of Economics and Statistics* 94 (1): 20–36. https://doi.org/10.1162/rest_a_00161.

Pedersen, Olaf. 1992. *Lovers of learning: a history of the royal danish academy of sciences and letters, 1742-1992.* Munksgaard.

Porta, Rafael La, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2008. The economic consequences of legal origins. *Journal of Economic Literature* 46 (2): 285–332. https://doi.org/10.1257/jel.46.2.285.

Rosés, Joan R, and Nikolaus Wolf. 2021. Regional growth and inequality in the long-run: Europe, 1900–2015. *Oxford Review of Economic Policy* 37 (1): 17–48. https://doi.org/10.1093/oxrep/graa062.

Ryder Jr, Harl E, and Geoffrey M Heal. 1973. Optimal growth with intertemporally dependent preferences. *The Review of Economic Studies* 40 (1): 1–31.

Sandra de Pleijt, Julius Kochnick, and Patrick Wallis. 2023. *Reading, writing, and wrenches. the contribution of human capital to the first industrial revolution: a reassessment.* Working paper.

Schulz, Jonathan F. 2022. Kin Networks and Institutional Development. *The Economic Journal* 132, no. 647 (April): 2578–2613. ISSN: 0013-0133. https://doi.org/10.1093/ej/ueac027.

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. https://doi.org/10.1093/qje/qjv025.

St. Andrews, University of. 2019. Universal short title catalogue. Https://www.ustc.ac.uk.

Weber, Max. 1930. *The protestant ethic and the spirit of capitalism.* London: George Allen & Unwin Ltd. https://doi.org/10.4324/9780203995808.

Wootton, David. 2015. *The invention of science: a new history of the scientific revolution.* Penguin UK.

A Appendix: Data

A.1 Important topics by clusters

In this section, we provide some additional detail into our k-means clusters that explains how we assigned them their names. Table 6 displays the top terms by cluster. As described in the main text, for each scholar, every term is assigned an importance weight 0–5. The topics are ranked by the mean weight for scholars in the cluster, with the top 5 per cluster displayed. Note that these average weights can also be interpreted in terms of the k-means centroids. Each cluster has a centroid in the Cartesian coordinate system with one axis per topic. For topic t, the mean weight for cluster c is the coordinate of the centroid for c on the t-axis. One notable feature of these clusters is that the cluster we label "Classics" has no particular strong associations. It seems to be a cluster of scholars interested in a broad range of topics, perhaps related to the Humanistic Revolution.

Table 6: Top 5 terms associated with clusters 1-5

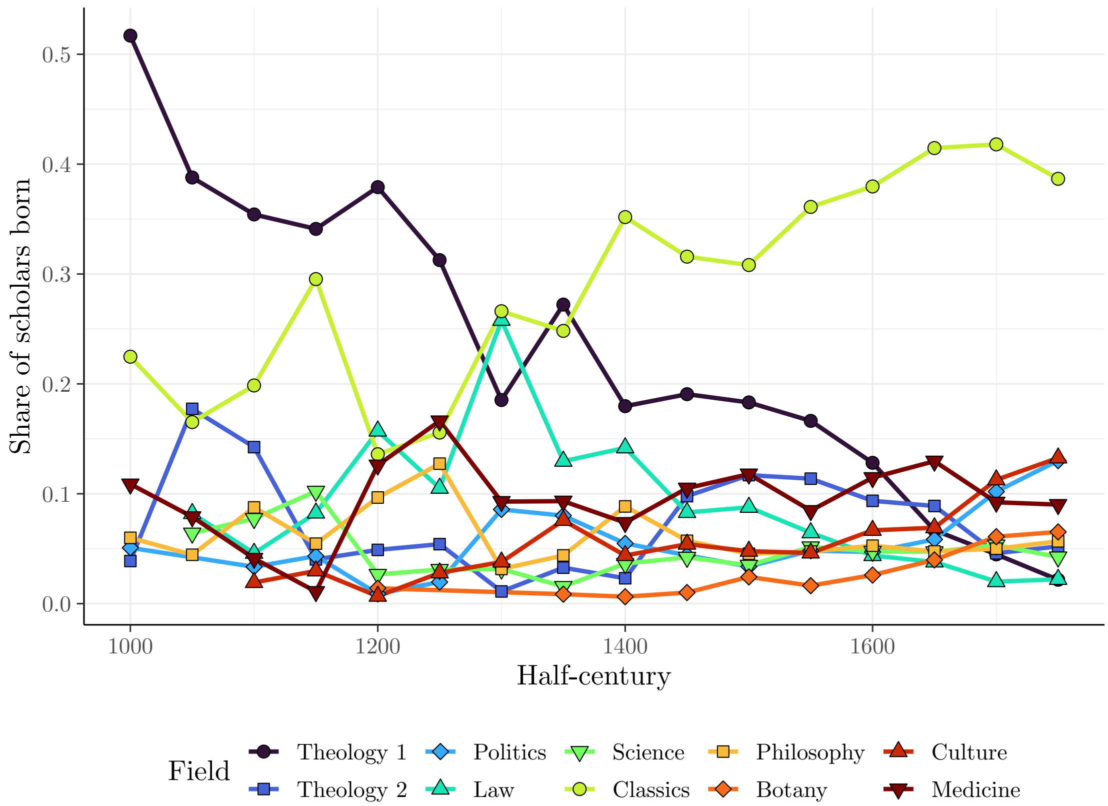
Cluster	Topic	Mean weight
Theology1	Catholic Church	4.69
Theology1	Theology, Doctrinal	1.91
Theology1	Theology	1.02
Theology1	Clergy	0.65
Theology1	Reformation	0.55
Theology2	Bible	3.36
Theology2	Theology	2.03
Theology2	Theology, Doctrinal	1.71
Theology2	Jesus Christ	1.01
Theology2	Bible.–Old Testament	0.95
Politics	Politics and government	4.51
Politics	Political science	0.76
Politics	Diplomatic relations	0.71
Politics	Economics	0.58
Politics	Catholic Church	0.58
Law	Roman law	4.55
Law	Law	2.05
Law	Canon law	1.54
Law	Civil law	1.14
Law	Digesta	0.81
Science	Mathematics	3.41
Science	Astronomy	2.68
Science	Geometry	2.06
Science	Science	1.79
Science	Physics	1.53

Table 7: Top 5 terms associated with clusters 6-10

Cluster	Topic	Mean weight
Classics	Rome (Empire)	0.24
Classics	Intellectual life	0.18
Classics	Law	0.18
Classics	Jesus Christ	0.14
Classics	Antiquities	0.13
Philosophy	Philosophy	4.27
Philosophy	Ethics	1.24
Philosophy	Logic	1.05
Philosophy	Science	0.90
Philosophy	Metaphysics	0.89
Botany	Botany	4.67
Botany	Plants	2.48
Botany	Natural history	1.87
Botany	Medicine	1.18
Botany	Botany, Medical	1.03
Culture	Travel	4.60
Culture	Antiquities	0.88
Culture	Manners and customs	0.74
Culture	Natural history	0.69
Culture	Voyages and travels	0.64
Medicine	Medicine	4.79
Medicine	Physicians	0.85
Medicine	Human anatomy	0.74
Medicine	Materia medica	0.67
Medicine	Surgery	0.62

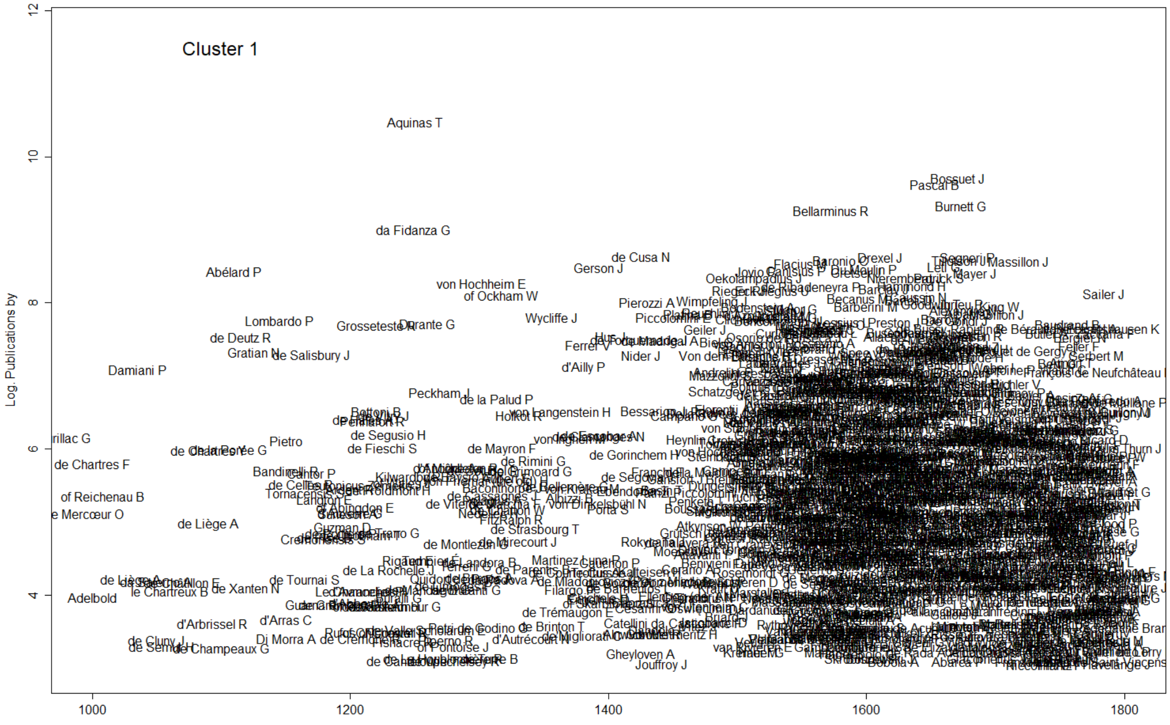
A.2 Academic fields over time

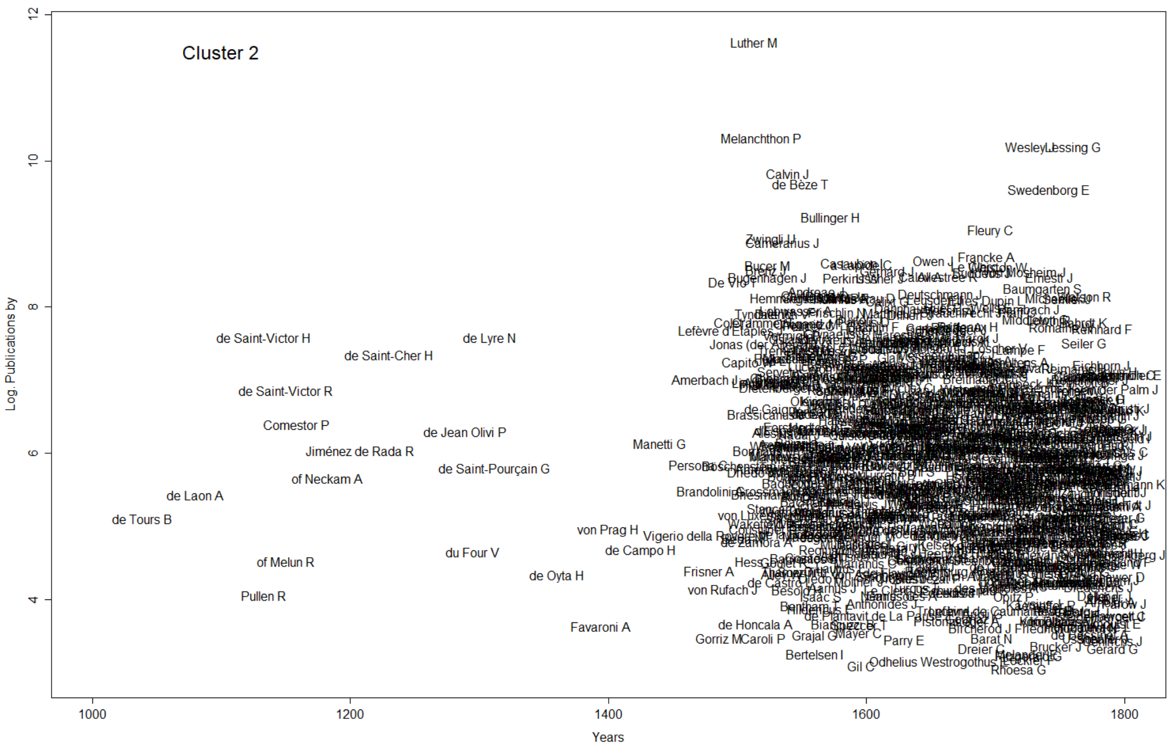
Figure 6: Shares of academic fields over time

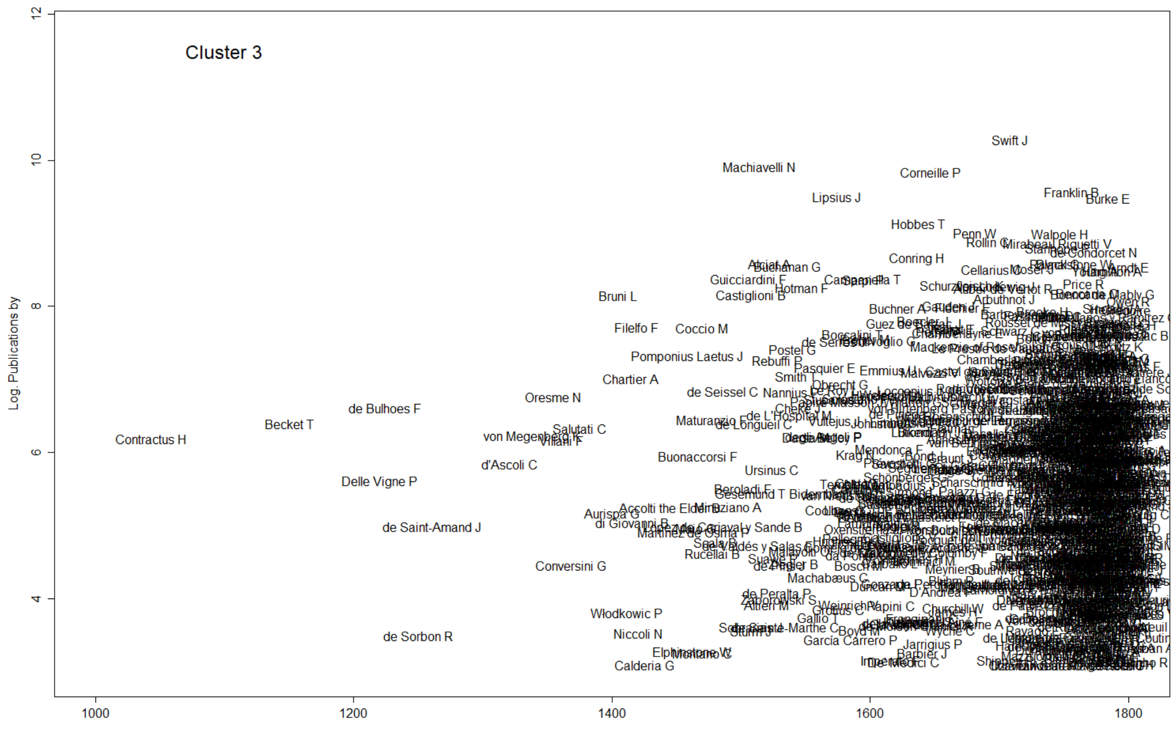
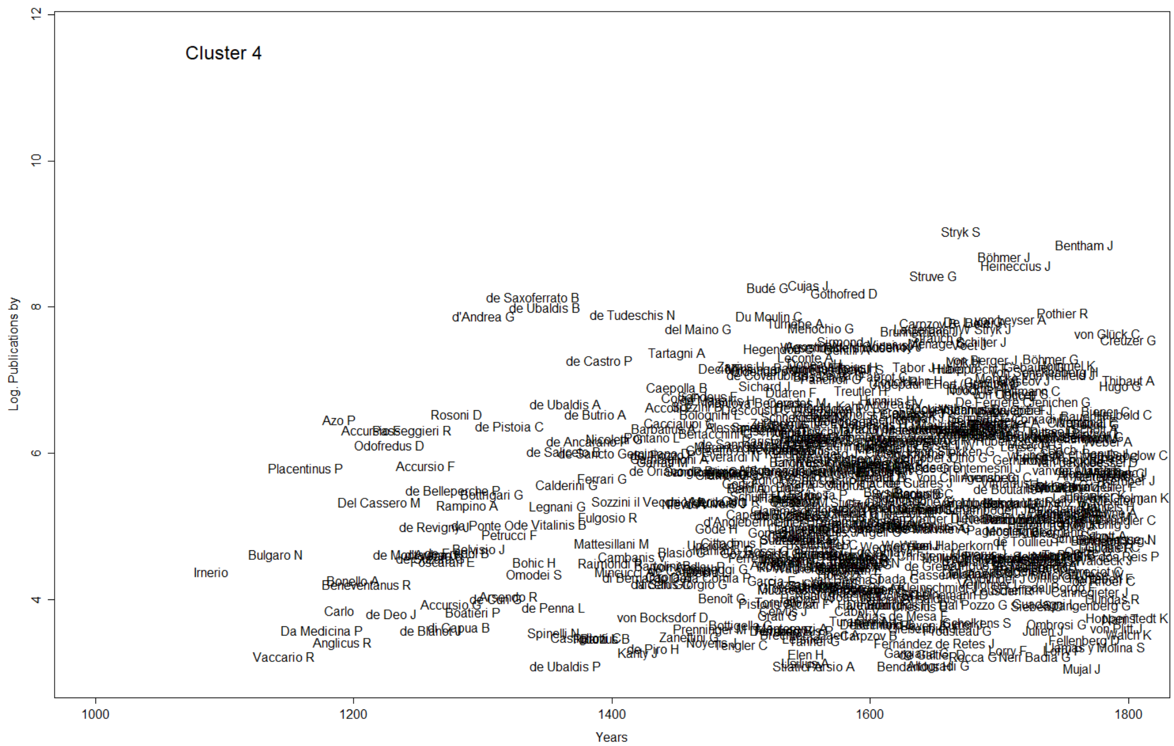


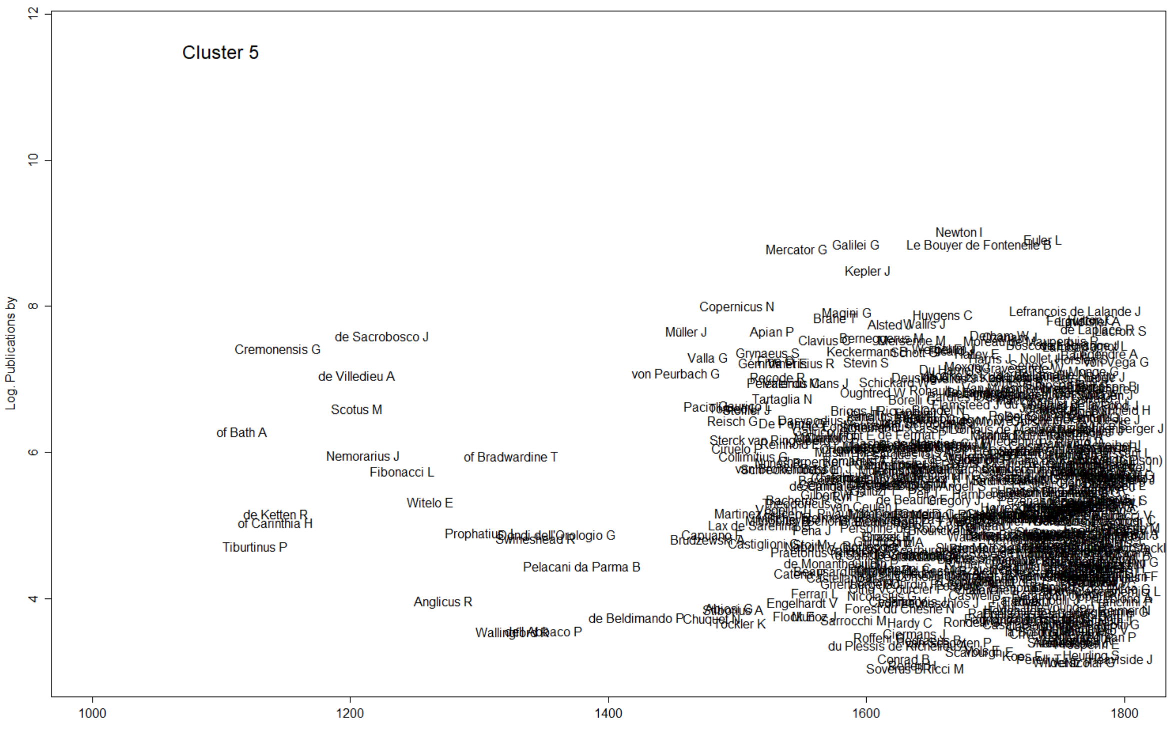
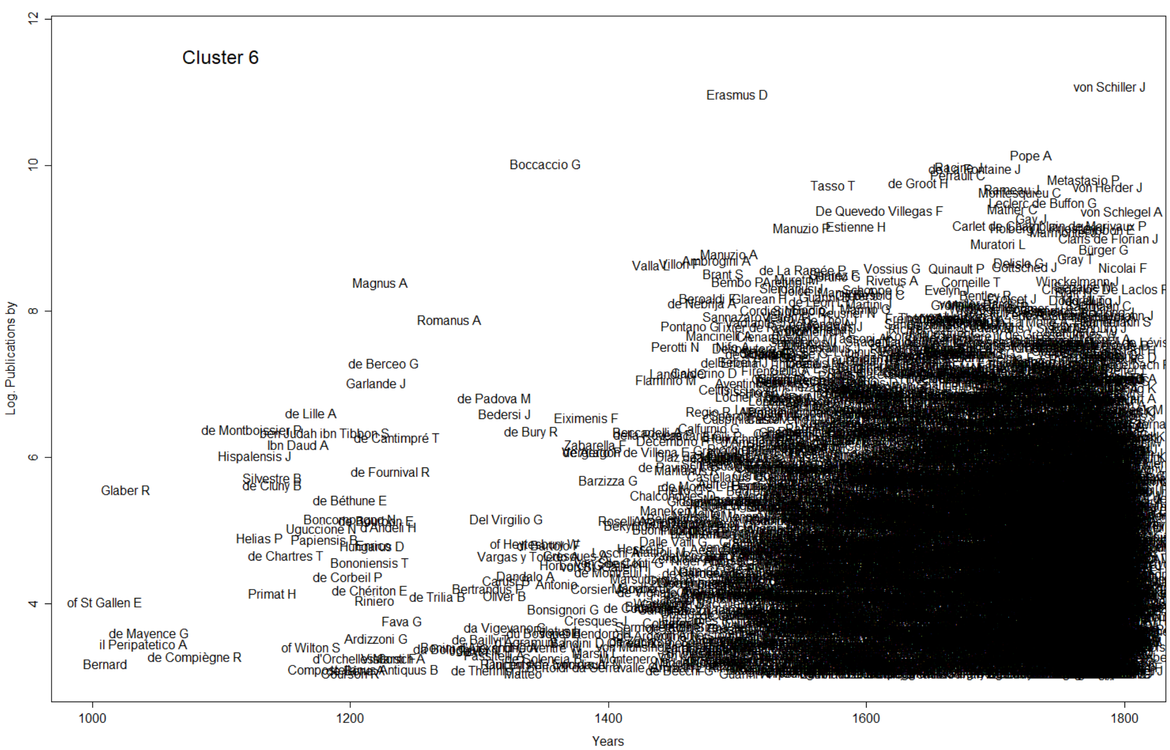
Note: Fields estimated by k-means clustering.

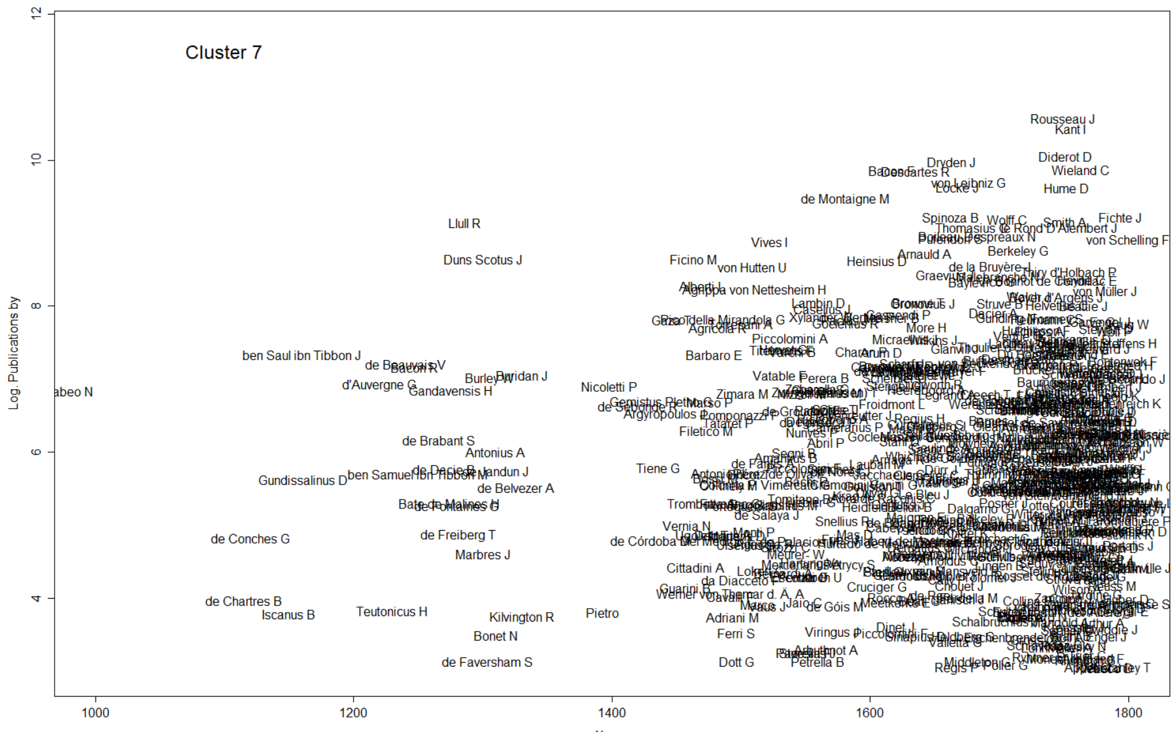
A.3 Publishing Scholars over Time by Cluster

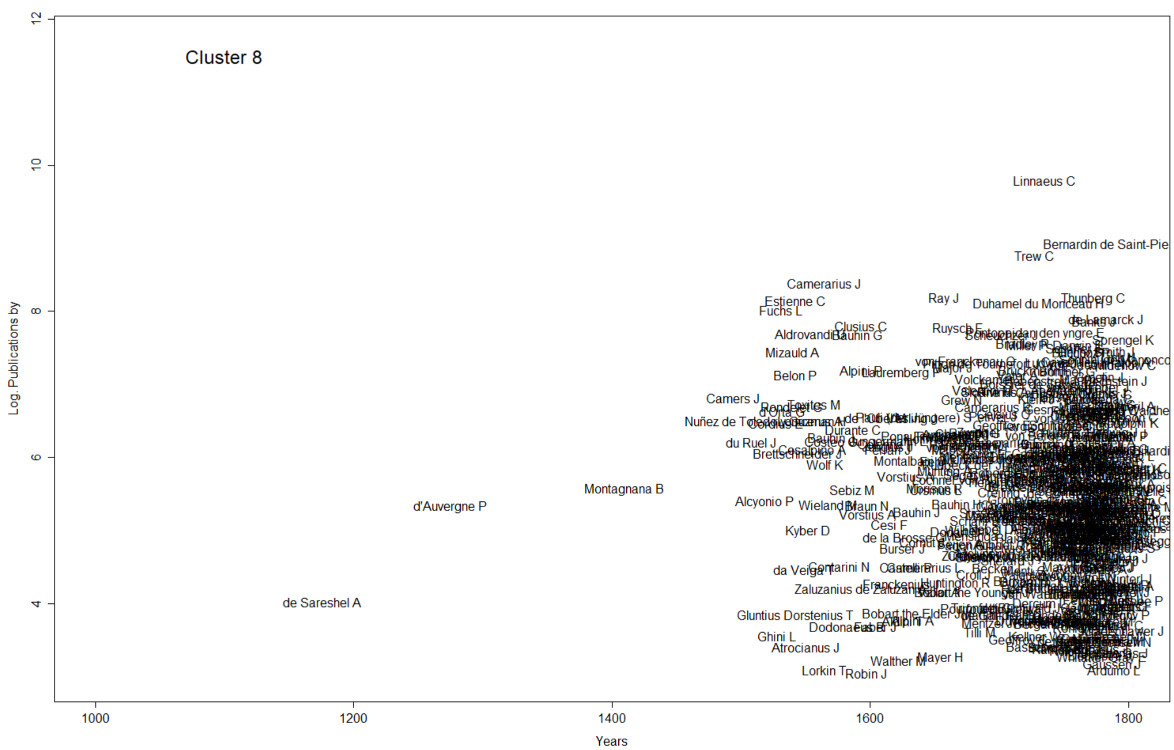


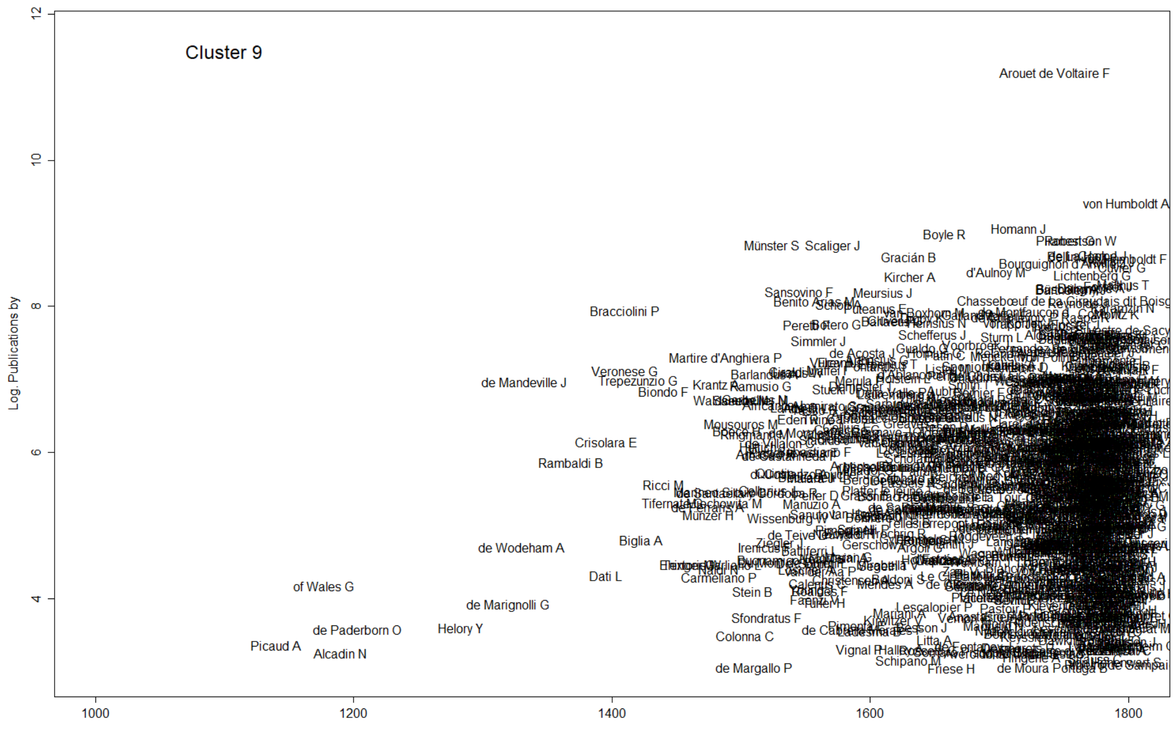


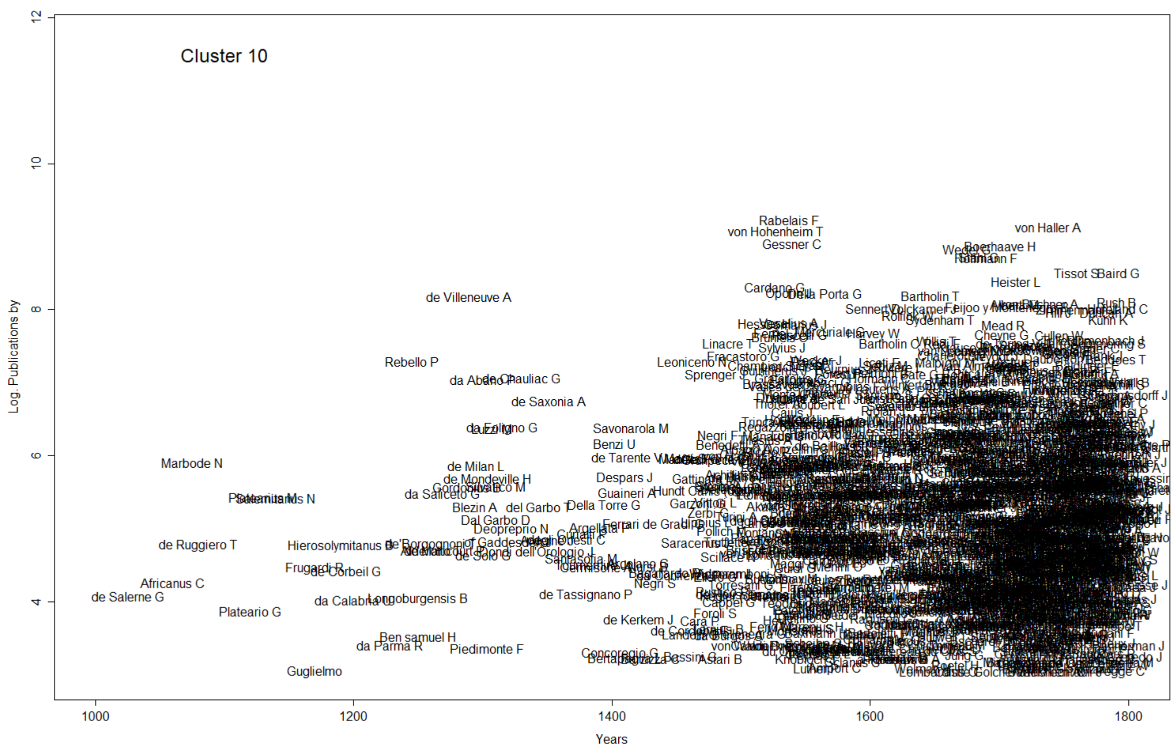

Cluster 3
Log. Publications by
Swift J
Machiavelli N
Corneille P
Lipsius J
Franklin B
Burke E
Hobbes T
Penn W
Walpole H
Conring H
Guicciardini F
Castiglioni B
Bruni L
Filelfo F
Coccio M
Pomponius Laetus J
Chartier A
Oresme N
de Bulhoes F
Becket T
Contractus H
von Megenberg K
Salutati C
d'Ascoli C
Delle Vigne P
de Saint-Amand J
Buonaccorsi F
Ursinus C
Conversini G
Wlodkowic P
Niccoli N
Calderia G
de Sorbon R
de Seissel C
1000
1200
1400
1600
1800

Cluster 4
Log. Publications by
Stryk S
Bentham J
Böhmer J
Heineccius J
Struve G
Budé G
Cujas J
Gothofred D
de Saxoferrato B
de Ubaldis B
d'Andrea G
de Tudeschis N
del Maino G
Du Moulin C
Rothier R
von Glück C
Creuzer G
de Castro P
Tartagni A
Thibaut A
Hugo G
de Ubaldis A
de Butrio A
Azo P
Accursius F
Rosoni D
de Pistoia C
Odofredus D
Placentinus P
Accursio F
Del Cassero M
Rampino A
Legnani G
Bulgaro N
Irnerio
Beneventanus R
Carlo
de Deo J
Da Medicina P
Anglicus R
Vaccario R
de Ubaldis P
von Bocksdorf D
Mujal J
1000
1200
1400
1600
1800
Years


Cluster 5
Log. Publications by
12
10
8
6
4
1000
1200
1400
1600
1800
Newton I
Mercator G
Galilei G
Le Bouyer de Fontenelle B
Euler L
Kepler J
Copernicus N
Huygens C
Lefrançois de Lalande J
Müller J
Apian P
de Sacrobosco J
Cremonensis G
Valla G
von Peurbach G
de Villedieu A
Scotus M
Tartaglia N
of Bath A
Nemorarius J
of Bradwardine T
Fibonacci L
Witelo E
Tiburtinus P
Pelacani da Parma B
Anglicus R
de Beldimando P
Ferrari L
Roffeni
du Plessis de Richelieu
Conrad B

Cluster 6
Log. Publications by
12
10
8
6
4
1000
1200
1400
1600
1800
Years
Erasmus D
von Schiller J
Boccaccio G
Pope A
Tasso T
de Groot H
Metastasio P
von Herder J
De Quevedo Villegas F
Manuzio P
Estienne H
von Schlegel A
Muratori L
Magnus A
Romanus A
Nicolai F
de Berceo G
Garlande J
de Padova M
Bedersi J
Eiximenis F
de Lille A
de Bury R
Hispalensis J
de Fournival R
Glaber R
Barzizza G
de Béthune E
Del Virgilio G
Helias P
de Chartres T
Bononiensis T
de Corbeil P
Primat H
of St Gallen E
Riniero
Fava G
Ardizzoni G
de Mayence G
il Peripatetico A
Bernard
de Compiègne R
of Wilton S


Cluster 7
Log. Publications by
1000
1200
1400
1600
1800
12
10
8
6
4
Rousseau J
Kant I
Diderot D
Wieland C
Hume D
Dryden J
de Montaigne M
Llull R
Duns Scotus J
Vives I
Ficino M
von Hutten U
Agrippa von Nettesheim H
Heinsius D
Arnauld A
Berkeley G
Fichte J
von Schelling F
ben Saul ibn Tibbon J
d'Auvergne G
Gandavensis H
Buridan J
Nicoletti P
Barbaro E
Piccolomini A
Vatable F
Froidmont L
de Brabant S
Antonius A
Gundissalinus D
de Belvezer A
Tiene G
de Conches G
de Freiberg T
Marbres J
Cittadini A
Cruciger G
de Chartres B
Iscanus B
Teutonicus H
Kilvington R
Bonet N
Pietro
Adriani M
Ferri S
Viringus P
de Faversham S
Dott G
Petrella B


Cluster 8
Log. Publications by
Years
1000
1200
1400
1600
1800
12
10
8
6
4
Linnaeus C
Bernardin de Saint-Pie
Trew C
Camerarius J
Estienne C
Fuchs L
Ray J
Thunberg C
Mizauld A
Belon P
Camers J
du Ruel J
Wolf K
Montagnana B
d'Auvergne P
Alcyonio P
Sebiz M
Kyber D
de Sareshel A
da Vergel
Zaluzianus de Zaluzan
Gluntius Dorstenius T
Ghini L
Atrocianus J
Walther M
Mayer H
Lorkin T
Robin J



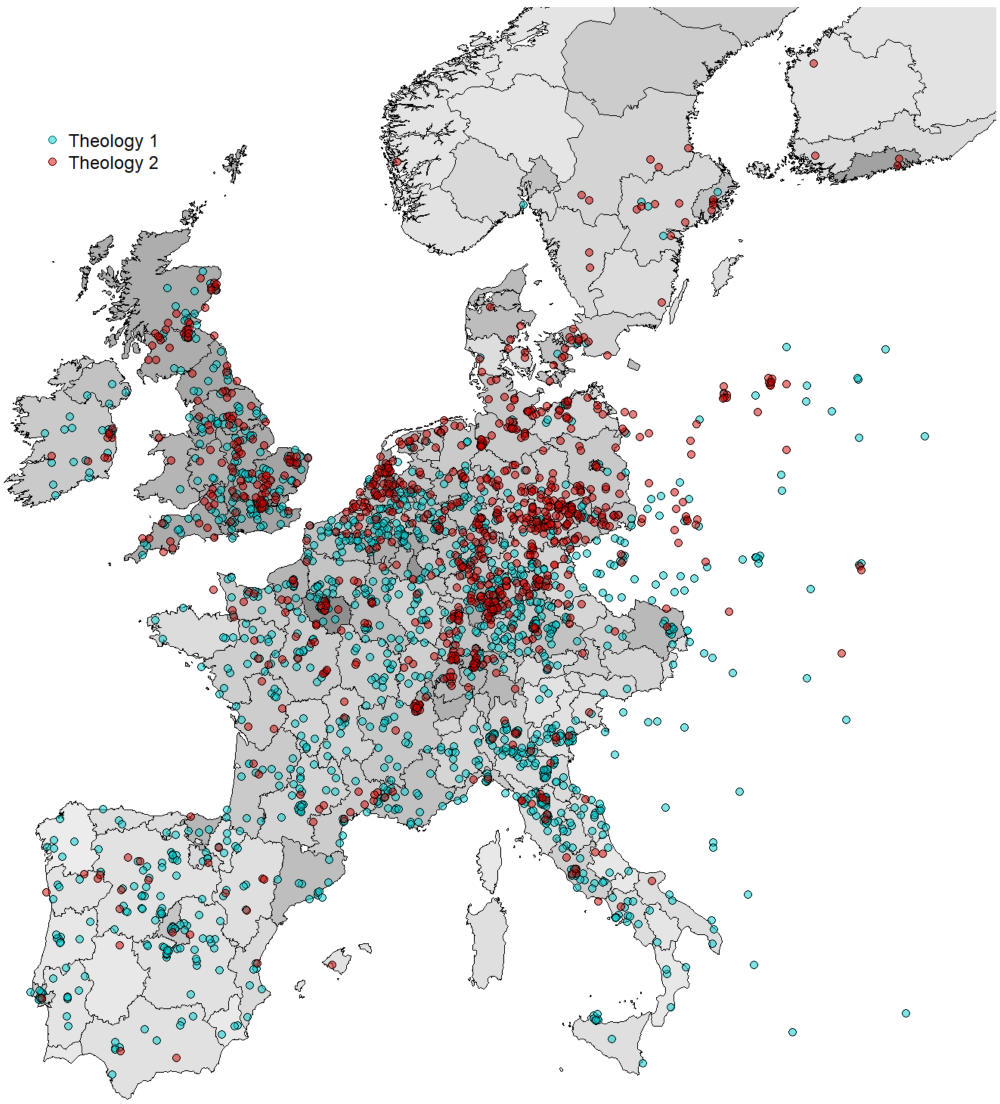


A.4 The Two Theology Clusters

As shown in Figure 7, the Theology 2 cluster is strongly concentrated in historically Protestant areas, whereas the Theology 1 cluster is less geographically clustered. Theology

2 contains notable Protestant Reformers; the top 3 scholars by publication are Luther, Melanchton, and Wesley. However, the correlation between the clusters and denomination is not perfect. For example, Theology 2 contains scholars born back to 1039, centuries before the Reformation.

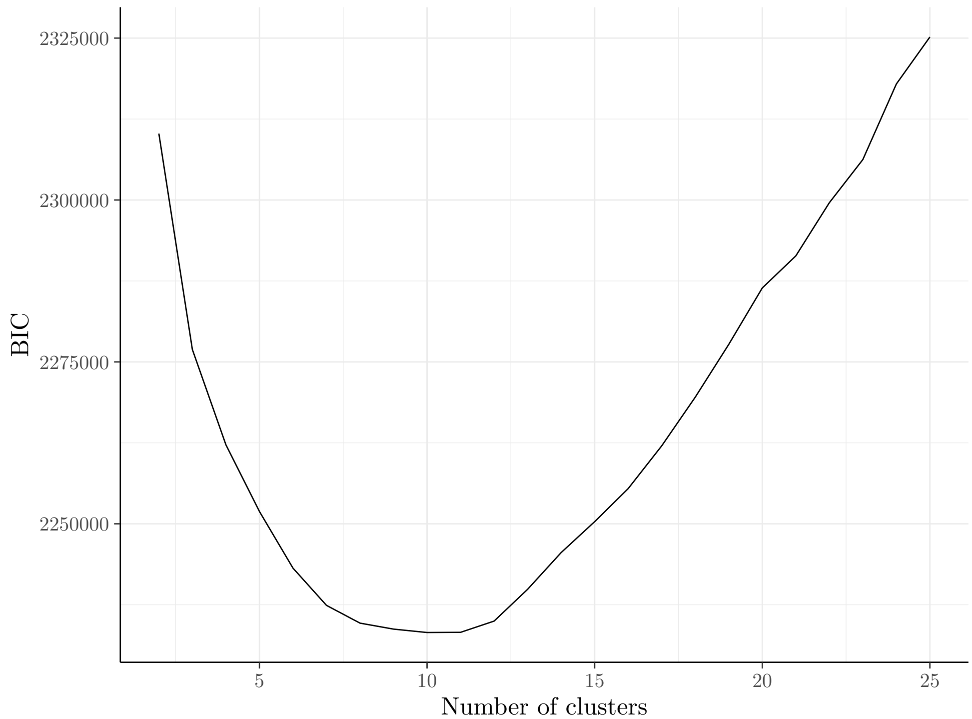
Figure 7: Birthplace of Theologians by Cluster



A.5 k-means algorithm

Figure 8 displays the BIC for the k-means algorithm estimated at different values of k. $BIC_k = TWSS_k + \log(I)Jk$, where $I = 16,149$ and $J = 1,360$. This is minimized at $k = 10$.

Figure 8: Bayesian Information Criterion for Different Values of k



Note: k-means estimated using the default R package which implements the Hartigan-Wong algorithm (Hartigan and Wong 1979). We initiate the k-means algorithm with 500 random guesses.

B Appendix: Regressions

B.1 Alternative weighting

In the main results, we weight each scholar by the fourth root of the number of publications. This assigns each scholar a weight from 1 to 18 and assumes Martin Luther is worth 18 obscure theologians. Below, we show the results for different methods of weighting: equal weights, the square root of the number of publications, and the number of publications. This is equivalent to weighting Luther equally with 1,334, or 111,660 obscure theologians. As shown in Tables 8 and 9, our results are robust to any of these alternative weighting methods.

Table 8: Unweighted regressions

	log GDP per capita, 1900				2015
	(1)	(2)	(3)	(4)	(5)
$log(n_{r,s})$	0.12***	0.12***	0.10***	0.10***	0.10***
	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)
Share Theology 1			0.05	-0.07	-0.04
			(0.23)	(0.19)	(0.20)
Share Theology 2			1.32*	0.25	-0.21
			(0.65)	(0.48)	(0.29)
Share Politics			0.43	0.08	-0.14
			(0.43)	(0.11)	(0.23)
Share Law			-0.78*	-0.25	0.06
			(0.42)	(0.20)	(0.18)
Share Science			1.28**	1.02**	0.40
			(0.46)	(0.43)	(0.28)
Share Philosophy			0.03	0.18	0.28
			(0.13)	(0.15)	(0.18)
Share Botany			0.85**	0.76***	0.61***
			(0.32)	(0.22)	(0.13)
Share Culture			-0.14	-0.45*	-0.32**
			(0.34)	(0.23)	(0.15)
Share Medicine			0.00	-0.46	-0.18
			(0.44)	(0.43)	(0.27)
N	172	172	172	172	165
Adj. R Squared	0.40	0.69	0.48	0.70	0.72
Country FEs		X		X	X

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses. The unit of observation is a NUTS2 region. Every scholar is assigned a field and to a birth region. GDP per capita from the Rosés-Wolf database on regional GDP Rosés and Wolf 2021. $\log(n_{r,s})$ is the log of the weighted total of scholars plus one. Shares are the share of the total scholars who are assigned to a given field. All regressions include as controls log ruggedness from Nunn and Puga (2012), log post-1500 calorie suitability index from Galor and Özak (2016), and the log of the region's area.

Table 9: Square root of number of publications weighted regressions

	log GDP per capita, 1900				2015
	(1)	(2)	(3)	(4)	(5)
$log(n_{r,s})$	0.09***	0.06**	0.08**	0.06**	0.04**
	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)
Share Theology 1			0.02	-0.06	-0.32
			(0.21)	(0.17)	(0.20)
Share Theology 2			0.17	-0.21	-0.09
			(0.28)	(0.22)	(0.17)
Share Politics			0.07	-0.09	-0.19
			(0.21)	(0.13)	(0.22)
Share Law			-0.75**	-0.29	-0.69***
			(0.29)	(0.22)	(0.19)
Share Science			0.51	0.34	0.36
			(0.33)	(0.30)	(0.21)
Share Philosophy			-0.12	-0.14	-0.05
			(0.18)	(0.14)	(0.28)
Share Botany			0.36	0.28	0.74***
			(0.25)	(0.29)	(0.23)
Share Culture			-0.04	-0.03	-0.34
			(0.24)	(0.20)	(0.25)
Share Medicine			-0.26	-0.45*	-0.30
			(0.28)	(0.25)	(0.31)
N	172	172	172	172	165
Adj. R Squared	0.40	0.66	0.46	0.67	0.35
Country FEs		X		X	

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses. The unit of observation is a NUTS2 region. Every scholar is assigned a field, a weight equal the square root of their number of publications, and to a birth region. GDP per capita from the Rosés-Wolf database on regional GDP Rosés and Wolf 2021. $\log(n_{r,s})$ is the log of the weighted total of scholars plus one. Shares are the share of the total scholars who are assigned to a given field. All regressions include as controls log ruggedness from Nunn and Puga (2012), log post-1500 calorie suitability index from Galor and Özak (2016), and the log of the region's area.

B.2 Controlling for academic institutions

As shown in Table 10 Column 1, there are positive associations between the presence of both academies and universities between 1000–1800. However, the associations are no longer significant after we include the total output of scholars born in the region. The presence of both institutions of course correlates with our measure of productivity (0.57 for universities, 0.54 for academies). Regardless, this suggests that the importance of universities or academies is primarially in producing scholars and scholarly output.

Table 10: Controlling for presence of academies and universities

	log GDP per capita, 1900				2015
	(1)	(2)	(3)	(4)	(5)
University	-0.03	0.07	0.00	0.07	-0.02
	(0.10)	(0.05)	(0.10)	(0.05)	(0.04)
Academy	0.02	0.05	0.06	0.07	0.04
	(0.08)	(0.04)	(0.05)	(0.04)	(0.03)
$log(n_{r,s})$	0.12**	0.08*	0.09*	0.07	0.09***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.02)
Share Theology 1			0.05	-0.06	-0.04
			(0.24)	(0.19)	(0.19)
Share Theology 2			0.99	0.05	-0.23
			(0.65)	(0.46)	(0.25)
Share Politics			0.26	-0.01	-0.17
			(0.38)	(0.14)	(0.20)
Share Law			-0.82*	-0.30	-0.00
			(0.41)	(0.21)	(0.23)
Share Science			1.11***	0.87**	0.34
			(0.37)	(0.37)	(0.24)
Share Philosophy			-0.06	0.02	0.16
			(0.13)	(0.17)	(0.18)
Share Botany			0.75**	0.66**	0.52***
			(0.27)	(0.27)	(0.15)
Share Culture			-0.05	-0.26	-0.29
			(0.36)	(0.31)	(0.17)
Share Medicine			-0.26	-0.61	-0.14
			(0.39)	(0.39)	(0.28)
N	172	172	172	172	165
Adj. R Squared	0.40	0.68	0.48	0.70	0.71
Country FEs		X		X	X

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses. The unit of observation is a NUTS2 region. Every scholar is assigned a field, a weight equal the fourth root of their number of publications, and to a birth region. GDP per capita from the Rosés-Wolf database on regional GDP Rosés and Wolf 2021. University is an indicator variables that are one if there are any active, non-corresponding scholars in a university in the region between 1000 and 1800 academy is the equivalent for academies. $\log(n_{r,s})$ is the log of the weighted total of scholars plus one. Shares are the share of the total scholars who are assigned to a given field. All regressions include as controls log ruggedness from Nunn and Puga (2012), log post-1500 calorie suitability index from Galor and Özak (2016), and the log of the region's area.

B.3 Place of activity

In our main regression analysis, presented in Table 2, we concentrate on GDP per capita at the scholar's birth location. Table 11 displays the results when scholars are aggregated based on their place of activity rather than their place of birth.

Table 11: Activity location regressions

	log GDP per capita, 1900				2015
	(1)	(2)	(3)	(4)	(5)
$log(n_{r,s})$	0.05	0.04*	0.06*	0.04	0.09***
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
Share Theology 1			0.12	-0.02	-0.08
			(0.23)	(0.15)	(0.10)
Share Theology 2			0.35	0.03	-0.28
			(0.59)	(0.26)	(0.23)
Share Politics			0.33	0.24	0.09
			(0.38)	(0.23)	(0.16)
Share Law			-0.78	-0.16	-0.17
			(0.62)	(0.34)	(0.32)
Share Science			0.23	0.35	-0.44*
			(0.52)	(0.48)	(0.22)
Share Philosophy			0.72*	0.45	0.07
			(0.40)	(0.27)	(0.75)
Share Botany			0.58	-0.21	0.58
			(0.76)	(0.60)	(0.73)
Share Culture			0.06	0.22	-0.26
			(0.37)	(0.50)	(0.25)
Share Medicine			-0.11	-0.11	-0.71***
			(0.40)	(0.13)	(0.18)
N	106	106	106	106	101
Adj. R Squared	0.33	0.73	0.33	0.72	0.65
Country FEs		X		X	X

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Robust standard errors in parentheses. The unit of observation is a NUTS2 region. Every scholar is assigned a field and to one or more activity regions (see text). GDP per capita from the Rosés-Wolf database on regional GDP Rosés and Wolf 2021. $\log(n_{r,s})$ is the weighted total of scholars. Shares are the share of the total scholars who are assigned to a given field. All regressions include as controls log ruggedness from Nunn and Puga (2012), log post-1500 calorie suitability index from Galor and Özak (2016), and the log of the region's area.

B.4 IV regressions

In our main regression analysis, presented in Table 3, we pool the three migration waves into one instrument. Tables 12, 13, and 14 display the results when each wave is used as a separate instrument.

Table 12: Byzantine scholars as an instrument for total scholarly output

	log GDP per capita, 1900		log total scholarship	log GDP per capita, 1900
	OLS	IV	1st-stage	OLS
$log(n_{r,s}+1)$	0.10***	0.16***		
	(0.03)	(0.04)		
1+ Byzantine botanist / scientist				1.27***
				(0.24)
1+ Byzantine scholars			1.32***	1.26***
			(0.39)	(0.37)
N	172	172	172	172
Adj. R Squared	0.68	0.66	0.55	0.55
Country FEs	X	X	X	X
1st stage F-stat.		26.34		

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors clustered by country in parentheses. The unit of observation is a NUTS2 region. Every scholar is assigned a weight equal the fourth root of their number of publications and to a birth region. GDP per capita from the Rosés-Wolf database on regional GDP (Rosés and Wolf 2021). $\log(n_{r,s,i})$ is the weighted total of scholars born in the region 1700–1800. 1+ refugee scholars is an indicator that is one if a refugee scholar was active or died in the region; see text for details. All regressions include as controls log ruggedness from Nunn and Puga (2012), the log post-1500 calorie suitability index from Galor and Özak (2016), and the log of the region's area.

Table 13: British Catholic scholars as an instrument for total scholarly output

	log GDP per capita, 1900		log total scholarship	log GDP per capita, 1900
	OLS	IV	1st-stage	OLS
$log(n_{r,s}+1)$	0.10***	0.16**		
	(0.03)	(0.07)		
1+ British Catholic botanist / scientist				0.73
				(0.45)
1+ British Catholic scholars			0.94***	0.77**
			(0.25)	(0.26)
N	172	172	172	172
Adj. R Squared	0.68	0.65	0.53	0.54
Country FEs	X	X	X	X
1st stage F-stat.		19.02		

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors clustered by country in parentheses. The unit of observation is a NUTS2 region. Every scholar is assigned a weight equal the fourth root of their number of publications and to a birth region. GDP per capita from the Rosés-Wolf database on regional GDP (Rosés and Wolf 2021). $\log(n_{r,s,i})$ is the weighted total of scholars born in the region 1700–1800. 1+ British scholars is an indicator that is one if a refugee scholar was active or died in the region; see text for details. All regressions include as controls log ruggedness from Nunn and Puga (2012), the log post-1500 calorie suitability index from Galor and Özak (2016), and the log of the region's area.

Table 14: Huguenot scholars as an instrument for total scholarly output

	log GDP per capita, 1900		log total scholarship	log GDP per capita, 1900
	OLS	IV	1st-stage	OLS
$log(n_{r,s}+1)$	0.10***	0.13***		
	(0.03)	(0.02)		
1+ Huguenot botanist / scientist				0.57*
				(0.29)
1+ Huguenot scholars			1.06***	0.89***
			(0.10)	(0.15)
N	172	172	172	172
Adj. R Squared	0.68	0.67	0.58	0.58
Country FEs	X	X	X	X
1st stage F-stat.		37.09		

Note: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors clustered by country in parentheses. The unit of observation is a NUTS2 region. Every scholar is assigned a weight equal the fourth root of their number of publications and to a birth region. GDP per capita from the Rosés-Wolf database on regional GDP (Rosés and Wolf 2021). $\log(n_{r,s,i})$ is the weighted total of scholars born in the region 1700–1800. 1+ refugee scholars is an indicator that is one if a refugee scholar was active or died in the region; see text for details. All regressions include as controls log ruggedness from Nunn and Puga (2012), the log post-1500 calorie suitability index from Galor and Özak (2016), and the log of the region's area.

B.5 A Model

We consider one region. Production combines physical and human capital, K and H, through a Cobb-Douglas function:

$$Y = AK^{\alpha}H^{1-\alpha}$$

$A > 0$ denotes total factor productivity, and $\alpha \in (0, 1)$ is the elasticity of output to physical capital. The function displays constant returns to scale (homogeneity of degree one) with respect to K and H.

Output Y can be used for consumption C, or investment in physical or human capital, I_K or I_H. The economy's resource constraint is

$$Y = C + I_K + I_H$$

$$\dot{K} = I_K - \delta_K K,$$

$$\dot{H} = I_H - \bar{\delta}_H H.$$

The depreciation rate $\delta_K \in [0, 1)$ is a parameter, while $\bar{\delta}_H$ is a function which depends on the stock of inspiration per unit of human capital S/H, with elasticity $-\eta$:

$$\bar{\delta}_H = \delta_H \left(\frac{S}{H}\right)^{-\eta}.$$

$\delta_H \in [0, 1)$ is a parameter. Behind the function $\bar{\delta}_H$ there is the idea that if the "stock of inspirations" is important, people benefits more from what has been studied in the past, and the actual human capital depreciates less. The stock of inspirations reflects the past history of human capital. It evolves according to

$$\dot{S} = I_H - \delta_S S, \tag{6}$$

with parameter $\delta_S \in [1, \delta_H)$. Inspiration is built from the same investment in human capital than actual human capital, but it depreciates at a slower rate. It thus particularly reflects all the human capital history of the region.

The endogenous depreciation rate of human capital introduces an externality. Although the region anticipates perfectly the path of $\bar{\delta}_H$, it fails to recognize the effect of its own action on its dynamics. The region maximizes the discounted flow of utility. Utility is defined over consumptions according to $u(C)$. The future is discounted at rate ρ. The maximization program of the region can be expressed with the following present-value Hamiltonian:

$$J = u(C)\exp(-\rho t) + \nu(I_K - \delta_K K) + \mu(I_H - \bar{\delta}_H H) + \omega(AK^{\alpha}H^{1-\alpha} - C - I_K - I_H)$$

The variable ω is the shadow price of income, while ν and μ are the shadow prices of physical and human capital. The optimal for the controls C, I_K and I_H should satisfy:

$$\frac{\partial J}{\partial C} = u'(C)\exp(-\rho t) - \omega = 0 \tag{7}$$

$$\frac{\partial J}{\partial I_K} = \nu - \omega = 0 \tag{8}$$

$$\frac{\partial J}{\partial I_H} = \mu - \omega = 0 \tag{9}$$

The differential equations for the shadow values of the state variables K and H are:

$$\frac{\partial J}{\partial K} = \dot{\nu} = \omega\alpha A K^{\alpha-1} H^{1-\alpha} - \nu\delta_K \tag{10}$$

$$\frac{\partial J}{\partial H} = \dot{\mu} = \omega(1-\alpha) A K^{\alpha} H^{-\alpha} - \nu\bar{\delta}_H \tag{11}$$

The transversality conditions are

$$\lim_{t\to\infty} K(t)\nu(t) = 0 \quad \text{and} \quad \lim_{t\to\infty} H(t)\mu(t) = 0$$

From Equations (8) and (9) $\nu = \mu = \omega$. Differentiating Equation (7) with respect to time and substitute for ω from (10) we get the standard condition for choosing consumption over time. Moreover, from (10) and (11) we find that the net marginal product of physical capital should equal the net marginal product of human capital:

$$\alpha A K^{\alpha-1} H^{1-\alpha} - \delta_K = (1-\alpha) A K^{\alpha} H^{-\alpha} - \bar{\delta}_H$$

This relation defines an implicit function $g(\cdot)$ relating $\frac{H}{K}$ to $\frac{S}{H}$:

$$\frac{H}{K} = g\left(\frac{S}{H}\right) \quad \Leftrightarrow \quad \alpha A \left(\frac{H}{K}\right)^{1-\alpha} - (1-\alpha) A \left(\frac{H}{K}\right)^{-\alpha} = \delta_K - \delta_H \left(\frac{S}{H}\right)^{-\eta}.$$

The function $g(\cdot)$ is unambiguously increasing in its argument. We can now rewrite output at time t as:

$$Y = AK^{\alpha}\left(K\, g\left(\frac{S}{H}\right)\right)^{1-\alpha} \quad = AK\left(g\left(\frac{S}{H}\right)\right)^{1-\alpha}$$

Output thus obeys an AK production function with an additional factor depending on the stock of inspirations per unit of human capital. Two regions with identical TFP and stock capital can thus differ in output: the one with the highest stock of inspirations will produce more.

In the long-run, however, this relation does not hold anymore. A balanced growth path is characterized by $\dot{C} = \dot{Y} = \dot{K} = \dot{H} = \dot{S}$. $\dot{H} = \dot{S}$ implies that $\bar{\delta}_H H = \delta_S S$, which leads to the long run stock of inspirations per unit of human capital

$$\frac{S}{H} = \left(\frac{\delta_H}{\delta_S}\right)^{\frac{1}{1+\eta}}$$

and

$$\frac{H}{K} = \frac{\delta_H}{\delta_K} \left(\frac{\delta_H}{\delta_S} \right)^{\frac{\eta}{1+\eta}} \frac{1-\alpha}{\alpha}.$$

Analyzing the stability properties of the balanced growth path would involve linearizing a system of equations including the dynamics of inspirations (6). Under the condition $\eta = 0$ (no inspirations), the growth rates of all variables immediately reach their balanced growth path values, and there is no transitory dynamics. In that case, the model behaves like a standard AK model. Using a continuity argument, the dynamics converge to the balanced growth path provided that η is small enough.

This approach can be generalized to accommodate more than one type of human capital (the academic fields).