# THE ACADEMIC MARKET AND THE RISE OF UNIVERSITIES IN MEDIEVAL AND EARLY MODERN EUROPE (1000–1800)

**David de la Croix**
IRES/LIDAM, UCLouvain, Belgium; CEPR, London

**Frédéric Docquier**
LISER, Esch-sur-Alzette, Luxembourg; University of Luxembourg, Luxembourg

**Alice Fabre**
Aix-Marseille Univ., CNRS, AMSE, Marseille, France

**Robert Stelter**
University of Basel, Switzerland; Max Planck Institute for Demographic Research, Rostock, Germany

**Abstract**
We argue that market forces shaped the geographic distribution of upper-tail human capital across Europe during the Middle Ages, and contributed to bolstering universities at the dawn of the Humanistic and Scientific Revolutions. We build a unique database of thousands of scholars from university sources covering all of Europe, construct an index of their ability, and map the academic market in the medieval and early modern periods. We show that scholars tended to concentrate in the best universities (agglomeration), that better scholars were more sensitive to the quality of the university (positive sorting) and migrated over greater distances (positive selection). Agglomeration, selection, and sorting patterns testify to an integrated academic market, made possible by the use of a common language (Latin). (JEL: N33, O15, I25)

*The editor in charge of this paper was Romain Wacziarg.*

Acnowledgements: We thank the editor and three anonymous referees for very helpful comments and suggestions. Technical help was obtained from Yves Croissant (the author of the package mlogit) and Alain Guillet (who reprogrammed the simulation command for a nested-logit model). The project benefited from detailed discussions with Ran Abramitzky, Thomas Baudin, Simone Bertoli, Matteo Cervellati, Gregory Clark, Zvi Eckstein, Sven Feldman, James Fenske, Martin Fernandez-Sanchez, Mattia Fochesato, Oded Galor, Paola Giuliano, Paula Gobbi, Simone Meraglia, Omer Moav, Simone Moriconi, Luca Pensieroso, Giovanni Peri, Itay Saporta, Yannay Spitzer, Gonzague Vannoorenberghe, Nico Voigtlaender, Romain Wacziarg, and Marlous van Waijenburg, and from comments at the following conferences and seminars: U. Manchester, WEHC Boston, ULB Brussels, TSE Toulouse, University of Luxembourg, UAB Barcelona, LISER (5th workshop on Economics of Migration), PSE Paris (Conference on Culture, Institutions and Prosperity), NYU Abu Dhabi, Deep Rooted Factors of Growth (Brown), CEPR Economic History Workshop (Tarragona), CEMIR Workshop on Migration Research (Munich), 12th Migration and Development Conference (Madrid), U Bologna (1st workshop on Economic History), IDC (Herzliya), Tel Aviv, Hebrew U (Jerusalem), Lille, UCLA, UC Davis, and Standford U. The first two authors acknowledge

E-mail: david.delacroix@uclouvain.be (de la Croix); Frederic.Docquier@liser.lu (Docquier); alice.fabre@univ-amu.fr (Fabre); robert.stelter@unibas.ch (Stelter)

*Journal of the European Economic Association* 2024 22(4):1541–1589
https://doi.org/10.1093/jeea/jvad061

**Teaching Slides**
A set of Teaching Slides to accompany this article is available online as Supplementary Data.

## 1. Introduction

Scholars and universities are widely believed to have played significant roles in the Rise of the West, as noted by Mokyr (2016) and Cantoni and Yuchtman (2014). We contend that the integration of an academic market during the pre-industrial era served as a formidable institution, enabling them to operate together. While establishing a definitive causal link across such a vast expanse of time and geographical area remains challenging, our simulations lend support to the hypothesis that universities played a crucial role in generating knowledge during the emergence of European dominance. This, in turn, potentially paved the way for the Enlightenment, humanistic movements, and scientific revolutions. Our results shed light on the importance of medieval roots in fostering scientific output, confronting qualitative studies on the subject with unique data sources, and sound estimates.

Universities are one of the most original creations of the Western Latin civilization during the Middle Ages, from the 11th century onwards.[1] They came into existence when society recognized that masters and students as a collective (*universitas* means community) had legal rights. Universities are voluntary, interest-based, and self-governed permanent associations (Greif 2006). As highlighted in Rashdall (1895), “such Guilds sprang into existence, like other Guilds, without any express authorisation of King, Pope, Prince, or Prelate. They were spontaneous products of the instinct of association that swept over the towns of Europe in the course of the 11th and 12th centuries.” Near the end of the 12th century, foreign law students at Bologna formed a union for the purpose of protection from discrimination by the town against foreign residents. At about the same time, teachers in Paris formed a corporation. Universities then began to spread across Europe, either through secession from existing ones (Cambridge from Oxford, Padua from Bologna, Orléans from Paris, etc.), or through creation *ex nihilo*. Some universities were founded from scratch by a higher authority (the University of Naples was arguably the first of this kind), but all followed

financial support from the EOS program of the Flemish (FWO) and French-speaking (FRS-FNRS) communities of Belgium (convention 30784531 on “Winners and Losers from Globalization and Market Integration: Insights from Micro-Data”). David de la Croix thanks 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.” Alice Fabre thanks the funding from the French government under the “France 2030” investment plan managed by the French National Research Agency (reference ANR-17-EURE-0020) and from Excellence Initiative of Aix-Marseille University—A*MIDEX. Robert Stelter acknowledges financial support from the Max Geldner foundation.

1. A few notable exceptions outside Europe: the Buddhist university of Nalanda in India, where both students and masters are known to come from distant places (Monroe 2000), and the University of Baghdad, which was destroyed by the Mongol invasion in 1258 CE.

the guild-like organizational principles of Bologna and Paris. Even at the Imperial Moscow University (established in 1755, charter of 1804), the rector was elected by his peers, not nominated by the emperor.

The European academic world in the medieval and early modern era provides a rich background for identifying location patterns within the upper tail of the skill distribution. The use of Latin facilitated mobility and, despite the political fragmentation of Europe, medieval universities were recognized for their independence and intellectual unity. The integration of the academic market was even formalized via the *licentia ubique docendi* (licence to teach everywhere), granted by the Church to the universities at the end of the 13th century, and conferring the right to teach at every university in Europe once a doctoral degree had been awarded (Hermans and Nelissen 2005). Understanding the mobility of academic scholars in that period matters because it potentially influenced the creation of knowledge in pre-industrial times, as well as technological and institutional progress.

Focusing on a period from 1000 CE to 1800 CE, our paper investigates whether location decisions were associated with distance and with measures of individual and institutional quality.[2] We distinguish three notions of quality. The *human capital of an academic scholar* is built from her/his achievements as seen today in the catalog of world libraries (Worldcat). The *notability of a university* in a given year is built from the human capital of its five best scholars who ceased their activity just before that year. The *simulated output of a university* is the aggregation of the human capital of all scholars who were predicted to work there in a given year.

Although the economic literature has looked at the characteristics of migrant workers at different periods in history, little is known about the mobility of upper-tail human capital in general, and about the internationalization of medieval and early modern European universities in particular.[3] To tackle this question, we develop a unique database that provides geolocalized information on the origin of thousands of academic scholars, on the location of universities, and on measures of individual human capital and institutional notability. We use it to estimate the effects of distance, the human capital of scholars, the notability of universities, and the attractiveness of European cities on location decisions. More specifically, we test (i) whether academic scholars tended to concentrate in the best universities in medieval Europe (*agglomeration*), (ii) whether those with more human capital were more likely to settle in more prestigious universities and/or in more attractive cities (*positive sorting*), and (iii) whether they were more mobile than others (*positive selection*).[4] We finally use our estimated location choice model to compute the potential gains in the output of

2. Although the official creation date of the first university (Bologna) is 1088 CE, many universities were active before they were formally recognized.

3. We do not observe a clear separation between local and international markets in academia, contrary to most other kinds of markets before the industrial revolution (Polanyi 1944).

4. In its common meaning, sorting is any process of arranging items systematically, and has two common, yet distinct meanings: ordering (arranging items in a sequence ordered according to a criterion) and categorizing (grouping items with similar properties). In the migration literature, *positive sorting* means that individuals with better attributes tend to concentrate in regions where returns are higher. In biology, positive natural selection is the force that drives the increase in the prevalence of advantageous traits. In

universities resulting from the agglomeration, positive sorting, and positive selection of academic scholars.

Our database builds on secondary sources (i.e. books and catalogs recovering information from institutional archives) and biographical dictionaries. It documents the mobility and the human capital of about 48,000 academic scholars over the whole period 1000–1800. The location choice set varies across years, as new universities were created or disappeared over time. On average, each scholar selected his/her optimal place of work out of about 100 possible locations. After excluding scholars with an unknown birthplace and universities with not enough information, the database used in our regression models includes about 3.5 million possible dyads (i.e. scholar–university pairs). By studying the mobility patterns of academic scholars at universities in the medieval and early modern periods, we capture a substantial part of upper-tail human capital. The two other—less numerous—groups were the members of scientific academies that developed in Europe in the 17th century (preceded by the Renaissance academies in Italy in the 16th century), and the scholars making a living at the courts of princes, kings, or bishops.

We estimate the mobility patterns using a multinomial logit model, and several variants accounting for sample biases, heterogeneous effects, and endogeneity issues. We show that agglomeration forces are at work: the destination choice of academic scholars depended on distance, on the notability of the university, and on the communal freedom enjoyed by the city (used as a proxy for local democracy). We also find robust evidence that better scholars were less sensitive to distance (positive selection) and more sensitive to the attractiveness of the university (positive sorting). Agglomeration and sorting patterns testify to the existence of a functioning academic market in Europe.[5]

Among the forces we consider, agglomeration and positive sorting are the most emblematic forces witnessing the competition among universities to attract talent. By attracting the best scholars at the same place, the academic market is a powerful engine to exploit the complementarities between scholars in the production function of universities and foster knowledge growth.[6] They played an important role when there were few universities. Agglomeration and sorting substantially helped universities to create knowledge at the dawn of the Scientific Revolution and during the subsequent European primacy. These effects became negligible later when the number of universities increased. By contrast, selection patterns tended to scatter talent across universities, and hardly influenced the aggregate production of knowledge.

our location choice model, we test whether better scholars are less sensitive to the distance from their birthplace, what we refer to as *positive selection*.

5. For Italy, historians, such as Cobban (1985) or Denley (2013), have documented the existence of a competitive academic market. Denley (2013), for instance, notes that the challenge for medieval Italian universities was "to attract sufficiently high-profile faculty (...) to remain open". The transition to salaried teaching staff was a direct and early consequence of a competitive system with, what he calls, its own "transfer season". In addition to the salary, several other mechanisms added to this competition, such as the privileges granted to scholars, copied from one institution to another, the organization of intellectual competitions, benefits in kind, and so forth.

6. As in the O-Ring theory of development (Kremer 1993) but not necessarily that strong.

Our paper speaks to three strands of literature. Firstly, we contribute to the literature on stagnation to growth and on the role of upper-tail human capital. Many authors have searched for the profound causes of the "Rise of the West" (e.g. Landes 1998; Galor and Moav 2002; Maddison 2007; Mokyr 2010; Galor 2011; Mokyr 2016).[7] For most of them, the self-reinforcing dynamics of technological and institutional progress played a key role. In particular, de la Croix, Doepke, and Mokyr (2018) argue that superior institutions for the creation and dissemination of productive knowledge help explain the European advantage in the medieval and early modern periods. The outstanding debate concerns the key forces that made these virtuous circles possible. There are currently no global quantitative analyses of the historical effect of upper-tail human capital on the dynamics leading to the Industrial Revolution. Recent country-level studies include Dowey (2017) for England, Squicciarini and Voigtländer (2015) for France, Dittmar and Meisenzahl (2020) and Cinnirella and Streb (2017) for Germany, and Blasutto and de la Croix (2023) for Italy. Squicciarini and Voigtländer (2015) show that the number of people who subscribed to Diderot's and d'Alembert's *Grande Encyclopédie* in 18th-century France predicts economic development later on, both at the city and county levels. Dittmar and Meisenzahl (2020) show that German cities that adopted better institutions following the Reformation grew faster and had more people recorded as famous in the German biography database.

The emergence of new scientific developments in the 16th century created a conflict with the traditional Aristotelian approach taught at universities, casting doubt on the precise role played by these institutions. Still, following Applebaum (2003), 87% of the scientists listed in the *Dictionary of Scientific Biography* born between 1450 and 1650 were university educated, and 45% of them were employed by universities. Beyond science, medieval universities may have contributed to the rise of the West through (i) the revival of Roman law, which was better suited to regulating complex economic transactions than the prevailing customary law, (ii) the translation of philosophical and scientific works from Classical Arabic and Greek, (iii) the diffusion of scientific thinking in Europe (e.g. Ockham's parsimony principle, Duns Scotus's logic, or Roger Bacon's empiricism), and a certain universality of curricula, (iv) the promotion by theologians of cultural norms such as the nuclear family, strict monogamy de la Croix and Mariani (2015), and the education of children (Thomas Aquinas), and (v) the interest in the natural sciences, reflected in the establishment of botanical gardens next to medical faculties.[8] A recent work by Dittmar (2019) lends credence to the idea of

7. Specifically, Galor and Moav (2002) explicitly refer to the universities: "Further, unlike the existing literature, investment in human capital increased gradually in the Pre-Industrial Revolution era due to a gradual increase in the representation of individuals who have higher valuation for offspring's quality. (...) In particular, in the Pre-Industrial Revolution era, the increase in the number and size of universities in Europe since the establishment of the first university in Bologna in the 11th century had significantly outpaced the growth rate of population."

8. Cantoni and Yuchtman (2014) show that university training in Roman law played an important role in the establishment of markets during the "Commercial Revolution" in medieval Europe. To establish this, Cantoni and Yuchtman determined the enrollment rates of German students at the universities of Bologna, Paris, Padua, Orléans, Prague, Heidelberg, Cologne, and Erfurt.

higher productivity of university scholars during the Renaissance. Dittmar computes the real wage of Italian professors during the Renaissance from archived payrolls, and shows that the premium of those involved in the new sciences increased after the adoption of the movable-type printing press. To our knowledge, this is the only paper other than ours focusing on university professors and using individual-level data. Closely related to our work, Schich et al. (2014) use birth and death locations of more then 150,000 notable individuals to investigate the cultural determinants of intellectual mobility and the dynamics of cultural centers over a period of 2,000 years.

Beyond the existence of universities and the role of elites, we stress what makes them operate better together, namely the integrated academic market. Higher education institutions and elites are present as soon as a civilization reaches a certain level of sophistication, but European universities were unique as they were bottom-up institutions operating in a continental market without many barriers (common language, political fragmentation of Europe, universality of curricula). This allowed top scholars to sort and concentrate, boosting thereby complementarities between them as well as the output of the whole academic system.

Secondly, our paper relates to the migration literature in general, and to historical migration in particular. Migration is a selective process, with some individuals choosing to leave their region of birth and others choosing to stay. Who moves and who stays depends on the costs and benefits of migration, which can vary across individuals for both systematic and idiosyncratic reasons. Two salient features of contemporary labor mobility are that well-educated people exhibit a much greater propensity to emigrate than the less educated, and they tend to agglomerate in countries/regions with high rewards to skill (Grogger and Hanson 2011; Beine, Docquier, and Ozden 2011; Kerr et al. 2016; Kerr et al. 2017). Migrant selection has also been examined in historical studies, most of them focused on the *Age of Mass Migration* to the United States, a period of unrestricted entry starting in 1850 and ending around 1920.[9] Abramitzky, Boustan, and Eriksson (2012) and Abramitzky, Boustan, and Eriksson (2014) and Spitzer and Zimran (2018) show that selection patterns are consistent with income-maximization models. In the 19th and early 20th centuries, migration to the United States was positively selected from some European countries and negatively selected from others. The differences in selection lined up with those in the relative returns to skill across sending countries, or with the easing or tightening of the liquidity constraints (Covarrubias, Lafortune, and Tessada 2015). Using data on servitude contracts from the 17th and 18th centuries, Abramitzky and Braggion (2006) found similar self-selection patterns (on health, physical strength, and literacy) of servants to the American colonies.

Thirdly, we shed light on the mobility patterns at the upper tail of the human capital distribution. Despite the potentially far-reaching implications for international

9. A few studies on intra-European migration support the positive selection hypothesis. Beltrán Tapia and de Miguel Salanova (2017) show that, in the late 19th and early 20th centuries, the literacy level was higher among internal migrants moving to the Spanish capital city than among those who remained in their provinces of origin.

knowledge creation and diffusion (Pierson and Cotgreave 2000; Breschi and Lissoni 2009; Miguelez and Moreno 2013; Trippl 2013), empirical evidence about the drivers and selection of scientists' mobility remains scarce. Existing studies show that, compared to college-educated migrants, scientists and inventors are less sensitive to distance and more sensitive to linguistic proximity, economic conditions, resources dedicated to R&D, and visa-related restrictions (Laudel 2003; Kerr 2008; Agrawal et al. 2011; Miguelez, Raffo and Fink 2013; Moed, Aisati, and Plume 2013; Grogger and Hanson 2015; Zhao et al. 2022). To the best of our knowledge, none of these studies have focused on the self-selection of scientists. One of the very few studies identifying selection effects among scientists is that of Gibson and McKenzie (2014). Using a survey on the mobility of researchers from the Pacific Islands, they show that current migrants produce substantially more research than similarly skilled return migrants and non-migrants. Hoisl (2007) also shows that mobility is generally found to be positively associated with inventor productivity as proxied, for example, by the education level of the inventor and the use of external sources of knowledge such as university research or scientific literature. Finally, Akcigit, Baslandze, and Stantcheva (2016) find that the international mobility of superstar inventors is influenced by tax policies.

The remainder of this paper is organized as follows. In Section 2, we present the data sources and define the key concepts used in our analysis. In Section 3, we describe the micro-foundations of our empirical model, present our main findings, and discuss their robustness. In Section 4, we simulate the model to draw its implications for the humanistic and scientific revolutions. The conclusion is in Section 5.

## 2. Data and Concepts

We collect a large sample of academic scholars (denoted by $i = 1, \ldots, I$) employed by the universities of Latin Europe over a time span that started around the year 1000 CE and lasted until 1800 CE.[10] The year 1800 CE is a convenient date to stop for several reasons. At a broad level, it spelled the end of the Malthusian pre-industrial era. At the university level, it corresponded to profound changes: All French universities were abolished by the Revolution in 1793, and would reappear in a different form later on, together with the new "Grandes Ecoles" created by Napoleon. In Prussia, the Humboldt reform of 1810 was also a game changer.

A key advantage of selecting scholars from universities to build our list is that the sample will be drawn from a clearly defined and well-established universe (unlike in the literature based on encyclopedia; see the "famous people" in de la Croix and Licandro (2015), the "creative people" in Serafinelli and Tabellini (2022), or the "notable people" in Schich et al. (2014) or Laouenan et al. (2022)). Another advantage is to allow us to identify cases of honorary members and remove them from the sample

10. Latin Europe means Europe minus the Muslim world and the Byzantine world.

(a 20th century example is Elena Ceausescu type of publication pattern—signing all publications of Romanian universities thanks to a position of power).

In this section, we first describe the institutional data sources used to identify academic scholars and their place of work. Secondly, we present the bibliographical data sources used to characterize the lifetime and place of birth of each academic scholar. In the third and fourth parts, we define an index of individual ability or human capital for each scholar, and go into a little more detail with regard to scholars with multiple affiliations. We finally explain how we construct our proxies for institutional notability and quality.

*Institutional Secondary Sources & Quality of Sampling.* The identification of academic scholars builds mostly on institutions' secondary sources of different types (see Online Appendix A). Ideally, we aim to cover the universe of scholars involved in university teaching and research before 1800 in Latin Europe. Although this universe is more precisely defined than in other studies of European scholars, its boundaries remain somewhat flexible. For example, according to biographies of Nicolaus Copernicus, he delivered lectures as a professor of astronomy to numerous students while in Rome. It is unclear whether this teaching took place within the walls of the university of Rome (Sapienza), and how long it lasted. This appears, however, to be the only time Copernicus taught students. Should we count Copernicus when measuring the notability of the Sapienza? Probably not, as it would overestimate the attractiveness of Rome during this period. Should we include the decision of Copernicus to go to Rome in our study? We did, but it does not matter much as he is only one among thousands of scholars.

Another dimension of flexibility concerns how we define a university. This seems simple *a priori*. We can rely on Frijhoff (1996) who provides a list of institutions granting doctorate degrees, together with their official foundation date. It is, however, meaningful to extend this list in two directions. One extension is to include important learning institutions, which were not formally universities. One example is the Herborn Academy (*Academia Nassauensis*), which was a Calvinist institution of higher learning in Herborn (Germany) from 1584 to 1817.[11] In addition, another relevant extension is to consider that universities were sometimes active before their official recognition as university. For example, the University of Amsterdam was officially founded in 1877, but its roots go back to 1632, when the *Athenaeum Illustre* was founded. For this reason, our period of analysis starts before the official creation date of the first university. Online Appendix A provides, for each university, official and effective years of creation, based on the scholars we can observe there. More details on foundation dates can be found in de la Croix and Vitale (2022).

For each university, we first checked whether there is an online historical database of professors. For example, the list of professors at the University of Groningen has already been established. The *Catalogus Professorum Academiae Groninganae*

11. Online Appendix A describes the institutions added to Frijhoff's list.

includes all full professors from 1614 onwards (see the website at http://hoogleraren.ub.rug.nl/). The website is still under development, but it shows the recent interest of universities themselves in looking at their past in a more systematic way. For those universities without such a database but with books of biographies of their professors, we encoded the contents of these books. For the remaining universities, we checked whether *matricula* (people registered at a given university) and *chartularia* (containing transcriptions of original documents related to the historical events of a university) exist. In some cases, the *matriculum* itself is of little use as the status of these people is not recorded (students, professors, etc.). We also used national biographies and other databases to complete the information needed. For example, for Jesuit universities, there is a biographical dictionary by Sommervogel (1890) listing all Jesuits having published material; as they are classified by place of activity, we can match the professors to the relevant universities. Moreover, for the late Middle Ages and the Renaissance, information can be retrieved from two recent projects, both aimed at collecting biographical and social data on those who graduated from medieval universities: the project "Repertorium Academicum Germanicum–The Graduated Scholars of the Holy Roman Empire between 1250 and 1550" and the project "Studium" for the University of Paris from the 12th century to the Renaissance. Both projects are currently under development.

We grouped universities into three categories, depending on the coverage of the sources (see Online Appendix A). We say the coverage is "comprehensive" when data collection was based on an existing website or book whose aim is to list all professors of a given institution. Coverage is "broad" when it is based on the combination of several sources, including books on the history of the university. The coverage is "partial" when the sample of scholars was informed by sources from other universities and general thematic biographies. Our benchmark analysis is based on universities with comprehensive and broad coverage. Notice that the quality of the coverage is not related to the prestige of the university. We have an excellent coverage of the University of Macerata—a small university in Italy, while there is no comprehensive list of professors for the University of Paris. A key requirement of our analysis is to cover almost all scholars with high human capital, and to include a large sample of unknown scholars as well. This requirement is met by encoding the academic scholars included in thematic biographies, such as Taisand (1721) for law, Eloy (1755) for medicine, Junius Institute (2013) for Protestant theology, Herbermann (1913) for important Catholic figures, and Applebaum (2003) for the key actors of the scientific revolution.

Over the whole time span 1000–1800, we identify 198 universities and teaching institutions. In the econometric analysis, we eliminate institutions that either have a partial coverage, as explained above, or are very small (fewer than ten scholars or fewer than 0.05 professors per year of existence). We thus obtain a working sample of 138 institutions (denoted by $k = 1, \ldots, K$). Each university $k$ is linked to a geo-referenced location. Accounting for the date of creation of each university, we estimate that these 198 institutions represent a total of 50,899 years of existence. A very comprehensive list of scholars can be obtained for the University of Heidelberg, which includes 1,222 scholars over 414 years of existence. Note that Heidelberg is not the largest university in our working sample; the data related to the University Bologna allow us to identify 3,293 scholars over the whole time span. However, Heidelberg is more representative of an average university than Bologna. Assuming Heidelberg is representative of all institutions, a back-of-the envelope calculation suggests that the order of magnitude of the universe of academic scholars for the medieval and early modern periods is around 150,238 (i.e. 1,222/414 scholars per year × 50,899 years of existence). Observing that scholars taught in 1.14 universities on average, the universe has about 134,073 unique persons.

So far, our bibliographical searches have allowed us to identify 47,897 academic scholars. These include very well-known professors as well as obscure scholars. We thus estimate that our current sample covers around 35.7% of the universe (i.e. 47,897÷134,073). This coverage is very likely to be higher for renowned scholars, as they are more likely to appear in the sources consulted, than for obscure scholars. Having obscure scholars in the sample is important to identify the characteristics of the famous ones—those who are more likely to play the academic market game. Including many obscure scholars in the analysis is thus a strength of our analysis.

*Biographical Individual Data.* We match each scholar's name with bio- and bibliographical dictionaries to identify their place of birth and, later, their quality. We exclude the small number of persons born outside a rectangle encompassing Europe, North Africa, and the Middle East (defined by latitudes $\in [28, 66]$ and longitudes $\in [-22, 51]$) because those would be outliers when computing distances. We also search online for Wikipedia and Worldcat pages to generate the ex-post indicators of human capital, as explained below.

One word about the quality of the bibliographical data. In many cases, it is quite high, as the secondary sources used—biographical dictionaries and university sources—were often compiled from archive materials. We should however warn the reader that for the earlier periods, we have chosen to adopt some approximations. A good example is the oculist *Benevenutus Hierosolymitanus*, also called Benevenutus of Jerusalem. His life is totally unknown to historians, but his book, *Ars probatissima oculorum*, was immensely popular and influential—having been translated into four languages already in medieval times. From other writings citing his work, historians infer he lived between 1100 and 1290. Assigning Jerusalem as his place of birth is disputed, but seems the likeliest option, given the knowledge of Middle Eastern cultures displayed in his writings (remember that Jerusalem was for some time a Latin kingdom (1099–1187)). He was also obviously acquainted with the medical school of Salerno, and he likely taught there (being called the physician from Salerno in one manuscript, namely the *Besançon Manuscript*). The most intriguing part concerns his relation with Montpellier, another famous medical school. In 1921, the Faculty of Medicine in Montpellier placed a marble slab in its entrance hall listing him among its early faculty members. There are some arguments to link Benevenutus of Jerusalem to Montpellier, but there remains a "considerable disparity between the fragility of the documentary basis for the Montpellier inscription and the robustness of the stone on, which it was engraved" (Kedar 1995).

Each individual at university $k$ is characterized by at most five dates: year of birth, year of death, first year of observation at university $k$, last year of observation at university $k$, and approximate date of activity at university $k$ (this corresponds to a date that is sometimes denoted by "fl.", from the Latin verb *floruit*, i.e. "they flourished").[12] From these dates, we define two dates, $t_i^b$ and $t_i^f$, which hypothetically bound the active life of each scholar. These dates are computed as follows:

$$t_i^b = \min\{\max\{\text{Year of Birth} + 30, \min_k[\text{first year of obs. at univ. } k],$$
$$\min_k[\text{approximate date at univ. } k]\}, \min_k[\text{last year of obs. at univ. } k], \text{Year of Death}\}, \tag{1}$$

$$t_i^f = \max\Big\{\min\{\text{Year of Birth} + 65, \text{Year of Death}\},$$
$$\max_k[\text{first year of obs. at univ. } k], \max_k[\text{last year of obs. at univ. } k],$$
$$\max_k[\text{approximate date at univ. } k]\Big\}. \tag{2}$$

For simplicity of exposition of the stylized facts, we divide the whole observation window into eight periods, denoted by $\tau \in \{0, 1, 2, 3, 4, 5, 6, 7\}$, corresponding to major historical events: from the urban revolution to the first universities (1000–1199), from the official foundation of Paris and Oxford in 1200 to the Black Death (1200–1347), from the Black Death to the invention of the movable-type printing press (1348–1449), from the printing press to the rise of Protestantism (1450–1526), from Protestantism to the beginning of the Thirty Years' War (1527–1617), from the Thirty Years' War to the revocation of the Edict of Nantes (1618–1684), from this revocation to the rise of Enlightened universities (1685–1733), and from Enlightened universities to 1800 (1734–1800).[13] We assign each scholar to a period $\tau$ based on $t_i^b$. The beginning date $t_i^b$ should be seen as a time when the individual can make location choices. The final date $t_i^f$ will be used to map the human capital achieved by a scholar to their universities. In the stylized facts, the period in which this end date falls determines the period for which we impute the quality of the scholar to their university.

Universities' scholars were almost always male, but we found a few females (See de la Croix and Vitale (2023) for a full list and an analysis of these data): Trotula de

12. The scholars for which we have no dates cannot be incorporated into the analysis.

13. The year 1527 corresponds to the foundation of the University of Marburg, the oldest Protestant university in the world. The Thirty Years' War war was of major importance for Germanic universities and the life expectancy of their scholars, as shown in Stelter, de la Croix, and Myrskylä (2021). Finally, in 1734, the University of Göttingen was founded to propagate the ideas of the European Enlightenment.

TABLE 1. Summary statistics for scholars by period.

| Periods $\tau$ | Number of observations | Number of universities | Birthplace (%) | Wikipedia (%) | Worldcat (%) |
|---|---|---|---|---|---|
| 0 (1000–1199) | 317 | 19 | 73.2 | 53.9 | 51.4 |
| 1 (1200–1347) | 1,922 | 33 | 63.5 | 20.7 | 20.6 |
| 2 (1348–1449) | 5,110 | 53 | 69.9 | 8.9 | 8.6 |
| 3 (1450–1526) | 7,225 | 73 | 65.9 | 11.0 | 15.8 |
| 4 (1527–1617) | 9,869 | 142 | 74.2 | 21.8 | 35.8 |
| 5 (1618–1685) | 8,215 | 158 | 73.9 | 21.3 | 38.3 |
| 6 (1686–1733) | 6,081 | 152 | 71.1 | 21.1 | 42.0 |
| 7 (1734–1800) | 9,158 | 166 | 76.8 | 32.2 | 52.6 |
| Total | 47,897 | 198 | 72.1 | 20.8 | 33.8 |

Notes: Column (1) defines the period. Columns (2) and (3) report the period-specific numbers of scholars and universities covered by the database. Column (4) gives the fraction of scholars whose birthplace is known. Columns (5) and (6) give the fraction of scholars who have a Wikipedia page and at least one recorded publication in Worldcat.

Ruggiero (11th century) and a few others in Salerno, Clotilde Tambroni, and a few others in Bologna, Beatriz Galindo in Salamanca, Ekaterina Romanovna in Moscow, and Dorothea Christiane Erxleben in Halle. Female scholars were a rare exception though. Novellà Calderini, for example, allegedly replaced her father repeatedly, teaching at Bologna veiled so that her beauty would not distract the students, according to the Italian Encyclopedia *Treccani*.

Table 1 shows the number of scholars per period, with some of their characteristics. We also report the number of universities per period, which increases steadily except from periods 4 to 6, especially when French Protestant "academies" had to close (Bourchenin 1882). On average, institutional data and bibliographical dictionaries allow us to identify the birthplace of 72.1% of university professors. For them, we can compute the cost distance $d_{ik}$ associated with each possible scholar–university dyad. Such a cost is defined as $d_{ik} = \ln(\text{cost}^{\min} + \text{cost}_{ij})$ where $\text{cost}_{ij}$ is computed using Özak (2010. Özak (2010, 2018)'s human mobility index and $\text{cost}^{\min}$ is the minimum cost incurred when having a position in one's own place of birth. The human mobility index is the cost-minimizing path between two locations. It depends on geographical, technological, topographic and terrain conditions, and performs better than standard great circle distance in gravity models. We assume $\text{cost}^{\min}$ is equivalent to the cost of walking within the old city of Rome between the Vatican City and the Colosseum (3.5 km).

In addition, 20.8% of our identified scholars have a Wikipedia page, and 33.8% have at least one recorded publication in Worldcat. Overall, these shares increase from periods 2–3 (the Middle Ages) to periods 4–7 (early modern period). The highest share is found in period 0, during which we encounter the most distinguished scholars. During period 1348–1449, we find many names of professors with no publications, either because they did not publish a lot, were never printed, or their publications did not survive.

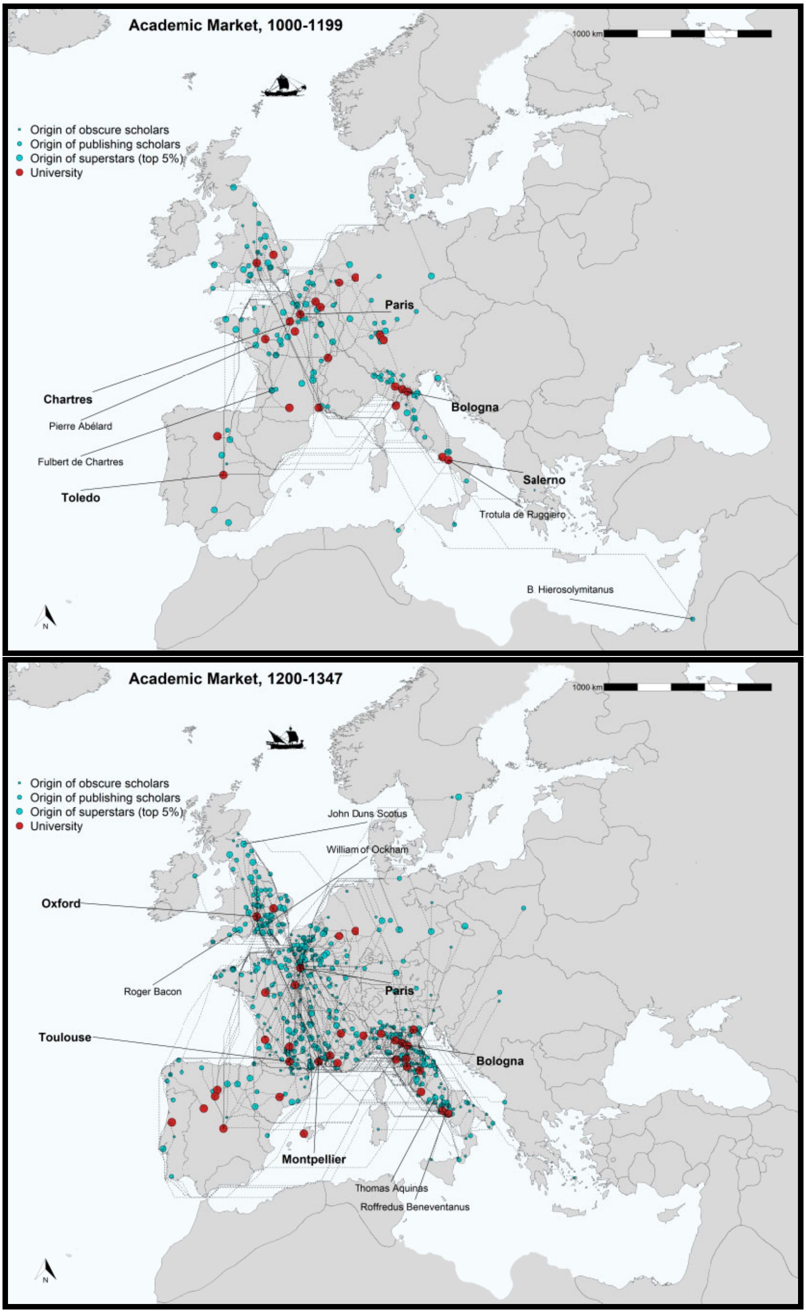
Figure 1 shows the university–scholar maps for all periods. Red dots correspond to universities. The top universities are labeled in bold. Blue dots represent scholars' birthplaces and again we have labeled some prominent names. Small blue dots refer to obscure scholars whose normalized ability index is equal to 0 (see below). The dashed lines link academic scholars to the university for which they taught. They represent the optimal (i.e. travel-time minimizing) route.

As the first two maps (1000–1199 and 1200–1347) show, universities emerged in the territory of the late Western Roman Empire. Paris clearly attracted scholars from all over Europe, from Portugal to Scotland and the south of Italy. The density of universities in Italy was already impressive. The period 1348–1449 saw a decline in the number of observations in France, probably due to the Hundred Years' War, combined with the Black Death. West German universities started to play a role, while Italy was very active. We can also see Greek scholars such as John Argyropoulos fleeing the expected fall of the Byzantine Empire (from Harris (1995)). The next period (1450–1526) has the same characteristics, but with more observations in Spain, Scotland, and southern Germany. The number of observations over the period 1527–1617 is high, with good coverage from Portugal to Poland: the portfolio of universities is expanding. The period 1618–1685 saw the development of Nordic universities, and a decline in movement in the south of Europe. A similar trend is observed for the period 1686–1733. The last period 1734–1800 is particularly rich in Germany, and universities expanded to the East. From Iceland comes Grímur Jónsson Thorkelin, who was professor of antiquities at Copenhagen University and is known for the first full translation of the poem Beowulf. On the whole, what can be seen on these eight maps corresponds closely to changes in economic primacy over time in Europe (Kindleberger 1996). More descriptive statistics can be found in Online Appendices B.1, B.2, and B.3.[14]

Using bibliographical data, we define two key concepts that characterize the notability of academic scholars and institutions, and can potentially influence location choices.

*Scholars' Human Capital.* Firstly, we construct an index of *ability or human capital of scholar* $i$, denoted by $q_i$. A standard way to measure human capital is to consider wages paid. While the latter might be the more standard measure of markets, it suffers from several limitations. Firstly, existing data on salaries are scarce, incomplete, or completely missing for several universities. Secondly, faculty compensation took a variety of forms, like payments directly from students, benefits in kind (rent, various

14. We include a breakdown of scholars by broad fields of knowledge. We were surprised to see "theology" decline from 25.5% to 12.2% between period 0 and period 3 (The Renaissance) and surge again at the occasion of the Reformation, peaking at 21.8% during period 6. It is interesting to contrast this result with the idea that the Reformation led to a secularization of the society. This secularization process is shown in Cantoni, Dittmar, and Yuchtman (2018) through the reallocation of students across fields in Germany (measured by degrees granted and first jobs). Such a reallocation did not seem to be matched by a similar process at the level of the teachers, or might be compensated by more theology in Catholic lands, under the lead of the Jesuits.




Period 1000-1199

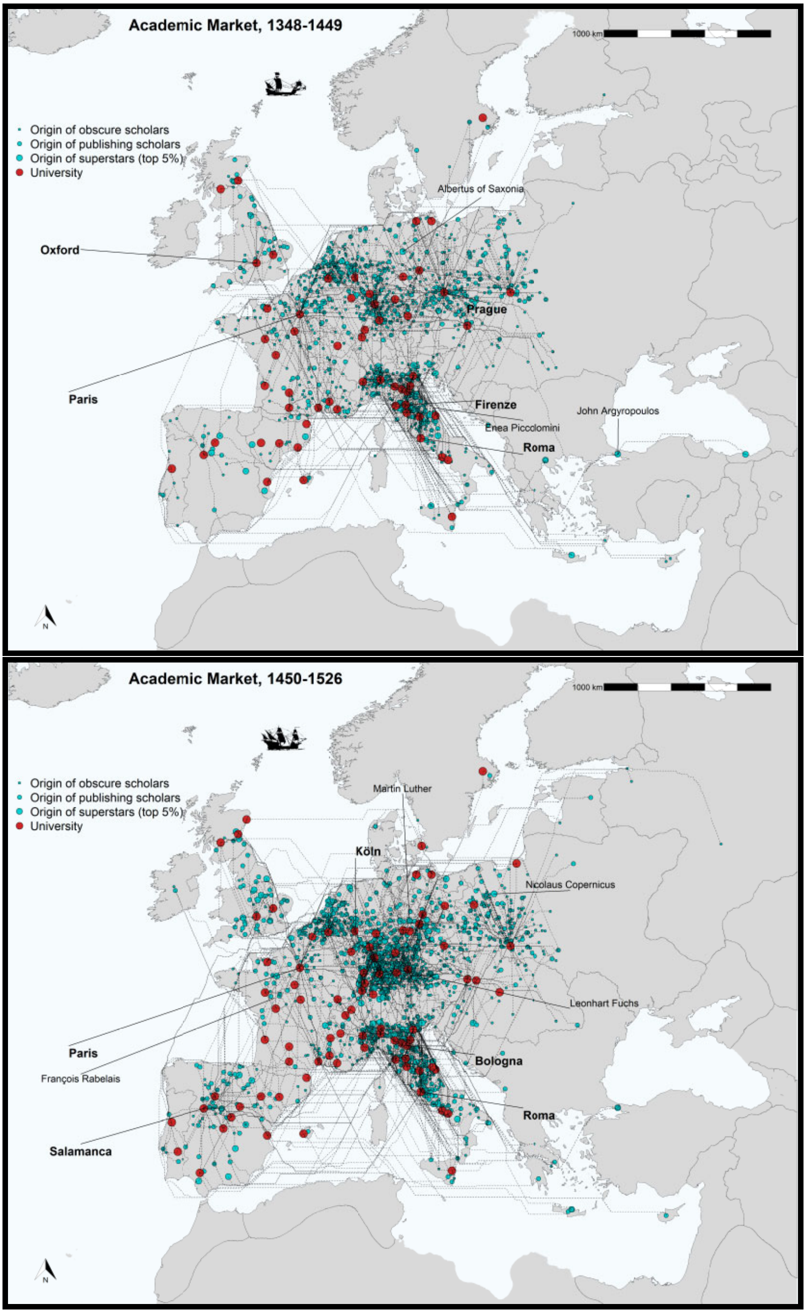
Places of top universities in bold

Includes names mentioned in the text

Dashed lines: cost-minimizing path

Period 1200-1347

FIGURE 1. Maps of scholar-university dyads by period (1/4).




Period 1348-1449

Places of top universities in bold

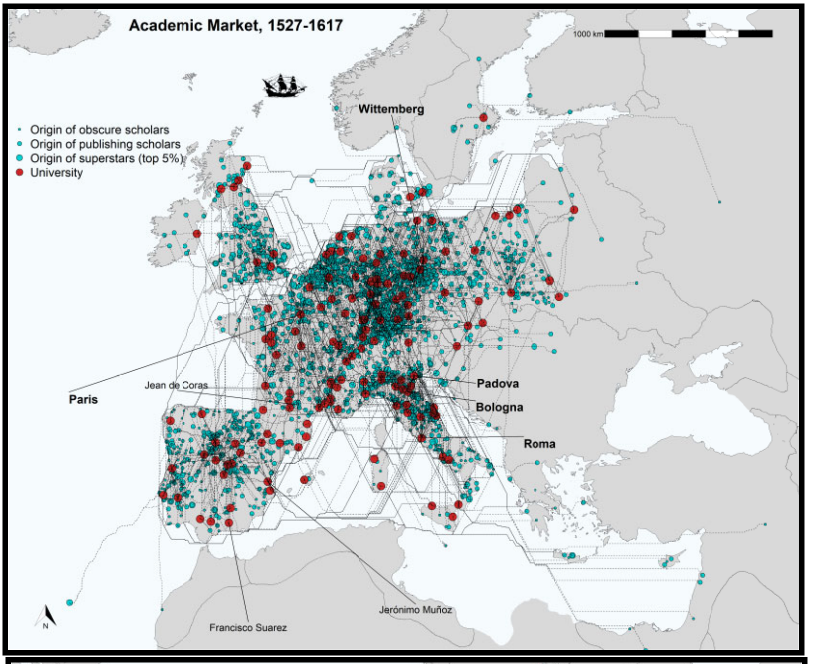
Includes names mentioned in the text

Dashed lines: cost-minimizing path

Period 1450-1526

FIGURE 1. Continued. Maps of scholar-university dyads by period (2/4).

Maps of scholar-university dyads by period (3/4)

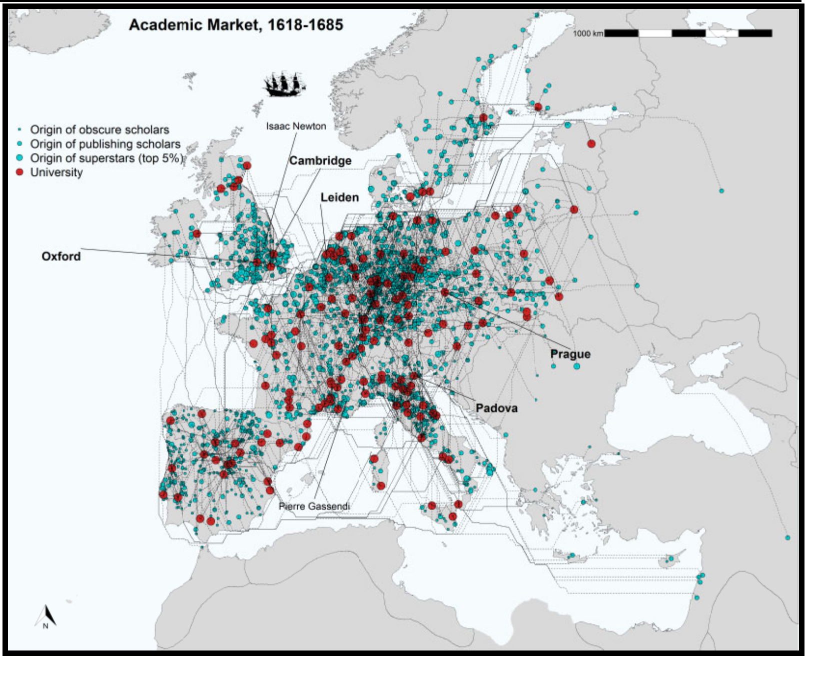



Period 1527-1617

Places of top universities in bold

Includes names mentioned in the text

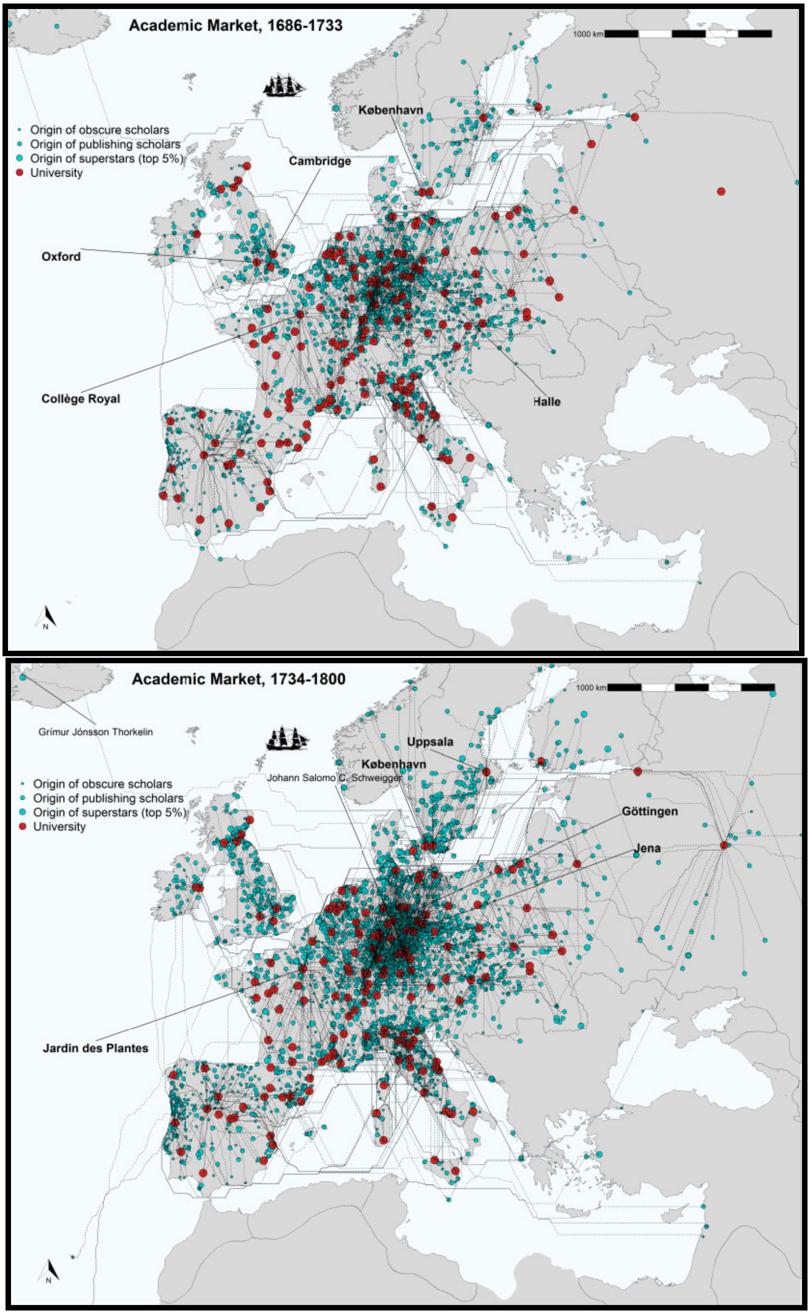
Dashed lines: cost-minimizing path




Period 1618-1685

FIGURE 1. Continued. Maps of scholar-university dyads by period (3/4).

Maps of scholar-university dyads by period (4/4)




Period
1686-1733

Places of top universities in bold

Includes names mentioned in the text

Dashed lines: cost-minimizing path

Period
1734-1800

FIGURE 1. Continued. Maps of scholar-university dyads by period (4/4).

expenses, allocations of wine, and grain like in Heidelberg (Drüll 1991), among others), fixed salaries subject to contingencies (recessionary times, will of the prince, etc.) or prestige. In addition, theology professors were often not paid for their teaching, and received *prebends* outside the university (Post 1932; Paquet 1958). All these challenges make it impossible to have sufficient informative, individual wage data. We discuss these issues in more detail in de la Croix et al. (2023), where we have compiled available wage data.

As an alternative, we measure human capital by individual notability as seen today in contemporary sources, Worldcat and Wikipedia. Worldcat provides a comprehensive measure of scientific output and citations, as books about the person are included in the measure. Wikipedia completes this measure by putting more weight on the mission of academics called, on today's terms, "service to society" (e.g. becoming an ambassador or a pope, or being canonized a saint). To combine the information provided by Worlcat and Wikipedia into one measure, we compute the first principal component of five indicators: (i) the log of the number of characters of the longest Wikipedia page across all languages (ranging from a minimum of 60 to 261,408), (ii) the log of the number of languages in which a Wikipedia page exists (ranging from a minimum of 1 to 219), (iii) the log of the number of works (by or about) in Worldcat (ranging from a minimum of 1 to 147,258), (iv) the log of the number of publication languages in Worldcat (ranging from a minimum of 1 to 56), and (v) the log of the number of library holdings in Worldcat (ranging from a minimum of 1 to 1,1232,873).[15]

In Online Appendix B.4, we present the results of the principal component analysis (see Online Appendix Table B.6). We use the first principal component as an index of scholar's human capital and normalize it such that a person with no Wikipedia page and no Worldcat entry will have a human capital of zero ($q_i = 0$). We also discuss the pros and cons of our measure of $q_i$—in particular, we show that our measure is very robust to the removal of Wikipedia indicators or to the breakdown of Worldcat publications into those by and about the person—and list the most famous scholars by period (see Online Appendix Table B.7). We find that the median value of $q_i$ from the set of positive $q_i$ (those with either a Wikipedia or Worldcat reference) shows no visible trend over time. It is likely that the writings of medieval scholars were lost compared to those of scholars active in the early modern period, yet this loss is compensated for by the accumulation of citations and new editions over time. In de la Croix et al. (2023), we reassuringly show that the correlation between wages paid and $q_i$ is positive but rather small, for the above reasons.

*Scholars with Multiple Affiliations (repeat movers).* Our database includes some scholars with multiple career spells or affiliations. Over the whole time span, 12.1% of our scholars are linked to more than one university, and the average number of

15. With regard to the first indicator, a correction for different languages length was performed, using the translations of the Gospel according to Saint Mark.

affiliations per scholar equals 1.14.[16] We denote by $S_i$ the number of universities where scholar $i$ spent time during their career (i.e. the number of career spells). The maximum number of recorded affiliations is 9. Among those with many affiliations, we find civil lawyers such as Antoine de Gouveia (nine affiliations) and Jean de Coras (seven affiliations). Jean de Coras (1513–1572) was a French jurist who taught at Padua, Toulouse, Ferrara, Valence, but also, according to Taisand (1721) at Orléans, Paris, and Angers (but we do not even know in which order).[17] Besides lawyers, Francisco Suarez (1548–1617) was a Spanish Jesuit philosopher and theologian who not only taught at Avila, Valladolid, Alcala, Salamanca, and Coimbra according to Herbermann (1913), but also at Paris and Rome according to Sommervogel (1890) (eight affiliations).

We refer to multi-affiliation scholars as *repeat movers*, and to those who have been only employed by a single university, as one-time movers. It is difficult to make any statement on the reasons for multiple moves. However, there is clear evidence that repeat movers are more likely to belong to the top of the distribution of human capital. Repeat movers are performing better than others at both the extensive and intensive margins. On average, 72.3% of repeat movers have at least one recorded publication, as opposed to 38.7% for one time movers.[18] Focusing on scholars with at least one publication, the average $q$ of repeat movers (4.139) is 25.7% greater than that of one-time movers (3.293). Combining both margins and keeping in mind that the minimal ability level is normalized to 0, the average ability index in the total population of repeat movers (2.992) is 2.35 times greater than the average ability index in the total population of one-time movers (1.274). The shares of repeat movers in the population in periods 0–7 are equal to 18.4%, 16.4%, 9.0%, 9.6%, 13.3%, 12.4%, 13.4%, and 11.3%, respectively. The greatest shares, observed in the first two periods, are likely due to a lower coverage of the population of obscure scholars.

*Institutional Notability.* In theory, we can compute a measure of annual quality for each university using the observed location and ability levels of all scholars identified in our database. In particular, for each year $t$, we can define the set of scholars affiliated with university $k$ as $\Lambda_{k,t} = \{i \,|\, \bar{p}_{ikt} = 1\}$ where $\bar{p}_{ikt}$ is a dummy equal to 1 if scholar $i$ was affiliated with university $k$ at year $t$ (implying that $t_i^b \leq t \leq t_i^f$), and equal to 0 otherwise. Nevertheless, given that sampling varies from one institution to the other, computations based on the total number of observed scholars are not directly comparable across places. Taking the means or medians of individual human capital would also be biased in favor of the least well covered universities.

16. If a scholar left a position and came back to the same institution after a while, we consider it as only one affiliation.

17. Jean the Coras might be known to the international audience as he instructed the famous trial of Martin Guerre. He wrote its best-known record, which was the basis for the movie *The Return of Martin Guerre* with Gérard Depardieu, which was nominated for Best Foreign Language Film by the U.S. National Board of Review of Motion Pictures in 1983.

18. Those shares are computed on the sample of scholars for whom the birth place is known.

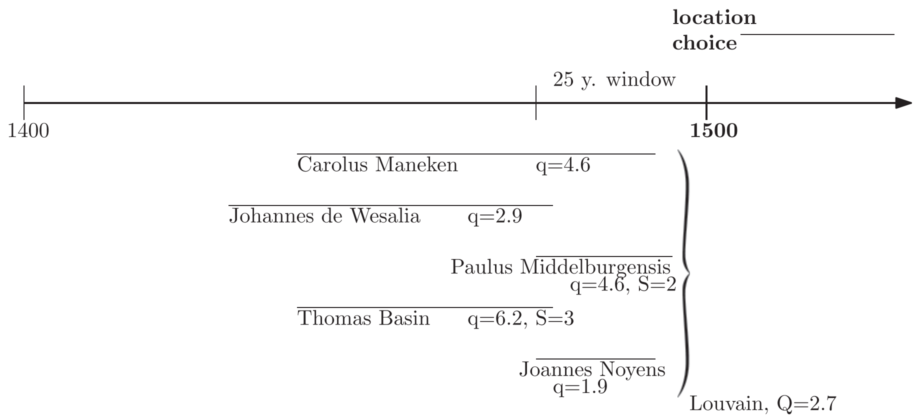



FIGURE 2. Computing the notability of the University of Louvain in 1500. For illustrative purpose, the computation of the notability of the University of Louvain in the year 1500 relies on the five Top scholars inactive in 1500 but active in the 25 years before (from 1475 to 1499). Two scholars, Paulus Middleburgensis and Thomas Basin, are repeat movers.

Hence, we introduce the concept of *notability of university* $k$ in year $t$ as a CES combination of the ability or human capital of the top five academic scholars having spent time there in a 25-year window preceding $t$, and having finished their career before $t$. Hence, these top five scholars are extracted from the sets $\Lambda_{k,t-25} \dots \Lambda_{k,t-1}$ and should not belong to $\Lambda_{k,t}$. Figure 2 shows an example. The notability index is denoted by $Q_{kt}$. To account for the partial presence of repeat movers, we weight the individual ability $q_i$ by $(1/S_i)^{\omega}$ where $S_i$ is the number of universities where scholar $i$ spent time during their career, and we define the adjusted ability level as $\bar{q}_i \equiv q_i \, (1/S_i)^{\omega}$. In our descriptive tables and benchmark regressions, we assume $\omega = 1$ (i.e. the ability of each multi-destination scholar is divided by their number of career spells).[19] We then denote by $(\bar{q}_{1kt}, \bar{q}_{2kt}, \bar{q}_{3kt}, \bar{q}_{4kt}, \bar{q}_{5kt})$ the ability of the top five academic scholars of university $k$ in year $t$, and we define the notability index as

$$Q_{kt} = \left( \frac{1}{5}\, \bar{q}_{1kt}^{\frac{\sigma-1}{\sigma}} + \frac{1}{5}\, \bar{q}_{2kt}^{\frac{\sigma-1}{\sigma}} + \frac{1}{5}\, \bar{q}_{3kt}^{\frac{\sigma-1}{\sigma}} + \frac{1}{5}\, \bar{q}_{4kt}^{\frac{\sigma-1}{\sigma}} + \frac{1}{5}\, \bar{q}_{5kt}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{3}$$

where $\sigma$ is the elasticity of substitution between scholars in producing notability.

Notice that it is not uncommon to base a ranking on top persons. RePEc has a ranking of institutions based on the top 10 scholars. Moreover, the notability of a university over time will be based on more persons than just 5. In total, our notability indexes are based on 11,516 scholars (among 46,406).

We use $Q_{kt}$ as a proxy for the attractiveness of the university. When making location decisions, it is unlikely that scholars were able to accurately quantify the quality of each university. However, they were aware of complementarity forces and

19. We will show below that our results are robust to the choice of $\omega$.

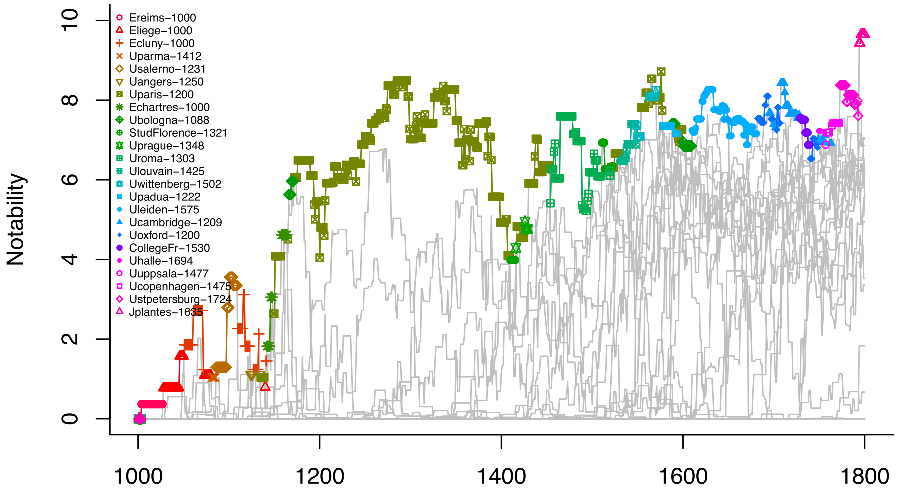



FIGURE 3. Institutional notability of top universities over the whole time span 1000–1800. We select 24 universities that topped the ranking during at least 1 year over the whole time span 1000–1800. We report the trajectory of their institutional notability, $Q_{kt}$, as computed from equation (3).

they observed the highest ability scholars of each university belonging to their choice set. Figure 3 shows the evolution of institutional notability of 24 universities, which topped the ranking of institutions during at least 1 year over the whole time span 1000–1800. The median notability indices of each university are provided in Online Appendix A.

Our ranking of the top institutions varies across the years. Prior to 1200, the cathedral schools of Reims, Liège, and Chartres, the monastery of Cluny, and the predecessors of the universities of Parma, Salerno, and Angers share the lead. After 1200, not surprisingly, the leading positions are taken by the two emblematic universities of Paris and Bologna. From the Black Death to the rise of Protestantism, we see the universities of Florence, Prague, Rome, and Louvain. After the Reformation, Wittenberg appears as the most notorious, followed some time later by Padua, Leiden, Cambridge, and Oxford. In the last century, we find the Royal College in Paris, the universities of Halle, Uppsala, Copenhagen, and St Petersburg (a university attached to the academy there). The Royal Gardens in Paris comes at the end of the parade. This ranking contains a few surprises. For example, the University of Cambridge does very well in the eighteenth century, contradicting the view that it was "an intellectual desert, in which a solitary man constructed a system of the world" (see Manuel (1968) about Isaac Newton in Cambridge).

One can evaluate the relevance of our ranking of universities by comparing it with rankings obtained using different methods. The Casati Law (Italy, 1858) sets rules for accrediting the pre-existing universities into the new Italian University system (Cottini, Ghinetti, and Moriconi 2019). It ranked universities into three categories, A-B-C depending on their quality. We can compare this ranking with our estimate of

$Q_{kt}$, averaged over the 18th century. The average for the nine universities ranked A is 3.231. The average for the eight universities ranked B is 1.018. And the two universities ranked C have a similar level of 0.582 (including the university of Macerata for which we harvested about 800 professors). This suggests that our approach is very likely to allow for a proper identification of the top institutions.

## 3. Empirical Analysis

We now turn our attention to the empirical analysis of the determinants of location choices. Economists have long recognized that spatial mobility decisions play a key role in the career choices of workers (e.g. Keane and Wolpin 1997; Neal 1999). Two types of models, spatial search and location choice, have been used to link mobility decisions to career choices. Spatial search and matching models formalize job search decisions across geographically segmented labor markets; they shed light on the effect of distance on the efficiency of a job search, on spatial heterogeneity in search frictions, and on the persistence of labor market disparities between regions (e.g. Manning and Petrongolo 2017; Schmutz and Sidibe 2019). Ideally, the estimation of matching models requires observing a large number of repeat movers with match-specific outcomes such as individual levels of earnings or employer's profit (e.g. Abowd, Kramarz, and Margolis 1999). This approach is unworkable for us, given the absence of data on match specific outcomes. Moreover, even if we had such outcomes, using only about 10% of the sample (the share of repeat movers) would be costly in terms of external validity of the analysis. Location choice models explain how different types of workers self-select into labor market areas by maximizing their current and expected future levels of income (e.g. Borjas 1987; Dahl 2002; Gallin 2004; Grogger and Hanson 2011). The latter framework is particularly relevant when focusing on the role of workers' attributes, and when match-specific outcomes, demand-side factors, and local matching frictions are unobservable. Hence, we opt for this type of framework.

In this section, we first explain the microfoundations and specificities of our location choice model (Section 3.1). We then estimate the determinants of location decisions with a standard multinomial logit model in Sections 3.2, and perform a series of robustness checks and heterogeneity analyses in Online Appendix C. The standard logit framework raises a number of econometric issues that might generate inconsistent estimates. Firstly and despite the fact that our database includes a large number of obscure scholars, renowned scholars are more likely to be recorded and information about place of birth is missing for a relatively large number of obscure scholars. In the benchmark regressions, these unknowns are eliminated from the sample. This raises sample selection issues that we address in Section 3.3. A related problem is due to the presence of scholars with multiple affiliations. Each $(i, k)$ dyad appears as one observation in the database and is assimilated to a career spell. This means that scholars with seven affiliations appear seven times, while those with a single affiliation appear only once. This also induces possible sample biases and raises the question of the relevance to model scholar $i$'s choice at stage $s$ independently from their other career spells $s'$. These issues are addressed in Section 3.4. Finally, the benchmark specification disregards the potential endogeneity of $q_i$, due to a *reflection problem* or arising from the fact that the ability of scholar $i$ is likely to be affected by their academic environment. We address these issues in Section 3.5.

### *3.1. A Microfounded Gravity Model*

We formalize the discrete location-choice problem of academic scholars in medieval and early modern Europe using a Random Utility Model (RUM), which provides the state-of-the-art microfoundations for most recent gravity models of migration. Our empirical model aligns with a substantial and expanding body of literature that examines factors influencing the decision to migrate and the subsequent choice of destination for both regular migrants (e.g. Grogger and Hanson 2011; Beine, Docquier, and Ozden 2011; Buggle et al. 2023) and asylum seekers and refugees (e.g. Hatton 2016; Hatton 2017; Dustmann, Vasiljeva, and Damm 2019; Hatton 2020). Most of the existing literature focuses on analyzing the determinants of aggregate migration flows. Research employing microdata is less extensive due to limited availability of comprehensive individual-level information encompassing various destinations. Notable exceptions include studies such as Foster and Rosenzweig (2002) or Chort and Senne (2015) and Chort and Senne (2018), who utilize microdata to compare households without direct exposure to migration against households with a family member residing abroad. Additionally, other studies specifically explore migration aspirations with microdata (e.g. Dustmann and Okatenko 2014; Bertoli and Ruyssen 2018; Manchin and Orazbayev 2018).

Our RUM leads to an empirical multinomial logit model, which is in line with Akcigit, Baslandze, and Stantcheva (2016), who study the international mobility of superstar inventors since 1977. Standard location choice models assume that the demand-side of the market is perfectly elastic. In our context, this means that the demand for academic scholars (or equivalently, the supply of academic positions) adjusts perfectly to supply. Although most universities have a fixed number of chairs, they also offer a set of other positions, which are easily adjusted (e.g. the fellows in Oxbridge, the *professores designati* in Copenhagen (Slottved 1978) , the *survivanciers* (designated successor) in Montpellier (Dulieu 1979). We account for potential demand-side factors by including “competition costs” whose size depends on the attractiveness of universities and cities as well as on the ability and “market value” of academic scholars.

Compared with the standard literature on the determinants of migration, and beyond the fact that we use unique micro-data, our approach has three specificities. Firstly, we use *geo-referenced* location data. Each scholar $i$ is assigned to a geo-referenced place of birth, whereas each university $k$ is linked to a geo-referenced position. Each scholar-university dyad is associated with a cost distance $d_{ik}$, measured with the human mobility index (see supra). Since the place of residence of academic scholars before moving to university $k$ cannot be observed, the distance between the place of birth and the university may capture the separation from family and friends (i.e. homesickness), the travel distance *per se*, or the costs of obtaining information about remote places. A striking example of the importance of distance is provided by Eloy (1755) and Michaud (1811) about *Septalius* (Lodivico Settala, 1552–1633). Born and living in Milan, he taught medicine at the nearby University of Pavia and received offers from: the King of Spain, the Duke of Bavaria, the Duke of Tuscany, the city of Bologna, and the Senate of Venice, all offers above what any local citizen could have dreamed of receiving. He enjoyed receiving them as tokens of well-deserved honors, but accepted none. He preferred the company of his 14 children to the luster of these foreign positions. Another clue to the preference for one's place of birth is the following. Among the 17,551 scholars with a known death place, 1,267 of them went back to their hometown before dying, although they held appointments in other places during their life. Another 3,211 were born, worked, and died at the same place.

Secondly, we exploit the *unbalanced panel dimension* of our database as some scholars made repeated choices. We do not necessarily know the timing of choices, but our database links several universities to some scholars. We assume an academic career is made of a maximum of $S$ spells indexed by $s$. At each stage of their career, each professor has to select their preferred location from the feasible university choice set. In practice, if scholar $i$ taught at $S_i$ universities, we include $S_i$ dyadic matches in the database. Robustness checks will be conducted in Section 3.4 to assess the role of repeat movers.

Thirdly, our discrete choice model allows for *varying choice sets*. As new universities are created (or abandoned) over time, the choice sets are individual specific depending on the universities that existed during the active life of the scholar. Each university has a founding date $t_0^k$ and an end date $t_1^k$, which we mostly take from Frijhoff (1996). Sometimes universities—or some schools, which would later become universities—existed before this official date. For example, the University of Paris was officially founded in 1200, but colleges and cathedral schools existed before that date. Gerard Pucelle (1117–1184), an Anglo-French scholar in canon law, taught at Paris from 1156 to 1167 (Arabeyre, Halpérin, and Krynen 2007), before becoming the Bishop of Coventry. We should thus lower the initial date $t_0^k$ for the University of Paris to match the first scholar who can be found there. More generally, the most ancient scholars in the database are Adelbold (965–1027), who taught, at the turn of the millennium, at the cathedral school in Liège, and Fulbert de Chartres (970–1029) who taught at the cathedral school in Chartres and at what would become the University of Angers (Rangeard and Lemarchand 1868).[20] This explains why our time span starts in the year 1000 CE. As far as individuals are concerned, we use the time interval $[t_i^b, t_i^f]$ defined in equations (1) and (2). Each scholar $i$ makes location choices at the start of the career, in year $t(i) \equiv t_i^b$. The portfolio available to individual $i$ is denoted by $K_{t(i)}$. We include a university $k$ in the choice set of individual $K_{t(i)}$ if $t_0^k \leq t(i) \leq t_1^k$.

20. Both Liège and Chartres had cathedral schools, which failed to morph into universities; see Jaeger (2013) on those early cathedral schools in Europe.

The utility that a professor $i$ obtains from locating at university $k \in K_{t(i)}$ at the stage $s \in S$ of their career is given by:

$$U_{iskt(i)} = V_{ikt(i)} + \epsilon_{isk} = \beta \boldsymbol{x}_{ikt(i)} + \epsilon_{isk}, \tag{4}$$

where $V_{ikt(i)} = \beta \boldsymbol{x}_{ikt(i)}$ represents the deterministic component of the indirect utility (net of moving costs), which depends on a vector of observable variables, and $\epsilon_{isk}$ is a vector of person-specific random taste shocks representing the unobservable determinants, which enter the utility function and are orthogonal to the deterministic component.

Assuming the random term $\epsilon_{isk}$ is independently and identically distributed as Extreme Value Type I (EVT-I), which implies that multiple career choices are independent, we can model the probability that university $k$ represents the utility-maximizing choice for professor $i$ at the stage $s$ of their career as the outcome of a standard multinomial logit model (McFadden 1974):

$$p_{iskt(i)} \equiv \text{Prob}\left[ U_{iskt(i)} = \underset{k' \in K_{t(i)}}{\text{Max}} \; U_{isk't(i)} \right] = \frac{\exp\left(\beta \boldsymbol{x}_{ikt(i)}\right)}{\sum_{k' \in K_{t(i)}} \exp\left(\beta \boldsymbol{x}_{ik't(i)}\right)}. \tag{5}$$

In this formula, the probability of going to a given place depends on the features of that place (the numerator) compared to the features of all the other places in the portfolio (the denominator). The property of the multinomial logit model is that the relative probability of choosing between two alternative options in $K_{t(i)}$ depends on the attractiveness of these two options only, that is, $\ln p_{iskt(i)} - \ln p_{isk't(i)} = \beta \boldsymbol{x}_{ikt(i)} - \beta \boldsymbol{x}_{ik't(i)}$, and is independent of the presence of other alternatives (IIA: Independence of Irrelevant Alternatives). In addition, the choice probabilities are independent across career spells as long as $\epsilon_{isk}$ and $\epsilon_{is'k}$ are assumed to be independently distributed. The latter assumption will be relaxed later.

As in the literature on migration, in which the location choice of migrants conditional on the decision to migrate is studied (Bertoli and Ruyssen 2018), our estimations are conditional on the choice of having an academic career. As we cannot observe the universe of scholars, including those not choosing to teach at universities, we cannot model the *ex-ante* problem of choosing between universities and other activities. The absence of a discernible trend in $q_i$ over different periods suggests that drastic shifts in scholars' selection into academia are unlikely. However, it is important to note that this decision is more multifaceted than a simple binary choice, as many scholars simultaneously held positions at universities and engaged in other occupations, such as serving as physicians or astronomers to the monarch, bishops, or judges. Our estimation thus rests on the independence of irrelevant alternatives property within the choice set $K_{t(i)}$, which implies that the relative probability of choosing between two alternative options in $K_{t(i)}$ depends exclusively on the attractiveness of these two options. Even if selection into academia would not affect the location choice of individuals having chosen to teach, it might affect our simulations if, for example, the total number of professors depends on the notability of universities. Hence, it is fair to acknowledge that our results remain partial equilibrium results.

Estimating the multinomial logit model in equation (5) requires specifying the analytical form of the deterministic component of the utility function as a function of observable individual ($q_i$), institutional ($Q_{kt(i)}$), and dyadic characteristics ($d_{ik}$). In the benchmark model, we consider $q_i$ as independent of their location choice. We also consider $Q_{kt(i)}$ as exogenous as it depends on scholars who are no longer active at $t$ (as illustrated in Figure 2). The endogeneity of individual ability will be treated later.

The deterministic component of the utility function captures the average benefits and the average cost for $i$ of locating at $k$, and is independent of the career spell $s$:

$$V_{ikt(i)} \equiv B_{ikt(i)}(.) - C_{ikt(i)}(.). \tag{6}$$

We model the benefits ($B_{ikt(i)}$) as an increasing function of the attractiveness of the city where the university is located (proxied by the indicator of local democracy from Bosker, Buringh, and Van Zanden (2013), $D_{kt(i)}$), as well as of the adjusted notability of the university ($Q_{kt(i)}$), as suggested by anecdotal evidence. For example, Navarro-Brotons (2006) discusses the case of Jeronimo Munoz, who moved from Valencia to Salamanca in 1578. Although Munoz was one of the best paid professors at the University of Valencia, his salary was considerably lower than those paid at universities in Castille. The prestige of the University of Salamanca, and its greater proximity to the seat of royal power, was probably also a factor in Munoz's decision to accept Salamanca's offer. Furthermore, the effect of $Q_{kt(i)}$ can vary with the ability of the professor as, for example, high-ability professors benefit more (or less) from expected interactions with high-ability colleagues (e.g. Stephan and Levin 2001; Kerr, Kerr, and Lincoln 2015a; Kerr, Kerr, and Lincoln 2015b; Kerr et al. 2017). We assume the following specification:

$$B_{ikt(i)} = a_0 + a_1 Q_{kt(i)} + a_2 D_{kt(i)} + a_3 q_i Q_{kt(i)}, \tag{7}$$

where all coefficients are predicted to be non-negative.

We model the cost of locating at university $k$ ($C_{ikt(i)}$) as an increasing function of the cost distance from the place of birth ($d_{ik}$) and of the competition for finding a job at university $k$ in year $t(i)$. The competition for finding a job reflects the demand side of the academic market. Again, anecdotal evidence suggests that the recruitment policy of the best universities included efforts to attract international talent. To give two examples, Eloy (1755) reports that Leonhart Fuchs (after whom the plant fuchsia was named), a professor at Ingolstadt in 1526, was offered 600 gold coins by the Duke of Tuscany, Como, to teach at the University of Pisa. Nadal (1861) discusses the case of the University of Valence, which was searching for a renowned legal scholar in 1583. They sent a messenger to convince a famous lawyer in Grenoble, Jean-Antoine de Lescure, to join the university. The latter reported that he would be willing to come for a salary of 1,500 pounds, provided his moving and house rental costs were covered by the university. They finally agreed on 1,200 pounds plus the house, partly paid by four merchants of the city. Later on, his colleague François Josserand became jealous of Lescure's treatment, threatened to go elsewhere, and obtained a pay rise.

We reasonably assume that the "competition cost" incurred by a professor increases with the attractiveness of the city ($D_{kt(i)}$), as well as with the (adjusted) notability of the university ($Q_{kt(i)}$). However, we also allow the latter "competition cost" to be negatively affected by the individual level of ability, as high-ability professors have a higher market value and receive more generous offers from top universities. In line with the literature on self-selection in migration (e.g. Beine, Docquier, and Ozden 2011; Grogger and Hanson 2011; Kerr et al. 2016; Kerr et al. 2017), we allow the cost of distance to be negatively affected by the individual level of ability. We assume the following specification:

$$C_{ikt(i)} = b_0 + b_1 Q_{kt(i)} + b_2 D_{kt(i)} + b_3 q_i Q_{kt(i)} + b_4 d_{ik} + b_5 d_{ik} q_i, \quad (8)$$

where $b_3$ and $b_5$ are predicted to be negative, whereas the other $b$s are expected to be non-negative.

Plugging equations (7) and (8) into equation (6) gives the expression for the net benefit of an $(i, k)$ employment match. However, in our empirical regressions, we extend the number of generic determinants of location choices ($\mathbf{x}_{ikt(i)}$) to account for the imperfect coverage of our database and for unobserved heterogeneity. We add a university fixed effect, $\gamma_k$, which captures both the unobserved pull factors associated with university/city $k$ that do not vary across years and the quality and extent of the sources used for each university. This yields:

$$V_{ikt(i)} \equiv \beta \mathbf{x}_{ikt(i)} = \beta_0 + \underbrace{\beta_1 Q_{kt(i)} + \beta_2 D_{kt(i)}}_{\text{Agglomeration}} + \underbrace{\beta_3 q_i Q_{kt(i)}}_{\text{Sorting}} + \underbrace{\beta_4 d_{ik}}_{\text{Distance}} + \underbrace{\beta_5 d_{ik} q_i}_{\text{Selection}} + \gamma_k, \quad (9)$$

where $\beta$ is a set of parameters that are common to all individuals and that can be estimated. Unlike standard (linear) regression models, the specification of the multinomial logit model depicted in equation (5) implies that the individual probability to take a position in a university $k$ depends on the characteristics of all universities and cities. Any change in one of these characteristics impacts the whole system.

*Identification*. In line with equations (7) and (8), the constant is given by $\beta_0 \equiv a_0 - b_0$. Coefficients $\beta_1 \equiv a_1 - b_1$ and $\beta_2 \equiv a_2 - b_2$ can be positive or negative and reflect the *agglomeration* (or dispersion) effects resulting from the attractiveness and competition effects. As the university fixed effect captures the mean level of agglomeration/dispersion forces throughout the entire time span covered by our sample, our estimation of these coefficients exploits the within (or demeaned) variations over time in the notability of universities and in the attractiveness of cities. Coefficients $\beta_1$ and $\beta_2$ are identified by the fact that the number of scholars who decide to take a position in university $k$ decreases when this university becomes worse in terms of quality (e.g., Spanish universities in periods 6 and 7) and when a city loses its communal freedom (e.g., Northern Italian cities after the Renaissance).

Just as $V_{ikt(i)}$, the other determinants of location choices are dyadic by construction (sorting and selection terms) or in nature (distance), which makes their identification and estimation possible. Coefficient $\beta_3 \equiv a_3 - b_3$ is positive if high-ability scholars tend to agglomerate at better universities (what we refer to as *positive sorting*) due to higher benefits or smaller costs; it is identified by the fact that the best scholars are more likely to take a position in university $k$ when its quality increases. Coefficient $\beta_5 \equiv -b_4$ is the standard *Distance* term capturing the expected negative effect of remoteness; it is identified by the fact that a given university attracts more scholars born in its vicinity than born far away. As for $\beta_5 \equiv -b_5$, it is positive if high-ability individuals are more mobile than lower-ability ones (what we refer to as *positive selection*); it is identified by the fact that renowned scholars are less sensitive to distance and are more likely to take a position in a remote university than obscure scholars.[21]

Finally, the multinomial logit expression (5) implies that variables that are not specific to a destination $k$, directly or through their interaction with individual characteristics, cannot be identified, as they would affect the net benefit of all $(i, k)$ employment matches symmetrically. This explains why our set of regressors in equation (9) includes neither purely personal characteristics (such as the ability of scholar $i$ *per se*) nor purely temporal phenomena (such as time fixed effects).

### *3.2. Results from the Multinomial logit Model*

Table 2 contains the results of a standard multinomial logit regressions for the whole time span 1000–1800. The sample includes scholars who have been members of universities with comprehensive and broad coverage (see Online Appendix A). The estimations are obtained by using the `mlogit` package of Croissant (2012), which allows for varying choice sets. These regressions characterize the location choices of 33,301 scholars with a mean number of career spells equal to 1.14, which gives a total of 37,964 individual observations. Denoting the number of elements in set $S$, by $\overline{\overline{S}}$, the mean number of institutions is equal to $E_i \overline{\overline{K_i}} = 86$ (the total number $\overline{\overline{\cup_i K_i}} = 138$). Our database includes 3.5 million possible dyadic matches. We focus here on the sign and significance of the agglomeration, distance, selection, and sorting terms. In all regressions, we control for institution fixed effects. The level of local democracy, $D_{kt}$, is obtained from Bosker, Buringh, and Van Zanden (2013) who created a binary variable equal to 1 when cities could organize themselves and claim a kind of self-rule that was often acknowledged by the sovereign in return for taxes or loyalty. The first occurrences of communal self-government were identified in the 11th and 12th centuries in Spain and Italy. They spread across the rest of Europe in the following centuries. For all dates $t$ at which $D_{kt}$ is not available, we impute the closest available data.

The regression in column (1) can be seen as a textbook gravity equation, including distance $d_{ik}$ and mass (in the fixed effect $\gamma_k$). This standard gravity regression shows

21. Positive selection and sorting can also arise if the utility function (4) is not additively separable.

TABLE 2. Multinomial logit regressions: results from a standard logit model.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| **Distance:** | | | | | | |
| $d_{ik}$ | −1.785*** | −1.782*** | −1.884*** | −1.780*** | −1.877*** | -1.888*** |
| | (0.006) | (0.006) | (0.008) | (0.006) | (0.008) | (0.007) |
| **Agglomeration:** | | | | | | |
| $Q_{kt(i)}$ (notability of $k$) | | 0.151*** | 0.153*** | 0.117*** | 0.121*** | 0.229*** |
| | | (0.004) | (0.004) | (0.005) | (0.005) | (0.004) |
| $D_{kt(i)}$ (democracy in $k$) | | 0.059* | 0.065* | 0.056* | 0.062* | 0.042*** |
| | | (0.034) | (0.034) | (0.034) | (0.034) | (0.015) |
| **Selection:** | | | | | | |
| $d_{ik}q_i$ | | | 0.052*** | | 0.050*** | 0.073*** |
| | | | (0.002) | | (0.002) | (0.002) |
| **Sorting:** | | | | | | |
| $Q_{kt(i)}q_i$ | | | | 0.017*** | 0.015*** | 0.012*** |
| | | | | (0.001) | (0.001) | (0.001) |
| $k$ FE | yes | yes | yes | yes | yes | no |
| N. Obs. | 37,963 | 37,963 | 37,963 | 37,963 | 37,963 | 37,963 |
| Log Likelihood | −81,842 | −81,243 | −80,992 | −81,145 | −80,911 | −87,581 |

Notes: Estimation of the multinomial logit model involving equations (5) and (9) using the package `mlogit` in R ($t - stat$ in parentheses). Columns (1)–(5) include university fixed effects. $^{*}p < 0.10$, $^{*}p < 0.05$, $^{***}p < 0.01$.

that the probability of observing a scholar-university match decreases with the cost distance between the birthplace and the university location. This effect remains strong in all specifications. The coefficient of distance is above unity, $\hat{b}_5 = -\hat{\beta}_5 > 1$, which is unsurprisingly greater than in the contemporary era. Focusing on the stock of international migrants in 2010, Beine, Docquier, and Ozden (2011) find a coefficient of 0.7 for all migrants and of 0.35 for college-educated migrants. Focusing on current academic researchers, Miguelez, Raffo and Fink (2013) find a smaller coefficient around 0.2. Agglomeration forces are added in column (2). Scholars are attracted by the notability of the university and the level of local democracy. Notice that due to the presence of university fixed effects, the agglomeration effects are identified through the variations in institutional notability and democracy over time, while the effect of distance is identified through the spatial variation in $d_{ik}$. In column (3), we add the interaction between distance and individual human capital $d_{ik}q_i$. This term is positive, which suggests that the most notable professors were more mobile than others (*positive selection*). In column (4), we interact the individual human capital index with the notability of the university. We find evidence of *positive sorting*: the most notable professors were more likely to settle in more prestigious universities. Putting all regressors together in column (5) shows that agglomeration, selection, and sorting are all significant.[22] Using the values of the log likelihoods, we can compute some simple LR tests: comparing (2) to (1), we can reject the null hypothesis that there is no agglomeration effect. Similarly, comparing (5) to (2), we reject the absence of

22. Similar results are found by Zanardello (2023) on modern-day Italy.

selection and sorting. To illustrate, university fixed effects are excluded in column (6); all results are similar.

To determine whether the coefficient of distance is stable over time, we also ran a specification with distance interacted with a period ($\tau$) dummy. This allows us to test whether the speed of travel improved before 1800. The eight estimated coefficients are: −1.335, −1.424, −1.816, −1.913, −1.817, −1.914, −1.898, −1.975. The coefficient is thus quite stable over the last 6 periods, but lower during the Middle Ages, and especially during the periods before the Black Death. The other coefficients are unaffected, except the effect of communal freedom, which is reinforced. The unexpected non-decreasing pattern in this coefficient reflects that there was little progress in the quality of roads until the 18th century (Bogart 2011), and little innovation in traveling by boat before the invention of steamboats in the 19th century. The lower cost of moving during the Middle Ages may reflect weaker national states, and also the lower density of universities in this period.

The coefficient of the interaction term $Q_{kt(i)}q_i$ captures the fact that high-quality scholars are more sensitive to the reputation of the university when solving the location-decision problem that they face, and/or that higher-quality universities reward scholars' quality more (i.e. higher wages per unit of quality). Remember that the few wages we observe are generally positively correlated with $q_i$ at different points in history (de la Croix et al. 2023). This is in line with our assumption that a higher level of $q_i$ usually translates into a higher level of remuneration.[23]

In a non-linear model, the coefficients cannot be interpreted in terms of predicted probability as the effect of a change in a variable depends on the values of all variables in the model. To put it differently, the effect depends on where we evaluate it: the derivatives of the choice probabilities are given by $\partial p_{isk}/x_{ik} = \beta p_{isk}(1 - p_{isk})$, which is largest when $p_{isk} = 0.5$. For this reason, our coefficients $\beta$ can only be interpreted as the effect of $x_{ik}$ on indirect utility. The results in Table 2 also indicate that the effect of positive selection is relatively small, albeit non negligible: when $q_i$ is around 10 (scholars at the top of the ability distribution), the utility loss due to distance is reduced by just 27%. By contrast, the effect of positive sorting is large: when $q_i$ is around 10, sorting increases the gains from settling in a more prestigious university or in a more attractive city by a factor of 2.24 as compared with obscure scholars (i.e. $q_i = 0$). Besides the standard distance term, agglomeration and positive sorting are

23. We may further want to include $q_i$ among the determinants of location-specific utility, allowing its coefficient to vary across alternatives. This is standard in the estimation of a multinomial logit model with variables that are individual but not alternative specific. Still under the assumption that wages are proportional to $q_i$, it would purge the estimated coefficient of $Q_{kt(i)}q_i$ from the confounding effect of differences in wages across universities. Including these choice-specific terms, we obtain $\overline{\overline{\cup_i K_i}} = 138$ more parameters to estimate. The estimated coefficients of these $q_i$s vary from one university to the other, as does the university fixed factor. They also sometimes have a negative value, which is hard to interpret in the context where scholars would be remunerated in proportion to their $q_i$. In this new specification, the interaction term $Q_{kt(i)}q_i$ is weakened but remains highly significant (0.007 (0.002) instead of 0.015 (0.001)) despite the inclusion of many terms correlated with $q_i$. The three agglomeration effects are barely affected.

TABLE 3. Multinomial logit regressions: robustness to selection and coverage.

| | Benchm (1) | Sample Unknown (2) | Adding partial cov. (3) | Limited cov. $\Lambda_{kt} \geq 20$ (4) |
|---|---|---|---|---|
| $d_{ik}$ | −1.877*** | −2.000*** | −1.839*** | −1.949*** |
| | (0.008) | (0.007) | (0.007) | (0.013) |
| $Q_{kt(i)}$ | 0.121*** | 0.135*** | 0.127*** | 0.037*** |
| | (0.005) | (0.005) | (0.005) | (0.008 ) |
| $D_{kt}$ | 0.062* | 0.079** | 0.069** | −0.094** |
| | (0.034) | (0.034) | (0.033) | (0.042) |
| $d_{ik}q_i$ | 0.050*** | 0.072*** | 0.046*** | 0.056*** |
| | (0.002) | (0.002) | (0.002) | (0.004 ) |
| $Q_{kt(i)}q_i$ | 0.015*** | 0.014*** | 0.012*** | 0.030*** |
| | (0.001) | (0.001) | (0.001) | (0.002) |
| FE | yes | yes | yes | yes |
| Number of Observations | 37,963 | 50,722 | 39,288 | 22,668 |
| Log Likelihood | −80,911 | −83,532 | −88,098 | −35,146 |

Notes: Estimation of the multinomial logit model (equation (5) using the package `mlogit` in R ($t - stat$ in parentheses). Column (1) recalls the benchmark results. In column (2), scholars with unknown birthplace are assigned the minimal distance (3.5 kms) from their university. In column (3), we include all universities with at least. In column (4), we exclude universities with less than 20 recorded scholars. All columns include university fixed effects. $^{*}p < 0.10$, $^{*}p < 0.05$, $^{***}p < 0.01$.

important forces governing the mobility decisions of academic scholars. Table C.9 in Appendix C shows that our results are globally preserved but vary across periods, fields of knowledge, or region.

### 3.3. *Sample Selection*

Our database does not include the universe of professors. This implies two sources of sample selection issues: (i) many obscure scholars are not included in the sample, and (ii) there is a considerable heterogeneity in the coverage of institutions. We assess whether our results are robust to sample selection. Results are reported in column (2) and column (3) of Table 3. We report our benchmark results in column (1).

As far as scholars are concerned, some are included in the sample but data on their place of birth are missing. This is usually the case for less well known professors. Indeed, among the scholars with a known birthplace, 52% have a positive $q_i$. This proportion drops to 14% for those with an unknown birthplace. Hence, our sample is likely to overweight top-quality professors (high $q_i$) and underweight the less well known (low $q_i$). This is a limitation because the co-existence of professors who are famous and those who are not is key to identifying selection and sorting patterns. To measure the importance of sample selection, we re-estimate the multinomial logit (5) by making the sample less selective. To do so, we use the identified scholars of unknown origin, and assume that they were born in the city of their university, implying $d_{ik} = \ln(\text{cost}^{\min})$ for them.

Column (2) shows the results obtained when assuming that all identified scholars from unknown origin are locals. This increases the sample size by one third, reinforces substantially the positive selection effect (increasing the coefficient by a factor of 1.44), and increases the agglomeration terms; positive sorting is not much affected. These results suggest that if we had observed the whole universe of scholars, which contains many more unknown people born locally, positive selection would appear stronger while leaving sorting unaffected. Hence, our benchmark estimates might give a lower bound on selection.

As far as institutions are concerned, our benchmark regression sample excludes universities with fewer than 10 scholars in total and those with partial coverage. We relax the latter constraint in column (3) of Table 3, which increases the sample by about 1,300 observations. Our empirical results are highly robust to these changes. In column (4), we restrict our working sample to universities with at least 20 scholars (instead of 10 in the benchmark). This amounts to reduce the sample size (by 40%) as well as the choice set of every scholar by removing some small universities. The effect of communal freedom turns negative, probably because some important cities for identifying this effect were removed, but selection and sorting mechanisms are reinforced.

### *3.4. Treatment of Repeat Movers*

Remember that 12.0% of our scholars are linked to more than one university and we count each dyadic match as one observation. This raises two potential issues. Firstly, the weight of repeat movers exceeds that of one-time movers. As the number of career spells increases with human capital, this reinforces the over-representation of renowned scholars in our database. Secondly, by assuming that career-spell-specific choices are independent, we ignore the possibility that movers may have had correlated preferences.

The first problem can be easily addressed by removing repeat movers from the regression sample, which eliminates many famous scholars, or by linking them to a single university. We do both and, when following the second option, randomly select one of their affiliations. Solving the problem of correlated career spells is more complicated. To account for it, we generalize the standard logit model by relaxing the hypothesis of independence of individual choices. The independence property can be unrealistic in many settings, especially in situations with repeated choices over time. We can expect unobserved factors that affect a decision maker to persist over time. In a multinomial logit, we cannot include individual fixed effects since they would not affect the probability that a university $k$ dominates another university $k'$. A more general deterministic component of utility can be written $V_{ikt(i)} = \beta_i \mathbf{x}_{ikt(i)}$, where $\beta_i$ is a vector of coefficients that is unobserved for each $i$ and varies randomly across professors, representing their tastes. This specification is the same as for the logit except that now the coefficients $\beta_i$ vary in the population rather than being fixed. In particular, the coefficient $\beta_i$ can be expressed as the sum of a population mean, $\bar{\beta}$, and an individual deviation, $\eta_i$, such that their utility of moving to destination $k$ is written $U_{iskt(i)} = \bar{\beta}\mathbf{x}_{ikt(i)} + \eta_i \mathbf{x}_{ikt(i)} + \epsilon_{isk}$. The last two terms of such a *Random-Parameter Logit* capture the unobserved portion of utility. In other words, the marginal effect on the latent dependent variable is individual specific. The same tastes are used by the decision maker for each career spell and the variance in $\beta_i$ induces correlation in utility across destinations and career spells.

How these parameters vary across individuals is unknown. The mixed logit model assumes that these parameters vary according to the population PDF $g\left(\beta_i|\theta\right)$, where $\theta$ represents the moments of the distribution such as the mean and the variance, which must be estimated. A fully parametric mixed logit model arises once $g\left(\beta_i|\theta\right)$ is specified. We assume that the coefficient vector is independent and normally distributed, $\beta_i \rightsquigarrow N\left(\bar{\beta}, \sigma^2\right)$. The unobserved portion of utility is correlated across destinations and career stages due to the common influence of $\eta_i$, which violates the IIA property of the standard logit (Revelt and Train 1998). The full parametric model can be estimated using the simulated maximum-likelihood procedure (Sarrias et al. 2016).

In column (2) of Table 4, we show that most of our results are highly robust to the exclusion of repeat movers. Compared to the benchmark specification of column (1), removing repeat movers increases the magnitude of the agglomeration and gravity terms. As for sorting and selection, their identification relies on the difference between famous and obscure scholars in the sensitivity of location choices to institutional quality and distance. Remember repeat movers exhibit an average ability index that is 2.35 times greater than the one-time movers; they account for 12.1% of our scholars and are linked to 2.16 universities, on average. Removing them from the sample decreases the number of observed dyads by 21.6% (from 37,963 to 29,748) and eliminates many famous scholars at the upper end of the ability distribution. The sorting term resists this change and its magnitude is strengthened compared to the benchmark. This is important because we will see it is the effect that is driving our simulation results in Section 4. By contrast, the selection term is reduced by more than half compared to the benchmark.

Instead of removing entirely the repeat movers, we keep them but associate them with only one of their affiliations (randomly chosen) in column (3). Compared with column (2), the drop in selection term is less pronounced; the coefficient is equal to 70% of the level obtained in the benchmark regression. This demonstrates once again that including famous and obscure scholars is key to identifying sorting and selection patterns. In particular, the magnitude of the selection term is strongly governed by the fact that the location choices of (high-ability) repeat movers are less sensitive to distance than those of lower-ability scholars.

In column (4) of Table 4, we relax the assumption of independent career choices for multi-destination scholars, and estimate a mixed logit model with individual-specific vectors of coefficients drawn from a normal distribution. The agglomeration, selection, and sorting mechanisms are preserved. Although the mixed logit entails six additional parameters (the s.e. of the six coefficients—not reported), a likelihood ratio test would reject the benchmark formulation in favor of the mixed logit formulation. The mixed

TABLE 4. Robustness to repeat movers, mixed Logit, alternative human capital measure, and nested logit.

| | Benchm | Repeat movers | | Mixed | Altern. | Nested |
|---|---|---|---|---|---|---|
| | | removed | linked to 1 university | logit | $\tilde{q}_i$, $\tilde{Q}_{kt(i)}$ | logit |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $d_{ik}$ | −1.877*** | −2.029*** | −1.934*** | −2.154*** | −1.876*** | −1.614*** |
| | (0.008) | (0.010) | (0.008) | (0.012) | (0.008) | (0.011) |
| $Q_{kt(i)}$ | 0.121*** | 0.141*** | 0.142*** | 0.124*** | 0.096*** | 0.103*** |
| | (0.005) | (0.006) | (0.006) | (0.006) | (0.005) | (0.005) |
| $D_{kt}$ | 0.062* | 0.146*** | 0.116*** | 0.095*** | 0.075** | 0.064** |
| | (0.034) | (0.040) | (0.038) | (0.037) | (0.034) | (0.031) |
| $d_{ik}q_i$ | 0.050*** | 0.022*** | 0.035*** | 0.055*** | 0.045*** | 0.047*** |
| | (0.002) | (0.003) | (0.003) | (0.003) | (0.002) | (0.002) |
| $Q_{kt(i)}q_i$ | 0.015*** | 0.021*** | 0.011*** | 0.018*** | 0.016*** | 0.012*** |
| | (0.001) | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| $\psi_1$ (low) | | | | | | 0.743*** |
| | | | | | | (0.012) |
| $\psi_2$ | | | | | | 0.814*** |
| | | | | | | (0.010) |
| $\psi_3$ | | | | | | 0.778*** |
| | | | | | | (0.008) |
| $\psi_4$ (high) | | | | | | 0.819*** |
| | | | | | | (0.008) |
| FE | yes | yes | yes | yes | yes | yes |
| Number of observations | 37,963 | 29,748 | 32,334 | 37,963 | 37,963 | 37,963 |
| Log Likelihood | −80,911 | −52,749 | −64,168 | −79,914 | −81,070 | −80,665 |

Notes: Estimation of the multinomial logit model (equation (5) using the package `mlogit` in R ($t-stat$ in parentheses). Column (1) recalls the benchmark results. In column (2), we exclude repeat movers. In column (3), we randomly assign repeat movers to a single university that they visited. In column (4), we estimate equation (5) with a mixed logit; the six variance parameters are estimated as well, four of them exhibit a variance that significantly differs from zero (variance of the coefficients of $d_{ik}$, $D_{kt}$, $d_{ik}Q$, and $Q_{kt}Q$). In column (5), human capital and notoriety are measured exclusively with "publications by" from Worldcat. In column (6), we estimate a nested logit model, with nest defined as quartiles of universities in terms of notability. The estimated coefficients $\psi_m$ are the within-nest dissimilarity parameters. $^{*}p < 0.10$, $^{*}p < 0.05$, $^{***}p < 0.01$.

logit has a disadvantage though: The estimates are obtained by simulation, while in the multinomial logit, a likelihood function is maximized. In addition, the results depend on the assumption regarding the distribution of the random parameters.

### 3.5. *Measurement and Endogeneity of* $\mathbf{q}_i$

To ensure the robustness of our findings regarding the measurement and endogeneity of scholars' human capital, we undertake several steps. Firstly, we demonstrate the robustness of our estimation results in Table C.8 in Appendix C, where we observe that our findings remain robust across different parametric assumptions, such as the choice of $\sigma$ and $\omega$.

Secondly, we address the potential issue of the *reflection problem* when measuring the human capital of scholars using Wikipedia and publications about them. This problem arises when a person's fame impacts the prestige of a university, drawing more attention to his predecessors and successors. For example, Martin Luther has influenced the prominence of the University of Wittenberg and the notability of all scholars who have been employed there. They might thus have longer Wikipedia pages, and more publications about them. To mitigate this concern, we compute an alternative measure of human capital based on "publications by" only, $\tilde{q}_i$, hence, excluding the Wikipedia features and the publications about. A new notability index of each university, $\tilde{Q}_k$, is then computed from this new measure of the human capital. We reestimate our location model with these new measures of human capital and notability, which is less affected by endogeneity than the one used in the benchmark. The results, presented in column (5) of Table 4, exhibit remarkable robustness to this modification, further strengthening the validity of our findings.

Thirdly, the most problematic endogeneity issue arises because the ability of each professor $i$ is measured by an index of human capital observed a long time after the end of their career ($q_i$), which is likely to be influenced by the quality of the university that was chosen. This means that we should distinguish between $\bar{q}_i$, the innate/exogenous level of ability, and $q_i$, the ex-post level of notability. Let us denote by $k^*$ the university chosen by a scholar. Ideally, we should use $\bar{q}_i$ to estimate the multinomial logit (5). However, we only observe $q_i$, and this ex-post level might be affected by $Q_{k^*t(i)}$, the notability of the chosen university. This implies that we do not observe the potential level of human capital if the individual had been working at a different university $k$. Assume for example that $q_i = \bar{q}_i + \theta Q_{k^*t(i)}$ and denote by $\overline{V}_{ikt(i)}$ the indirect utility level obtained after replacing $q_i$ by $\bar{q}_i$ in equation (9).

In theory, the multinomial logit implies that university $k$ dominates university $k'$ if $\overline{V}_{ikt(i)} + \epsilon_{isk} > \overline{V}_{ik't(i)} + \epsilon_{isk'}$, which only depends on the characteristics of individual $i$ and universities $k$ and $k'$. In practice, we are unable to model $\overline{V}_{ikt(i)}$ and $\overline{V}_{ik't(i)}$ properly because our measure of individual human capital is $k^*$-specific (i.e. influenced by $Q_{k^*t(i)}$). The endogeneity of $q_i$ implies that the difference in utility is measured with additional noise: university $k$ dominates university $k'$ if

$$\overline{V}_{ikt(i)} + \epsilon_{isk} > \overline{V}_{ik't(i)} + \epsilon_{isk'} + \theta Q_{k^*t(i)} \Delta_{ikk't(i)}, \tag{10}$$

where $\Delta_{ikk't(i)} \equiv \beta_4(Q_{kt(i)} - Q_{k't(i)}) + \beta_6(d_{ik} - d_{ik'})$ results from the two interaction terms that are affected by our noisy measure of individual human capital in equation (9). The term $\theta Q_{k^*t(i)} \Delta_{ikk't(i)}$ in equation (10) is correlated across destinations, due to the presence of $Q_{k^*t(i)}$. Hence, the inability to observe $\bar{q}_i$ leads to the violation of the IIA property.

To mitigate this problem, we estimate a nested logit model (McFadden 1978) where nests are defined as groups of universities sharing similar levels of notability ($Q_{kt(i)} \approx Q_{k't(i)}$) during the years of activity of individual $i$. We partition the choice set $K_{t(i)}$ into four groups of alternatives, $K_{mt(i)}$ with $m = (1, 2, 3, 4)$ for the top, mid-high, mid-low, and bottom universities. Our partition is based on the notability index observed in the 4th and 5th periods. Each university belongs to exactly one nest. Building on Ortega and Peri (2013) and Bertoli and Moraga (2015), we assume that the individual random taste shock is a mixture of a location-specific and a nest-specific term:

$$\epsilon_{isk} = \psi_m \upsilon_{isk} + (1 - \psi_m)\upsilon_{im},$$

where $\psi_m \in [0, 1]$ is the weight associated with the location-specific term, $\upsilon_{isk}$, which is assumed to be independently and identically distributed as EVT-I; and $\upsilon_{im}$ is an error term that is specific to the $m^{th}$ nest ($k \in K_{mt(i)}$), whose distribution depends on $\psi$ such that the marginal distribution of $\epsilon_{isk}$ is also EVT-I (Cardell 1997). Parameter $\psi_m$ also determines the mutual correlation in the realizations of the nest-specific error term. We have $\psi_m = \sqrt{1 - \rho_m}$, where $\rho_m$ represents the correlation coefficient within nest $m$. Hence, $\psi_m$ is a dissimilarity parameter. The higher $\psi_m$, the smaller the weight of the nest-specific component and the smaller the within-nest correlation of error term. When $\psi_m = 1$ for all $m$, the nested logit boils down to the standard multinomial logit ($\epsilon_{isk} = \upsilon_{isk}$).

The nested logit model assumes a generalized version of the EVT-I distribution, such that (i) the mean error varies across nests, and (ii) alternatives within a nest exhibit mutually correlated error terms (but the same mean). On the contrary, the error terms of two alternatives belonging to different nests are uncorrelated but have different means. In our context, this difference in the means captures the component of the error term $\theta Q_{k^*t(i)} \Delta_{ikk't(i)}$ and hence corrects for the endogeneity bias. It reflects the influence of the chosen university on individual quality. Within a nest, this component is close to zero because $\Delta_{ikk't(i)} \approx 0$. Notice that this technique to correct for the endogeneity bias is possible only because the $q_i$ always appears interacted with a variable for which we can build nests, and never appears alone (it cannot explain location choice alone as it is not destination specific).

The probability of individual $i$ choosing university $k \in K_{mt(i)}$ is equal to the product of the probability of choosing alternatives in nest $K_{mt(i)}$ and the probability of choosing exactly $k$ in $K_{mt(i)}$ (Heiss et al. 2002). It is given by

$$p_{iskt(i)} = \frac{\exp\left(\beta \boldsymbol{x}_{ikt(i)}/\psi_m\right)}{\exp\left(IV_{mt(i)}\right)} \times \frac{\exp\left(IV_{mt(i)}\psi_m\right)}{\sum_{m'} \exp\left(IV_{m't(i)}\psi_m\right)} \,\tilde{}\tilde{}\, \forall t, \tag{11}$$

where $IV_{mt(i)} = \ln \sum_{k' \in K_{mt(i)}} \exp(\beta \boldsymbol{x}_{ikt(i)}/\psi_m)$ is the inclusive value of each nest $K_{mt(i)}$, representing the rescaled measure of attractiveness of the nest for individual $i$ (i.e. the expected value of the utility individual $i$ obtains from the alternatives in nest $K_{mt(i)}$).

Results are provided in column (5) of Table 4. Compared to the benchmark, the effects of selection and agglomeration are similar despite the fact that part of the agglomeration force is likely to be captured by the nest-specific error term. Sorting is slightly weakened but remains highly significant and important compared to the agglomeration effect: When $q_i$ is around 10, positive sorting increases the gains from settling in a more prestigious university by a factor around 2.5. Note that we reject the assumption of no nests, either through a likelihood ratio test (lr=493), or by testing whether the correlations within nests are zero, or equivalently $\psi_m = 1\ \forall n$ (Wald=919.7, p-val=0.000). We also reject that the degree of correlation inside each nest is the same, $\psi_m = \psi_{t(i)}\ \forall m$ (Wald=62.111, p-val=0.000).

## 4. Implications for the Scientific Revolution

Market forces significantly impact individual choice probabilities. In Online Appendix C, we analyze the simulated and counterfactual choice probabilities of two scholars, Roffredus Beneventanus (1170–1243) and Thomas Aquinas (1225–1274), both born in the same region during the initial period with a limited choice set. Despite their shared background, however, they differ in their levels of human capital. We examine the impact of agglomeration, positive selection, and positive sorting in shaping their location decisions. We emphasize the crucial role of selection and sorting effects for highly talented scholars, compelling them to congregate in the most prestigious universities (see Online Appendix Table C.10). To demonstrate the overall implications, we extend our analysis beyond individual cases and employ our estimated model to simulate the contributions of agglomeration, positive selection, and positive sorting to the total annual university output.

We construct a proxy for the total output of university $k$ in year $t$, denoted by $Y_{kt}$, which is an aggregation of the human capital of all scholars predicted to work there. Denoting the number of active scholars in year $t$ by $\Lambda_t = \{i | t_i^b \leq t \leq t_i^f\}$, we define $Y_{kt}$ as a CES (constant elasticity of substitution) combination of the ability levels of its predicted members:

$$Y_{kt} = \left( \sum_{i \in \Lambda_t} \hat{p}_{ikt} q_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{12}$$

where $\hat{p}_{ikt}$ is the weight given to professor $i$ at university $k$ in year $t$. We set it equal to the simulated probability that $i$ goes to $k$ from the multinomial logit model—like the probabilities shown in Online Appendix Table C.10 for Roffredus Beneventanus and Thomas Aquinas. It differs from the actual probability $\bar{p}_{ikt}$, which is equal to 1 if scholar $i$ was affiliated with university $k$ at year $t$. Parameter $\sigma$ represents the elasticity of substitution between academic scholars' human capital in producing scientific knowledge, assumed to be equal to the elasticity of substitution between the top-5 scholars in producing notability in equation (3).

The simulated output $Y_{kt}$ should be interpreted as including advancement to knowledge, quality of teaching, and service to society (such as supplying cautious physicians, rigorous lawyers to the local courts, or well-educated priests and pastors to parishes). Then, we compare the total simulated output,

$$Y_t = \sum_k Y_{kt}, \tag{13}$$

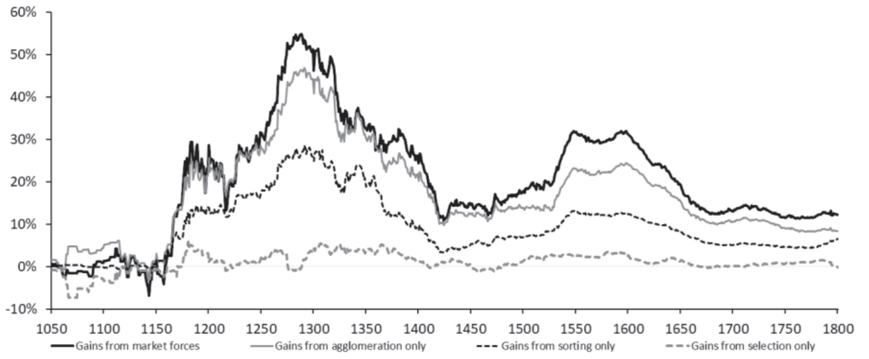
in the benchmark—we use the term "benchmark" for the simulation and for the regression model—with the counterfactual output levels obtained after neutralizing the agglomeration, positive selection, and sorting terms (separately or jointly). Hence, the important point here is not the level of $Y_t$ in itself, but the gap between $Y_t$ with and without academic market forces. We define the university output gain due to market forces as the difference between the predicted academic output and the counterfactual level, expressed as percentage of deviation from the counterfactual level.

When the elasticity of substitution tends to infinity ($\sigma = \infty$), we have perfect substitutability between scholars. The total output is the sum of individual human capital, independent of location (represented by the $\hat{p}_{ikt(i)}$): $\lim_{\sigma \to \infty} Y_t = \sum_k \sum_{i \in \Lambda_t} \hat{p}_{ikt(i)} q_i = \sum_{i \in \Lambda_t} q_i$. Hence, there is no gain to expect from market forces. By contrast, when $\sigma$ is finite, there is a complementarity relationship between academic scholars. The smaller $\sigma$, the greater the knowledge gain from agglomerating high-ability scholars at the same university, and the agglomeration of the highest ability scholars leads to output gains. In our benchmark regressions and simulations, we use a CES production function with $\sigma = 2$, in line with the definition of the notability of the university in equation (3).

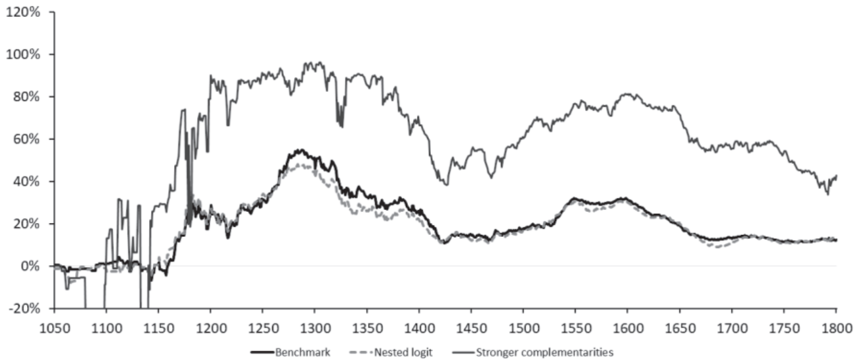
Figure 4 shows the university output gains obtained with the benchmark model (i.e. standard multinomial logit (ML) model with $\sigma = 2$) in the top Panel (a), and with alternative estimates in the middle Panel (b). In Panel (b), we compare the total output gain obtained with the benchmark model, with that obtained under the nested logit variant, and under the high-complementarity variant (i.e. multinomial logit model with $\sigma = 1.2$). In the bottom Panel (c), we get back to the benchmark model and compute the output gains by restricting the sample of scholars to those who taught in Science and/or Medicine, as reported in our institutional and/or bibliographical data sources. Note that a scholar might be reported to have taught in more than one "field."

Focusing on the benchmark model in Panel (a), the black curve shows the output gains from academic market forces (i.e. agglomeration, selection, and sorting jointly). These gains are obtained after neutralizing the effect of agglomeration, positive selection, and positive sorting on $\hat{p}_{ikt(i)}$ in the multinomial logit model while keeping the distance term and the university fixed effects (i.e. basic gravity). The other curves in gray show the gain from agglomeration, sorting, and selection forces in isolation. We find that market forces increase the total output of Europe by about 54% before the Black Death, by about 33% before the invention of the printing press, and by 30% after the rise of Protestantism, at the beginning of the Scientific Revolution.

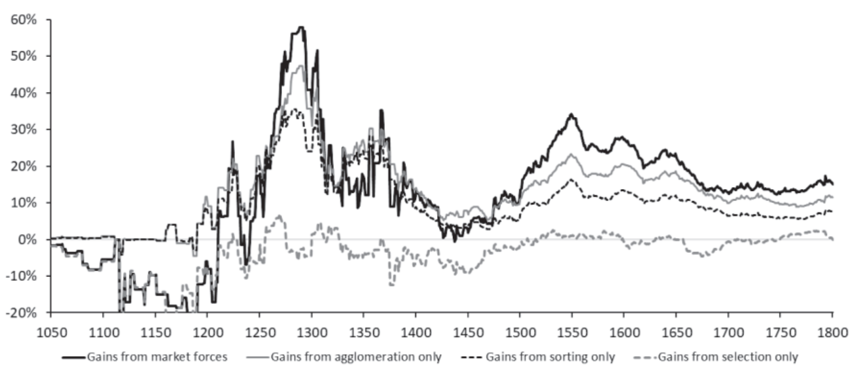
It is worth noticing that academic market forces do not necessarily increase total simulated output, as appearing in the first century. Their joint effect on output depends on the correlation between the notability of universities and the level of city/university amenities (captured, in our regressions, by the university fixed effect and the level of local democracy). When the correlation is high, the effects of notability and amenities point in the same direction; the best scholars agglomerate in the best universities. When the correlation is lower, agglomeration and sorting can result in the concentration of




(a) Effect of agglomeration, selection and sorting (benchmark ML with $\sigma = 2$)




(b) Gains from market forces (benchmark ML with $\sigma = 2$, NL with $\sigma = 2$, ML with $\sigma = 1.2$)




(c) Effect of agglomeration, selection and sorting in Science and Medicine (benchmark ML with $\sigma = 2$)

FIGURE 4. University output gain from market forces by year (1000–1800). The top panel shows the trajectory of the university output gain from market forces and the isolated effects of agglomeration, selection, and sorting. Observed and counterfactual output levels are computed using equation (13). The middle panel shows the sensitivity of the total output gain to the estimation techniques. We compare the gain obtained under the benchmark ML with $\sigma = 2$, the nested logit with $\sigma = 2$, and the multinomial logit with $\sigma = 1.2$. The bottom panel shows the trajectory of the university output gain from market forces when restricting the sample of scholars to those who taught in Science and/or Medicine.

talent in second-best universities, which reduces total academic output. This is at least the case if the intensity of agglomeration and sorting forces is limited.

The effects of academic market forces become weaker over the period 1650–1800. They increase total output by about 10%–15% over this period. Overall, we find that agglomeration and sorting induce smaller output gains when the individual choice set is large. For this reason, their effect on academic output diminished in the 17th centuries, when the number of universities almost doubled. What is specific to this period is the presence of many universities with a large number of scholars having published something ($q_i > 0$), which was not highly influential. Shutting down agglomeration redistributes superstars to the advantage of these less prestigious universities, thus increasing the output of the many average people there. When decomposing the total output gains in its three components, we find that agglomeration and positive sorting play an important role, especially in earlier times, when there are few universities. By contrast and in line with our empirical findings, positive selection hardly influences the total simulated output.[24] It is worth noting that the sum of the three components taken in isolation significantly exceeds the total output gains in most years. This is because our model is non linear, the effect of positive sorting is greater when agglomeration forces are not accounted for, and vice versa.

Another reason for the waning of academic market forces after 1650 might come from a sample selection effect. Over the period 1650–1800, we see the creation of academies of arts and sciences in many cities all over Europe (McClellan 1985). These academies do not usually teach but they organize research in the emerging fields of knowledge, which are not developed in the universities yet. There are many scholars who are active in both types of institutions, universities, and academies, but we also find many brilliant minds who are only present in academies (such as Voltaire or Leibniz).

To confirm our intuition that top scholars are less likely to be member of a university after 1650, we looked at European astronomers and mathematicians having given their name to a crater on the Moon (for details on how this naming was achieved; see Online Appendix D). Before 1650, almost 70 % of them are university professors. After 1650, it goes down to 42 % (but 88% are members of some academy). This suggests that a growing proportion of prominent scholars could have started affiliating with academies beginning in the 17th century, thereby influencing scientific output beyond university settings. There are multiple reasons preventing us from quantifying the influence of academies in explaining changes in academic market dynamics post-1650. The primary hurdle lies in precisely delineating the universe of academies and correctly pinpointing scholars within them. Additionally, the complexity arises from scholars simultaneously belonging to universities and multiple academies. Moreover,

24. We have seen in Online Appendix Table C.10 that positive selection tends to scatter talents across universities by increasing the menu of options for the highest ability scholars, but the total effect of this increased dispersion is small.

certain scholars hold foreign membership in an academy, which suggests that affiliation itself does not necessarily indicate a definitive location choice.[25]

The middle panel of Figure 4 shows that very similar findings are obtained when using estimates from the nested logit model with $\sigma = 2$. By contrast, our results are quantitatively sensitive to the value of $\sigma$. As stated above, the gains from market forces would disappear if we had taken $\sigma = \infty$, as the allocation of scholars (represented by the $\hat{p}_{ik\tau(i)}$) across places would not matter. Although the benefits of sorting, however, only slightly exceed those obtained in the benchmark, the positive effects of agglomeration are magnified under the high-complementarity variant with $\sigma = 1.2$. In addition, with this variant, we find smaller variations in output gains across years. This means that the huge gains from agglomeration in the top universities are not compensated for by losses in average- and low-quality universities in the more recent periods. Overall, when combining all mechanisms, the simulated output increases by 90% in the 13th and 14th centuries, and by 80% in the first half of the 17th century.

As introduced earlier, the advancements in fields such as law, theology, philosophy, and arts within universities have played a crucial role in enhancing institutional quality (Cantoni and Yuchtman 2014) and shaping cultural norms (de la Croix and Mariani 2015). These developments may have fostered an environment conducive to the emergence of markets and the dissemination of scientific thinking, which have been instrumental in explaining the Rise of the West. Over time, scientific knowledge has increasingly translated into inventions and innovations, driving economic growth.

To assess more precisely the importance of markets for science, we run another simulation by restricting the sample to physicians and scientists (their weight in the sample is described in Online Appendix B.3). The lower panel of Figure 4 presents the results. We observe that market forces have led to a nearly 60% increase in scientific output in Europe before the Black Death, followed by a roughly 30% increase after the rise of Protestantism. These effects mirror the outcomes observed at the onset of the Scientific Revolution, indicating a significant positive influence. However, a negative effect is noted during the 11th and 12th centuries, when the correlation between university prominence and the availability of city/university amenities was lower.

To further understand the role of market forces, Figure 5 maps the winners (in green) and losers (in red) due to market forces in the period 1686–1733. The surface of each circle represents the difference in simulated output between the benchmark case and the basic gravity case. An easily understandable case is Lund vs. Copenhagen. With market forces, scholars born in Sweden are more likely to locate in Copenhagen, which has a high notability rather than in Lund, which is just average, while without agglomeration and sorting forces, they are content with Lund. We also note that Rinteln is a big loser in Germany, being surrounded by many good universities such as Leipzig

25. Illustratively, take the case of Ernst Gottfried Baldinger, who traversed academic institutions such as the University of Jena (1768–1773), the University of Göttingen (1773–1782), and the University of Marburg (1786–1804). Concurrently, he became a regular member of the *Academia Electoralis Moguntina Scientiarum* in Erfurt (1766) and of the Leopoldina (1770) while also assuming foreign membership in the Bavarian Academy of Sciences and Humanities (1775).

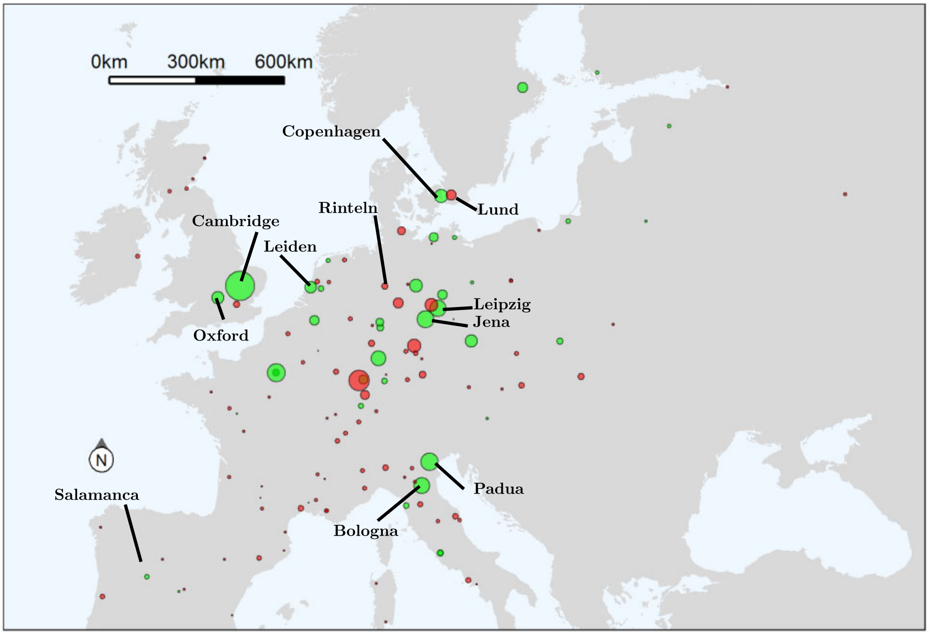



FIGURE 5. Winners (green) and Losers (red) from Market Forces in 1686–1733. For each university, we use equation (12) to compute the difference in simulated output between the benchmark case and the basic gravity case. The surface of each circle is proportional to this difference in absolute value, and the color indicates whether this difference is positive (winners in green) or negative (losers in red).

and Jena. It is also noteworthy that the South of Europe is not doing so poorly, although the bigger gains are in the North. Renowned Southern universities still attract talents (Salamanca, Padua, Bologna, and Rome). In a sense, without market forces, the fate of the South would have been worse.

When looking at universities, which were permanently closed down over the period 1700–1900, many of them were losers from the market in the last three periods. Altdorf (closed in 1809), Bamberg (closed in 1803), Cahors (closed in 1751), Cervera (closed in 1821), Dorpat (closed in 1710), Harderwijk (closed in 1811), Pont-à-Mousson (closed in 1768), Rinteln (closed in 1809), Siguenza (closed in 1807), and Valence (closed in 1793) are in this category. A few winners were closed too: Erfurt (closed in 1816) and Frankfurt (Oder) (closed in 1811).

Overall, our results show that agglomeration and sorting effects in the academic market contribute to fostering university output. The sizes of the agglomeration and sorting effects before the middle of the 16th century are quantitatively significant. Thanks to these effects, university output increases by 50% when considering conservative complementarity forces, and almost doubles when considering greater complementarity forces. In addition, we do not model any cumulative effect of knowledge creation. Hence, our 50% should be understood as a lower bound.

Several economic historians claim that labor markets were relatively complete and competitive in Medieval Europe: “Given the low reproductive success of the urban population there had to be a constant flow of labor from the country to the city (Clark 2008). The records of a 1292 tax levied by Philip the Fair on the commoner households of Paris show that 6% were foreigners: 2.1% English, 1.4% Italian, 0.8% German, 0.7% Flemish, 0.6% Jewish, and 0.4% Scottish” (Sussman 2006, Clark 2008). We can compare these numbers with the origin of the scholars of Paris University in the first two periods of our sample (1000–1347). Based on the 443 persons with known origin, we obtain that 57% of the scholars were born in France (in its 2020 limits), 20% are British, 3% are from Germany, 9% are from Italy, and 5% are from the Low Countries—the data for this period are mostly based on Courtenay (1999), Gorochov (2012), and Genet (2019). Although the mobility of ordinary people seems quite high already, the mobility of university scholars is higher by an order of magnitude.

The importance of market forces seem particularly relevant in the years preceding and coinciding with the dawn of the Scientific Revolution, a period commonly defined as spanning Copernicus’s and Newton’s times, that is, 1543–1687 (Applebaum 2003). In the last two centuries before the Industrial Revolution, these effects decrease significantly or even become non-existent. Hence, although we provide no causal evidence of such a link, our simulations lend credence to the hypothesis that universities might have been key to triggering the rise of this new science. This view is corroborated by the analysis of the gains from the market at the local level. In our simulations, the universities gaining the most from agglomeration and sorting forces in the period 1450–1526 are Rome, Bologna, Padua, Paris, and Louvain. In the period 1527–1617, one can add Cambridge and Leiden to the list. Those were indeed leading universities for the Scientific Revolution.

## 5. Conclusion

In European universities, students were educated by a plurality of masters, and schools were open to students and scholars from all parts of Europe. In this paper, we map the European academic market in the medieval and early modern times. We build an original database of thousands of scholars from university sources to study the location pattern of scholars over the time span 1000–1800. The quality of scholars is measured using information provided by Worldcat and Wikipedia. Using a multinomial logit, we show that scholars tend to agglomerate in the best universities, and that this phenomenon is more pronounced within the upper tail of the talent distribution: Better scholars are more sensitive to the quality of the university (positive sorting), and migrate over greater distances (positive selection). Agglomeration and sorting patterns influenced the distribution of upper-tail human capital across Europe, and contributed to fostering university output at the dawn of the Scientific Revolution.

Agglomeration, sorting, and selection testify to market forces at work. They appear when there is a competition between universities to attract scholars, or among scholars to land the best positions available. This contrasts with a common but mistaken view that markets are a modern phenomenon, but our findings are in line with the qualitative evidence put forward by historians such as Denley (2013) who describes the emergence in Italy of "an efficient and sometimes cut-throat academic market, with its own 'transfer season,' clearly defined hierarchies, rocketing salaries for the top players, and a mentality of academic celebrity that fed off it." At the European level, two features may have helped the integrated academic market to develop. First, political fragmentation, together with competition between church and state, prevented a centralized control by the political sphere of universities. Second, the use of Latin as a *lingua franca*, which persisted late into the early modern period and allowed scholars to teach anywhere at low cost.

Our simulations suggest that the presence of a functioning academic market in Europe helped universities to produce more at the dawn of European primacy. This might have paved the way for the enlightenment, humanistic, and scientific revolutions. We thus provide some quantitative support to the views developed by historians, such as Huff (2017)'s approach to the Scientific Revolution, comparing the West to China and the Islamic World. Huff suggests that the origins of the stronger support given to scientific inquiry in the West during the early modern period can be traced back to medieval times when European institutions were reconstructed. In this context, he sees the rise of European universities in the Middle Ages and their long-run contribution to the Scientific Revolution as highly significant.

## References

Abowd, John, Francis Kramarz, and David Margolis (1999). "High Wage Workers and High Wage Firms." *Econometrica*, 67, 251–333.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson (2012). "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration." *American Economic Review*, 102(5), 1832–56.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson (2014). "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy*, 122, 467–506.

Abramitzky, Ran and Fabio Braggion (2006). "Migration and Human Capital: Self-Selection of Indentured Servants to the Americas." *Journal of Economic History*, 66, 882–905.

Agrawal, Ajay, Devesh Kapur, John McHale, and Alexander Oettl (2011). "Brain Drain or Brain Bank? The Impact of Skilled Emigration on Poor-Country Innovation." *Journal of Urban Economics*, 69, 43–55.

Akcigit, Ufuk, Salomé Baslandze, and Stefanie Stantcheva (2016). "Taxation and the International Mobility of Inventors." *American Economic Review*, 106(10), 2930–2981.

Applebaum, Wilbur (2003). *Encyclopedia of the Scientific Revolution: From Copernicus to Newton*. Routledge, New York, NY.

Arabeyre, Patrick, Jean-Louis Halpérin, and Jacques Krynen (2007). *Dictionnaire historique des juristes français (XIIe-XXe siècle)*. Presses Universitaires de France, Paris.

Michaud, Joseph-François. (1811). *Bibliographie universelle ancienne et moderne*, 1811–1862, Michaud frères, Paris.

Beine, Michel, Frederic Docquier, and Caglar Ozden (2011). "Diasporas." *Journal of Development Economics*, 95, 30–41.

Beltrán, Tapia, Francisco J. and Santiago de Miguel Salanova (2017). “Migrants’ Self-Selection in the Early Stages of Modern Economic Growth, Spain (1880–1930).” *Economic History Review*, 70, 101–121.

Bertoli, Simone and Jesús Fernández-Huertas Moraga (2015). “The Size of the Cliff at the Border.” *Regional Science and Urban Economics*, 51, 1–6.

Bertoli, Simone and Ilse Ruyssen (2018). “Networks and Migrants’ Intended Destination.” *Journal of Economic Geography*, 18, 705–728.

Blasutto, Fabio and David de la Croix (2023). “Catholic Censorship and the Demise of Knowledge Production in Early Modern Italy.” *The Economic Journal*, 133, 2899–2924.

Bogart, Dan (2011). “Did the Glorious Revolution Contribute to the Transport Revolution? Evidence from Investment in Roads and Rivers 1.” *Economic History Review*, 64, 1073–1112.

Borjas, George J. (1987). “Self-Selection and the Earnings of Immigrants.” *American Economic Review*, 77(4), 531–553.

Bosker, Maarten, Eltjo Buringh, and Jan Luiten Van Zanden (2013). “From Baghdad to London: Unraveling urban Development in Europe, the Middle East, and North Africa, 800–1800.” *Review of Economics and Statistics*, 95, 1418–1437.

Bourchenin, Pierre Daniel (1882). *Étude sur les académies protestantes en France au XVIe et au XVIIe siècle*. Grassart, Paris.

Breschi, Stefano and Francesco Lissoni (2009). “Mobility of Skilled Workers and Co-Invention Networks: An Anatomy of Localized Knowledge Flows.” *Journal of Economic Geography*, 9, 439–468.

Buggle, Johannes, Thierry Mayer, Sakalli Seyhun Orcan, and Mathias Thoenig (2023). “The Refugee’s Dilemma: Evidence from Jewish Migration out of Nazi Germany.” *Quarterly Journal of Economics*, 138, 1276–1345.

Cantoni, Davide, Jeremiah Dittmar, and Noam Yuchtman (2018). “Religious Competition and Reallocation: The Political Economy of Secularization in the Protestant Reformation.” *Quarterly Journal of Economics*, 133, 2037–2096.

Cantoni, Davide and Noam Yuchtman (2014). “Medieval Universities, Legal Institutions, and the Commercial Revolution.” *Quarterly Journal of Economics*, 129, 823–887.

Cardell, N. Scott (1997). “Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity.” *Econometric Theory*, 13, 185–213.

Chort, I. and J. N. Senne (2018). “You’ll be a Migrant My Son: Accounting for Migrant Selection within the Household.” *Economic Development and Cultural Change*, 66, 217–263.

Chort, Isabelle and Jean-Noël Senne (2015). “Selection into Migration within a Household Model: Evidence from Senegal.” *World Bank Economic Review*, 29, 247–256.

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 (2008). *A Farewell to Alms: A Brief Economic History of the World*. Princeton University Press, Princeton, NJ.

Cobban, Alan B. (1985). “Elective Salaried Lectureships in the Universities of Southern Europe in the Pre-Reformation Era.” *The Bulletin of the John Rylands Library, University of Manchester*, 67, 662–687.

Cottini, Elena, Paolo Ghinetti, and Simone Moriconi (2019). “Higher Education Supply, Neighbourhood Effects and Economic Welfare.” CESifo Working Paper No. 7483.

Courtenay, William J. (1999). *Parisian Scholars in the Early Fourteenth Century: A Social Portrait*, *Cambridge Studies in Medieval Life and Thought: Fourth Series*, Vol. 41. Cambridge University Press, Cambridge.

Covarrubias, Matías, Jeanne Lafortune, and José Tessada (2015). “Who Comes and Why? Determinants of Immigrants Skill Level in the Early XXth Century US.” *Journal of Demographic Economics*, 81, 115–155.

Croissant, Yves (2012). “Estimation of Multinomial Logit Models in R: The Mlogit Packages.” Note available from R-Forge.

Dahl, Gordon B. (2002). “Mobility and the Return to Education: Testing a Roy Model with Multiple Markets.” *Econometrica*, 70, 2367–2420.

De la Croix, David, Frédéric Docquier, Alice Fabre, and Robert Stelter (2023). “Are Scholars’ Wages Correlated with their Human Capital Index?” *Repertorium Eruditorum Totius Europae*, 10, 10–16.
De la Croix, David, Matthias Doepke, and Joel Mokyr (2018). “Clans, Guilds, and Markets: Apprenticeship Institutions and Growth in the Preindustrial Economy.” *Quarterly Journal of Economics*, 133, 1–70.
De la Croix, David and Omar Licandro (2015). “The longevity of famous people from Hammurabi to Einstein.” *Journal of Economic Growth*, 20, 263–303.
De la Croix, David and Fabio Mariani (2015). “From Polygyny to Serial Monogamy: A Unified Theory of Marriage Institutions.” *Review of Economic Studies*, 82, 565–607.
De la Croix, David and Mara Vitale (2022). “A Timeline of Medieval Universities.” *Repertorium Eruditorum Totius Europae*, 8, 1–5.
De la Croix, David and Mara Vitale (2023). “Women in European Academia before 1800 - Religion, Marriage, and Human Capital.” *European Review of Economic History*, 27(4), 506–532.
Denley, Peter (2013). “’Medieval’, ’Renaissance’, ’Modern’. Issues of Periodization in Italian University History.” *Renaissance Studies*, 27, 487–503.
Dittmar, Jeremiah (2019). *“Economic Origins of Modern Science: Technology, Institutions, and Markets.”* London School of Economics.
Dittmar, Jeremiah and Ralf Meisenzahl (2020). “Public Goods Institutions, Human Capital, and Growth: Evidence from German History.” *Review of Economic Studies*, 87, 959–996.
Dowey, James (2017). *Mind Over Matter: Access to Knowledge and the British Industrial Revolution*. Ph.D. thesis, The London School of Economics and Political Science (LSE).
Drüll, Dagmar (1991). *Heidelberger Gelehrtenlexikon: 1652–1802*. Springer Berlin Heidelberg, Berlin, Heidelberg.
Dulieu, Louis (1979). *La médecine à Montpellier, vol II: La Renaissance*. Les presses universelles, Avignon.
Dustmann, C. and A. Okatenko (2014). “Globalization, Brain Drain, and Development.” *Journal of Development Economics*, 110, 52–63.
Dustmann, Ch., K. Vasiljeva, and A. P Damm (2019). “Refugee Migration and Electoral Outcomes.” *Review of Economic Studies*, 86, 2035–2094.
Eloy, Nicolas F. J. (1755). *Dictionnaire Historique De La Médecine: Contenant Son Origine, Ses Progrès, ses Révolutíons, ses Sectes & son Etat chez différens Peuples*. Bassompierre, Liège.
Miguelez, Ernest, Julio Raffo, and Carsten Fink (2013). “The Global Race for Inventors.” VOXEU Column https://cepr.org/voxeu/columns/global-race-inventors. Accessed November 12, 2023.
Foster, Andrew and Mark R. Rosenzweig (2002). “Household Division and Rural Economic Growth.” *Review of Economic Studies*, 69, 839–869.
Frijhoff, Willem (1996). “Patterns.” In *A History of the University in Europe. Vol. II: Universities in Early Modern Europe (1500–1800)*, edited by Hilde de Ridder-Symoens , chap. 2. Cambridge University Press.
Gallin, Joshua Hojvat (2004). “Net migration and state labor market dynamics.” *Journal of Labor Economics*, 22, 1–21.
Galor, Oded (2011). *Unified Growth Theory*. Princeton University Press, Princeton.
Galor, Oded and Omer Moav (2002). “Natural Selection and the Origin of Economic Growth.” *Quarterly Journal of Economics*, 117, 1133–1191.
Genet, Jean-Philippe (2019). “Projet Studium Parisiense.” Université Paris 1.
Gibson, John and David McKenzie (2014). “Scientific Mobility and Knowledge Networks in High Emigration Countries: Evidence from the Pacific.” *Research Policy*, 43, 1486–1495.
Gorochov, Nathalie (2012). *Naissance de l’université. Les écoles de Paris d’Innocent III à Thomas d’Aquin (v. 1200-v. 1245)*. Honoré Champion, Paris.
Greif, Avner (2006). “Family Structure, Institutions, and Growth: The Origins and Implications of Western Corporations.” *American Economic Review*, 96(2), 308–312.
Grogger, Jeffrey and Gordon H. Hanson (2011). “Income Maximization and the Selection and Sorting of International Migrants.” *Journal of Development Economics*, 95, 42–57.
Grogger, Jeffrey and Gordon H. Hanson (2015). “Attracting Talent: Location Choice of Foreign-Born PhDs in the United States.” *Journal of Labor Economics*, 33(S1), 5–38.

Harris, Jonathan (1995). *Greek Emigres in the West 1400–1520*. Porphyrogenitus, Camberley.
Hatton, Timothy J. (2016). “Refugees, Asylum Seekers, and Policy in OECD Countries.” *American Economic Review*, 106(5), 441–445.
Hatton, Timothy J. (2017). “Refugees and Asylum Seekers, the Crisis in Europe and the Future of Policy.” *Economic Policy*, 32, 447–496.
Hatton, Timothy J. (2020). “Asylum Migration to the Developed World: Persecution, Incentives, and Policy.” *Journal of Economic Perspectives*, 34(1), 75–93.
Heiss, Florian (2002). “Structural Choice Analysis with Nested Logit Models.” *Stata Journal*, 2, 227–252.
Herbermann, Charles George (1913). *The Catholic Encyclopedia*. Encyclopedia Press, Incorporated, New York.
Hermans, Jos M. M. and Marc Nelissen (2005). *Charters of foundation and early documents of the universities of the Coimbra Group*. Leuven University Press, Louvain.
Hoisl, Karin (2007). “Tracing Mobile Inventors–The Causality Between Inventor Mobility and Inventor Productivity.” *Research Policy*, 36, 619–636.
Huff, Toby E. (2017). *The Rise of Early Modern Science: Islam, China, and the West*. Cambridge University Press.
Jaeger, C. Stephen (2013). *The Envy of Angels: Cathedral Schools and Social Ideals in Medieval Europe, 950-1200*. University of Pennsylvania Press, Philadelphia.
Junius Institute (2013). “Post Reformation Digital Library—Scholastica.” Available at https://www.prdl.org. Accessed November 12, 2023.
Keane, Michael P. and Kenneth I. Wolpin (1997). “The Career Decisions of Young Men.” *Journal of Political Economy*, 105, 473–522.
Kedar, Benjamin (1995). “Benvenutus Grapheus of Jerusalem, an Oculist in the Era of the Crusades.” *Korot*, 11, 14–41.
Kerr, Sari Pekkala, William Kerr, Çağlar Özden, and Christopher Parsons (2016). “Global Talent Flows.” *Journal of Economic Perspectives*, 30(4), 83–106.
Kerr, Sari Pekkala, William Kerr, Çağlar Özden, and Christopher Parsons (2017). “High-Skilled Migration and Agglomeration.” *Annual Review of Economics*, 9, 201–234.
Kerr, Sari Pekkala, William R. Kerr, and William F. Lincoln (2015a). “Firms and the Economics of Skilled Immigration.” *Innovation Policy and the Economy*, 15, 115–152.
Kerr, Sari Pekkala, William R. Kerr, and William F. Lincoln (2015b). “Skilled Immigration and the Employment Structures of US firms.” *Journal of Labor Economics*, 33, S147–S186.
Kerr, William R. (2008). “Ethnic Scientific Communities and International Technology Diffusion.” *Review of Economics and Statistics*, 90, 518–537.
Kindleberger, Charles P. (1996). *World Economic Primacy: 1500-1990*. Oxford University Press, Oxford.
Kremer, Michael (1993). “The O-Ring Theory of Economic Development.” *The Quarterly Journal of Economics*, 108, 551–575.
Landes, David S. (1998). *The Wealth and Poverty of Nations: Why are Some so Rich and Others so Poor*. WW Norton, New York.
Laouenan, Morgane, Palaash Bhargava, Jean-Benoît Eyméoud, Olivier Gergaud, Guillaume Plique, and Etienne Wasmer (2022). “A Cross-Verified Database of Notable People, 3500BC-2018AD.” *Scientific Data*, 9, 290.
Laudel, Grit (2003). “Studying the Brain Drain: Can Bibliometric Methods Help?” *Scientometrics*, 57, 215–237.
Maddison, Angus (2007). *The World Economy Volume 1: A Millennial Perspective Volume 2: Historical Statistics*. OECD, Paris.
Manchin, Miriam and Sultan Orazbayev (2018). “Social Networks and the Intention to Migrate.” *World Development*, 109, 360–374.
Manning, Alan and Barbara Petrongolo (2017). “How Local are Labor Markets? Evidence from a Spatial Job Search Model.” *American Economic Review*, 107(10), 2877–2907.
Manuel, Frank E. (1968). *A Portrait of Isaac Newton*. Harvard University Press, Cambridge MA.
McClellan, James E. (1985). *Science Reorganized: Scientific Societies in the Eighteenth Century*. Columbia University Press, New York.

McFadden, Daniel (1974). "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics (Chapter 2)*, edited by P. Zarembka . Academic Press: New york, pp. 65–119.

McFadden, Daniel (1978). "Modeling the Choice of Residential Location." *Transportation Research Record*, 673, 72–77.

Miguelez, Ernest and Rosina Moreno (2013). "Do Labour Mobility and Technological Collaborations Foster Geographical Knowledge Diffusion? The Case of European Regions." *Growth and Change*, 44, 321–354.

Moed, Henk, M'hamed Aisati, and Andrew Plume (2013). "Studying Scientific Migration in Scopus." *Scientometrics*, 94, 929–942.

Mokyr, Joel (2010). *The Enlightened Economy: An Economic History of Britain 1700-1850*. Yale University Press, New Haven.

Mokyr, Joel (2016). *A Culture of Growth: The Origins of the Modern Economy*. Princeton University Press, Princeton.

Monroe, Paul (2000). *Encyclopaedia of History of Education*. Cosmo, New Delhi.

Nadal, Joseph Cyprien (1861). *Histoire de l'Université de Valence, et des autres établissements d'instruction de cette ville, etc*. Impr. E. Marc Aurel, Valence.

Navarro-Brotons, Victor (2006). "The Cultivation of Astronomy in Spanish Universities in the Latter Half of the 16th Century." In *Universities and Science in the Early Modern Period*. Springer, pp. 83–98.

Neal, Derek (1999). "The Complexity of Job Mobility Among Young Men." *Journal of Labor Economics*, 17, 237–261.

Ortega, Francesc and Giovanni Peri (2013). "The Effect of Income and Immigration Policies on International Migration." *Migration Studies*, 1, 47–74.

Özak, Ömer (2010). "The Voyage of Homo-Economicus: Some Economic Measures of Distance." Department of Economics, Brown University.

Özak, Ömer et al. (2018). "Distance to the Pre-Industrial Technological Frontier and Economic Development." *Journal of Economic Growth*, 23, 175–221.

Paquet, Jacques (1958). "Salaires et prébendes des Professeurs de l'université de Louvain au XVe siècle." *Studia universitatis "lovanium"*, 2, 1–36.

Pierson, Alice Sharp and Peter Cotgreave (2000). "Citation Figures Suggest that the UK Brain Drain is a Genuine Problem." *Nature*, 407, 13.

Polanyi, Karl (1944). *The Great Transformation*, Vol. 2. Beacon Press, Boston.

Post, Gaines (1932). "Masters' Salaries and Student-Fees in the Medieval Universities." *Speculum*, 7, 181–198.

Rangeard, Pierre and Albert Lemarchand (1868). *Histoire de l'Université d'Angers*. Barassé, Angers.

Rashdall, Hastings (1895). *The Universities of Europe in the Middle Ages*. Clarendon Press, Oxford.

Revelt, David and Kenneth Train (1998). "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." *Review of Economics and Statistics*, 80, 647–657.

Sarrias, Mauricio et al. (2016). "Discrete Choice Models with Random Parameters in R: The Rchoice Package." *Journal of Statistical Software*, 74, 1–31.

Schich, Maximilian, Chaoming Song, Yong-Yeol Ahn, Alexander Mirsky, Mauro Martino, Barabasi Albert-Laszlo, and Dirk Helbing (2014). "A Network Framework of Cultural History." *Science*, 345, 558–562.

Schmutz, Benoit and Modibo Sidibe (2019). "Frictional Labour Mobility." *Review of Economic Studies*, 86, 1779–1826.

Serafinelli, Michel and Guido Tabellini (2022). "Creativity Over Time and Space: A Historical Analysis of European Cities." *Journal of Economic Growth*, 27, 1–43.

Slottved, Ejvind (1978). *Lærestole og lærere ved Københavns Universitet 1537-1977*. Samfundet for dansk Genealogi og Personalhistorie, Copenhagen.

Sommervogel, Carlos (1890). *Bibliothèque de la Compagnie de Jésus*. Oscar Schepens, Brussels.

Spitzer, Yannay and Ariell Zimran (2018). "Migrant Self-Selection: Anthropometric Evidence from the Mass Migration of Italians to the United States, 1907–1925." *Journal of Development Economics*, 134, 226–247.

Squicciarini, Mara P. and Nico Voigtländer (2015). "Human Capital and Industrialization: Evidence from the Age of Enlightenment." *Quarterly Journal of Economics*, 130, 1825–1883.
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.
Stephan, Paula E. and Sharon G. Levin (2001). "Exceptional Contributions to US Science by the Foreign-Born and Foreign-Educated." *Population Research and Policy Review*, 20, 59–79.
Sussman, Nathan (2006). "Income Inequality in Paris in the Heyday of the Commercial Revolution." Working paper.
Taisand, Pierre (1721). *Les Vies des plus Célèbres Jurisconsultes de Toutes Les Nations, Tant Anciens que Modernes, Savoir, Latins ou Romains, Francois, Espagnols, Italiens, Allemans, Anglois, Hollandois*. Sevestre, Paris.
Trippl, Michaela (2013). "Scientific Mobility and Knowledge Transfer at the Interregional and Intraregional Level." *Regional Studies*, 47, 1653–1667.
Zanardello, Chiara (2023). "Market forces in Italian academia today (and yesterday)." *Scientometrics*, 128:651–698.
Zhao, Xinyi, Samin Aref, Emilio Zagheni, and Guy Stecklov (2022). "Return Migration of German-Affiliated Researchers: Analyzing Departure and Return by Gender, Cohort, and Discipline using Scopus Bibliometric Data 1996–2020." *Scientometrics*, 127, 7707–7729.

## Supplementary Material

Supplementary data are available at *JEEA* online.

*Journal of the European Economic Association* 2024 22(4):1541–1589
https://doi.org/10.1093/jeea/jvad061