

Katholieke Universiteit Leuven
Faculty of Arts
Research Group of Early Modern History



Université catholique de Louvain
Faculty of Economic, Social and Political Sciences and Communication
Economics School of Louvain
Louvain Institute of Data Analysis and Modeling in economics and statistics
Institute of Economic and Social Research

# Bridging Data Science and Intellectual History: Knowledge Networks in and around the Old University of Louvain (1425-1797)

Doctoral Dissertation submitted by

Rossana Scebba

in fulfillment of the requirements
for the joint degree of Doctor in History (KU Leuven) and Economics and Management (UCLouvain)

Thesis committee:

| | | |
|---|---|---|
| Prof. | Margherita Fantoli (Supervisor) | Literary Theory and Cultural Studies, KU Leuven |
| Prof. | David de la Croix (Supervisor) | IRES/LIDAM, UCLouvain |
| Prof. | Violet Soen (Co-supervisor) | Early Modern History, KU Leuven |
| Prof. | Maïka de Keyzer (Chair) | Medieval History, KU Leuven |
| Prof. | Dirk van Miert | Huygens Instituut voor Nederlandse Geschiedenis en Cultuur |
| Prof. | Fabio Mariani | IRES/LIDAM, UCLouvain |
| Prof. | Mark Depauw | Ancient History, KU Leuven |
| Prof. | Romain Wacziarg | UCLA Anderson School of Management |

Leuven, Belgium
September, 2026

**Use of Generative AI**

I did use the generative AI assistance tools ChatGPT and Claude.ai.
Generative AI was used for generating programming code. Specifically, it was used to generate and debug Python and R code for data cleaning and analysis; I checked, tested, and re-elaborated the code myself, and the analytical decisions are my own.
Generative AI was used for mere language assistance, as English is not my first language.
The text/code/images in this thesis are my own (unless otherwise specified) and generative AI has only been used in accordance with the KU Leuven guidelines and appropriate references have been added. I have reviewed and edited the content as needed and I take full responsibility for the content of the thesis.

# Acknowledgements

This thesis is the culmination of four years of a PhD I was very privileged to undertake. Many people accompanied me along the way, and I am grateful to all of them.

First, thanks to the people who made it possible: my supervisors, Margherita, Violet and David. It was a sunny day in Leuven, on 27 June 2022, when I first saw the three of you together, at my interview for the position. I doubt any of you could have anticipated the challenges that this double-degree interdisciplinary project would put in your path, but I am thankful that you never faltered before them. It has been an honor to learn from you and to work with you. Thank you for your generosity, your patience and your trust.

To my committee and jury members, Dirk, Fabio, Mark and Romain: thank you for your feedback, which has considerably improved this work and will continue to shape the research still in progress.

To the professors, researchers and fellow PhD students at KU Leuven and UCLouvain whose time there overlapped with mine.

To Francesco and Rosaria, who introduced me to academic research.

And finally, to my dear ones, for your unconditional love and support.

# Contents

**Introduction** **xi**

**Overlap with previous publications** **xv**

**1 Integrating Library and Prosopographical Data in the Early Modern Publication Network of the University of Louvain** **1**
1.1 Introduction . . . . . 3
1.2 Data preparation: disambiguating the *Collectio academica antiqua* 6
1.3 Networking library metadata: State of the art . . . . . 10
1.4 Network representation of the Caa . . . . . 12
1.5 Censorship in the Caa network . . . . . 17
1.6 Defying the Erasmus bias? Comparing network centrality with RETE's human capital index . . . . . 18
1.7 Scholars' sociability in early modern Louvain's intellectual landscape . . . . . 21
1.8 Conclusions . . . . . 25

**2 Bibliographic Metadata as Relational Data: A Cross-Disciplinary Methodological Reflection** **27**
2.1 Introduction . . . . . 29
2.2 What is in a Catalog? Anatomy of Bibliographic Metadata . . 32
2.3 From Structured to Relational Data . . . . . 36
2.3.1 Identifying Historical Actors: Disambiguation and Authority Alignment . . . . . 36
2.3.2 Geolocating Place Names . . . . . 39
2.3.3 Classifying Content: Assigning Topical Terms . . . . . 45
2.4 Modeling Catalog Data for Analysis . . . . . 48
2.4.1 Adopting a Network Perspective . . . . . 48

2.4.2 The Multilayer Approach . . . . . 50
2.4.3 Insights from Quantitative Economic History . . . . . 56
2.5 Conclusions . . . . . 58
Appendices . . . . . 60
2.A Mapping of Roles . . . . . 60
2.B Per-layer Network Analysis . . . . . 61
2.C Inter-layer Network Analysis . . . . . 62

**3 Pre-modern academic institutions as conduits for the potential spread of ideas 63**
3.1 Introduction . . . . . 64
3.2 Data and Methodology . . . . . 71
3.2.1 88,792 scholars . . . . . 71
3.2.2 Definition of the affiliation network $\mathbb{G}$ . . . . . 73
3.2.3 Ideas, inventors and exposed scholars . . . . . 79
3.2.4 Epidemiological model . . . . . 81
3.3 Structural estimation . . . . . 84
3.3.1 Methodology . . . . . 84
3.3.2 Auxiliary model 1: Botanical Realism . . . . . 84
3.3.3 Auxiliary model 2: Mathematical Astronomy . . . . . 87
3.3.4 Results . . . . . 89
3.4 Further empirical assessments . . . . . 91
3.4.1 Scholasticism and Protestantism . . . . . 92
3.4.2 Anti-Judaism and the Persecution of Jews . . . . . 98
3.4.3 Finding Atlantis: A True Story of Genius and Madness 100
3.5 Counterfactual experiments . . . . . 101
3.5.1 Placebo inventors of Botanical Realism . . . . . 102
3.5.2 Removing components of the network . . . . . 105
3.6 Conclusions . . . . . 112
Appendices . . . . . 114
3.A Descriptive Statistics . . . . . 114
3.A.1 Completeness of the data . . . . . 114
3.A.2 Universities, academies, scholars over time . . . . . 118
3.B Definition of active periods . . . . . 118
3.C Contextual information . . . . . 121
3.C.1 Botanical Realism and botanic gardens . . . . . 121
3.C.2 Mathematical Astronomy and astronomical observatories 122

3.D Empirical Assessments . . . . . 126
3.D.1 Descriptive Statistics . . . . . 126
3.D.2 Botanical Realism . . . . . 126
3.D.3 Mathematical Astronomy . . . . . 128
3.E Structural estimation . . . . . 130
3.F Tests for the proportionality of Hazard Functions . . . . . 132
3.G Example of scholastic reasoning . . . . . 133
3.H Augmenting Rubin's data . . . . . 133
3.I Additional results with alternative ideas: Practical Surgery and Civil Law . . . . . 135
3.J Placebo Inventors of Botanical Realism . . . . . 137
3.J.1 Centrality of the Placebo inventors . . . . . 140
3.K Placebo networks and the role of the Jesuits . . . . . 141

**4 Lov'Ego: Printers and their Partnerships in the Academic Book Trade of Early Modern Louvain (1501–1797) 145**
4.1 Introduction . . . . . 146
4.2 Lov'Ego: Ego-networks of Selected Louvain Printers (1501-1797) 149
4.3 The Printer-Scholar: Academic networks of Martens and Rescius 161
4.4 The Jobbing Printer: University Commissions for Johannes Masius (d. 1616) . . . . . 168
4.5 The Printer-Propagandist: The Jansenist Cause of Nempaeus and Denique . . . . . 170
4.6 Conclusions . . . . . 177
Appendices . . . . . 180
4.A Detecting Johannes Masius (d. 1616) in Caa/STCV . . . . . 180

**5 Network-mediated screening under uncertainty: Evidence from early modern academic publishing 183**
5.1 Introduction . . . . . 184
5.2 Related literature . . . . . 185
5.3 Historical context . . . . . 187
5.3.1 The printing industry and academic authors . . . . . 187
5.3.2 The Low Countries . . . . . 189
5.4 Data . . . . . 193
5.5 Empirical strategy . . . . . 196
5.5.1 A random utility model of the printer's choice . . . . . 196
5.5.2 Variables construction and specifications . . . . . 198

5.6 Results . . . . . 200
5.6.1 Robustness checks . . . . . 209
5.7 Discussion and conclusions . . . . . 213
Appendices . . . . . 216
5.A Supplementary tables and figures . . . . . 216

**6 General conclusion** **221**

# Introduction

For centuries, the dominant image of the scholar has been one of solitude, detached from the world in an ivory tower. It celebrated the individual thinker but rarely the community that sustained him.

Yet many scholars, especially those attached to universities and academies, lived within a vibrant social fabric sustained by institutions.

Medieval and early modern professors teaching at the same university shared lecture halls, took part in institutional governance, and in some cases lived under the same roof. Academies, from the literary circles of sixteenth-century Italy to the scientific societies of the Enlightenment, provided additional settings for sustained intellectual exchange. Against this backdrop, the circulation of printed books carried knowledge further than ever before. But the production of books was itself a collaborative enterprise, one that bound scholars to printers and publishers in yet another layer of community.

Thus, even when knowledge was created individually rather than collaboratively, it emerged within a collective, multilayered setting that made it possible in the first place.

This dissertation studies such communities through the lens of networks: academic networks of scholars affiliated with universities and academies, and print networks of authors, printers, publishers, editors, and censors whose names appear together on the title pages of early modern books. It sits at the intersection of computational history and quantitative economic history, in the traditions, respectively, of the humanities and of the social sciences. As such, it attempts to operate a synthesis between the two disciplinary approaches: sometimes by combining them in hybrid ways, sometimes by enriching the perspective of one with insights from the other.

It deploys computational methods throughout, including network visualization and analysis, descriptive statistics, and more advanced econometric tools, and gradually incorporates different dimensions of the

study of knowledge production, moving from the institutional fabric of academic communities to the commercial and confessional dynamics of the early modern book trade.

Both traditions have increasingly drawn on similar empirical materials, such as bibliographic metadata, prosopographical records, and geospatial data, yet they have developed largely in parallel. Digital humanities have pioneered the use of network visualization, text mining, and linked data to explore cultural production at scale, while quantitative economic history has favored research designs grounded in causal identification, structural modeling, and linked administrative datasets. This thesis does not claim to resolve the differences between these traditions, but it aspires to bridge them by showing that the same data can be made to speak to both, and that each tradition offers tools the other lacks.

The Old University of Louvain serves as both the anchoring case study and the point of departure for broader inquiry. Established by papal charter in the duchy of Brabant in 1425 and abolished by French decree in 1797, Louvain became one of the leading universities of northern Europe, attracting students and scholars from across the continent. Its intellectual output, preserved in curated library collections such as the *Collectio academica antiqua* and documented in the bibliographic repertory of the Short Title Catalog Flanders, constitutes a remarkably rich record of early modern academic life. The individuals behind this output left traces in title pages, imprint statements, and paratextual attributions. These traces, when systematically extracted and linked to prosopographical databases such as the *Repertorium Eruditorum Totius Europae*, make it possible to reconstruct the webs of collaboration and exchange that underpinned knowledge production in the early modern Low Countries and beyond.

The thesis is organized in five chapters. They have been conceived as self-contained essays, but they follow a deliberate progression that moves from methodological groundwork to empirical application, from the local scale of Louvain to the continental scale of European academia and back, gradually combining the different dimensions of these knowledge networks.

Chapter 1, coauthored with Margherita Fantoli, takes the metadata of the *Collectio academica antiqua* as its starting point. We disambiguate and harmonize the collection's authorship and imprint attributions, converting the library holdings into relational data and modeling them as a multi-layer network. Using a rolling-window approach to ensure time-consistent metrics,

we track the network's evolution across three centuries. The analysis highlights how centrality varies across role categories, addresses potential sampling bias inherent in a curated collection, and links shifts in scholars' network positions to prosopographical evidence.

Chapter 2 relies on the Louvain case to offer a cross-disciplinary methodological reflection on the use of bibliographic metadata as relational data. Drawing on practices from both computational history and quantitative economic history, the chapter examines how catalog records, originally designed for item-level description, can be transformed into data structures suitable for longitudinal and network-based analysis. It argues that a multilayer network representation offers the most coherent way to preserve the role-differentiated forms of participation encoded in bibliographic sources, rather than collapsing them into flat co-occurrence projections.

Chapter 3, coauthored with David de la Croix and Chiara Zanardello, shifts to the continental scale. We reconstruct the European academic affiliation network from approximately 88,800 scholars active between the eleventh and eighteenth centuries and propose an epidemiological model of idea diffusion through institutional co-presence. The link activation probability, the parameter governing the likelihood that face-to-face interaction leads to knowledge transfer, is estimated via indirect inference, benchmarked against the historical spread of botanical realism and mathematical astronomy. The estimated parameter is then applied to study the diffusion of scholasticism and its association with the adoption of Protestantism and anti-Judaic persecution. Counterfactual experiments reveal a network that is globally robust yet locally fragile: while alternative paths nearly always exist, early diffusion depends on narrow institutional channels, and the survival of ideas is partly contingent on chance.

Chapter 4, coauthored with Violet Soen, returns to Louvain to examine the ego-networks of five printers operating at pivotal moments across the early modern era. Introducing *Lov'Ego*, a curated dataset combining records from the *Collectio academica antiqua* and the Short Title Catalogue Flanders, the chapter traces three distinct configurations of printer-scholar partnership: humanist-oriented networks anchored in the *Collegium Trilingue* (Martens and Rescius), commission-based structures characteristic of jobbing printers (Masius), and spiritually driven affiliations shaped by the Jansenist controversy (Nempaeus and Denique).

Chapter 5 formalizes the intuition developed in the preceding chapter

under the lens of a classical economic problem by asking whether printers relied on their accumulated social ties to screen unfamiliar authors, and whether this reliance intensified during periods of institutional disruption. Combining bibliographic data from the Short Title Catalogs of the Netherlands and Flanders with academic affiliation records from the *Repertorium Eruditorum Totius Europae*, I construct a database of over 5,000 printer-author dyads spanning 1476-1800 and estimate a conditional logit model of the printer's choice among potential collaborators. The Dutch Revolt serves as a natural experiment: I find that the predictive power of shared affiliation between a printer's prior collaborators and a new candidate author is near zero before the 1560s, peaks around the Fall of Antwerp in 1585, and fades as stability returns, consistent with networks functioning as screening devices under uncertainty.

Taken together, the five chapters illustrate how the same underlying materials can sustain analyses at very different scales and levels of formalization. The thesis does not argue that computational methods can replace archival scholarship or close reading; it argues that they can complement them, and that the payoff is greatest when the traditions of the humanities and the social sciences are brought into dialog.

# Overlap with previous publications

The following parts of the dissertation are based on submitted, accepted or published articles that I (co)authored. Where a chapter is said to be "based on" a publication, the text is largely identical to that publication, with reworking kept to a minimum (e.g., a few extra footnotes, and formatting for integration into the dissertation):

1. Chapter 1 is based on Scebba and Fantoli (2026): Rossana Scebba and Margherita Fantoli, "Integrating Library and Prosopographical Data in the Early Modern Publication Network of the University of Louvain (1501–1797)," in Violet Soen, Wouter Druwé, Wim François, and Ralph Dekoninck, eds., *Students, Scholars and Their Books at the University of Louvain (1425–1797)*, Vol. 17 of Lectio Series Studium Lovaniense, Turnhout: Brepols, pp. 395–424.

2. Chapter 2 is based on Scebba (2026): Rossana Scebba, "Bibliographic metadata as relational data: A cross-disciplinary methodological reflection," *Journal of Cultural Analytics*, 11 (1), https://doi.org/10.22148/jca.1787.

3. Chapter 3 is based on de la Croix, Scebba and Zanardello (submitted to the *Journal of Political Economy*; circulated in an earlier draft as 2025 CEPR Discussion Paper 20569, "Flora, Cosmos, Salvatio: Pre-modern Academic Institutions and the Spread of Ideas").

4. Chapter 4 is based on Scebba and Soen (submitted and pending acceptance at *Nieuwe Tijdingen*).

# 1

# Integrating Library and Prosopographical Data in the Early Modern Publication Network of the University of Louvain (1501–1797)

*This chapter is co-authored with Margherita Fantoli.*

*This paper presents a potential use case of the Collectio academica antiqua, a cultural heritage collection from the Old University of Louvain (1425-1797). We first disambiguate and harmonize the collection's authorship and imprint attribution metadata. Next, we convert the library holdings into relational data and model them as points of connection among the actors involved in their production. The resulting graph is processed as a multi-layer multiplex network, which we make available for interactive visualization. We analyze this network using a rolling-window approach, which ensures time-consistent metrics and generates a sequence of data snapshots, allowing us to track the network's evolution over time. Our analysis highlights the centrality of censors, emerging as the most prominent role in the network. Additionally,*

*we address the sampling bias of the collection, particularly the overrepresentation of Erasmus. Finally, we examine the diachronic evolution of scholars' influence in early modern Louvain's intellectual network using centrality measures, linking these findings with prosopographical chronicles.*

## 1.1 Introduction

After the Catholic University of Louvain split in 1968, the rare books from the Old University of Louvain (1425–1797) were inherited by both KU Leuven and UCLouvain. Decades later, the two sister universities have been joining efforts to reunite this intellectual heritage into one and the same digital gateway. A key milestone of this renewed cooperation is the reassembly and digitization of parts of the *Collectio academica antiqua* (Caa hereafter) on the online platform *Lovaniensia*.[1] As described in the *Collectiewijzer Erfgoedbibliotheken*,[2] the Caa comprises the works from and around the Old University of Louvain. Held by the KU Leuven Libraries Special Collections, the Caa reflects a deliberate curatorial approach. It represents the Old University period within the broader *Collectio academica*, a collection dedicated to comprehensively documenting Louvain University's history through publications. As was the case for most of the *studia generalia* of the premodern period, the foundation of a university in Louvain provided a breeding ground for intellectual interactions. Establishing a higher learning center laid out a structured setting for academic debate and dissemination of knowledge. After all, bringing literacy, wisdom, and prosperity to Louvain, the duchy of Brabant, and nearby regions was one of the aims expressed by the foundational bull. Scholars in Louvain—whether as professors, lecturers, or visiting academics—were in a privileged position to engage in teaching and research.

With the introduction of the printing press in and around Louvain in the late fifteenth century (Adam, 2017, p. 14), scholarly work increasingly took the form of printed publications, part of which survived and represents the cultural heritage of the Old University, embodied in the Caa. The defining criterion for inclusion in the Caa is the association of a book's creators—authors, publishers, editors, dedicatees, and the likes—with the university. This emphasis on individuals connected to the Old University aligns with efforts to reunite its dispersed academic heritage, given the tumultuous history of the University's Central Library. Founded in 1636, the library faced repeated disruptions that scattered its collections, jeopardizing its integrity and historical continuity. After invading Louvain in 1794, French

[1] See: https://www.lovaniensia.be/ and https://vlaamse-erfgoedbibliotheken.be/en/node/4030.

[2] See: https://collectiewijzer.be/s/nl/col/cweb-10854.

republican institutions 'expertly plundered the library', as documented by historians Chris Coppens, Mark Derez, and Jan Roegiers in their 2005 book on the history of the Louvain University Library, and confiscated parts of it to make it property of the regime (Coppens et al., 2005, p. 71). Shortly after, with the Old University's dissolution in 1797, the ownership of the remaining holdings was transferred to the municipality of Louvain (Smets, 2016, p. 167). The two World Wars brought further devastation, with arson attacks during both conflicts resulting in the loss of invaluable materials.[3] Following the split of the University towards the end of the 1960s, its library holdings, including rare books and manuscripts, were divided between the newly founded KU Leuven and UCLouvain, in first instance prioritizing donor wishes, in second instance maintaining at least one copy of duplicate works at each location, and eventually resorting to an *odd-even* division for the remaining items, yet holding multi-volume editions together.

In this chapter, we aim to explore the intellectual network of the Old University of Louvain by examining trends in the metadata of the Collectio academica antiqua, particularly authorship and imprint attributions. This approach is not without precedent, as a growing body of scholarship is rediscovering metadata, catalogs, and full texts as a rich source of information for network analysis and other computational approaches to history. Examples of this metadata include book titles, authors' names, printers, dedicatees, publication dates, and locations, as well as more specific information from bookbinding, title pages, and paratexts (Constantinidou et al., 2025, p. 107). While the use of library metadata as resources for empirical investigations might offer valuable insights into book history and knowledge production, it comes with important caveats, such as survival and sampling biases. The survival of books and their metadata depends on various factors, including popularity, preservation practices, and historical events. As a result, curated collections might not fully represent the complete range of works originally produced. Additionally, working with metadata requires careful harmonization, as cataloging rules often prioritize original

[3]On the first fire of Louvain during WW1: 'All the works of art, with the exception of a few scorched and damaged professors' heads, and the entire book collection went up in flames on the night of August 25–26, 1914. Not one of the books and manuscripts was saved', from Coppens et al. (2005), p. 165. On the second fire: 'on 16 May 1940 the Library was targeted with artillery grenades from the hills to the east of the town and after the Germans entered the town the following morning, the book store with its nearly one million volumes was ablaze', from Coppens et al. (2005) p. 319.

data over standardization. This can lead to internal inconsistencies, making the rigorous harmonization of the Caa collection's metadata a key contribution of this paper. We detail this process in Section 2.

Keeping in mind the caveats discussed above, the Caa collection presents a unique opportunity to explore the intellectual network of the Universitas Lovaniensis. While not a perfect reflection of the past, the collection is arguably representative of the scholarly community of Louvain at the time, given the inclusion criterion discussed above. The metadata holds the potential to map the web of relationships among individuals involved in the university's intellectual landscape. This is what we strive to achieve in the second part of the paper, starting from section 3. We model the various individuals involved in the creation and distribution of the Caa holdings as the *nodes* or vertices of the network and establish a tie or *edge* among any pair of nodes in case of co-occurrence on the same publication. Given that the Caa spans nearly three centuries, we apply a rolling-window approach to capture the evolving dynamics of the network over time. This method allows us to compute key metrics while accounting for changes in scholarly relationships across different periods, as detailed in Section 4. Sections 5, 6, and 7 present our network analysis results. We acknowledge the limitations that this approach may bear. As stated by the historian Charles Wetherell, '[b]ecause historians are plagued by an incomplete historical record and imperfect understandings of past social relations, [historical social network analysis] remains an inherently problematic enterprise' (Wetherell, 1998). However, the richness of the Caa metadata and of the collection itself, in conjunction with the state-of-the-art techniques in data science and social network analysis, offers an unprecedented opportunity to integrate a computational approach to history in the fashion of the digital humanities, corroborating with quantitative evidence the network perspective to higher learning emerging from recent research avenues in the field of history of knowledge (Goeing et al., eds, 2021). Section 8 concludes.

## 1.2 Data preparation: disambiguating the *Collectio academica antiqua*

Library metadata often requires extensive cleaning and harmonization before it can be effectively explored and described.[4] This section addresses the specific challenges we encountered with the Caa collection and the steps we took to parse and index the data, which form a key contribution of this work and lay the foundation for the subsequent network analysis. Our study is based on an export of the Caa collection metadata received in July 2023 from the KU Leuven Libraries[5] in the form of a MARC21 XML export.[6] This widely used format for library cataloging combines the standardized set of rules MARC21 with the structured data representation of XML.[7] Each MARC21 XML entry represents a single physical item—in our case, a Caa library holding—identified by a unique control number (tag 001). Within each entry, a series of elements and sub-elements organized using MARC21 fields and subfields describe the various attributes of the holding, such as the title, author(s), subject, and publication information. Table 1 provides an example record to illustrate the format's technicalities. We extracted information from authorship, paratextual and imprint attributions (respectively tags 100, 110, 600, 700, 710, and tags 245, 264, and 953), with the aim of transforming these attributions into standardized representations of individual identities. While the goal of metadata is to faithfully record holdings' information, it might lack internal consistency.[8] Data harmonization requires not only

[4] As Dirk van Miert points out in his 'Afterword' to the 86th volume of the Huntington Library Quarterly, '[f]or a decade, we have been advancing our software engineering, but we have neglected to properly curate our data because we lack the funding to carry out the non-innovative work to index, standardize, and link certain types of data, such as named entities', Van Miert (2023), p. 416

[5] The libraries plan to share the metadata in a repository which will be made available once it is ready for release.

[6] See https://www.loc.gov/marc/bibliographic/.

[7] MARC (an acronym that stands for Machine-Readable Cataloging) is a commonly adopted standard to describe bibliographic materials, such as books, journals, and audiovisual materials, whereas XML (eXtensible Markup Language) is a markup language that allows the representation of structured data using tags and attributes.

[8] 'For various reasons, it is important to comprehend that supporting quantitative, data-intensive research is not the original or intended goal of analytical bibliography', from Lahti

**Table 1.1.** Content description of selected tags and subfields of an example record of the Caa metadata

| Subfield | Content description | Example |
|---|---|---|
| **Tags 100, 600, 700, 710** | | |
| a, v | Personal name | *Estienne, Robert* |
| b | Numeration | *I* |
| c | Titles, words, and places associated with a name | *Paris* |
| d, f | Dates associated with a name | *1502–1560?* |
| e | Relator term (i.e., role in the Caa holding) | *printer* |
| g | Miscellaneous information | *Also active in Geneva 1553–1560* |
| **Tag 245** | | |
| a | Main title (title page) | *Meditationes Graecanicae in artem grammaticam. authore Nicolao Clenardo [. . . ] Emendatiores* |
| **Tag 264** | | |
| a | Place of publication | *Lutetiae* |
| c | Date of publication | *1550* |
| **Tag 953** | | |
| a | Period of publication | *1541–1600* |

consolidating the personal names extracted from metadata fields but also eventually linking said names and related biographical data to actual historical figures. Name disambiguation tackles the issue of identifying whether different name forms refer to the same individual or multiple individuals. Two types of occurrences are possible: either a single individual possesses multiple name variants (e.g., due to changing conventions or Latinized forms), or different individuals share the same name. For instance, the namesakes Thomas Stapleton (1535–1598) and Thomas Stapleton (1622–1694): one was an English Catholic who sought exile in Louvain, where he taught theology, the latter was educated in the Irish College, was professor in canon and civil law and was elected rector multiple times between 1661 and 1681. In our sample, however, we found more prevalent evidence of the former issue. For instance, theologian Martinus Steyaert (1647–1701) is also referenced as Martin, while Adriaan van Baarland and Adrianus Barlandus (1486–1538) represent Dutch and Latin variant spellings for the same Louvain professor. While for the reader it might seem obvious that these name

et al. (2019a), p. 3.

variants refer to the same individual, they must be translated into intelligible input for software analysis (i.e., being linked through a shared identifier).

To address these ambiguities, we used a systematic approach. First, we leveraged a combination of personal names, associated dates, and roles within book production to identify individuals with multiple variant names, using a clustering function coupled with a careful manual inspection. Second, we turned to authority files and biographical data as a means of resolving these ambiguities and establishing greater certainty in our entity resolution process. Specifically, we relied on the ODIS database and the CERL *Thesaurus* to ensure accurate identification of individuals across different naming conventions.[9] By assigning unique identifiers to each disambiguated individual, we transitioned from a collection of names to a more structured representation of historical figures, laying the foundation for our network representation. Overall, our rigorous disambiguation process consolidated variant spelling and aliases, resulting in 4300 uniquely identified individuals and corporate entities. However, since our analysis focuses on individuals, we filtered out corporate entities from the sample early in the process. We observe that around 97% of the persons involved in our subset of the Caa metadata lived in the Early Modern Period, but we find that 3% consists of long-dead scholars, mainly classical authors like Aristotle (384–322 BCE), Publius Cornelius Tacitus (c. 56–117), and Lucius Annaeus Seneca (c. 4 BCE–65 CE) but also, in fewer cases, medieval eruditi. We also find 56 printer's widows, whom the catalog mostly mentions expressly as widows in most of the cases, despite there being instances of imprint attributions using maiden names. This falls in a peculiar case of disambiguation we addressed in our data cleaning process. Leveraging the ODIS and CERL authority files, we

[9]The *ODIS* (*Online Database for Intermediary Structures*) research infrastructure is a comprehensive research tool hosted by KADOC, with a focus on historical and cultural legacy of civil society in Flanders and Belgium; a subgroup of ODIS entries specifically deal with the Old University of Louvain, and are compiled by a series of experts stemming from KU Leuven Archives, KU Leuven Libraries and researchers at the Early Modern History Research Group of the Faculty of Arts at KU Leuven. The *CERL* (*Consortium of European Research Libraries*) Thesaurus is a comprehensive database that links variant spellings, Latinized forms, and pseudonyms common in early modern European printing. In addition to linking variant names, both CERL and ODIS often include biographical notes and genealogies. These pieces of information proved essential in solving ambiguities and confidently linking our records to the correct individuals. More on ODIS and on the CERL Thesaurus, respectively, at: https://www.odis.be/hercules/_en_overODIS.php, and https://www.cerl.org/resources/cerl_thesaurus/main.

can identify 46 of them by their name before marriage.

A rapid exploration of the holdings' publication dates reveals that the earliest attested publication goes back to 1501. Given the introduction of the printing press in Louvain dates back to 1473, it is documented that academic publishing had already taken place by 1474 (Adam, 2009), with the original *incunabula* collection partly lost during both World Wars (Coppens et al., 2005, p. 159-160, and p. 313), then restored via donations and purchases, and then again torn into two parts, after the split of the University in 1968. What survives of these early specimens of print matter requires different preservation conditions, so the KU Leuven University Libraries store it in a separate collection of incunables. Although the Caa is often described as a rare book collection, its metadata reveal that some books date from after 1801 yet are thematically connected to the Old University (e.g., related to its history). For the purposes of our analysis, we establish the end of the timeframe to the year 1797, the year of the abolishment of the Old University of Louvain. In Figure 1, we plot the frequency of Caa rare books published per decade between 1501 and 1797. The stock of Caa publications records an initial peak around the 1640s–1650s and an all-time high in the 1780s. The second peak may reflect the heightened activity in the production of pamphlets for the opposition against Joseph II (Judge, 2018, p. 37).

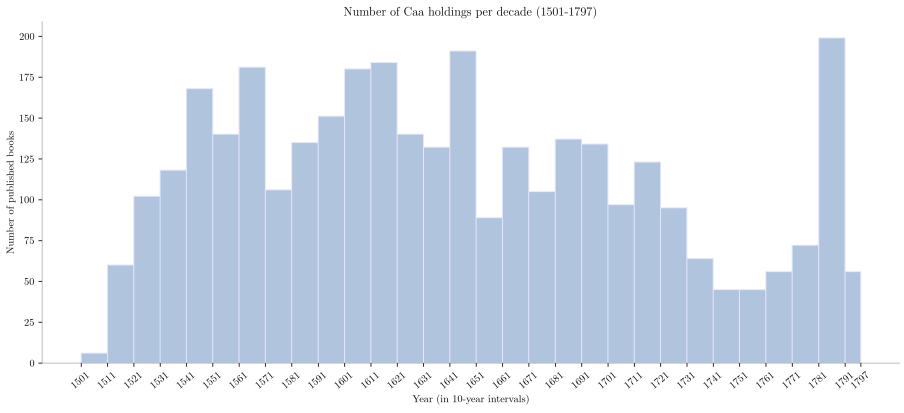



**Figure 1.1.** Decade frequency of Caa holdings (1501–1797). This chart visualizes the volume of printed materials within the Caa collection, with each bar representing a decade.

Figure 2 displays the publication places extracted from the metadata and disambiguated using Wikidata. While Louvain, Antwerp, and Brussels

together account for over 40% of the Caa holdings' places of publication, the collection extends beyond the Low Countries to include cities such as Paris, Lyon, Leiden, Venice, and Basel. This geographical diversity suggests that the scope of the collection is influenced more by the contributors' affiliation with the Old University of Louvain than by the publication location itself. Our analysis of Caa authorship and imprint attributions also reveals that 44% of the library holdings involve one actor in authorship and another in production or distribution, with 13% including a dedication as well. Overall, we observe a predominance of individual authorship: on average, only 1 in 4 holdings involve more than one author. Academic scholars appear in the collection primarily as authors, but they are also occasionally found as dedicatees of Caa rare books alongside prominent clerical figures, monarchs, and *stadtholders*.

## 1.3 Networking library metadata: State of the art

The notion of networks as complex webs of relationships embedded in social systems is not unfamiliar to traditional historical scholarship, yet it is mostly employed as a metaphor to evoke the intricate pattern of connections within a community (Ruffini, 2008; Davies, 2015; Lemercier, 2015a; Preiser-Kapeller, 2020; Rollinger, 2020; Grandjean, 2021). However, the digital turn in the humanities has made it possible to move beyond this abstract understanding. Through data-driven approaches, researchers can rely on Social Network Analysis (SNA) to provide the formal representation of a network starting from a specific case study. This is an unprecedented advance that offers the possibility of shifting from loosely speaking of networks as figures of speech to elaborating tangible, graph objects: real-world networks. SNA draws its tenets from graph theory, a branch of mathematics dealing with the study of graphs, i.e., complex systems of reciprocally connected entities (Wasserman and Faust, 1994; Hanneman and Riddle, 2005). Through the lens of SNA, we can gain some insights into the properties of the system we have chosen to outline as a network, both at the global and the local level (that is, at the level of single entities).

The present chapter explores an application of SNA to the field of history, also known as Historical Social Network Analysis (HSNA). Specifically, we aim to reconstruct the historical network of the publishing world surrounding

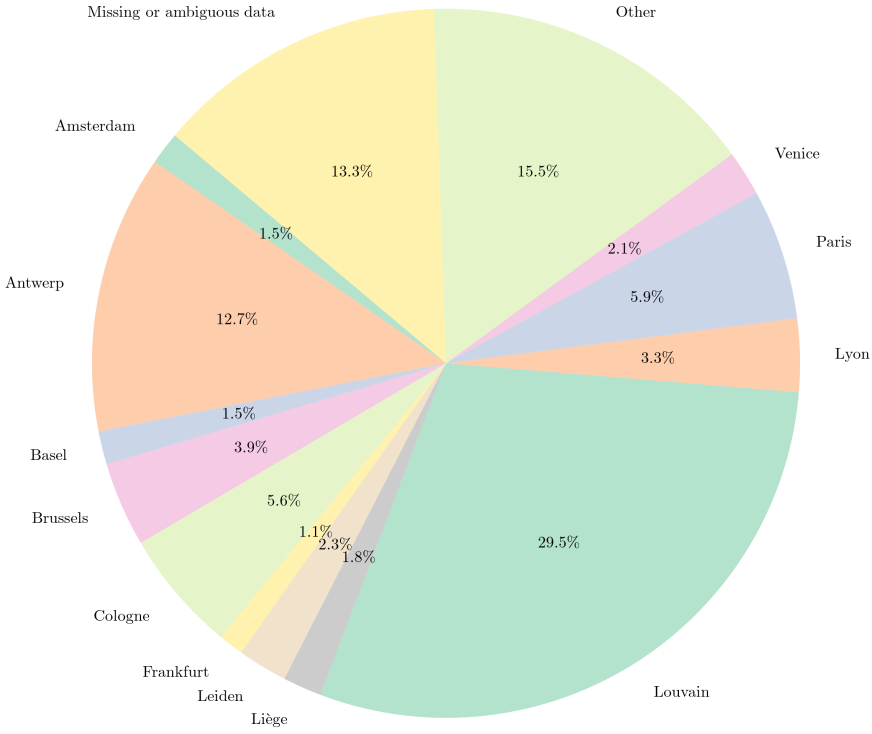



**Figure 1.2.** Places of publication of the Caa holdings, 1501–1797.

the Old University of Louvain. Our approach utilizes library metadata, a rich source of information that is being rediscovered by a growing body of scholarship, alongside catalogs and full texts, for network analysis and other computational approaches to history. In their seminal 2019 paper, Finnish data scientist Leo Lahti and colleagues champion a rediscovery of these types of data. Through the concept of *bibliographic data science*, Lahti and coauthors highlight the potential of library catalogs and bibliographic collections as high-quality research material for quantitative explorations in book history (Lahti et al., 2019a, p. 2). The authors gather metadata from four bibliographies (both national collections and union catalogs) spanning from the sixteenth to

the nineteenth century and uncover some Europe-wide trends regarding the rise of octavo format and 'vernacularization' around the Enlightenment period.

Beyond Lahti and colleagues' work, which leverages metadata fields related to publication time and place, language, and physical dimensions, authorship, paratextual and imprint attributions have emerged as a precious resource for computational approaches to book history, particularly to reconstructing historical networks. Namely, Michael Gavin's study mapping the growth of late seventeenth-century English criticism; John Ladd's exploration of dedication practices due to historical events like political upheavals and religious changes during the Early Modern Period; Yann Ciarán Ryan and Mikko Tolonen's work on the printing industry around the times of the Scottish Enlightenment (Gavin, 2016; Ladd, 2021; Ryan and Tolonen, 2024a). In the works by Gavin, Ladd, and Ryan and Tolonen, network actors are linked based on the co-occurrence of their names in book title pages, often implying joint collaboration in the production of those books. We use a similar approach to the above-mentioned works, starting from the metadata from the Caa collection as source material. Another related compelling contribution in the field is the *Six Degrees of Francis Bacon* project, which conceptualizes a network of early modern persons. Unlike the previous examples, this network does not stem from authorship and imprint attribution but from biographical data (i.e., the Oxford Dictionary of National Biography), and links are inferred from joint mentions in said biographies using natural language techniques (Warren et al., 2016). Similarly to this project, we provide an interactive visualization of the network reconstructed from the Caa metadata elaboration, discussed in the next section.

## 1.4 Network representation of the Caa

To create our network representation, we first structure the parsed metadata into a relational database consisting of three interconnected tables: persons, books, and activities (Figure 3). The activities table is key insofar as it captures the various ways individuals engage with books—authorship, production/distribution, dedication—representing broader categories derived

from the original metadata roles.[10] This infrastructure allows us to represent each Caa holding as a point of connection among its creators, resulting in a two-mode or bipartite network. In a multi-modal network, the entities or *nodes* can be of several types. In our case, they represent either books or individuals involved in their production. Persons are connected to as many books as they contributed to based on authorship, paratextual and imprint attributions from the Caa metadata. In this setting, edges link exclusively books to persons (never books to books, nor persons to persons) and carry information about the specific role of each contributor to each book. To capture the multifaceted nature of involvement, we conceptualize our bipartite network as a multilayer graph with distinct layers for each role. For instance, all the edges representing authorship belong to one layer, while those representing editing belong to another. This multilayer structure will allow us to isolate role-specific relationships throughout our analysis. To focus on connections among individuals, we project the initial two-mode graph into a one-mode network. In this projected network, links between individuals arise from their shared participation in a given book.[11]

Figure 4 presents a network visualization of the projected multilayer graph, where nodes are the individuals involved in the creation of the Caa holdings, and the edges represent their book collaborations. The edges are undirected and weighted to capture the fact that the same pair of nodes may have collaborated more than once. In that case, the larger the number of shared contributions, the thicker the edge will appear. An interactive version of the network, created by elaborating the Caa metadata using Gephi and uploaded to the Retina web application, can be accessed either by scanning the QR code or clicking on the figure itself.[12] The Retina dashboard allows the user to freely zoom the network in and out, hover on the nodes to display their label, but, most importantly,

[10]We acknowledge that our thematic grouping of roles inevitably simplifies the complexities of early modern book production. However, grouping roles variants allows us to construct meaningful thematic layers within the network, at the expense of some nuance.

[11]This approach aligns with Geraerts and Filho, who employ one-mode projections to analyze social dynamics within historical confessional networks. Drawing on parish records, they construct a multi-modal network of persons, religious events like baptisms and marriages, and physical places, ultimately projecting it into a one-mode network in order to focus on the relationships among individuals, based on their attendance to the same religious events. See Geraerts and Vasques Filho (2024).

[12]Bastian et al. (2009); more on Retina at: https://ouestware.gitlab.io/retina/1.0.0-beta.1/.

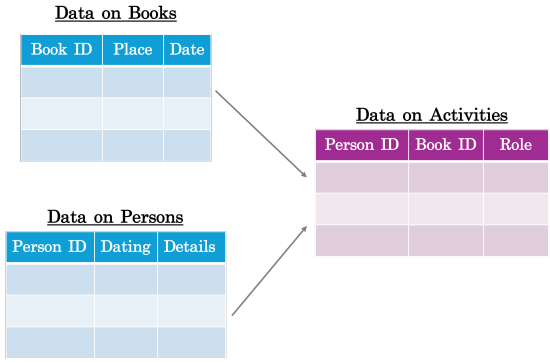



**Figure 1.3.** A visual representation of the infrastructure of our relational database.

input names of interest in the search bar and highlight their connections to other vertices in the network.

Figure 5 shows the authorship layer of the Caa network. Specifically, this graph includes all nodes that exist in this layer—those who were involved, to varying degrees, in authoring at least one of the library holdings. In terms of edges, the graph contains only those where both members of the dyad acted as authors. Authorship involvement is used as a proxy to identify scholars who were significant to the intellectual landscape of the Old University. Overall, the network we built represents the publishing ecosystem of the Caa, encompassing connections that are not necessarily direct and contemporary collaboration acts but more along the lines of intangible scholarly contacts across time—as underscored by the presence of long-dead authors.

However, we acknowledge that the graph covers an extensive timeframe of almost three centuries, effectively flattening its temporal dimension. The temporal aspect, in our case study, lies in the publication year of the rare books of the collection—a record widely available for most of the metadata. To analyze the structure of a network that unfolds across time, we need to move beyond a static representation and consider the dynamic aspect of evolution. As network scientists emphasize, ignoring the temporal dimension risks misinterpreting network metrics and drawing inaccurate conclusions (Blonder et al., 2021, p. 958-959). This is particularly relevant for directed graphs, as well as for studies where the flow of information matters, and preserving the chronological order of path traversals is key.[13] But even for

[13]On the topic, see Scholtes (2017); van Vugt (2017); Roller (2023b). Ingo Scholtes is

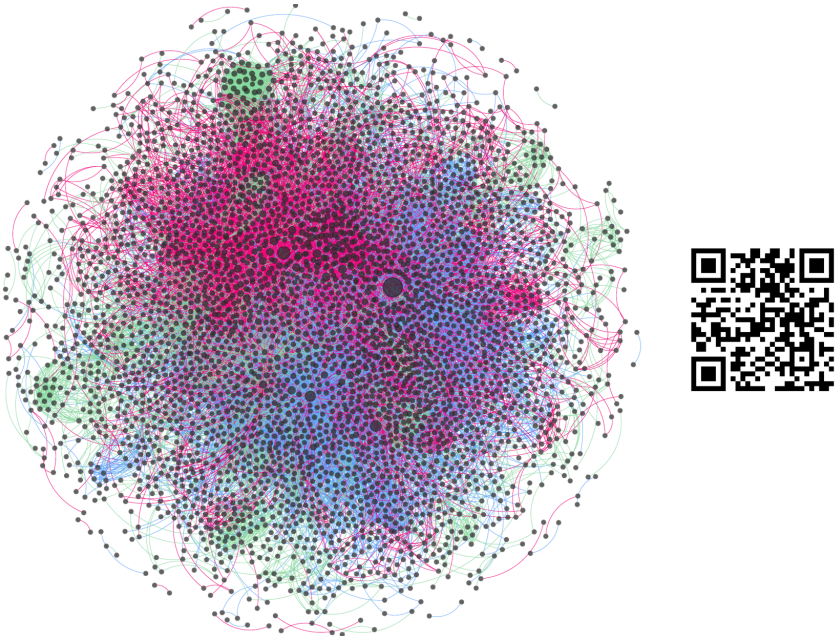


**Figure 1.4.** Snapshot of the network derived from the Caa metadata (1501–1797). The size of the nodes depends on the number of connections. Edges' colors blue, magenta, and green represent shared book collaboration taking place during the sixteenth, seventeenth, and eighteenth centuries, respectively. The position of the nodes in the graph is determined by the application of the Fruchterman-Reingold spring-based layout.

undirected networks like ours, the temporal order of interactions matters. Neglecting time implies treating all book collaborations as if they happened simultaneously, regardless of when they occurred, and as if they always existed. Since our goal is to investigate the temporal buildup of nodes' influence, as well as the evolving patterns of collaborations, we process the graph as a dynamic entity.

Two main approaches exist: time-ordered networks and time-aggregated networks. The former allows for the analysis of time-ordered paths in networks representing ties between nodes in a sequential manner. The latter consists of breaking down the main network into several subgraphs, each of which covers a time window defined by the user. Meaning that only the edges

---

also the developer of PathpyG, an open-source package designed to enable advanced network analysis for temporal graph data. More at: https://www.pathpy.net/.

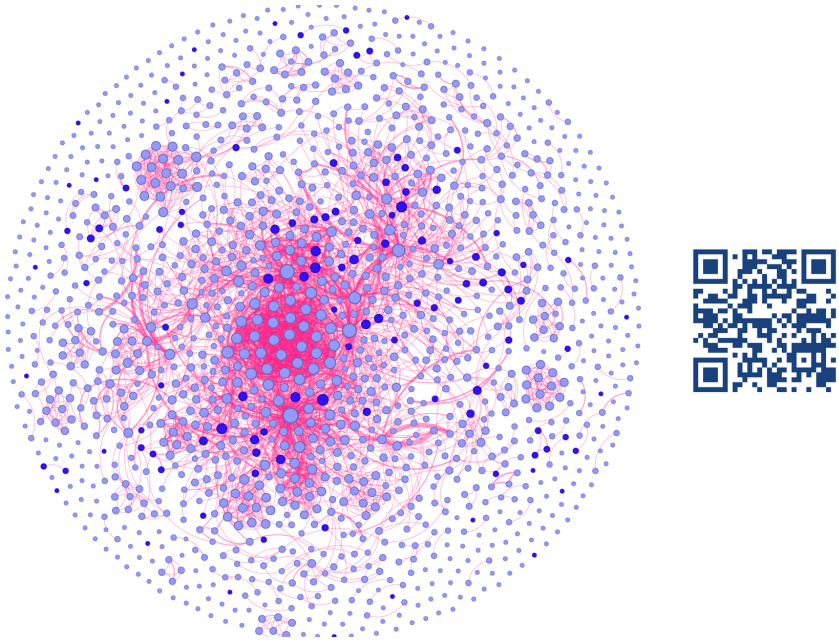


**Figure 1.5.** Snapshot of the Authorship layer of the network derived from the Caa metadata (1501–1797). The size of the nodes depends on the number of connections. Nodes in dark blue highlight classical or medieval authors from the rest, amounting to about 8%.

occurring during the selected window will be present. Each of the subgraphs is thus reduced to a static network associated with a specific time frame, which can, as such, be analyzed using standard SNA routines. In our case, however, we face an additional challenge: the inherent sparsity of our data. Each edge in our network is associated with a single year, not an interval, and using a year-by-year slicing approach would yield fragmented and thinly populated subgraphs. This would fail to capture the long-term influence of intellectual interactions. To address this, we use a rolling window of 25 years with an increment of one year. Each subgraph centers on a specific year, incorporating data from the 12 preceding and succeeding years, creating a temporally averaged context. This approach shares some similarities with the time-aggregated network methodology in the sense that we collapse our network to multiple subgraphs as the time window rolls along the full period. However, our design uniquely addresses two challenges: data sparsity and the need for maintaining continuity in context. By centering each subgraph on a

specific year and incorporating data from the 12 years preceding and succeeding that focal year, we create a temporally averaged context for each year. This technique mitigates the data sparsity and ensures a denser and more informative representation of the network as an intellectual landscape. The rolling window allows us to compute local metrics for each node in each subgraph.[14] Once generated and mapped to their central years, these subgraphs enable us to visualize the temporal evolution of node metrics. Ultimately, we can calculate their averages, producing a panel dataset with a cross-sectional dimension (nodes) and a time-series dimension (periods for which metrics are available).[15]

## 1.5 Censorship in the Caa network

In analyzing the multifaceted dynamics of the early modern intellectual community, we explore how different types of involvement compare in terms of the centrality of nodes. We leverage the multilayer structure of our network and aggregate the main centrality metrics (degree, betweenness, closeness, and eigenvector centrality) by layer (Figure 6). We find that nodes involved in the censorship layer tend to consistently score the highest centralities over time. These results confirm the leverage censors possessed in the regulation of the early modern book industry. Historical scholarship specialized in censorship in the Low Countries highlights in several instances that censors often found themselves in the position of negotiators, their tasks far more complex and strategic than merely banning books (Cammaerts, 2023, p. 12). They skillfully navigated a delicate equilibrium between the demands of religious institutions and the creative expression of authors (Vanden Bosch, 2017a, p. 4). In this perspective, our data-driven elaboration confirms the relevance of the role of censors in the cultural landscape of the publishing world of the Old University of Louvain. In a network representing the intellectual community based on printed publications, it is to be expected that censors would emerge

[14]These metrics and their evolution could in principle be tested for robustness against alternative window widths. We do not carry out this check here, and flag it as a direction for future work.

[15]Not all nodes in the cross-sectional dimension have metrics available for every subgraph year in the time-series dimension. This is a direct consequence of the fact that the presence of nodes in the network is variable, and depends on their earliest and latest attested publication date.

as critical hubs that 'connect' different segments of the network.[16]




**Figure 1.6.** Centrality metrics aggregated by layer. In clockwise rotation, we show degree (centrality metric based on the number of connections), betweenness centrality (capturing how much a node acts as a bridge connecting different parts of the network), eigenvector centrality (assessing a node's importance based on the importance of its connections, and closeness centrality (quantifying how quickly a node can reach all other vertices).

## 1.6 Defying the Erasmus bias? Comparing network centrality with RETE's human capital index

We argue that the centrality metrics in the authorship layer of our network might serve as proxies for understanding scholars' social presence or

[16]The censorship layer is likely incomplete, since approbations are unevenly recorded in the Caa metadata. In principle one could supplement the material from other sources, in the fashion of meticulous book-historical research, but here we limit ourselves to the information directly available in the catalog metadata.

sociability during the era. Although these metrics are not intended to be a measure of authority, they mirror the concept of individual quality or human capital as developed by the *Upper Tail Human Capital and the Rise of the West* project, hosted at UCLouvain, entailing the creation of a new database of European scholars, the *Repertorium Eruditorum Totius Europae* (RETE) database (Catoire et al., 2021, p. 55). The RETE project gathers data from secondary sources to quantify the human capital of scholars and literati at premodern higher education institutions and attempt to quantify their human capital—that is, the economic value of an individual's skills.[17] For each scholar, a human capital index is computed by integrating data on scholars' library footprint (from VIAF) and general recognition (from Wikipedia) using Principal Component Analysis (PCA).[18] While innovative, this index may not fully capture the multifaceted nature of historical human capital due to the use of contemporary standards. For instance, relying on Wikipedia for recognition may introduce an anachronistic element. Wikipedia's coverage and the length of articles can be influenced by modern perspectives and might not accurately reflect the historical significance or the impact of a scholar's work in their own time.

The centrality metrics emerging from our network provide a more focused window into the scholarly network of the Old University, as reflected in the curatorial approach of the Caa collection. Thus, these metrics might complement the broader scope of the RETE index. After defining the set of Louvain scholars present in both the Caa network and the RETE dataset, we test whether our network-based centrality metrics correlate with the human capital index from RETE. Because we are interested in comparing the ranks of the data points and not their absolute value, we use Spearman's correlation.[19]

Table 2 shows the computed correlation coefficients measuring the strength and direction of the monotonic relationship of the centrality measures in relation to the human capital index. The coefficients may vary between –1 and 1, indicating respectively a perfectly negative and a perfectly positive correlation. Overall, we find a positive correlation with RETE's

[17] More on the RETE project at: https://ojs.uclouvain.be/index.php/RETE/index

[18] PCA is a widely used technique for constructing indices, which reduces the dimensionality of data while retaining the most important components.

[19] The two methods used to compute rank correlation differ in computation, as Spearman's directly compares the rank of each observation. For a thorough overview, see: Wilcox (2017), pp. 573-75.

**Table 1.2.** The rank correlation matrix between the human capital index and network-based centrality measures

| Centrality metric | Human capital index (RETE) Rank correlation |
|---|---|
| Degree centrality | 0.30 |
| Betweenness centrality | 0.21 |
| Closeness centrality | 0.31 |
| PageRank | 0.36 |
| Eigenvector centrality | 0.32 |

human capital index, ranging from 0.21 to 0.36. Upon closer examination of the correlation matrix, we identify the most significant discrepancy with RETE's measure in the case of humanist Desiderius Erasmus of Rotterdam (1467/1469–1536). Specifically, Erasmus has a higher score in the human capital index but a relatively lower value across all network centralities computed. This outcome suggests that the collection is not as biased as one might expect, considering that the original collection was destroyed and the library grew through donations and acquisitions, potentially favoring notable figures like Erasmus. Given his prominence in European intellectual history, one could have anticipated Erasmus to be overrepresented in the Caa, potentially skewing the data. Having more books by Erasmus in the collection would naturally result in more connections in the network, leading, on average, to higher centrality. However, this does not appear to be the case. Erasmus exhibits relatively lower centrality across all network metrics when compared to his RETE's human capital index, suggesting that the Caa collection does not disproportionately favor his academic output. This finding indicates that, despite the challenges posed by the destruction and reconstruction of the library, the Caa collection may be more balanced in its representation of the intellectual network of Louvain than we initially assumed.[20]

[20]We acknowledge that this is a relatively strong claim resting on a single test. Extending the comparison to a panel of well-known scholars across periods, fields, and confessional orientations, with subgroup rank correlations against RETE, would turn this single check into systematic evidence. We regard the present test as a preliminary application in that direction.

## 1.7 Scholars' sociability in early modern Louvain's intellectual landscape

Because of our rolling window approach, we can compute a longitudinal series of centrality metrics for each node active in the network at some point. This allows us to track the evolution of the 'relevance' in the network of some selected nodes within the scholars' subgraph. Among the various network-based measures of centrality, we propose that normalized degree—which essentially captures how many connections a vertex has with other nodes in the network relative to the maximum possible number of connections it could have in the graph—may be a direct indicator of the social aspect of books in an academic scope. This metric measures the intensity and scope of collaborations associated with the Caa collection from which we draw the network. A scholar who frequently authors numerous works with a few other scholars will have a low degree centrality. Conversely, a scholar who publishes a few books but collaborates with many co-authors will score a high degree centrality. While not constituting a measure of authority, the degree centrality within the network may proxy the social presence or scholar sociability during that era. We show below a set of plots tracking the evolution of degree centrality for some scholars of influence in the history of the University of Louvain. It must be noted that while each scholar might exist in the temporal network from the start of their activity,[21] their permanence in the network mirrors the presence of said scholar in the Caa and does not necessarily coincide with the end of the activity or death.

Figure 7 plots the evolution of degree centrality for some of the giants of the history of the Old University of Louvain, primarily Desiderius Erasmus and Justus Lipsius (1547–1606). Erasmus was in Louvain for several years during the first half of the sixteenth century. He did matriculate at the university but was never a formal professor; his permanence in Louvain was a period of intense achievements. Not only did he give a strong impetus for the creation of the Collegium Trilingue in 1517, but he also laid the seeds for future collaborations with printers from the Low Countries. While these connections mattered for his career, they ended up being crucial to other scholars as well, either indirectly, via shared collaboration with Erasmus, or even directly. For instance, the volume edited by François, Soen, Dupont, and Robiglio (2020) underlines

[21] If the start of activity is not explicitly mentioned in the metadata, we infer the year of beginning of activity from the date of earliest attested publication in the Caa.

how the editio princeps of Thomas More's (1478–1535) groundbreaking work, Utopia, in 1516 would not have been possible without Erasmus connecting More to the Louvain printer Dirk Martens (1446–1534), with whom Erasmus had published several works (François, 2020, p. 12). Erasmus starts with a high degree—the highest among his fellow nodes in the scholars' network—but this pattern decreases over time. From the second half of the sixteenth century, Erasmus's supremacy seems to be surpassed by Lipsius's.

The humanist Justus Lipsius taught in Louvain at the start and toward the end of his academic career (Roegiers, 1997, p. 26 and p. 31). Lipsius's tenure at the University of Louvain contributed to the revival of the institution at the end of the sixteenth century, a tumultuous period for the history of Louvain due to revolts, military occupations, and epidemics of illnesses. In this context, Lipsius was able to pivot the Old University toward humanistic studies, classical scholarship, and Stoic philosophy, leaving a legacy on academic and philosophical thought for generations to come. The evolution of his social presence in the Caa network corroborates the surge in popularity he experienced from the last decades of the 1500s, and his persistence reaches the end of our timeline, meaning he kept being published well after his death.

The Spanish humanist Juan Luis Vives (1492–1540) worked as a tutor in several European courts and noble houses, was close to Erasmus during his permanence in Louvain and via correspondences, and supported the creation of the Collegium Trilingue (Tournoy et al., 1993, , p.12 and p.62). He is the third most popular scholar among those depicted in Figure 6, touching a peak at over 0.2. From the end of the sixteenth century onwards, his degree decreases, but Vives continues being 'present' in the publishing world of Louvain until the year of the abolishment of the Old University.

Figure 7 also includes the evolution of the degree centrality of Nicolaus Clenardus (1495–1542)—a linguist who focused on the study and promotion of languages including those taught at the Collegium Trilingue along with Arabic—and of Cornelius Valerius (1512–1578)—who, like Clenardus, was a linguist and humanist, but unlike Clenardus, taught at the Collegium Trilingue. Clenardus scores higher values than Valerius, but Valerius has a longer-lasting presence in the network.

Figure 8 provides an insight into the demise of scholasticism and the advent of the 'new science' by plotting the degree centrality of Aristotle alongside those of outstanding Louvain scholars whose area of expertise would fall in the hard sciences by modern standards. Indeed, Reinerus Gemma

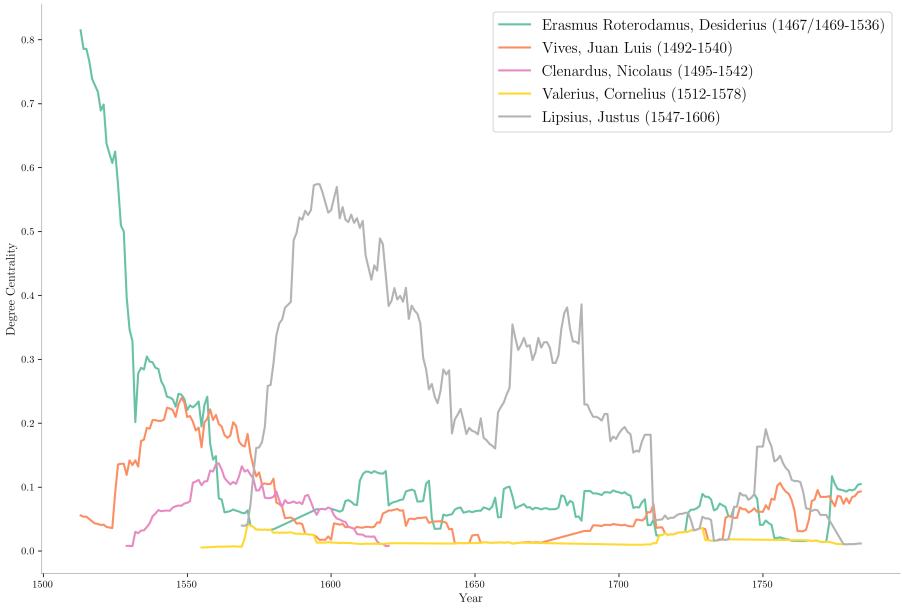



**Figure 1.7.** Evolution of degree centrality for some prominent Louvain scholars and giants of the humanist period. Computed using the rolling window approach on the Authorship layer.

Frisius (1508–1555) greatly contributed to the development of the sciences of cartography and geography through the method of triangulation, as well as the creation of maps and other scientific instruments. His successor to the mathematics and medicine chair, Adrianus Romanus or Adriaan van Roomen (1561–1615), on the other hand, significantly advanced the field of algebra. The inclusion of the degree evolution of Frisius and Romanus serves mainly as a point of comparison to comment on the trajectory of Aristotle's popularity. The most compelling finding is given by the trajectory of the centrality of the ancient philosopher Aristotle. While Aristotle himself was not associated with the University of Louvain, his works profoundly influenced the curriculum and intellectual discourse there—which explains why Aristotle is consistently present in the authorship layer of the network (Geudens and Masolini, 2016, p. 1 and p. 5). Aristotelian philosophy was a cornerstone of the education system, deeply affecting the teaching of theology, philosophy, and the natural sciences. Compared to other long-dead authors, like Euclid (active around 300 BCE) and Thomas Aquinas (c. 1225–1274), Aristotle significantly stands out in the plot. The Old University was a stronghold of Scholasticism, but

this state of affairs was destined to change. Starting during the Renaissance, the decline of Aristotelian philosophy persisted and intensified during the Scientific Revolution and the Reformation, and by 1650, Aristotle stopped being published for the education of the Artes, marking a return to the practice of dictating classes. Figure 8 confirms the loss in popularity of Aristotle, who was surpassed by Gemma Frisius from 1550 circa and who ceased to be mentioned in Caa holdings by 1660. The plot shows that social presence in the network can go beyond a scholar's lifetime and offers an additional point of comparison to appreciate Aristotle's importance and decline, as discussed above.

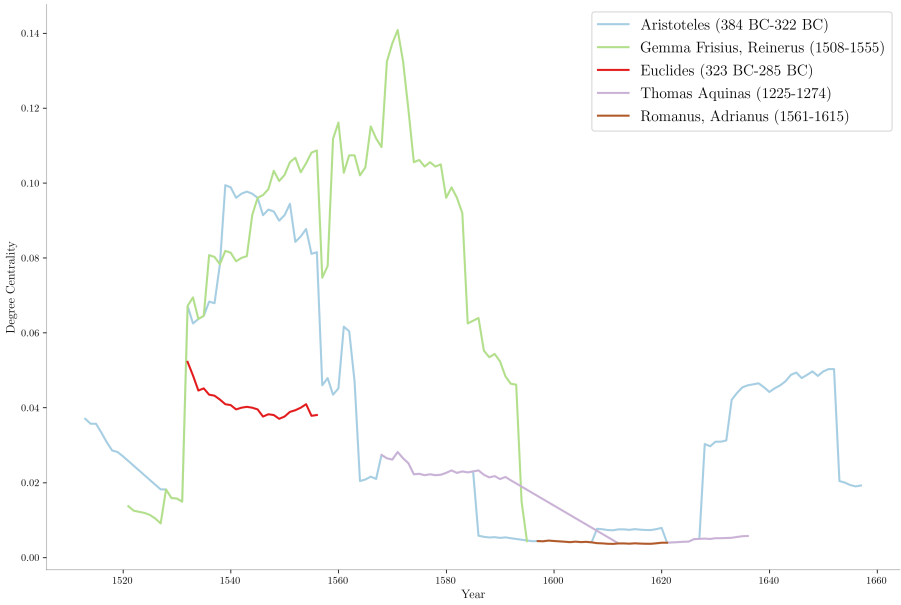



**Figure 1.8.** Evolution of degree centrality. Classical authors and Louvain 'scientists'. Computed using the rolling window approach on the Authorship layer.

Figure 9 depicts the degree evolution of some scholars active in the contemporary rather turbulent theological landscape. Among these figures, Cornelius II Jansenius of Ypres (1585–1638), the founder of Jansenism; Robertus Bellarminus (1542–1621), Jesuit theologian from Italy which taught at the Jesuit College of Louvain between 1570 and 1576, who vigorously defended teachings of the Church against Protestant reformers; Leonardus Lessius (1554–1623), professor at Louvain for several years, contributing

enormously to theology and moral philosophy; Michael Baius (1513–1589), whose revolutionary theological position preconfigured Jansenism and cost him orthodoxy accusations and a papal condemnation; Johannes Malderus (1563–1633), professor of scholastic theology in Louvain and later bishop of Antwerp, concerned with reforming the clergy within his diocese while also opposing to the Calvinists at the latter stages of the Dutch Revolt (1568–1648). Figure 9 suggests that the social presence of Jansenius took off around the 1630s, with a steep spike that reached its peak around 1650. Afterward, Jansenius's degree plummets around the 1640s, a trend that closely mirrors the timing of the ensuing Jansenist controversy. More specifically, the drop seems to occur after the publication of Jansenius's Augustinus and its subsequent condemnation of heresy by Pope Urban VIII (r. 1623–1644). Another interesting pattern is that regarding the evolution of Malderus's degree, growing slowly until the 1640s and experienced a steep increase from the second half of the seventeenth century and throughout the eighteenth century. Lastly, we attest that Baius's degree centrality in the scholars' network remains relatively low and constant, overlapping more or less with the theologian's lifetime.

## 1.8 Conclusions

In the present paper, we addressed inconsistencies in the Caa collection's authorship, paratextual and imprint metadata, including variant spellings and name conventions, and ensured an accurate indexing of individuals and publication locations through a rigorous data-cleaning process. Subsequently, we modeled this metadata as a multilayer co-occurrence network to explore the intellectual community of the Old University of Louvain. Our findings show that censors played a crucial role in the network, consistently exhibiting high centrality scores, reflecting their significant influence in the early modern book industry. This is evidenced by the fact that, within a network based on printed publications, censors naturally emerge as key connectors. Additionally, our results indicate that, despite being a curated sample rather than a naturally accumulated one, the collection is not as biased as one might expect. This is demonstrated by the moderate prominence of Erasmus relative to his large presence in the library and his widespread notability. The evolving degree centrality of selected scholars over time further highlights the intellectual shifts within early modern Louvain, illustrating the decline of

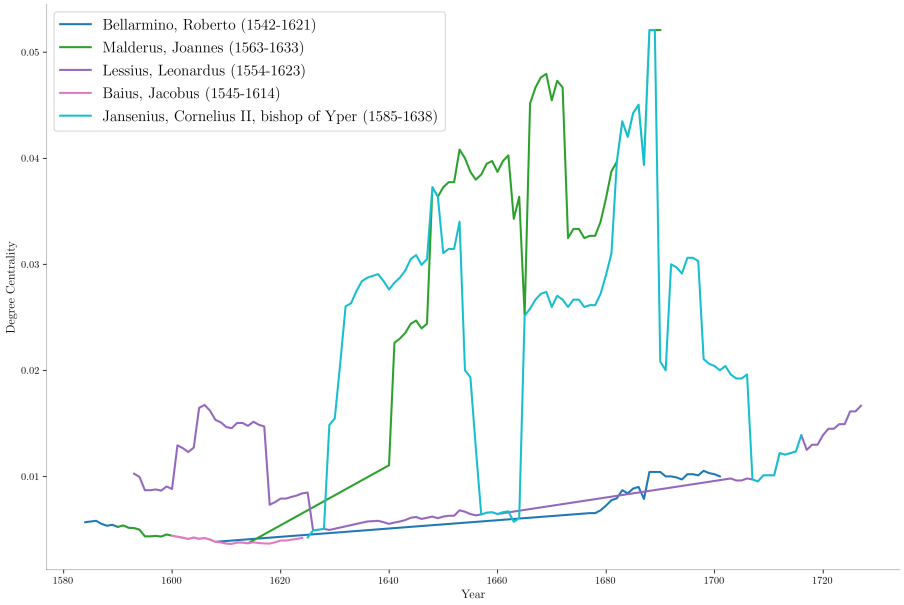



**Figure 1.9.** Evolution of degree centrality. A focus on $16^{th}$ century Louvain theologians. Computed using the rolling window approach on the Authorship layer.

Aristotelian philosophy and the rise of humanism and emerging scientific disciplines. We also shed light on the enduring legacy of Louvain humanists, whose influence persisted long after their lifetimes. Lastly, the interactive network visualizations generated from the Caa metadata analysis provide the ability at one's fingertips to browse through the publication network of the Old University of Louvain, a new lens to explore the intricate web of publishing relationships at the Old University of Louvain.

# 2

# Bibliographic Metadata as Relational Data: A Cross-Disciplinary Methodological Reflection

*In this paper, I reflect on the growing cross-disciplinary convergence of the use of bibliographic metadata as empirical material in digital humanities and computational history as well as in quantitative economic history. I discuss the main challenges involved in preprocessing metadata from library catalogs, comparing techniques from the abovementioned fields. A case study based on the Collectio academica antiqua of the Old University of Louvain serves to demonstrate how its metadata can be parsed, disambiguated, and reorganized into relational format, that is, into linked tabular datasets describing different aspects of the collection's records. Building on this foundation, I examine the use of bibliographic data as empirical material in historical network analysis, with particular attention to the assumptions underlying co-occurrence representations. I show that role-differentiated participation encoded in the original metadata is not naturally accommodated by flat network projections, and I argue that a multilayer network representation provides a coherent way to preserve this heterogeneity. I use this to assess how differences in*

*participation across roles relate to the institutional context in which publications were produced. I also highlight both the analytical limits of bibliographic metadata when used in isolation and the gains that arise when relational representations derived from catalog data are integrated with complementary sources.*

## 2.1 Introduction

Traditionally compiled by librarians for purposes of bibliographic control and collection management, metadata from union catalogs and curated library collections are increasingly being repurposed as empirical material in historical research both in the fields of digital humanities and computational history as well as in the field of quantitative social sciences. Digital humanities scholars have been at the forefront of this reappropriation. Building on the framework introduced by Lahti et al., "Bibliographic Data Science and the History of the Book," who coined the term bibliographic data science to describe the systematic use of catalog metadata for historical inquiry, a growing body of work has shown how information encoded in title pages, imprint statements, dedications, approbations, and physical descriptions can support large-scale analyses of intellectual and book history. Key contributions include studies drawing on large-scale book catalogs to examine cultural production, canon formation, book pricing and the circulation of classical and vernacular texts (Lahti et al., 2015; Tolonen et al., 2019, 2021; Hill et al., 2019; Tiihonen et al., 2024; Fantoli et al., 2025), as well as network analyses of print culture that trace influence through dedications, printers' relationships, and publishing communities (Gavin, 2016; Gittel, 2021; Valleriani et al., 2022; Ryan and Tolonen, 2024b,a; Heßbrüggen-Walter, 2024). A parallel strand of scholarship draws on metadata of epistolary sources, as in Hotson and Wallnig, eds (2019) on the Republic of Letters or Roller (2023b) on the circulation of Reformation ideas through correspondence.

In parallel, researchers in the social sciences—particularly, though not exclusively, applied economic historians—have also begun to employ bibliographic data to addressing topics such as the decline of Islamic science and preindustrial city growth (Chaney, 2023, 2024), teacher-directed scientific progress in early modern England (Koschnick, 2025b), the diffusion of ideas through railroad networks in nineteenth-century German-speaking regions (Chiopris, 2024), the relationship between limited academic career prospect and the rise of dissenting religious print (de Pleijt and Koschnick, 2025), and the role of the Republic of Letters in Britain's Industrial Revolution take-off (Cervellati et al., 2025). This convergence around the same materials reveals both the expanding research potential of bibliographic data and the persistence of distinct disciplinary conventions in how these data are employed. The digital humanities have often developed without full

integration with the quantitative social sciences (Lemercier, 2015b, p. 273), which helps explain why the two traditions, despite similar empirical aims, have adopted different, though occasionally overlapping, methodological conventions.

The way the very same bibliographic data are mobilized in the two literatures reveals this divergence most clearly. While not all digital humanities or computational history projects using these materials adopt network methods, the network perspective remains one of the most common modelling frameworks paired with bibliographic metadata in the field. Although catalog data do not inherently encode relationships, they are often interpreted as networked systems by treating the co-occurrence of agents within the same publication event as evidence of collaboration or shared involvement. This approach has been productive for mapping and exploration, though its analytical development remains uneven. In this perspective, skepticism toward digital approaches among historians and humanists trained in textual or archival traditions is not without foundation (Gregory, 2014, p. 2). Yet this limitation is increasingly being addressed more broadly, with Roller (2023a) insisting that quantitative methods be tied to explicit research questions and historiographical theories that establish testable relationships between measurable concepts of interest.

By contrast, in quantitative economic history, recent trends shaped in part by methodological norms associated with the so-called "credibility revolution" (Angrist and Pischke, 2010) have reinforced expectations about explicit research design and the use of linked datasets (Cantoni and Yuchtman, 2021, p. 216). Within this framework, bibliographic data are routinely employed in research designs that integrate multiple data sources to produce theoretically grounded interpretations of historical phenomena. Economic history thus offers a useful point of comparison for digital humanities and computational history because it faces similar preprocessing and interpretative challenges and shows how formal analytical modeling can expand the scope of the same bibliographic metadata corpora.

The paper builds on this cross-disciplinary comparison by examining how metadata from library catalogs can be de-structured, parsed, and reorganized in relational form and how the resulting data can support analytical modeling within a network framework. In doing so, it responds to calls for closer alignment between quantitative analysis and explicit theoretical frameworks (Roller, 2023a). Related work in the digital humanities has pursued more

explicitly analytical modeling strategies, including multimodal combinations of network methods and text analysis (Hill et al., 2023) and temporal modeling approaches Roller (2023b). More broadly, the present study aligns with the commitment articulated by Lahti et al. (2019b) to treat bibliographic metadata as a substantive source for historical inquiry.

Against this background, this paper is guided by three related research questions. First, how can bibliographic catalog records, originally designed for item-level description, be transformed into relational data structures suitable for longitudinal and network-based analysis? Second, what modeling assumptions are implicitly introduced when bibliographic metadata are interpreted as co-occurrence networks, and how do these assumptions shape the interpretation of relationships inferred from catalog data? Third, what analytical leverage is gained by preserving role-differentiated participation through a multilayer network representation compared with flat co-occurrence projections commonly used in digital history?

To address these questions, I focus on the metadata of the *Collectio academica antiqua*, the cultural heritage collection of the Old University of Louvain (1425-1797), and I examine how relational modeling choices affect the identification of systematic patterns in academic print production. The central claim of the paper is that bibliographic data encode role-differentiated forms of participation that are not naturally accommodated by standard, flat co-occurrence network representations. When such data are modeled as such, distinct forms of involvement in book production are implicitly collapsed. I argue that, if one adopts a network perspective, a multilayer representation offers the most coherent way, within a co-occurrence framework, to preserve this role differentiation instead of flattening it into an undifferentiated relation. This modeling choice is justified because it preserves distinctions encoded in the original sources and because these distinctions might correspond to historically meaningful differences in modes of participation to book production. I use my case study of the old academic collection of Louvain to address whether holding multiple production roles is associated with a higher likelihood that a publication involves university professors.

This analysis also makes clear that recovering features of Louvain's academic publishing ecosystem requires integrating bibliographic metadata with external sources beyond the catalog itself. However rich, bibliographic data *alone* offer limited leverage for addressing complex historical questions. Their analytical potential is substantially expanded when they are linked to

complementary sources, such as prosopographical—the study of background characteristics—or institutional datasets. The comparison with quantitative economic history reinforces this point: In that field, bibliographic data typically function as one component within a broader data infrastructure combining multiple sources. The paper proceeds in five sections. Section 2 outlines the structure of metadata from union catalogs and library collections and discusses how key bibliographic elements can be identified, extracted, and transformed into a longitudinal and relational format. Section 3 details the conceptual and practical steps involved in preprocessing the resulting data for systematic quantitative analysis. Section 4 examines the assumptions involved in interpreting catalog metadata as network data and shows how a multilayer co-occurrence representation allows role-differentiated participation to be modeled and combined with external sources for analytical use. Section 5 concludes.

## 2.2 What is in a Catalog? Anatomy of Bibliographic Metadata

Digital library catalogs are curated databases describing and indexing the holdings of one or more libraries. The structured information stored in these catalogs is known as bibliographic metadata. Metadata refer to attributes about catalog entries, rather than their full content, and support organization, classification, and discovery. For printed book collections, such metadata include elements such as title, publication date and place, names of agents involved in the creation and distribution, the language, and the physical description of each holding. Internally, library catalogs rely on two complementary layers: an encoding format and a content standard. The encoding determines how metadata is structured for storage and exchange, whereas the content standard governs what information is recorded and how it is expressed. Most libraries use MARC (MAchine-Readable Cataloging) as their encoding format. Developed at the Library of Congress by Henriette Avram and later formalized as MARC 21, this schema makes cataloging data legible to computers across libraries and associates text strings with numbered fields that are conventionally linked to specific bibliographic properties. Despite its limitations and its limited alignment with linked-data infrastructures (Tennant, 2002), MARC 21 continues to underpin most

catalog systems. Even when library records are exported to relational tables or Extensible Markup Language (XML) formats such as MARCXML or Metadata Object Description Schema (MODS), they retain the underlying MARC 21 structure. Researchers must therefore parse fields, indicators, and subfields according to MARC conventions. Tools such as pymarc assist with extraction and conversion but still require familiarity with MARC 21 logic.[1]

While encoding formats define *where* information is stored, content standards determine *what* information is entered and how consistently. These standards shape decisions such as how titles are transcribed, how imprints are formatted, and how names, roles, and subject information are recorded. General collections typically follow Resource Description and Access (RDA), rare book collections often rely on more specialized descriptive standards such as Descriptive Cataloging of Rare Materials (DCRM), and archival holdings use standards such as Describing Archives: A Content Standard (DACS). As a result, records encoded in the same MARC 21 infrastructure may display distinct descriptive logics depending on the content standard applied. In other words, it is only in light of the cataloging conventions that one can interpret which MARC fields are actually available for analysis.

I work with the Collectio academica antiqua, a curated collection of early modern printed works associated with the Old University of Louvain, the first university founded in the historic Low Countries, in the Brabant region (present-day Belgium) in 1425, from which the contemporary sister universities of Katholieke Universiteit Leuven (KU Leuven) and Université Catholique de Louvain (UCLouvain) descend, following its abolition in 1797.[2] The collection gathers works linked to the university, its affiliated scholars, its illustrious alumni, and its institutional history. Although printing activity in Louvain is attested from 1473 onward, incunabula at the KU Leuven

[1] pymarc is a Python library for reading, writing, and parsing MARC21 records. For the documentation, see https://pypi.org/project/pymarc/.

[2] The institutional history is more complex: the original university, known as the Old University of Louvain, was suppressed under French revolutionary rule in 1797, then briefly refounded as the secular State University of Louvain in 1817 before being abolished again in 1835, and finally reestablished as the Catholic University in 1834. In 1968, amid the linguistic tensions between Flemish and Francophone communities that have long shaped Belgian public life, it was split into the Dutch-language Katholieke Universiteit Leuven, remaining in Louvain (in the province of Flemish Brabant), and the French-language Université Catholique de Louvain, relocated to the newly built campus of Louvain-la-Neuve (in the province of Walloon Brabant).

Libraries are held in a separate preservation collection, so the Collectio academica antiqua begins artificially in 1501 and so does not reflect the chronology of local printing (Scebba and Fantoli, 2026). The catalog records of the *Collectio academica antiqua* are encoded in MARC 21 and follow the RDA-based cataloging standard used at the KU Leuven Libraries, supplemented by internal guidelines. For early printed materials, catalogers also draw on the Short Title Catalog Vlaanderen (STCV) rule set, which provides descriptive guidance for rare books (i.e., books printed before 1801).

Table 1 summarizes the MARC 21 fields most commonly encountered in early modern and rare book contexts. It generalizes from the data structure of the *Collectio academica antiqua* and from comparable early modern catalogs. Whether these fields become analytically useful depends on the research question as well as on the content standards and cataloging practices that shaped the metadata. This dependence is evident when early modern catalogs are compared with contemporary circulating collections. For example, Petras et al. (2006) use metadata from modern library collections to construct time-period directories that allow users to navigate holdings by historical era and place. Their approach relies on subject strings containing explicit chronological and geographic subdivisions, which are common in contemporary catalogs but generally absent from early modern collections such as the *Collectio academica antiqua*.

This closer examination of the technical structure of catalog data underscores that working with bibliographic metadata requires an explicit understanding of their internal organization and their transformation into relational form. It also makes immediately clear that the extracted data often require harmonization. Because cataloging practices prioritize item-level description over cross-record consistency, records frequently lack internal coherence, particularly in the identification of persons, corporate entities, and place names (Padilla, 2025, p. 20). Unless a catalog relies on robust authority control or controlled vocabularies, these elements tend to appear in multiple, nonaligned forms. As a result, they must be interpreted, cleaned, and restructured before they can serve as input for empirical modeling. The next section outlines the conceptual and practical decisions involved in converting catalog records into harmonized relational data.

**Table 2.1.** The table lists selected MARC 21 fields with their Library of Congress definitions and brief content description.

| Tag | LoC Definition | Content |
|---|---|---|
| **Agents** | | |
| 100 | Main Entry–Personal Name | Personal name primarily responsible for the work (e.g., author) |
| 110 | Main Entry–Corporate Name | Institutional body primarily responsible for the work (e.g., university press) |
| 600 | Subject Added Entry–Personal Name | Person discussed or referenced in the work (e.g., biography subject) |
| 610 | Subject Added Entry–Corporate Name | Corporate entity discussed or referenced (e.g., religious order, university) |
| 700 | Added Entry–Personal Name | Additional individual associated with the work (e.g., editor, translator) |
| 710 | Added Entry–Corporate Name | Additional institutional body involved (e.g., publishers, sponsors) |
| **Content** | | |
| 245 | Title Statement | Full title of the work, including subtitles |
| 650 | Subject Added Entry–Topical Term | Keywords or subject headings describing the topic |
| 655 | Index Term–Genre/Form | Material type or genre designation (e.g., ephemera, pamphlets, and other genre or form designations) |
| **Paratext and imprint** | | |
| 500 | General Note | Miscellaneous notes, often paratextual (e.g., dedications, colophons) |
| 260/264 | Publication, Distribution, etc. (Imprint) | Publisher information, place, and date of publication |
| 041 | Language Code | Language(s) in which the work is written |
| **Access information** | | |
| 852 | Location | Physical location and call number of the holding institution |
| 856 | Electronic Location and Access | URL or link to digital version or online resource |

## 2.3 From Structured to Relational Data

Transforming catalog metadata into a form suitable for quantitative modeling involves a series of preprocessing steps. This stage raises questions that are increasingly shared across disciplines. Scholars in both digital humanities and social sciences working with historical data are confronting the same challenges: how to identify and reconcile references to historical persons, places, and subjects across historical data, including bibliographic metadata. Digital humanists have long grappled with these issues in prosopographies, bibliographies, and textual corpora (Ehrmann et al., 2023), whereas social scientists are more recently engaging with similar problems (Arora et al., 2024). Both traditions are converging toward workflows that combine domain knowledge with scalable, semi-automated tools in order to prepare historical data for analysis. In this context, extracting and parsing structured bibliographic metadata is a crucial step because converting MARC's heterogeneous field structure into an analytically usable form requires normalizing it into relational data, with entities separated into tables and linked through keys. In what follows, I show how these workflows unfold in practice, using the *Collectio academica antiqua* as a case study across three key stages: actor identification, spatial referencing of place names, and content classification.

### 2.3.1 Identifying Historical Actors: Disambiguation and Authority Alignment

Catalogs often record the names of both personal and corporate entities associated with a given holding, in line with the MARC 21 schema.[3] These entries may include additional identifying information such as life dates or floruit, numeration, pseudonyms, religious affiliation or noble titles, and places of activity, as well as the role under which the individual appears in the publication. One of the most challenging steps in working with bibliographic metadata is the identification and disambiguation of historical actors based on these name strings. This difficulty arises from variant spellings, inconsistent orthography, and uneven role attribution across entries. These

[3]Corporate entities usually designate institutions such as religious orders, universities, academies, government bodies, or, in some cases, printing houses and workshops, which may be cataloged as such rather than as individual printers.

challenges may be more or less pronounced depending on the uniformity of cataloging conventions and the degree of authority control implemented, and this may apply to curated collections and union catalogs alike. To construct meaningful relational data, it is essential to reconcile these inconsistencies and unambiguously identify historical individuals. The disambiguation step ensures that relationships in the data are constructed around unified entities instead of fragmented or duplicated name strings. This process relies on accompanying biographical attributes to distinguish homonyms and resolve pseudonyms. In the case of homonyms, these additional attributes can help differentiate distinct individuals with similar names, while in the case of pseudonyms or variant forms referring to the same person, shared dates or roles may allow for accurate consolidation even when string similarity alone is insufficient. Ideally, at the end of this reconciliation process, one should be able to assign a persistent identifier to each person or corporate entity.

In practice, two main strategies are typically employed: a bottom-up approach, based on string similarity and clustering, or a top-down strategy that aligns name strings with external authority files. Both techniques can be employed at different stages of the workflow, and often both might benefit from a semi-automated approach. For instance, fuzzy string matching and clustering algorithms can be used to group likely name variants, which are then manually reviewed for validation. Similarly, a name search can be launched against a selected external authority file to identify potential matches, but the final assignment of identifiers should remain under human supervision.

In my pilot study using the *Collectio academica antiqua*, I implemented a structured pipeline to extract, standardize, and cluster personal and corporate names from MARC 21 records. I began by identifying the relevant MARC 21 tags containing names of interest, then used pymarc to extract their content into a flat, tabular format while retaining the original tag associated with each entry. I grouped all name strings from subfield $a into a single column, regardless of the MARC field they originated from. I applied the same principle to other associated attributes–such as numeration, life dates, and roles–by creating standardized columns based on subfield content, across different MARC tags.

At this stage, I cleaned the strings to ensure consistency: I removed leading and trailing spaces, standardized the formatting of numeration and dates, and eliminated unnecessary punctuation. While reviewing the grouped values, I also checked for misplacements (e.g., numeration incorrectly stored in a name

subfield) and reassigned them to the appropriate column when necessary.

I then processed the resulting dataframe in OpenRefine, which allows for semi-automated clustering with high user oversight.[4] I applied different clustering techniques and manually evaluated the suggested matches. I recommend working on a duplicate column and proceeding in successive passes: first clustering on the combination of surname, given name, numeration, and parsed life dates (possibly formatted as YYYY–YYYY or as YYYY in case only one date is available); then reapplying clustering without the dates; and lastly filtering by common life dates to identify remaining variants that earlier steps may have missed. After this bottom-up procedure, I adopted a top-down approach by matching the clustered names against an external authority file. Given that my case study mainly involved early modern persons active in Europe, the Consortium of European Research Libraries (CERL) Thesaurus proved particularly useful.[5] I relied on the dedicated Python library, cerl, which is a wrapper for the CERL Thesaurus API.[6] However, other reconciliation workflows are equally viable. Many thesauri, including Virtual International Authority File (VIAF), Gemeinsame Normdatei (GND), and Wikidata, offer dedicated reconciliation services within OpenRefine or provide public APIs, which allow users to build custom pipelines for automated queries. Crucially, querying authority files often returns more than one potential match per name. These results must be carefully reviewed and manually approved, ideally by inspecting supporting metadata such as life dates, places of activity, or roles. This step, requiring human validation, is essential to ensure accurate reconciliation, especially when dealing with common names or minimal contextual information.

External authority files offer three main advantages. First, they often have already resolved duplications across records by grouping known name variants. Second, they enrich the dataset with additional information such as

[4] OpenRefine is an open-source data cleaning and transformation tool designed for working with messy or semi-structured data. For more, see OpenRefine, https://openrefine.org/.

[5] The CERL Thesaurus is maintained by the Consortium of European Research Libraries and aggregates historical name authorities for persons and corporate entities active in book production. For more on this, see https://data.cerl.org/thesaurus/.

[6] For more on the Python library cerl, see https://pypi.org/project/cerl/. The library was developed by Andreas Walker. A minor compatibility fix is required to use it with recent versions of urllib3.

birthplaces, deathplaces, or institutional affiliations, which can be repurposed in later stages of our own analysis. Third, if the metadata needs to be merged with other datasets, records reconciled against the same thesaurus can be matched more reliably and seamlessly by relying on the same identifier. Of course, authority files come with some limitations, namely, they are inevitably incomplete, and many lesser-known historical figures remain unlisted. In such cases, these individuals will likely be unmatched in the thesaurus. Still, the majority of complex name variation tends to cluster around well-documented individuals, precisely those who are more likely to appear in authority files. In practice, researchers may need to consult multiple thesauri. To bring consistency to the resulting patchwork of matches, it becomes essential to generate a consistent internal identifier system, one that accounts for entities matched across authority files, those identified through clustering, and those that remain unmatched.

In the context of my pilot study on the *Collectio academica antiqua*, the disambiguation process began with 5,018 name strings extracted from MARC 21 records. I first applied a bottom-up procedure using OpenRefine's clustering functions, combined with biographical information such as life dates and numeration. I then adopted a top-down approach by matching these entities to external authority files. I ended up identifying a total of 4,192 distinct individuals. Among these identified individuals, 47.5% (2,033) were matched to the CERL Thesaurus and a further 5% (213) to other authority sources such as Wikidata or Onderzoekssteunpunt en Databank Intermediaire Structuren (ODIS). The remaining 2,032 (47.5%) had no external match and were retained as internally disambiguated entities.

Disambiguation remains one of the most time-consuming steps in preparing bibliographic metadata, yet it is essential. Without it, individual identities would not be consistently established, and relationships would rest on unreliable associations. Crucially, it is what gives the data meaning for relational study. The process can also create synergies: matching entities to authority files often yields additional information that proves useful even beyond the immediate project.

### 2.3.2 Geolocating Place Names

Bibliographic metadata is typically rich in geographic information because it often records the place of publication, manufacture, and distribution for

each library holding. As with personal and corporate names, place names are typically transcribed as they appear on the title page. This practice, once again, prioritizes fidelity to the original source over internal consistency of the catalog. As a result, the same city may appear under multiple variants. These may reflect different languages, historical spellings, political jurisdictions, or typographical inconsistencies. To make this information usable for relational analysis, place names must be standardized and geolocated. This task is not fundamentally different from disambiguating historical actors, but it is often somewhat easier due to the more stable nature of place identities. Nonetheless, caution is essential. Even with present-day place names, it is surprisingly easy to misidentify a location because many places can share the same name within or across countries. This issue may be less common for cities that are on average better known and thus more consistently recorded, like printing centers. Still, historical geolocation poses a challenge. The process often requires close attention to historical context and supporting metadata.

For harmonizing place names, best practice is to reconcile each string against a standardized entry in a historical urban gazetteer or geographic authority file. These resources contain a wide range of name variants, including multilingual and obsolete forms, which significantly facilitate disambiguation and geolocation. Among the most comprehensive options are Wikidata, GeoNames, the CERL Thesaurus, the RBMS/BSC Latin Place Names File, and the World Historical Gazetteer, all of which are particularly well suited to historical case studies. In my pilot study of Louvain's Collectio academica antiqua, I processed 631 valid place-name strings, reconciled into 166 distinct locations. Table 2 reports, for the ten most frequent printing locations, the number of distinct place-name variants and the corresponding number of imprints.

After harmonization, locations can be geographically referenced. Although coordinates typically refer to a single point rather than the full spatial extent of a city or region, they still enable meaningful spatial representation. For example, in Schich et al. (2014) and de la Croix and Scebba (2024), coordinates are used to visualize city attractiveness based on the mobility of, respectively, notable individuals and medieval and pre-modern scholars affiliated to European academic institutions.

Once coordinates are available, they can be plotted on vectorized shapefiles or digital atlases, which represent geopolitical or cultural regions as spatial polygons. In historical analyses, it is particularly important to work

**Table 2.2.** Top 10 unified locations by number of imprints.

| Place name | Number of variants | Number of imprints |
|---|---|---|
| Louvain | 64 | 1,143 |
| Lyon | 18 | 129 |
| Antwerp | 68 | 493 |
| Ingolstadt | 5 | 9 |
| Mainz | 5 | 17 |
| Amsterdam | 14 | 62 |
| Brussels | 30 | 176 |
| Hanover | 1 | 11 |
| Rome | 7 | 31 |
| Cologne | 16 | 215 |

with maps that reflect time-consistent boundaries. Using modern administrative boundaries for early periods can misrepresent political geography and bias spatial patterns. A wide array of historical vectorized digital atlases or shapefiles are available to support this kind of analysis. Both computational humanists and applied economic historians are increasingly contributing projects in this direction. However, the limited integration between disciplines means that humanists are often unaware of the spatial datasets used by economists, and vice versa. These resources may vary in geographic coverage and temporal resolution, where some provide century-scale snapshots, while others offer yearly updates with shifting polity boundaries. Below, I present a selection of the most commonly encountered and widely used resources across computational humanities and quantitative economic history.

- The Centennia Historical Atlas: Academic Research Edition offers high-resolution, time-sensitive boundaries of European polities between the late Middle Ages and the nineteenth century (Reed, 2016).
- The Cliopatria database records overlapping and successor polities from 3400 BCE to the present (Bennett et al., 2025).
- EurAtlas provides digital cartographic reconstructions of European boundaries spanning two millennia.[7]

[7]Developed by Christos and Marc-Antoine Nüssli. For more on EurAtlas, see https://euratlas.com/.

- Developed by economist Victor Gay and hosted by Harvard Dataverse, the Third Republic France Geographic Information System (TRF-GIS) maps French administrative constituencies spanning the period 1870–1940.[8]

- The IPUMS Mosaic project, as the name suggests, provides a patchwork of harmonized historical census microdata and corresponding geographic boundary files for selected countries in the European area at different points in time.[9]

- The Historical GIS Collection at ETH Zurich, compiled by computational sociologist Ramona Roller, supplies static polygon boundaries for territories within the sixteenth-century Holy Roman Empire, along with attributes such as foundation dates and confessional status.[10]

- The project (Re)counting the Uncounted, led by historian and digital humanist Rombert Stapel, produced the Historical Atlas of the Low Countries, a detailed GIS dataset designed to anchor premodern population censuses to historically accurate administrative units.[11]

A comparative overview of these resources is provided in Table 3, which summarizes their temporal and geographic coverage, resolution, and accessibility. While the Centennia, Cliopatria, and EurAtlas shapefiles offer broad historical coverage—either pan-European or global—and span from antiquity to the modern era, they were primarily designed as general-purpose tools and are widely used in applied economic history. By contrast, the other atlases are more narrowly focused, both in geographic scope and in time frame. Figure 1 plots publication centers of the *Collectio academica antiqua*'s holdings over time, employing time-consistent and evolving boundaries.

[8] For more on the TRF-GIS, see https://dataverse.harvard.edu/dataverse/TRF-GIS.

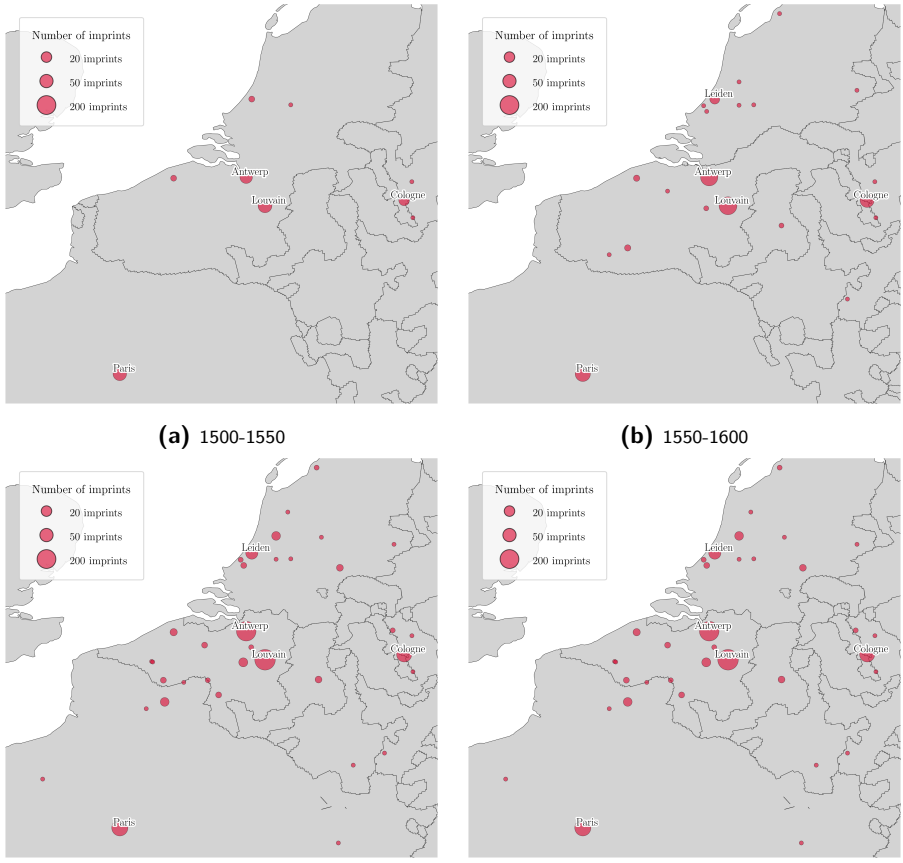
[9] For more on the IPUMS Mosaic project, see https://mosaic.ipums.org/historical-gis-datafiles.

[10] For more on the Historical GIS Collection at ETH Zurich, see https://www.research-collection.ethz.ch/handle/20.500.11850/472583.

[11] For more on (Re)counting the Uncounted, see https://datasets.iisg.amsterdam/dataverse/recountingtheuncounted.

**Table 2.3.** Comparison of selected historical shapefiles by time span, spatial scope, temporal resolution, and accessibility.

| Name | Coverage (Time) | Coverage (Geography) | Temporal Resolution | Open Access |
|---|---|---|---|---|
| Centennia Historical Atlas | ca. 1000–2000 CE | Europe | Variable, up to one-tenth of a year | No |
| Cliopatria | 3400 BCE–present | Global | Variable | Yes |
| EurAtlas | 1–2000 CE | Europe | 100-year steps | No |
| TRF-GIS | 1870–1940 | France | Annual | Yes |
| IPUMS Mosaic | 1770–2003 | Europe and selected countries | 30-year steps (Europe-wide); finer for some states | Yes |
| ETH Zurich HGIS | ca. 1500–1600 | Holy Roman Empire | Static (single date) | Yes |
| Historical Atlas of the Low Countries | 1350–1850 | Low Countries | Variable (per census) | Yes |




**(a)** 1500-1550 **(b)** 1550-1600

**(c)** 1600-1650 **(d)** 1650-1700

**Figure 2.1.** The maps show printing centers in the Low Countries and surroundings at fifty-year intervals. Bubble area is proportional to the number of printed holdings from the *Collectio academica antiqua* associated with each city. Historical boundaries are time-varying and drawn from the Centennia Historical Atlas.

### 2.3.3 Classifying Content: Assigning Topical Terms

Library catalogs typically include some indication of what each book is about. In the MARC 21 schema, this information tends to be stored in tags 650 and 655, which refer respectively to topical words and genre categories. However, this metadata might be incomplete, inconsistently applied, or skewed by cataloging practices. While a close reading approach might allow for a direct understanding of the subject of each book, this method is not feasible in large-scale, data-driven studies. In these cases, a systematic and ideally automated workflow for classifying the content of library holdings becomes essential. Most of these approaches draw on a key element: textual input, most commonly the book title, which in early modern imprints is often lengthy and content-rich. Although full-text content could, in theory, offer even richer input, the focus here is on completing and standardizing existing catalog metadata to enable structured comparisons across records. Assigning accurate subject metadata is an essential preprocessing step for higher-level analyses tracking thematic trends over time, identifying disciplinary clusters, or studying the diffusion of ideas across communities.

Automated classification techniques broadly fall into two categories: supervised and unsupervised methods. Supervised methods rely on a labeled dataset, that is, a subset of records for which the correct classification (e.g., field or topic) is already known. Machine-learning models are trained to learn these associations and predict labels for unclassified data. An example of supervised topic modeling in the case of bibliographic metadata is introduced in Koschnick, and it involves fine-tuning a pretrained transformer model such as DistilBERT (Sanh et al., 2019) on a dataset of titles or abstracts paired with subject headings. The model learns to classify each input into one of the predefined topics, based on patterns in the text. Unsupervised methods, by contrast, do not require labeled input. Topic modeling techniques, such as Latent Dirichlet Allocation (LDA) (Blei et al., 2003), attempt to infer hidden thematic structures by analyzing word co-occurrence patterns in a corpus. Similarly, algorithms such as k-means clustering can group texts based on similarity in their vector representations without any prior knowledge of categories. Applications of k-means clustering to bibliographic metadata include the analysis of genre-indicating subtitles in German literature (Gittel, 2021) and the definition of academic fields from publication titles (Curtis and de la Croix, 2025).

While unsupervised methods can help identify thematic patterns, they are often less appropriate when the goal is to complete or standardize existing subject metadata. In such cases, including the present paper, supervised approaches are favored because they directly learn from cataloged examples and preserve alignment with established classification schemes.

In my case study of the *Collectio academica antiqua*, I used a supervised approach. I identified a topical classification term for 38.4% of the holdings (1,450 out of 3,777 records). This subset served as the labeled data for training. I then proceeded to derive a higher-level classification by grouping subject headings in more general categories. To assign topics to the books in the *Collectio academica antiqua* that lacked subject metadata, I fine-tuned a DistilBERT model to classify titles based on the labeled records. Each title was mapped to one of three broad disciplinary categories—theology, humanities, or sciences—based on the existing catalog labels. Because the model was pretrained on English text, I first translated titles to ensure compatibility. The model was then trained over four epochs; that is, it went through the entire training dataset four times to refine its predictions.

To evaluate the model, I set aside 20% of the labeled data (290 titles) as a validation set. Model performance differs markedly across categories. As displayed in Table 4, precision, recall, and F1 scores are high for theology and humanities, the two dominant classes in the labeled data, indicating that the classifier learns stable and interpretable patterns for these categories. By contrast, the sciences category is severely underrepresented in the training set, and the model fails to predict this class in the validation data, resulting in zero recall for this category. As a consequence, overall accuracy is driven by the two majority classes, while macro-averaged performance metrics provide a more informative summary under class imbalance.

**Table 2.4.** The table reports precision, recall, and F1 scores by class for the validation set of the *Collectio academica antiqua*. Support indicates the number of titles per category in the held-out validation data. Performance is strong for theology and humanities, the two majority classes, while the sciences category is severely underrepresented and not predicted by the model, reflecting the limits imposed by class imbalance rather than classifier instability.

| | **Precision** | **Recall** | **F1-score** | **Support** |
|---|---|---|---|---|
| Humanities | 0.793 | 0.836 | 0.814 | 110 |
| Sciences | 0.000 | 0.000 | 0.000 | 6 |
| Theology | 0.897 | 0.897 | 0.897 | 174 |

Figure 2 illustrates this imbalance, reporting both raw counts and row-normalized percentages in order to make performance comparable across unevenly sized classes. Each row of the matrix represents the actual category of the entries from the validation set, while each column shows the predicted category assigned by the model. If the model worked perfectly, each class count would fall on the diagonal, where predicted labels match the actual ones, and the off-diagonal entries, which indicate misclassification, would be empty. Most errors occurred between theology and humanities, which reflect their historical overlap. The sciences category is severely underrepresented in the training data, and the model fails to predict this class in the validation set, reflecting the limits imposed by class imbalance. Once the model was trained, I used it to assign topics to the remaining set of titles that lacked subject metadata. These predicted labels were then combined with the original metadata to generate a complete, standardized topical classification across the entire collection.

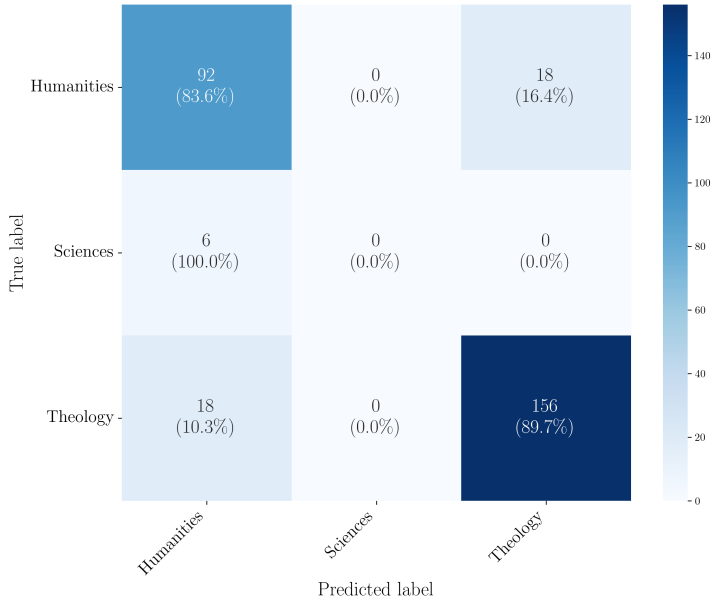



**Figure 2.2.** Confusion matrix for the validation set of the labeled holdings in the Collectio academica antiqua. Each row shows the true topic based on catalog metadata, and each column shows the model's predicted topic. Values along the diagonal (off-diagonal) indicate correct (incorrect) classifications.

## 2.4 Modeling Catalog Data for Analysis

This section returns to the research questions outlined in the Introduction by examining how different modeling choices condition the interpretation of bibliographic metadata when represented as relational and network-based data. Having outlined the steps required to transform bibliographic metadata into analyzable data, I now turn to their analytical modeling, focusing on network representations. First, I revisit the assumptions and limitations of interpreting catalog metadata as a graph. Then, drawing on a co-occurrence network derived from the *Collectio academica antiqua*, I show how the role attribute encoded in catalog metadata is not naturally accommodated by a single-layer projection and how a multilayer framework offers a coherent way to preserve this heterogeneity when adopting a co-occurrence representation. Lastly, I turn to quantitative economic history as a comparative analytical framework for the use of bibliographic metadata.

### 2.4.1 Adopting a Network Perspective

Catalog records are not inherently networked systems. However, they do encode structured associations among entities—authors linked to books, books linked to subjects, people linked to other people via shared roles in a publication, and so on. A network representation emerges only when we interpret these associations under a specific set of assumptions. For example, if two agents appear together on a title page, one might interpret that co-occurrence as a connection between them. Many studies in the computational humanities adopt this strategy. For instance, Gavin (2016) and Hill et al. (2019) on early English literary criticism and book trade, respectively; Greteman (2021) on the English print network from its origins through the eighteenth century; Ladd (2021) on early modern dedications; Valleriani et al. (2022) on networks of printers and publishers and their influence on the evolution of scientific knowledge; Ryan and Tolonen (2024a,b) on Scottish Enlightenment publishing; and Heßbrüggen-Walter (2024) on interdisciplinarity in seventeenth-century German dissertations. All of these works use the co-occurrence of individuals in print artifacts as an interpretative lens to reconstruct intellectual, textual, or social communities. This raises two broader methodological questions: Under what interpretive assumptions can bibliographic metadata be treated as network data, and

what are the limits of such representations? The *Collectio academica antiqua* provides a particularly fitting test for these questions because it documents a well-bounded institutional world (i.e., an academic print production orbiting around the university center of Louvain) where the logic of collaboration can be observed directly in the paratext and imprint data. This discussion directly addresses the second research question mentioned early on in the paper by making explicit the interpretive assumptions under which bibliographic metadata can be treated as network data and by delineating the limits of co-occurrence-based representations.

First and foremost, library catalog data were never intended to capture explicit relationships between historical actors. Their purpose is providing bibliographic description and information management. Nonetheless, converting a collection's metadata into a co-occurrence network means establishing links based on the structured associations present in the records (people, texts, places, and topics connected via shared bibliographic entries). This reinterpretation is not without pitfalls. A key risk is assuming that co-occurrence automatically implies a meaningful social relationship. Simply sharing a title page does not guarantee contemporaneity, collaboration, or even mutual awareness, especially in cases like posthumous editions or honorary dedications.

Yet, despite this ambiguity, the connections we derive from bibliographic metadata are grounded in historical evidence, that is, the sources themselves: title pages, imprint statements, colophons, and dedications. They are not arbitrary, ad hoc constructions imposed by the researcher. While the meaning of each connection may be open to interpretation, the very co-occurrence in the production of a shared work constitutes material evidence derived directly from the historical record (Hill et al., 2019). At the same time, bibliographic metadata can sometimes support a stronger definition of network ties than simple co-occurrence. For instance, using bibliographic metadata from the Sphaera corpus, a curated collection of 359 astronomy and cosmology textbooks published between 1472 and 1650,Valleriani et al. (2022) reconstruct a network of book producers by defining "awareness relationships" between early modern printers and publishers. Such relationships arise when two similar editions are attributed to different producers, typically indicating imitation or participation in the same print run. Here, the bibliographic fingerprint of editions and additional metadata on printers' and publishers' biographical timelines are used to establish, respectively, chains of editions and contemporaneity, yielding ties

that can be interpreted more robustly than co-occurrence alone.

### 2.4.2 The Multilayer Approach

With these caveats in mind, bibliographic metadata have been used to construct various kinds of networks depending on the entities and relationships of interest. One of the most common projections is a person-to-person network where an edge represents two individuals' joint participation in the same publication. This co-occurrence network is frequently employed to study scholarly or intellectual communities. Even this straightforward representation involves important modeling choices. A first decision concerns an edge's directionality. Defining edges as directed or undirected depends on the interpretative aim. For example, cases reflecting unilateral acknowledgment like links from authors to dedicatees might call for directed edges, whereas undirected edges more accurately capture the symmetric nature of shared participation in other contexts. A second choice concerns encoding some data as node attributes (e.g., relevant life dates, biographical information, and religious or institutional affiliation), or as edge attributes (e.g., the publication year and the number of shared works).

However, there is one critical dimension that resists both types of attribute encoding: the role of each person within the book. This is because roles are not fixed properties of individuals. Rather, a node may be an author in a publication and a dedicatee in another. Nor do they characterize the tie itself, because each tie reflects individual involvement in a shared work, not the work as a whole. For example, the connection between a printer and an author cannot be reduced to either an attribute of the person or an attribute of their tie.[12] Basically, individuals' roles in each publication are asymmetric and two-way, which disqualifies both node and edge attributes as adequate carriers of this information. Capturing these asymmetries is essential for understanding early modern print economies, where the same person could occupy different positions across publications.

Hence, I argue that the most coherent solution is to model the co-occurrence network as a multilayer graph. Following the definition offered by Kivelä et al. (2014), a multilayer network extends beyond nodes and edges to include *layers* as core components, allowing each node to appear in multiple layers

[12] An exception arises when focusing solely on a specific mode of co-occurrence, such as author–dedicatee pairings; see Ladd (2021)

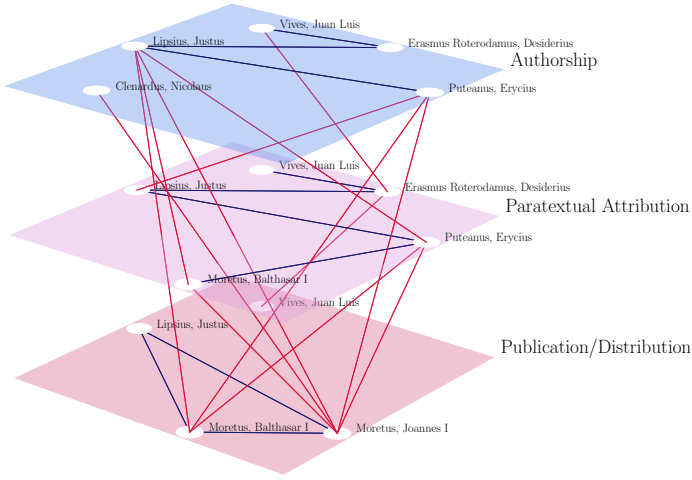



**Figure 2.3.** Multilayer person-to-person co-occurrence network constructed from the bibliographic metadata of the *Collectio academica antiqua*. Each layer isolates one type of role-based connection. Nodes represent individuals (only those with the highest degree are shown for legibility). Intra-layer edges (blue lines) and inter-layer edges (red lines) connect individuals who co-occur in at least one book, respectively in the same or in different roles. Plots generated with `pymnet` (Nurmi et al., 2024).

and enabling edges to connect any pair of node-layer instances. Additionally, multiple *aspects*, or features—that is, different sets of layers—can be modeled on top of each other (Kivelä et al., 2014, p. 209). In practice, an aspect defines a distinct dimension of variation in the data, such as the type of activity, the temporal slice, or the institutional context, so that multiple aspects can coexist within one unified, multilayer structure.

In the present case, I only focus on a single aspect, the co-occurrence dimension in the catalog, where each layer corresponds to a distinct role and where nodes may participate in one or more role-layers depending on their involvement in each publication. This framework allows for both inter-layer edges (i.e., connections stemming from shared works in different roles) and intra-layer edges (i.e., ties between individuals sharing the same publication *and* role). This representation makes it possible to measure how role multiplicity structures collaboration and to distinguish between within-role cohesion and cross-role integration.

Figure 3 offers a visual example[13] of what a multilayer co-occurrence network can look like.[14] I built a person-to-person co-occurrence network from the bibliographic metadata of the *Collectio academica antiqua*. The layers are defined from the roles recorded in the metadata, which I grouped into broader categories as detailed in Table A.1 in Appendix A.1.[15] The figure illustrates the methodological shift from descriptive catalog lists to an explicitly relational view, where each role layer isolates one channel of participation. Although the full network comprises several thousand individuals and ten role layers, the figure focuses on the fifty most active nodes and four main layers to improve legibility. In this representation, layers are the loci where nodes appear when a given role emerges in at least one shared publication. Thus, if an individual has different roles in different works, they appear in multiple layers. The layer classification preserves the co-occurrence principle: Edges are still defined by co-occurrence in publications. Dashed lines denote inter-layer edges and solid lines denote intra-layer edges.

Thinking of co-occurrence in multilayer terms allows for several modes of analysis. One can examine each layer separately to study the network of collaborations within a single role. This is particularly useful when intra-layer ties dominate or when the aim is to identify role-specific structural patterns. In practice, within-role co-occurrence in early modern data tends to be sparse. In the *Collectio academica antiqua* network, fewer than one-quarter of the ties link individuals who share the same role in a publication; more than 75% of connections are between people in different roles (inter-layer links; see Table

[13]Multilayer visualizations are admittedly hard to read, which is one reason they remain uncommon in applied work that relies mostly on network visualizations. Their potential, however, lies less in the visualization than in the data structure they preserve and the analytical infrastructure one can build on it.

[14]Multilayer networks are not to be confused with multiplex networks. As discussed in Kivelä et al. (2014), the latter are a special case of the former in which inter-layer links only connect identical nodes across layers (diagonal coupling), and layers are often—though not necessarily—node-aligned (i.e., all nodes exist in every layer). In the present case, we refer specifically to a multilayer network, as the very nature of co-occurrence entails inter-layer edges between different nodes, and nodes are not guaranteed to appear in every layer.

[15]This grouping is a practical necessity for visualization and interpretive clarity, yet it is not without consequences. Aggregating fine-grained roles into broader categories reduces the apparent number of ties and slightly alters degree distributions, since individuals active under multiple specific designations (e.g., “printer” and “publisher”) are counted as one.

A.3 in the Appendix). This is consistent with the low density observed in each individual layer (Table A.2), which in turn reflects the historical reality that collaborative roles, like for instance co-authorship, were relatively rare in early modern print culture. If we were to apply this method to modern scientific publications, we would expect to see a much denser authorship co-occurrence layer, whereas the early modern dataset under analysis shows that authors typically connect to others through complementary roles (e.g., author to printer, author to dedicatee, etc.) rather than directly to other authors. This dominance of inter-layer ties underscores how production and intellectual labor were intertwined: Most connections bridge material and scholarly functions rather than occur within a single occupational sphere.

Second, the layers can also be collapsed into a single projection when applying classical metrics such as centrality or modularity, though doing so sacrifices the interpretive nuance that multilayer distinctions offer.

Third, and most distinctively, one can analyze patterns that involve the interplay between layers, something not possible in a single-layer projection. In this case, the layers linked to material production (printers, publishers, and booksellers) provide a controlled setting to study how the combination of roles across publications relates to the university ecosystem from which much of the corpus originated.[16] Focusing on the production side is particularly revealing because the *Collectio academica antiqua* primarily contains academic works—dissertations, textbooks, and reprints of classical authors—commissioned for or produced within the orbit of the Old University of Louvain. In such a context, holding multiple production roles can be read as an indicator of alignment with the university's printing economy, where a small number of trusted workshops frequently combined several functions under one roof. Analyzing these overlaps therefore allows one to test whether functional integration within print workshops corresponded to closer institutional or intellectual integration with the university world itself.

To examine this mechanism empirically, I traced the involvement of academic actors across the corpus. The Studium.AI research infrastructure is developing a comprehensive list of the students enrolled at the Old University

[16]Unlike in the general multilayer representation, in the production-side analysis I rely on the original, fine-grained role designations instead of the aggregated macro-categories. The distinction between "printer," "publisher," and "bookseller" is historically meaningful in the early modern context and captures variations in economic function that would be obscured by broader grouping.

**Table 2.5.** The table reports the association of multi-role production agents with professors.

| Group | $\text{Share}_{\text{profs}}$ | $\text{Share}_{\text{Louvain}}$ | Observations |
|---|---|---|---|
| Single-role | 0.734 | 0.631 | 1,589 |
| Multi-role | 0.826 | 0.770 | 2,403 |
| Difference (multi − single) | 0.092 | 0.139 | |
| $z$ statistic | 6.96 | 9.53 | |
| $p$ value | $3.35 \times 10^{-12}$ | $1.61 \times 10^{-21}$ | |

of Louvain, but academic scholars can already be identified through the Repertorium Eruditorum Totius Europae (RETE) project, a pan-European prosopographical database documenting the university careers of medieval and early modern professors.[17]

Through the disambiguation process described in Section 3, I linked each person in the *Collectio academica antiqua* to a unique CERL and Wikidata identifier. This, combined with a manual search, allows the matching of individuals appearing in the collection to scholars from the RETE database. Given RETE's pan-European scope, the linkage allowed me to identify which books involve professors in general and which specifically involve professors affiliated with Louvain. The integration with RETE data adds a socio-institutional dimension to the bibliographic co-occurrence network: Some nodes now carry information on whether the person held a university position, thus allowing the relationship between print production and academia to be tested empirically. I find that around three-fourths of the books of the Collectio academica antiqua see the participation of a university professor, specifically Louvain professors in most of the cases but also professors active at other European universities (around 7%). These figures confirm that although the collection is institutionally rooted in Louvain, its intellectual network extended beyond the city. After all, the transregional dimension of the book trade in the Low Countries is a well-established fact in the literature (De Ridder and Soen, 2020).

I then tested whether production agents who held multiple roles, for instance those appearing as both printer and bookseller in different publications, are more frequently associated with books linked to professors. The test is based on a simple comparison of proportions. For each book, I classified the involved

[17] For a provisional interface of the project, see https://studium-ai.org/. For more on the RETE project, see https://ojs.uclouvain.be/index.php/RETE/about.

production agents as single-role or multi-role and recorded whether the book involves at least one professor. I computed the share of professor-linked books among multi-role and single-role agents and assessed whether the difference is larger than expected by chance using a difference-in-proportions ($z$) test.[18]

Results in Table 2.5 show that books involving multi-role production agents are significantly more likely to involve professors. The share of professor-linked books is about 83% for multi-role agents, compared with 73% for single-role agents, a difference of 9 percentage points ($z = 6.96$, $p = 3.35 \times 10^{-12}$). The association is even stronger for books involving professors affiliated with Louvain, where the corresponding shares are 77% and 63%, yielding a difference of 14 percentage points ($z = 9.53$, $p = 1.61 \times 10^{-21}$). These patterns illustrate the kind of association between production structure and academic involvement that a role-aware representation brings into view.[19] In this sense, the case study demonstrates how modeling choices are not merely technical, but substantively shape the kinds of institutional and historical patterns that can be recovered from catalog data.

The same reasoning applies when considering how many distinct layers each person appears in, sometimes called the actor's "multiplexity," or the number of role dimensions in their activity. In the Collectio academica antiqua network, only about 12% of individuals participate in more than one role layer (see Table A.4 in the Appendix). Those who do are often individuals who also served in another capacity (e.g., editors of other work, authors who also wrote dedications, etc.). Table A.5 in the Appendix shows the share of multi-role individuals in each layer. Unsurprisingly, roles like authorship and paratextual attributions account for a large fraction of the multiplex individuals, whereas very few censors, translators, or illustrators in this dataset also took on additional roles. The prevalence of paratextual connections (e.g., authors who share a printer also often share an editor or a

[18]The test compares proportions across two independent groups using a large-sample normal approximation to the difference in means of a binary outcome. Given the sample sizes involved, the approximation is appropriate.

[19]Taken in isolation, this comparison rests on a count of roles per agent, which a flat network could admittedly encode as a node attribute. What only the layer structure can reveal is the distinction between within- and cross-role co-occurrence: in my case study, more than three-quarters of all ties are inter-layer, linking individuals who share a work under *different* roles—a fact that cannot be recovered from a flattened projection. The structure moreover opens questions such as which roles co-occur with which, or whether particular role combinations emerge in certain conditions.

contributor) is partly a reflection of how thoroughly this particular catalog records those secondary roles. By contrast, the censorship role appears underrepresented in the network. Historically, from the 1520s onward, works printed in places like Louvain often included an approbation or censor's note (Cammaerts, 2024b, p. 241), but in the data derived from the Collectio academica antiqua, such connections are few. This likely indicates that many approbations were not captured in the catalog metadata, not that censorship was absent. A comparison with a dataset where censorship information is fully recorded could yield a denser censorship layer and alter the overall network structure. This example highlights how the depth and focus of cataloging practices (which information was recorded or omitted) directly influence the networks we can extract. In this perspective, a comparative analysis using this analytical lens could substantiate how differences in cataloging practices shape resulting network structures.

### 2.4.3 Insights from Quantitative Economic History

Some studies in history and literary scholarship that use bibliographic metadata often combine meticulous preprocessing pipelines with network visualizations and descriptive commentary, typically focused on centrality measures. These exercises undoubtedly elevate our understanding of the collections and catalogs used as material, yet, as Lemercier and Zalc (2019) observe, such work frequently remains at the stage of mapping and visual speculation, without fully mobilizing networks as formal analytical tools. The patterns observed in Section 4.1 (e.g., the sparsity of within-role links or the rarity of multi-role connections) highlight both the potential and the limitations of co-occurrence networks derived solely from catalog data. On one hand, such networks can reveal hidden structures in the catalog or collection studied; on the other, relying on networked bibliographic data exclusively, without complementary data or further analytical techniques, may limit the types of historical questions we can answer. To overcome this limitation, bibliographic network studies can draw on frameworks that have demonstrated how relational data can be embedded in explicitly articulated analytical designs, most notably, quantitative economic history.

In recent years, bibliographic metadata have begun to feature in applied economic history, often alongside other sources. These studies go beyond descriptive and exploratory scopes and use a range of empirical tools beyond

network analysis. Importantly, bibliographic metadata are not treated as econometric data but instead as surviving historical evidence, interpreted and combined with other sources within composite datasets tailored to specific historical questions, for example, building on library catalogs, biographical sources, and manuscript data Chaney (2023, 2024) constructs a georeferenced dataset of authors from which he derives metrics such as author counts, the share of authors engaged in scientific topics, and number of authors' deaths. He uses this resource to address two distinct questions, respectively, that of the decline of scientific output in the medieval Islamic world and that of the enhancement of preindustrial city growth estimates.

Chiopris (2024) analyzes how spatial connections, namely the introduction of the railroad network in nineteenth-century German-speaking regions, shaped the creation and diffusion of ideas. Using comprehensive library catalogs from the German Collective Library Consortium, the author measures the emergence and spread of new ideas, broadly defined as novel publication topics, new words, or new combinations of existing concepts derived from enriched catalog metadata.

Koschnick (2025b) links author-level publication data from the English Short Title Catalogue (ESTC) to collegiate records from Oxford and Cambridge Universities to examine the diffusion of ideas between teachers and students from 1600 to 1800. In another study, de Pleijt and Koschnick (2025)combine ESTC data with information on scholars' socioeconomic backgrounds, and using sentiment analysis on the titles of printed works, they find that limited career prospects for less advantaged academics coincided with a rise in dissenting religious publications.

Finally, Cervellati et al. (2025) use metadata from the epistolary union catalog Early Modern Letters Online (EMLO) to study the role of the Republic of Letters in explaining why Britain experienced a sustained economic take-off at the time of the Industrial Revolution. These contributions do not leverage networks as an analytical lens, but they do use bibliographic metadata as part of broader quantitative frameworks to address questions about scholarly output, institutional change, and idea diffusion. What they also show is that bibliographic metadata can sustain formal empirical strategies once linked with complementary datasets, to the end of the pursuit of a specific research question.

The disciplinary silo between quantitative economic history and digital or computational history has left analytical strategies developed in the former

largely unexplored in the latter. From this perspective, large-scale patterns revealed by bibliographic metadata can be connected to deeper historical questions when analyzed with the appropriate combination of domain knowledge and formal methods. A crucial step in this direction involves enriching catalog metadata with external sources, based on the research question at hand. The multilayer model proposed in Section 4 might operationalize this integration as follows. The bibliographic co-occurrence network can be conceived as an initial aspect or feature, that is, a first set of layers, onto which additional layer sets could be stacked to encode relations for the same entities drawn from complementary sources. This approach facilitates the incorporation of external knowledge into digital history analyses and lays the groundwork for applying more sophisticated network techniques, like, for example, peer-effects models.

## 2.5 Conclusions

In an effort to provide perspective on research materials and methods from neighboring fields, this paper has highlighted a growing convergence of computational history and quantitative economic history around shared sources and tools, most notably bibliographic metadata and network analysis.

Using a role-differentiated co-occurrence network representation of the Collectio academica antiqua of the Old University of Louvain, I demonstrated how different involvement types in book production (authorship, publication and distribution, dedication, censorship, and so on) can be modeled without being flattened into a single type of tie. The interplay of inter- and intra-layer connections reveals structural patterns that reflect both historical publishing practices and cataloging norms. Yet the exploration of my case study also shows that when bibliographic metadata is used in isolation and interpreted solely through network visualizations, the analytical power remains limited.

This underlines two distinct but complementary points. First, when bibliographic metadata are modeled as networks, analytical coherence depends on preserving the internal structure of the records rather than flattening heterogeneous forms of participation into undifferentiated ties. Second, even when such structure is preserved, bibliographic metadata alone offer limited leverage for addressing complex historical questions. Their analytical potential is fully realized only when relational representations derived from catalog data are linked to complementary sources that provide

institutional, social, or biographical context. Together, these points respond to the research questions posed at the outset by clarifying both the conditions under which bibliographic metadata can be meaningfully modeled as relational data and the limits of such modeling when used in isolation.

In this sense, the contribution of the present article is threefold: a methodologically reflective case study, a role-aware relational modeling strategy, and a cross-field perspective that situates bibliographic metadata at the intersection of humanities and social science research.

# Appendix

## 2.A Mapping of Roles

This subsection presents the grouping strategy used to translate detailed role descriptors into broader layer categories for multilayer network construction. The table clarifies how fine-grained roles, such as "writer of preface", "censor", "engraver", and others, are aggregated into interpretable layers like Paratextual Attributions, Censorship, and Engraving/Illustration. The grouping reflects a deliberate analytical choice designed to balance representational fidelity with visual and conceptual clarity. Other researchers may opt for alternate classifications according to their research aims.

**Table 2.A.1.** The table describes the mapping of granular roles to broader layer categories.

| Layer | Granular roles |
|---|---|
| Authorship | author; author original work; dubious author |
| Paratextual Attributions | writer of preface; compiler; adapter; readapted by; author in quotations or text abstracts; narrator; commentator for written text; author of introduction, etc.; corrector; compiler of index; eulogist; collaborator; contributor |
| Translation | translator |
| Publication/Distribution | printer; editor; bookseller; publisher; auctioneer; patron |
| Dissertant | dissertant |
| Promotor | praeses; thesis advisor |
| Censorship | approbation; censor |
| Engraving/Illustration | engraver; illustrator; woodcutter; artist |
| Dedication | dedicatee; honoree; depicted; addressee |
| Unknown | other; responsible party; collection from; missing or unspecified role |

## 2.B Per-layer Network Analysis

**Table 2.B.1.** The table reports per-layer counts of nodes, intra-layer edges, and density.

| Layer | Nodes | Intra-layer edges | Density |
|---|---|---|---|
| Authorship | 467 | 1,103 | 0.0101 |
| Censorship | 172 | 718 | 0.0488 |
| Dedication | 376 | 1,120 | 0.0159 |
| Dissertant | 42 | 50 | 0.0581 |
| Engraving/Illustration | 71 | 68 | 0.0274 |
| Paratextual Attributions | 419 | 1,574 | 0.0180 |
| Promotor | 4 | 2 | 0.3333 |
| Publication/Distribution | 948 | 1,200 | 0.0027 |
| Translation | 35 | 31 | 0.0521 |
| Unknown | 277 | 670 | 0.0175 |

**Table 2.B.2.** The table reports global multilayer network statistics, including node counts, layer combinations, and the number of intra- and inter-layer edges.

| Metric | Value |
|---|---|
| Total intra-layer edges | 6,536 |
| Total inter-layer edges | 20,487 |
| Total edges overall | 27,023 |
| Total unique nodes across all layers | 4,119 |
| Total node-layer combinations | 4,905 |

## 2.C Inter-layer Network Analysis

**Table 2.C.1.** The table reports descriptive statistics of degree of multiplexity in the *Collectio academica antiqua* network.

| Min | Q1 | Median | Mean | Q3 | Max | Single role | Multiple roles |
|---|---|---|---|---|---|---|---|
| 1 | 1.0 | 1.0 | 1.2 | 1.0 | 7 | 3,595 (87.3%) | 524 (12.7%) |

**Table 2.C.2.** The table reports the shares of multiplex nodes present in each layer.

| Layer | Share |
|---|---|
| Authorship | 0.618 |
| Unknown | 0.529 |
| Paratextual Attributions | 0.471 |
| Publication/Distribution | 0.328 |
| Dedication | 0.313 |
| Censorship | 0.095 |
| Translation | 0.088 |
| Promotor | 0.034 |
| Engraving/Illustration | 0.017 |
| Dissertant | 0.006 |

# 3

# Pre-modern academic institutions as conduits for the potential spread of ideas

*This chapter is co-authored with David de la Croix and Chiara Zanardello.*

*While good ideas can emerge anywhere, it takes a community to develop and disseminate them. In premodern Europe (1084-1793), there were approximately 200 universities and 150 academies of sciences, home to thousands of scholars from the Middle Ages to the First Industrial Revolution. By inferring co-presence from institutional affiliations, we simulate how ideas would spread from one scholar to another across the European academic network. We find that the implied exposure patterns align with observed urban developments: examples include botanic gardens, astronomical observatories, and Protestantism. Scholars' mobility and multiple affiliations sustain the diffusion, and counterfactual simulations underscore the bridging role played by scientific academies. We also show that the spread of ideas through the affiliation network was locally fragile but globally robust, pointing towards academia as being a connective infrastructure underlying early European development.*

## 3.1 Introduction

In Europe during the Middle Ages and Early Modern period, more than one hundred thousand scholars were engaged in the production, dissemination, and development of various forms of knowledge. These scholars did not operate in isolation: two key institutions, universities and academies, facilitated their interaction. These institutions organized teaching and research within self-governed communities of scholars, as stressed in Rashdall (1895) and McClellan (1985). From medieval universities such as Paris and Bologna, to the Royal Society centuries later, formal institutions brought together scholars to develop and disseminate scientific discovery. These institutions provided scholars not just with employment, but with a physical proximity that facilitated interaction.

Over this period, Europe experienced successive waves of transformation from the medieval urban revolution to the Enlightenment through humanism, the Reformation, and the Scientific Revolution. Universities and academies were closely entwined with these foundational shifts. Whether such institutions were central to Europe's intellectual success, or instead functioned largely as privileged and resource-consuming corporations (Smith, 1776), remains an open issue. We address this question by studying one mechanism through which they could matter: the diffusion of ideas through face-to-face interactions within academic institutions.

Specifically, we propose a structural, microdata-driven model to measure how ideas would spread from one scholar to another based on co-presence at the same academic institution. Similar to the approach taken by Becker et al. (2024) in their analysis of the pre-WWII period, we create an affiliation network where connections between scholars are based on their overlapping presence at a university or an academy. To simulate the transmission of ideas, we combine an epidemiological approach with the network structure (Banerjee et al., 2013; Koher et al., 2016; Fogli and Veldkamp, 2021). Our framework fits within the class of network diffusion models, such as that formalized in Bramoullé and Genicot (2024). This is the first analysis to reconstruct the European academic affiliation network at such a long-run and granular scale, allowing us to simultaneously capture the continent-wide dynamics of idea diffusion and the local, individual-level conditions under which ideas spread—or fail to do so.

Our network is dynamic and spans 1084 to 1793. Our timeframe starts

with the establishment of Irnerius's school of jurisprudence in Bologna (c. 1050 - after 1125) and concludes with the French Revolutionary Convention (1793). The *nodes* of the network are premodern scholars. A connection, or an *edge* in the network, is established between any pair of scholars who share at least one year of concurrent affiliation at the same institution and work within a broadly similar field. We assume that ideas spread through networks like infections, with scholars transmitting their inventions to peers with a certain probability, the *link activation probability*. This parameter governs the overall diffusion of ideas through the network by regulating the likelihood that an interaction between two connected scholars leads to exposure. A value close to zero suggests that face-to-face interactions rarely lead to knowledge transfer (e.g., due to non-academic discussions), while a high value indicates they are an effective channel for diffusion.

Ideally, we would estimate this parameter by comparing the model's predictions to empirical moments that reflect the historical spread of ideas. However, a key challenge is that we do not directly observe ideas as they propagate. Unlike studies that trace the trajectory of a single idea—such as the study by Xue (2025) on Wang Yangming's influence in Chinese texts, or Giorcelli et al. (2022) on the spread of Darwinian theory—or those that focus on specific linguistic or regional contexts, like Chiopris (2024) on 19th-century ideas in the German library consortium, our setting lacks detailed records of idea diffusion. Contemporary work, such as that of Ahmadpoor and Jones (2017), leverages citation and patent data to track knowledge flows over time: no comparable measures exist for Early Modern Europe. Nevertheless, we observe outcomes plausibly linked to idea adoption, which we use to conduct indirect inference.

Our approach is to estimate the link activation probability by simulating the diffusion of ideas through the affiliation network and constructing measures of exposure at the scholar, institution, and city levels. We then correlate these simulated exposures with historical outcomes in what we term auxiliary models, drawing on the framework of indirect inference (see Smith (2008)). We benchmark our diffusion model on two breakthroughs of the Scientific Revolution that have both clear historical importance and observable city-level proxies.

The first auxiliary model is the rise of botany as an independent discipline (*flora*), driven by Leonhart Fuchs (1501–1566), a professor at the Universities of Tübingen and Ingolstadt. Fuchs' work emphasized direct observation of nature,

culminating in the publication of a comprehensive herbal featuring accurate plant illustrations and medicinal descriptions. We label this shift "Botanical Realism". The heightened interest in botany across Europe that followed is evidenced by the spread of botanic gardens. We calculate each city's simulated exposure to Botanical Realism and use a proportional hazard model to estimate the speed at which a botanic garden is being established. We find that higher exposure is associated with a faster establishment of botanic gardens, especially in cities with a smaller scholarly mass susceptible to be exposed.

The second auxiliary model focuses on the mathematical reform of astronomy (*cosmos*), particularly the foundational role of Johannes Regiomontanus (1436–1476) in the advancement of trigonometry and astronomy. We group his innovations, which were later instrumental to the work of Copernicus, Kepler, and Galileo, under the label "Mathematical Astronomy". Regiomontanus held positions in Vienna (Universität Wien, 2024), Bratislava (Gabriel, 1969), Padua (Facciolati, 1757), and Rome (Renazzi, 1803), which facilitated the transmission of his ideas. We measure exposure to his innovations and analyze the correlation with the creation of astronomical observatories across Europe. Again using a proportional hazard model, we find that higher exposure is associated with a faster establishment of observatories, with the effect strongest in cities with a smaller eligible scholarly mass.

Admittedly, the historical foundation of botanical gardens or of astronomical observatories may have depended on many local factors beyond intellectual exposure, such as institutional resources and patronage, whose systematic reconstruction is not possible. The correlations we document between network exposure and the speed at which these institutions are established should therefore not be read as exposure alone determining the outcome, but as indicating that exposure paved the way to act earlier, *ceteris paribus*.

We then estimate the link activation probability by maximizing the joint likelihood of these two auxiliary models. The estimated value is 0.45. This value generates exposure patterns that align with related historical outcomes and is consistent with an important role of social and institutional connections in the transmission of ideas. In that sense, our results echo both apprenticeship-based models in which the master-apprentice face-to-face communication is key (de la Croix et al., 2018), but also contemporary insights on the persistent importance of face-to-face interactions in innovative environments like the Silicon Valley

(Atkin et al., 2022).

With the estimated link activation probability, we use the model in two additional applications by linking exposure to scholasticism (*salvatio*) to the probability of cities adopting Protestantism and the spread of anti-Semitic ideas. Scholasticism was pioneered by Petrus Lombardus (c. 1100 – 1160), a professor in Paris, and it was the dominant approach to philosophy and theology in the Middle Ages. Followers of scholasticism used logical reasoning to explore theological questions, and this method was adopted in many universities. Over time, however, it became increasingly detached from the practical concerns of believers and devolved into abstract debates, a decline often cited in historical literature (Chaunu, 2014; Barrett, 2023). Exposure to scholasticism may have been associated with an intellectual climate receptive to Reformation ideas, possibly reflecting a backlash against intellectualized theological debate and a stronger emphasis on scripture (Chaunu, 2014). To document this association, we simulate the diffusion of scholasticism through the affiliation network. We capture the exposure to scholasticism across universities in 1508 and infer the exposure of nearby cities. Using a linear probability model, we show that a higher exposure to scholasticism correlates with the adoption of Protestantism. In a similar vein, we extend the classical finding that pogroms against Jews were more likely to occur following major shocks, such as the Black Death (Becker and Pascali, 2019; Jedwab et al., 2019; Voigtländer and Voth, 2012), by showing that the anti-Judaic climate embedded in scholastic thought acted as a complementary force in this process.

To assess the features of the network that are most critical for idea diffusion, we turn to exploring counterfactual scenarios. In this analysis, we compare observed outcomes with the hypothetical scenarios that would have emerged under alternative conditions.[1] In the first counterfactual experiment, we reassign the intellectual origin of an idea to different individuals and ask whether it would still propagate across Europe. While some peripheral institutions are not well enough connected to guarantee the survival of an idea, we find that in most cases, ideas still reach the entire network within a

[1]It is important to emphasize that our focus on counterfactual scenarios is primarily methodological. We use counterfactual scenarios to gain a deeper understanding of the academic network's intrinsic properties, rather than to definitively predict what might have occurred in these hypothetical situations. This approach aligns with the spirit of classic works like Fogel's 1964 seminal study on the impact of railroads on American economic development, which emphasized the importance of counterfactual reasoning for historical analysis.

couple of centuries. However, the speed and route of diffusion vary, emphasizing the non-ergodic nature of the process and the importance of initial conditions. In the second experiment, we compare the role of academies to that of universities. When we remove academies from the network, there is a notable drop in the geographical reach of ideas, especially those that originated in remote or disrupted areas. Academies, with their international memberships, often served as critical diffusion channels, bridging regions and sustaining intellectual exchange during crises such as university closures during the Thirty Years' War. In the third experiment, we simulate the removal of entire countries or regions—such as the British Isles, France, or the Italian and Iberian Peninsulas—to assess their systemic importance. Surprisingly, even when major regions are excluded, most ideas continue to spread widely across Europe. This suggests that most regions were not essential to the flow of knowledge, underscoring the network's remarkable resilience to local shocks.

Taken together, these counterfactuals reveal an affiliation network that is globally robust yet locally fragile. While alternative paths almost always exist, early diffusion often depends on narrow institutional channels, making the survival of ideas partly contingent on chance.

The primary interest of the present work lies in the network formed by universities and academies, its structure, connectivity, and capacity for diffusion. The ideas whose spread we model are instrumental to reveal these properties, and as such are treated in the present article as necessarily stylized abstractions of much richer intellectual traditions studied extensively elsewhere. In this perspective, our article is not intended as a contribution to the history of knowledge in the strict sense. The argument is therefore best understood as a thought experiment about mechanisms of diffusion, one that deliberately simplifies by attributing ideas to single originators and treating observed outcomes as reflections of network reach among many other determinants.

Our paper relates to several strands of the literature. First, we contribute to the growing body of work on the diffusion of knowledge in historical contexts. Existing studies have explored complementary channels such as printed books (Chiopris, 2024), correspondence networks (Cervellati et al., 2025; Roller, 2023c), translations (Abramitzky and Sin, 2014), and mentorship influences (Koschnick, 2025b). Other contributions combine multiple channels to reconstruct broader diffusion processes (Becker et al.,

2020, 2025). By contrast, we emphasize academic institutions as aggregating centers,[2] fostering repeated opportunities for intellectual exchange among their members.[3] This could occur via physical proximity but also through institutionally-organized epistolary exchanges, as practiced in some academies. Focusing on institutional proximity offers two key advantages. On the one hand, observing whether scholars were affiliated with the same institution at the same time allows us to infer potential exposure to ideas, independently of whether those ideas were ultimately adopted or endorsed. On the other, gaps in the network are, at least in part, observable. Gaps occur when universities are underrepresented in our sources, causing some professors to be missing. In contrast, gaps in correspondence networks, such as letters that have not survived, are fundamentally unobservable and thus impossible to systematically evaluate.

Second, our paper engages with the history of science and upper tail human capital, from broad studies on the emergence of science and the scientific method (Needham, 1964; Wootton, 2015), to analyses of scientists' roles in society (Ben-David, 1971; Hanlon, 2025). Mokyr (2005; 2016; 2011b), Ó Gráda (2016), and Almelhem et al. (2023) explore the roots of the industrial revolutions through the accumulation and application of useful knowledge during the Scientific Revolution and Enlightenment. In addition, historical scholarship highlights the impact of highly skilled individuals (Meisenzahl and Mokyr, 2012), specialized engineers (Hanlon, 2025; Maloney and Valencia Caicedo, 2022), and key inventors (Hallmann et al., 2022) in driving Britain's technological edge. Our analysis mostly covers propositional knowledge, complementing these studies by tracing the intellectual roots from which practical knowledge later emerged—for example, when early modern advances in atmospheric science associated with Torricelli and Boyle paved

[2] Aggregation centers were crucial for intellectual exchange in the medieval and premodern world. As shown by Brunt and García-Peñalosa (2022), cities played a similar role in fostering knowledge diffusion by concentrating individuals and increasing opportunities for encounters and idea transmission.

[3] A striking example of this dynamic is the relationship between the abovementioned Regiomontanus and Polish astronomer Martinus Bylica de Ilkusz, who met at the University of Padua in 1463. Their long-lasting intellectual bond—with Bylica amending the Regiomontanus manuscripts—mirrored that of their mentors, respectively, Georg Peuerbach and Martinus Król (Domonkos, 1968), and underscores how institutional settings also fostered scholarly connections across generations.

the way for practical applications in the eighteenth century.[4]

Other works show how the institutional context matters: classical composers thrived in more liberal environments (Borowiecki, 2013), patent systems shaped the direction of innovation (Moser, 2005), and external shocks—such as the U.S. Civil War, spurred targeted technological responses in Britain's textile industry (Hanlon, 2015). Our approach also aligns with Akcigit et al. (2018) in its focus on tracking "individuals, their productivity, and their interactions over time" (Akcigit et al., 2018, p.2). Our analysis also relates to work on academic superstars, which shows that the removal of central figures can disrupt or reconfigure knowledge networks (Azoulay et al., 2010, 2019). While the focus of those papers is on publication output, the underlying question—how resistant to disruption academic networks are—is also relevant to our framework.

Finally, we also engage tentatively with the literature on the rise of the West, and the Europe-China differences (Greif et al., 2025). Indeed, having described the strength of European academia in terms of connectivity and resilience, we believe it could have been instrumental to Europe's success during the Early Modern period. This remains speculative, however, as we lack comparable data on academic networks in other parts of the world, and rely only on anecdotal evidence. In 1798, Thomas Malthus (1766–1834), a fellow at the University of Cambridge from 1793, published a treatise on population and development (Malthus, 1807). He developed the idea that population growth tends to outpace food production, leading to inevitable constraints on development. In 1818, Malthus became a Fellow of the Royal Society. Malthus' view had an immense influence on political economy in the following decades. Still today, his ideas are modeled and debated (André and Platteau, 1998; Ashraf and Galor, 2011). At about the same time, Hung Liang-Chi (1744–1809), a high official of the Chinese imperial administration, developed similar ideas. The ideas were particularly relevant for understanding Chinese dynamics in the 19th century. Still, Hung Liang-Chi largely disappeared from the record and was only rediscovered in the 20th century (Silberman, 1960). How can we understand the differences in the fates of these two ideas? The approach we develop in this paper can be applied to explain Malthus' success. Malthus was integrated into the broad European academic network, where his ideas could

[4]More generally, Koschnick (2025a) provides evidence, through automated text analysis, of knowledge spillovers between propositional and practical knowledge in England, 1600–1800.

spread. Hung Liang-Chi belonged to an administration in which ideas were developed by individuals but not subject to broad dissemination and discussion.

The paper is organized as follows. In Section 3.2 we present the methodology, including the construction of the database and how it is mapped into an affiliation network. We also present the epidemiological model to simulate the diffusion of ideas. In Section 3.3 we detail the structural estimation of the model parameter, based on *flora* and *cosmos*. In Section 3.4 we present further empirical assessments, based on *salvatio*. The counterfactual experiments are detailed in Section 3.5. Section 3.6 concludes.

## 3.2 Data and Methodology

We now present our methodology, starting with the compilation of the database of professors, followed by the definition of the temporal network and the epidemiological model that we use to describe how ideas flow.

### 3.2.1 88,792 scholars

Our database of scholars comprises information on 88,792 individuals spanning the period 1000-1800. The data were collected manually from approximately 700 distinct sources. Unlike other studies that rely on ex-post recognition of scholars—such as that derived from Wikipedia/Wikidata (see Laouenan et al. (2022) and Serafinelli and Tabellini (2022))—our selection is based on membership lists or secondary sources related to key higher education institutions. These institutions fall into three categories: universities (referenced in Frijhoff (1996); see also de la Croix et al. (2024)), scientific academies (as cataloged in McClellan (1985), and further discussed in Zanardello (2024)), and various other institutions with links to universities, including Italian Renaissance academies mentioned in The British Library (2021), and other higher education entities that conferred academic degrees.

Medieval universities primarily focused on four disciplines: theology, law, arts and humanities, and medicine. The faculty of arts provided foundational education to grammar school pupils, many of whom became teachers themselves, contributing to rising literacy rates among the general population. Some students progressed to higher faculties, preparing for professions in other fields. The faculty of medicine trained medical practitioners, the faculty of laws produced future administrators with specialized knowledge in canon

or civil law, and the faculty of theology trained teachers for episcopal schools, where ordinary parish priests were instructed (Pedersen, 1992). Academies, emerging later in the 17th and 18th centuries, were created to foster new areas of research not traditionally covered by universities (McClellan, 1985; Applebaum, 2003). These ranged from informal groups of amateur naturalists or local historians to prominent official societies that gathered leading scholars, published journals, and formed networks of corresponding members, known collectively as the Republic of Letters (Mokyr, 2016).

To compile the list of scholars from each academy and university, we mostly relied on secondary sources, mainly books on the history of these institutions and their members, which were themselves based on primary records. For universities, we aimed to include scholars involved in teaching, covering a range of positions from royal chairs in France to fellowships in England. Further details on the inclusion criteria for university scholars can be found in de la Croix et al. (2024), while global statistics are available in de la Croix (2021) and different issues of the *Repertorium Eruditorum Totius Europae*. For academies, this process was generally straightforward, since comprehensive membership lists are often available. Our data on academies have already been utilized in works such as Blasutto and de la Croix (2023) for Italian academies, de la Croix and Goñi (2024) for analyzing father-son pairs across academies and universities, and Zanardello (2024) for evaluating the impact of different fields of study within academies. Academies often have several membership categories, including ordinary, corresponding, and honorary members. Corresponding members, though not present at academy meetings, contributed from a distance. Honorary members often included local dignitaries like bishops, wealthy merchants, and governors, who supported and protected the academies. To prevent skewing our results due to the inclusion of these sometimes prominent figures, we excluded anyone holding honorary membership or those who were clearly not scholars or intellectuals (e.g., Napoleon, who was elected to the Académie des Sciences in 1797).

Appendix 3.A reports descriptive statistics. Among the 88,792 scholars, 26,549 lack any personal information (no place of birth, vital dates, or Wikipedia link), while another large group is known only by place of birth. Complete information on all three indicators exists for 17,200 scholars. Overall, 19,579 scholars, about one-fifth of the dataset, have a Wikipedia page. Hence, for academia, our dataset is much more comprehensive than

Wikipedia-based samples such as Laouenan et al. (2022). Around 80% of scholars are associated with a single field, most commonly the humanities, followed by the sciences (including medicine), theology, and law. For 6,939 scholars the field is unknown, and 3,744 are classified as "honorary," reflecting administrative or ceremonial roles and will be excluded from the analyses. Among multi-field scholars, the most common combinations are theology and humanities (6,707) and humanities and sciences (3,382). Start years of affiliation are known in most cases (90,416), typically corresponding to university appointments or elections to academies. End dates are less systematically recorded, as sources often report when a professorship begins but not when it ends. Academy membership can generally be assumed to last until death, whereas no comparable assumption can be made for university posts. For 12,965 scholars only an approximate activity date can be assigned, and for 1,162 no dating information is available. More details about activity dates in Appendix 3.B.

### 3.2.2 Definition of the affiliation network $\mathbb{G}$

We now look at the data on scholars and their affiliation to institutions with reference to network theory, a powerful tool for studying the spread of information over time and space (Jackson, 2008; Goyal, 2023). We model the affiliation network as a graph, where *nodes* represent scholars and *edges* denote their contemporaneous presence at the same institution. This network is derived from an initial bipartite representation, consisting of two types of nodes: scholars and institutions. In the bipartite version, edges connect scholars to the institutions they were affiliated with. Since our focus is on scholar-to-scholar interactions, we project this bipartite graph onto a single-mode network of scholars, where an edge between two nodes represents their concurrent affiliation with the same institution, as shown in Figure 3.1. Premodern universities were small, and it is reasonable to assume that professors had regular opportunities for interaction. Appendix 3.A presents some descriptive statistics for the 10 institutions with the highest number of scholars in 1793, differentiating by type (university vs academy). The average university has fewer than 70 professors. Academies hosted a higher number of scholars on average, reaching almost 80 in Halle and 123 in London.

Our analysis spans the foundation of a new school of jurisprudence, which would later become the University of Bologna, in $\underline{t} = 1084$, to the French

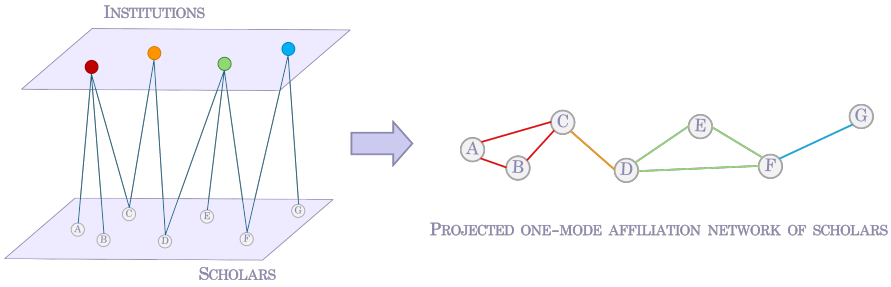



**Figure 3.1.** Intuitive representation of network projection: from a bipartite or two-mode graph to the one-mode affiliation network of scholars. Diagram inspired by Geraerts and Vasques Filho (2024).

Convention in $\bar{t} = 1793$, which led to the abrupt closure of all universities and academies on the territory of the new Republic. During this timeframe, the network's nodes (scholars) and edges (connections) existed only within specific periods defined by the duration of each scholar's activity and their affiliations with institutions.

More formally, given two scholars $i_s$ and $i_v$, the link between $i_s$ and $i_v$ lasts as long as $i_s$ and $i_v$ share an overlapping period of affiliation at the same institution. This implies that the collection of edges is dynamic over time: edges serve as channels for the spread of ideas, appearing and disappearing only while scholars are active at the same institution. In contrast, nodes (scholars) exist in the network as long as they are active, i.e., affiliated with one or more institutions as in our main database. Figure 3.2 shows the evolution of the number of active scholars over time, showing overall exponential growth, in particular after 1650, with the emergence of academies. Appendix 3.B gives more details on how activity periods are computed.

The affiliation network reveals several important characteristics, visualized in Figure 3.3. Before the rise of academies, the network at any given time typically consisted of a series of mostly disconnected clusters, each representing a single university. These clusters in turn comprised *cliques* (fully connected clusters) of scholars operating in the same field at the same university. Occasionally, cliques overlapped, highlighting scholars active in multiple fields. As scholars move, occasional links form among different university clusters, creating pathways between otherwise separate regions of the network. With the emergence of academies, however, connections between

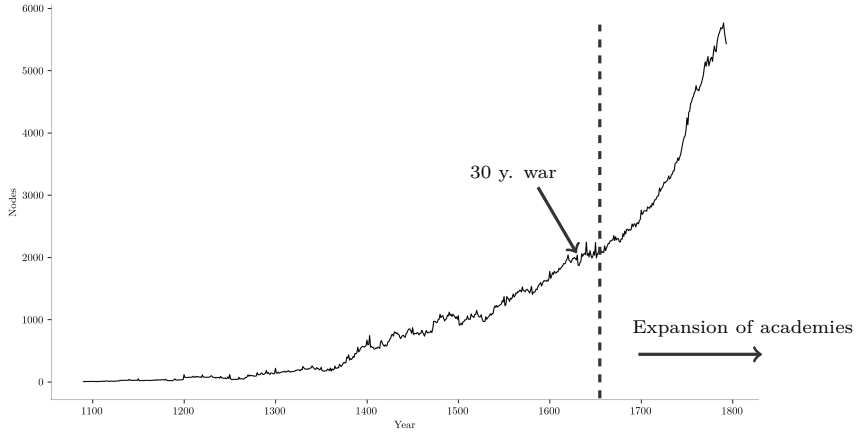
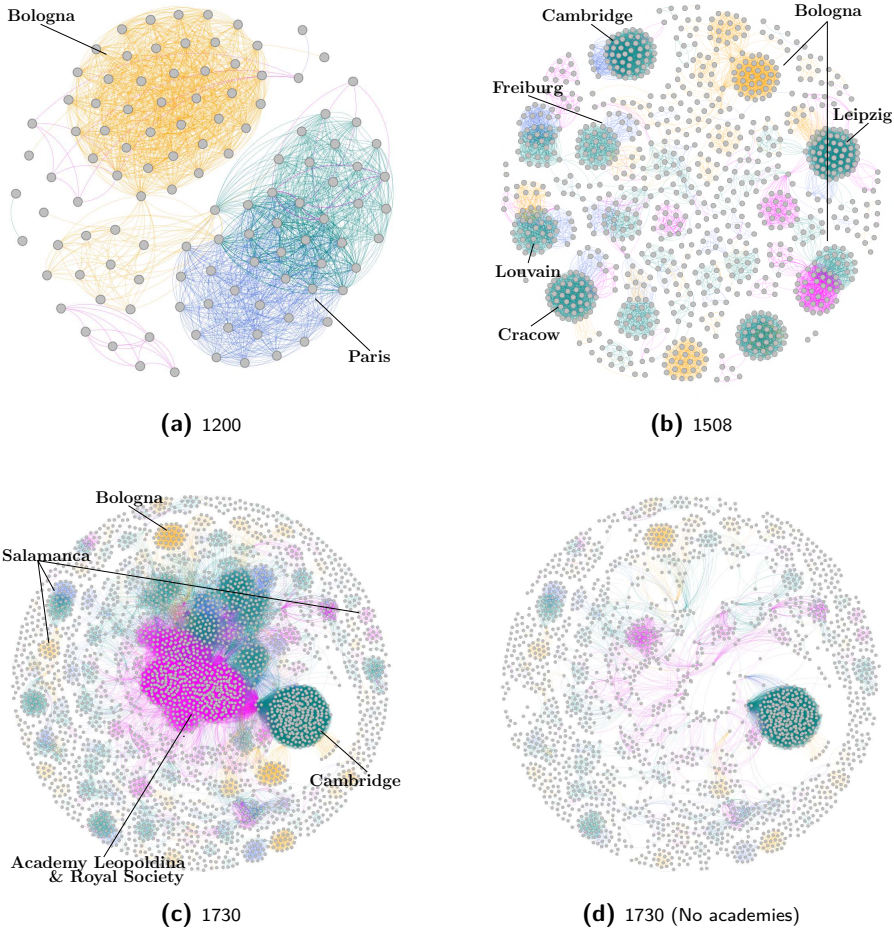



**Figure 3.2.** Number of active scholars in the network, 1084-1793.

scholars multiplied significantly. Academies often appointed foreign members, serving as bridges between previously isolated university clusters: Figure 3.3d depicts the 1730 network in a scenario where academies are removed. From the 1650s to the end of the timeframe, the cluster- and clique-based appearance of the early network transformed into a densely interconnected web, and this is likely due to the academies.

This effect is also visible in the share of nodes connected to the main network, known as the giant component—the largest connected subgraph in which any two nodes are mutually reachable—as shown in Figure 3.4. In the early period, around 1200, the affiliation network was nearly fully connected, reflecting a small number of institutions and high scholar mobility. From the 14th century onward, network connectedness followed a general downward trend, punctuated by occasional spikes. These spikes typically occurred when a mobile scholar bridged otherwise disconnected parts of the network. The size of the giant component reached a low point during the Reformation, which curtailed mobility across religious lines, as discussed in de la Croix and Morault (2026). Although the Thirty Years' War (1618–1648) did not drive a further decline, it prolonged the fragmentation, with the giant component remaining relatively small. Connectivity recovered with the rise of academies, as the share of nodes in the giant component increased from 40% in 1650 to 90% by 1793.

In Table 3.1, we report network statistics at the beginning of each century, between 1200 and 1700, following Goyal et al. (2006), to assess whether our




**(a)** 1200 **(b)** 1508

**(c)** 1730 **(d)** 1730 (No academies)

**Figure 3.3.** Snapshots of the affiliation network in 1200, 1508, and 1730 (with and without academies). Edge colors broadly denote the disciplines: theology (blue), law (orange), humanities (teal), and sciences (magenta). Isolated nodes not represented.

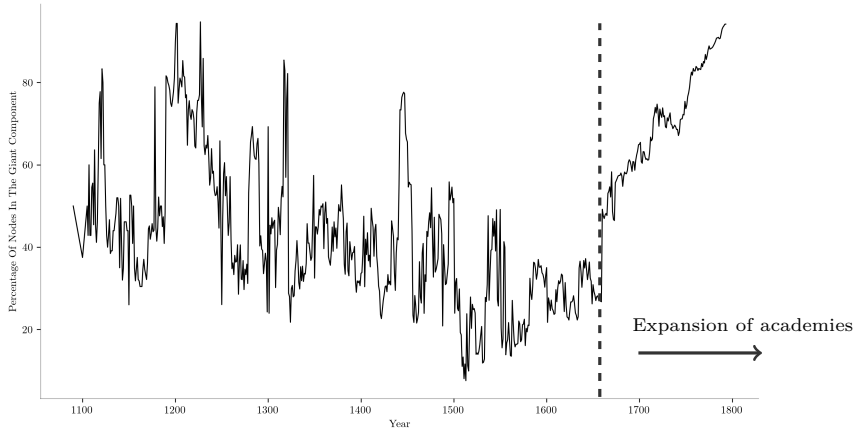



**Figure 3.4.** Percentage of active scholars in the giant or connected component over time, 1084-1793.

**Table 3.1.** Network statistics for the affiliation network, 1200-1700

| Year | 1200 | 1300 | 1400 | 1500 | 1600 | 1700 |
|---|---|---|---|---|---|---|
| Number of active scholars | 116 | 218 | 672 | 1065 | 1776 | 2757 |
| Avg. degree | 30.03 | 17.08 | 42.14 | 26.54 | 42.24 | 55.68 |
| Size of giant component | 104 | 151 | 226 | 552 | 622 | 1792 |
| % of giant component | 89.66 | 69.27 | 33.63 | 51.83 | 35.02 | 65 |
| Size of second-largest component | 6 | 29 | 139 | 75 | 244 | 51 |
| Avg. clustering coefficient | 0.88 | 0.92 | 0.93 | 0.91 | 0.89 | 0.90 |
| Avg. distance (giant) | 2.16 | 3.36 | 1.95 | 4.14 | 5.51 | 4.31 |

Note: The degree of a node is the number of distinct scholars with which the node is connected. The Avg. degree is the mean of the degrees of all connected scholars in the network. The clustering coefficient of a node is the fraction of all possible links among a node's neighbors that are actually present. The Avg. clustering coefficient is the mean of the clustering coefficients of all connected scholars in the network. The distance between two scholars the length of the shortest path between them. The Avg. distance is the mean distance among all connected pairs of scholars in the giant component.

network exhibits properties of an emerging small world. The graph consistently displays a high clustering coefficient (above 0.88 in all periods), well above the levels expected in a random graph with similar size and density, where clustering would typically approximate to the average degree divided by the number of nodes–and hence close to zero in sparse networks of this scale. At the same time, average path lengths (or distances) within the giant component remain low–between 2.16 and 5.51–and stable over time, even as the number of scholars increases significantly. While the network is highly fragmented in earlier centuries, with the giant component capturing as little as 33.63% of scholars in 1400, its expansion to over 60% by 1700 reflects a growing integration of academic communities across institutions. In other words, in the earlier centuries, the affiliation network resembles an archipelago of small "islands" (i.e., universities) which are internally well connected (as shown by the very high clustering coefficients) but connected to each other only via the occasional mobility of scholars. Over time, particularly by 1700, these isolated islands began to form bridges–thanks to multiple simultaneous affiliations, which were due to the emergence of academies. The result is an increasingly interconnected network, where the size of the giant component grows to nearly two-thirds of all scholars, while clustering remains high and path lengths stay short. The size of the second-largest component further underlines this aspect: it increases up to 1600 but drastically decreases in 1700, thanks also to the arrival of scientific academies. This shift marks the emergence of a small-world structure out of what was once a fragmented archipelago.

This affiliation network is the structure on which we will simulate the spread of ideas. Each idea is assumed to originate in a specific year and from a particular scholar, who can transmit it to their neighbors at each time step, but under certain conditions. Each idea belongs to a broad field that reflects the main disciplines of premodern times, and can spread only among scholars active within that field. Our assumption is that if a scholar is working in science[5] they are presumed to engage productively with peers in medicine or applied science.[6] One may argue that in the past there were all-round scholars who were equally fluent, for example, in both science and philosophy. Still, we decide to assign specific fields (albeit broadly defined rather than finely specialized) to

[5]Science includes mathematics, logic, physics, chemistry, biology, astronomy, earth sciences, geography, and botany.

[6]Applied science includes engineering, architecture, and agronomy.

each idea. We acknowledge that attributing an intellectual tradition to a single originator, on the one hand, and assuming ideas circulate within fields, on the other, are deliberate simplifications of a more nuanced and complex reality. They are, however, necessary given the constraints of our data and model.

### 3.2.3 Ideas, inventors and exposed scholars

Ideas originate from inventors, i.e., scholars identified in the historical literature as the principal proponents of a new idea or paradigm at a specific point in time. As already noted above, this assumption compresses what was in reality a collective process, but it provides the diffusion model with a well-defined initial condition. Inventors can propagate their ideas to their peers, who, once exposed, can further propagate them to others. Upon being exposed to the idea, scholars can pass it along to their own neighbors without needing to maintain an enduring direct link with the original inventor. We identify inventors from two major historical encyclopedias: English (2005) for ideas spread before 1500 C.E., and Applebaum (2003) for ideas diffused between the invention of the printing press (circa 1450s) and the French Revolution. From these sources, we identify some significant ideas that changed the course of history, prioritizing those for which historical outcomes are available. For each idea, we pinpoint the inventor, as detailed in Section 3.3.2, Section 3.3.3, and Section 3.4.1.

We simulate the spread of three main ideas—*flora*, *cosmos*, and *salvatio*—two from the Scientific Revolution and one from the Middle Ages, and compare simulated exposure to observed outcomes. For the Scientific Revolution, we draw on Applebaum (2003), focusing on the earliest ideas for which European-level outcome data is available. In botany, Applebaum (2003) identifies Botanical Realism, foundational to Fuchs's 1542 herbarium; we link exposure to this idea with the creation of botanic gardens. In astronomy, the idea is attributed to Regiomontanus, whose Mathematical Astronomy (developed from 1450, published in 1496) formalized Ptolemaic models for future research. Its corresponding outcome is the founding of astronomical observatories.

For the Middle Ages, we find key academic ideas using the index in English (2005). These include alchemy, anatomy (including practical surgery), astrology, computus, civil law, economic thought, cartography, humanism, music, optics, political theory, punctuation, and the Scholastic method. To balance the two scientific ideas above, we select one from a different domain:

theology. Scholasticism, rooted in the work of Lombardus, emphasized rigorous logic and dialectical reasoning to reconcile faith with reason. It fostered a systematic approach to inquiry that shaped both scientific and philosophical thought. At some point, it coincided with a theological backlash, especially from movements like Protestantism that emphasized scriptural authority over rational deduction. The associated outcome in our empirical analysis is the probability of a city becoming Protestant.

A key challenge, common to all three ideas, is determining when the scholar first developed the concept. To define this moment, we try to establish two dates for each idea: (i) the publication date, which refers to when the scholar first published a work on the topic, and (ii) the inception date, which is the year when the scholar first conceived the idea and likely began discussing it with colleagues. We identify the inception date manually, by reviewing biographical information and related historical context.

In our model, ideas spread via interactions among scholars, and these interactions can only occur when scholars are alive. Therefore, using the inception date rather than the publication date allows us to better capture the dynamics of idea dissemination through the scholarly network, as it reflects the period when discussions and exchanges of the idea were possible. Our preferred date is therefore the inception date, though we rely on the publication date when the inception date is unavailable.

Overall, ideas are more likely to spread if certain conditions are met: (a) scholars have long lifespans, giving them more time to spread their ideas; (b) there is a high density of scholars at a given institution, creating more opportunities for intellectual exchange, and (c) scholars move between institutions, which facilitates the dissemination of ideas across different scholarly communities. However, the spread of each idea may vary significantly depending on specific factors, including the centrality of the inventor within their peer network, their affiliations with large institutions, and the timing of the idea's inception.

To briefly clarify some terminology: we use inventor as shorthand for the main proponent or originator of an idea—not necessarily an inventor in the traditional sense, but often someone who developed, articulated, or popularized a concept. Similarly, idea is used broadly to encompass various intellectual contributions, including theses, paradigms, and methodological approaches, which differ in scope, complexity, and impact.

### 3.2.4 Epidemiological model

In our framework, the random nature of the spread of ideas within institutions is essential. The process mirrors infectious disease contagion: contact between colleagues creates the probability of transmission, but not its certainty. We hence adopt an epidemiological approach to model the spread of ideas from a scholar to another (Fogli and Veldkamp, 2021; Banerjee et al., 2013). There is a fixed number of nodes, $N$, each representing a scholar. Time is discrete, with $t \in \{\underline{t}, \ldots, \bar{t}\}$, $\underline{t}$ and $\bar{t}$ being the start and end dates of our analysis. At each date, a node can be susceptible or infectious.[7] A contact between two nodes appears as an undirected link in the network at a given time. Interactions are represented by an adjacency matrix $A_t = [a_{sv}]_t$ of dimension $N \times N$, with each element $a_{sv}$ taking value 1 if scholar $s$ and $v$ are connected at time $t$, and zero otherwise. Connections will depend on whether $s$ and $v$ are working in the same field at the same time in the same institution (more on this later). We represent a temporal network $\mathbb{G}$ by a set of adjacency matrices $A_t$.

The state of the world is described at each date by a vector $I_t = [i_s]_t$ of length $N$. We only have binary entries in $I_t$, with $i_s = 1$ if scholar $s$ is infected, and $i_s = 0$ otherwise. Initially, there is no idea, and nobody is infected. At some date $t_0$, an initial "inventor" has an idea. We thus have $[i_s]_t = 0$ for all $t < t_0$, and $[i_{s^\star}]_{t_0} = 1$, where $s^\star$ is the inventor.

Following the binary nature of the state vector, we use Boolean arithmetic, i.e., element-wise addition and scalar multiplication are replaced by the logical "or" and "and", respectively (Koher et al., 2016). Dynamics are then represented by:

$$I_{t+1} = A_t I_t + I_t \tag{3.1}$$

To understand this formula, consider the $s$ scholar. If they are alive at period $t$, their infection status at $t+1$ is given by $\sum_v a_{sv} i_v$. With Boolean arithmetics, this term is equal to 1 if there is at least one $v$ such that $a_{sv} = 1$ ($s$ has met $v$) and $i_v = 1$ ($v$ is infected). If, instead, $s$ is either unborn or dead at time $t$, $a_{sv} = 0\ \forall v$, and their infection status does not change.[8]

[7] Here, our model closely relates to Banerjee et al. (2013), since infection does not equate adoption of the idea.

[8] While the model could in principle track how many times $s$ has been exposed to infected neighbors, as suggested by Bramoullé and Genicot (2024), we adopt a simplified binary-state process: infection occurs upon the first effective contact. Subsequent contacts with infected peers do not accumulate and have no further effect on the "intensity" of infection.

## 3. Pre-modern academic institutions as conduits for the potential spread of ideas

We also assume that once exposed to an idea, a scholar cannot forget it. Hence the "recovered" state of the epidemiological model is not relevant here.

So far we have assumed that ideas are transmitted upon contact with probability 1. If instead, there is a link activation probability $\alpha \in [0,1]$,[9] we define a stochastic operator $\Omega^d(A)$ (following Koher et al. (2016)) which acts element-wise on the adjacency matrix: for $a_{sv} = 0$, we have $\Omega^d(a_{sv}) = 0$; for $a_{sv} = 1$, we have $\Omega^d(a_{sv}) = 1$ with probability $\alpha$ and $\Omega^d(a_{sv}) = 0$ with probability $1 - \alpha$. Each potential transmission is evaluated independently on each edge: a susceptible scholar becomes infected through contact with an infected neighbor with probability $\alpha$, based on an independent draw.

Dynamics of the state vector $I$ are now represented by:

$$I^d_{t+1} = \Omega^d(A_t) I^d_t + I^d_t \tag{3.2}$$

where $d$ is an index of simulations (draws). Since each 1 in $A_t$ independently survives with probability $\alpha$, the expected value of the stochastic contact matrices is:

$$\mathbb{E}[\Omega^d(A_t)] = \alpha A_t.$$

Such a specification increases the computational effort and allows for interactions between topological effects (those coming from the structure of the network) and probabilistic effects.

We now define levels of exposure. These levels are expected levels, given the stochastic nature of the simulations. The expected scholar $s$ exposure $[\bar{i}_s]_t \in [0,1]$ to a given idea is obtained by averaging individual exposure over $D$ simulations:

$$[\bar{i}_s]_t = \frac{1}{D} \sum_{d=1}^{D} [i^d_s]_t \tag{3.3}$$

Next, let us define the set of members of institution $k$ at time $t$ in field $\varphi$ as $V(k,t,\varphi)$. Institution $k$ Eligible Mass in field $\varphi$ is:

$$M^{k,\varphi}_t \;=\; \#V(k,t,\varphi). \tag{3.4}$$

It is the number of scholars at each point in time susceptible to receiving an idea in field $\varphi$ at institution $k$.

[9] Rather than assuming automatic transmission, we model a probabilistic approach to idea diffusion, reflecting plausible uncertainty and selectivity observed in historical intellectual exchanges–a logic similar to the stochastic imitation dynamics in Brunt and García-Peñalosa (2022).

Institution $k$ Prevalence is then obtained as an average individual exposure over the set of individuals $s$ belonging to $V(k,t,\varphi)$, at time $t$:

$$P_t^{k,\varphi} = \left( \sum_{s \in V(k,t,\varphi)} [\bar{i}_s]_t \right) / M_t^{k,\varphi} \tag{3.5}$$

Eligible Mass and Prevalence will be used in the proportional hazard models of Sections 3.3.2 and 3.3.3 to assess how these measures at a certain date are correlated with the emergence of botanic gardens or observatories.

Furthermore, by multiplying Eligible Mass by Prevalence, we obtain institution $k$ exposure as

$$S_t^{k,\varphi} = M_t^{k,\varphi} \times P_t^{k,\varphi} \tag{3.6}$$

So far, the above measures only concern cities that are the headquarters of an institution. To define a measure of exposure for all cities in Europe, we assume that headquarters' exposure affects nearby cities. We accordingly define expected city $c$ exposure $S_t^{c,\varphi} \geq 0$ by averaging over nearby institutions, weighting by inverse distance:

$$S_t^{c,\varphi} = \sum_k w_{ck} S_t^{k,\varphi} \tag{3.7}$$

The weights $w_{ck}$ are derived from the inverse distance between all the institutions in our database and the cities in our samples (precise details about the cities data are provided in each experiment). Considering the inverse distance means that the further a city is from an exposed institution, the lower the influence that reaches the urban center. An institution fully influences cities within 10 kilometers: essentially the city hosting that institution. Beyond 10 kilometers, the influence power decreases linearly, up to 1000 kilometers. After this threshold, we assume that the institution's influence loses all its power, reaching a weight of zero, and thus it cannot influence any city beyond 1000 kilometers.[10]

[10]The results are robust to alternative functional forms, such as exponential decay or varying distance thresholds.

## 3.3 Structural estimation

### 3.3.1 Methodology

How fast ideas spread in the affiliation network depends crucially on the link activation probability $\alpha$. Estimating this probability is difficult because we do not observe the spread of ideas directly; we only observe some outcomes of the ideas, after some time. These outcomes are related to exposure to ideas through statistical models, described below. We will use two of these outcomes to estimate $\alpha$. This setup resembles indirect inference (Smith, 2008), where one chooses $\alpha$ to maximize the likelihood of the observed data, given the structural model's output. In practical terms, we take the following steps.

1. We fix $\alpha$ at gridded values over $[0, 1]$ at intervals of 0.05.
2. For each value of $\alpha$, we simulate the spread of two key ideas through the affiliation network and the epidemiological model.
3. We compute Eligible Mass and Prevalence of cities with a university or academy, given by Equations (3.4) and (3.5), for these two ideas.
4. We estimate two auxiliary statistical models correlating the simulated exposure with observed outcomes (detailed below).
5. We record the log-likelihoods as the sum of the individual likelihoods at each point:
$$\ell_{\text{total}}(\alpha) = \ell_1(\alpha) + \ell_2(\alpha)$$
This approach treats the two outcomes as conditionally independent given $\alpha$.
6. We maximize the combined log-likelihood
$$\hat{\alpha} = \arg\max_{\alpha} \ell_{\text{total}}(\alpha).$$

We now describe the two auxiliary models referred to above in item 4.

### 3.3.2 Auxiliary model 1: Botanical Realism

During the Scientific Revolution, there were major advancements in botany, which grew from being primarily a descriptive field into a more systematic and experimental science—a shift we call Botanical Realism. A key figure in this transition was Leonhart Fuchs, a German physician and botanist. He is best known for his book *De historia stirpium commentarii insignes* (Notable

commentaries on the history of plants). Printed in Basel in 1542, this work laid the foundation for modern botany. Fuchs not only provided visual representations of 511 plant species, but he also included his own critical observations on their uses and characteristics, highlighting differences from ancient texts (Applebaum, 2003). More contextual information is available in Appendix 3.C.1. Fuchs was based in Tübingen for the majority of his life, where he taught medicine and botany at the local university between 1535 and 1566 (Conrad, 1960). Prior to that, he was a professor at the University of Ingolstadt from 1522 to 1533 (Schwinges and Hesse, 2019). Despite his fame, he was not a mobile scholar, and he declined prestigious teaching offers from Denmark and Italy (Applebaum, 2003).

In this first empirical assessment, we examine the potential correlation between Prevalence $P_t^{k,\varphi}$ of Botanical Realism and the establishment of botanic gardens. The founding of a botanic garden is admittedly an imperfect proxy for the spread and adoption of Botanical Realism, as it required material resources and a willingness to invest that we cannot directly observe. It is, however, the best observable outcome available at the European scale, and one for which the link to intellectual exposure is plausible: scholars interested in a more empirical approach to botany likely fostered favorable conditions for the establishment of gardens.

Using our epidemiological approach with link activation probability $\alpha$, the diffusion of the idea begins with Leonhart Fuchs and spreads to his colleagues at the University of Tübingen from 1542 onwards. Exposed scholars can in turn contaminate their own peers with probability $\alpha$, and so the idea propagates through the network until 1793 (the end of our timeframe). For each value of $\alpha$ on the grid, we average over $D = 5,000$ simulations, and we compute for each headquarter $k$ the Prevalence $P_t^{k,\varphi}$ of Botanical Realism as defined in Equation (3.5).

We use the sample of cities that hosted at least one institution, either a university or an academy before 1800, as recorded in our database, resulting in a total of 241 cities, 230 of which have non-null Eligible Mass $M_t^{k,\varphi}$ of scholars susceptible to Botanical Realism. We also gather information on the existence and founding dates of European botanic gardens from Montreal Botanic Garden, (1886). Figure 3.D.1 in Appendix 3.D.2 illustrates this sample of cities along with their Prevalence $P_t^{k,\varphi}$ of Botanical Realism in 1600, 1700, and 1800.

In what follows, we analyze whether exposure to Botanical Realism

accelerates the founding of botanic gardens across cities, using the timing of first establishment as our outcome. Garden creation is a one-time irreversible event: the relevant question is not whether a city eventually gets a garden but when. A Cox proportional hazards model is therefore natural, as it restricts attention at each point in time to cities still at risk, and accommodates time-varying covariates such as our measures of Prevalence $P_t^{k,\varphi}$ and Eligible Mass $M_t^{k,\varphi}$.[11]

Following the Cox model, the instantaneous hazard $h_k(t)$ that a city $k$ founds its first botanic garden at time $t$, conditional on covariates, is

$$h_k(t) = h_0(t) \exp(\beta_p P_t^{k,\varphi} + \beta_m M_t^{k,\varphi} + \beta_{pm} P_t^{k,\varphi} M_t^{k,\varphi} + \beta_{pop} Pop^k + \beta_d Dist^k) \tag{3.8}$$

where $h_0(t)$ is the baseline hazard. We include time-invariant regressors such as population in 1500, $Pop^k$, and the distance from Tübingen, $Dist^k$. In a simple gravity model of diffusion, distance from Tübingen captures the spread of the idea through pathways other than our affiliation network. Our regressors of interest are time-varying and include the Eligible Mass of scholars, $M_t^{k,\varphi}$, that counts the number of botanists, physicians, and scientists susceptible to be exposed to Botanical Realism at time $t$ in city $k$,[12] the Prevalence of Botanical Realism, $P_t^{k,\varphi}$ (see Equation (3.5)), and their interaction term. The interaction allows for heterogeneous effects by institutional size: a negative sign implies that the smaller the Eligible Mass, the stronger the association between Prevalence and the 'hazard' of founding a botanic garden; a positive sign implies the opposite.

We consider the period 1500–1793. The first botanic garden in our data is observed in 1520 (Pavia). Fuchs' invention dates to 1542, and our affiliation data runs through 1793. Cities that do not build a botanic garden during this period are treated as right-censored.[13] We estimate the baseline hazard using the Nelson–Aalen estimator. The Cox model's partial likelihood assumes no two events occur simultaneously. In our data, nine years record the founding of

[11] Precisely because timing is the object of interest, a logit, which collapses the outcome to a binary, or a panel regression, which cannot cleanly handle exit from the risk set, are ill-suited.

[12] These scholars work in fields such as medicine, botany, mathematics, physics, chemistry, astronomy, and applied sciences such as agronomy and engineering.

[13] Because botanic gardens are established in only 68 out of 230 cities before 1793, we do not stratify the baseline hazard by city. Instead, we cluster standard errors at the city level.

two botanic gardens, introducing ties that we handle using the Efron method, confirming robustness with the exact method.

**Table 3.2.** Cox Proportional Hazards Model – Botanical Realism and botanic gardens

| | Dependent variable: Hazard of botanic garden founding | | | | | |
|---|---|---|---|---|---|---|
| $\alpha$ | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 1 |
| Prevalence $P_t^{k,\varphi}$ | | 27.91 | 7.45*** | 2.96*** | 1.93*** | 1.67*** |
| | | (156.66) | (2.16) | (1.00) | (0.63) | (0.52) |
| (ihs) Eligible Mass $M_t^{k,\varphi}$ | 0.54*** | 0.57*** | 0.88*** | 0.90*** | 0.89*** | 0.89*** |
| | (0.07) | (0.19) | (0.10) | (0.10) | (0.10) | (0.10) |
| $P_t^{k,\varphi}$ | | -26.69 | -4.02*** | -1.38*** | -0.89*** | -0.75*** |
| $\times$ (ihs) $M_t^{k,\varphi}$ | | (177.85) | (0.77) | (0.28) | (0.18) | (0.15) |
| (ihs) Distance to | -0.15*** | -0.19*** | -0.12 | -0.11 | -0.11* | -0.11* |
| Tübingen | (0.06) | (0.07) | (0.10) | (0.07) | (0.07) | (0.06) |
| Log Likelihood $\ell_1(\alpha)$ | -343.18 | -342.51 | -335.76 | -335.28 | -335.32 | -335.28 |
| (ihs) Pop in 1500 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 67620 | 67620 | 67620 | 67620 | 67620 | 67620 |

*Notes*: *p<0.1; **p< 0.05; ***p<0.01. Robust standard errors are reported in parentheses. Similar models were run for every $\alpha$ between 0 and 1 at an interval of 0.05. Here, we only report results for 0, 0.1, 0.3, 0.5, 0.7, and 1.
Distance to Tübingen is computed using the Haversine formula. All models include the (ihs) population in 1500. All the variables with (ihs) are transformed to the inverse hyperbolic sine.

Table 3.2 shows the results for the Cox proportional hazard models at different levels of $\alpha$. We run similar specifications for every $\alpha$ between 0 and 1 at an interval of 0.05 to obtain the log likelihoods which enter the joint maximization described in Section 3.3.

### 3.3.3 Auxiliary model 2: Mathematical Astronomy

In the 15th and 16th centuries, growing interest in experimental science led astronomers to challenge Ptolemaic models and refine them through observation and mathematics. This shift, characterized by advances in trigonometry, geometry, and the use of decimals, laid the groundwork for the later Copernican revolution. A central figure in this transformation was Regiomontanus, a pseudonym of Johannes Müller. He studied the original works of Ptolemy through his mastery of Greek and mathematics, and around 1454 began collaborating with his mentor Georg Peurbach (1423–1461) at the University of Vienna on new methods for solving plane and spherical

trigonometry problems, including the use of sine and tangent functions. Regiomontanus also created extensive trigonometric tables with values calculated to decimal units, which remained influential for centuries, and published *Theoricæ novæ planetarum*, his collaboration with Peurbach, in 1472, after Peurbach's death. He can thus be considered a pioneer of Mathematical Astronomy (Applebaum, 2003). More details are available in Appendix 3.C.2.

In this second empirical assessment, we examine whether cities more exposed to what we label Mathematical Astronomy—the transition towards higher mathematical precision in astronomical practice—were faster to establish astronomical observatories. Advances in trigonometry and decimal calculation created demand for more precise instruments and dedicated observation sites, providing a direct link between the diffusion of Mathematical Astronomy and the founding of observatories. Regiomontanus himself opened an instrument shop that specialized in building and printing works related to Mathematical Astronomy (Applebaum, 2003). We collect the names and foundation dates of observatories from Howse (1986).

For each value of $\alpha$ on the grid, the diffusion of Mathematical Astronomy begins with Regiomontanus, who shared it with his colleagues in Vienna, Bratislava, Padua, and Rome. Exposed scholars can in turn contaminate their own peers with probability $\alpha$, and so the idea propagates through the network until 1793. Averaging over $D$ simulations, we compute for each institution $k$ the yearly Prevalence $P_t^{k,\varphi}$ of Mathematical Astronomy as defined in Equation (3.5), restricting to scholars working in fields such as mathematics, logic, physics, chemistry, biology, astronomy, geography, and botany. Figure 3.D.2 in Appendix 3.D.3 depicts the resulting exposure patterns in 1600, 1700, and 1793.

We estimate the hazard of establishing an observatory for each city using a Cox model identical to that used for Botanical Realism, replacing distance from Tübingen with distance from Vienna. We focus on the same period, 1500–1793, dating the spread of Regiomontanus' ideas from 1454. The first observatory in our data was established in Kassel in 1560. Cities that do not establish an observatory by 1793 are treated as right-censored.[14]

Table 3.3 shows the results for the Cox proportional hazard models at different levels of $\alpha$. We run similar specifications for every $\alpha$ between 0 and

[14] As with Botanical Realism, tied events are handled using the Efron method, with the exact method as a robustness check.

**Table 3.3.** Cox Proportional Hazards Model – Mathematical Astronomy and observatories

| | Dep. var.: Hazard of astronomical observatory creation | | | | | |
|---|---|---|---|---|---|---|
| $\alpha$ | 0 | 0.1 | 0.3 | 0.5 | 0.7 | 1 |
| Prevalence $P_t^{k,\varphi}$ | | 6.39*** | 2.49*** | 1.83*** | 1.66*** | 1.62*** |
| | | (1.94) | (0.73) | (0.53) | (0.48) | (0.45) |
| (ihs) Eligible Mass $M_t^{k,\varphi}$ | 0.46*** | 1.11*** | 1.11*** | 1.12*** | 1.11*** | 1.06*** |
| | (0.07) | (0.18) | (0.19) | (0.19) | (0.20) | (0.20) |
| $P_t^{k,\varphi}$ | | -3.69*** | -1.44*** | -1.08*** | -0.97*** | -0.89*** |
| $\times$ (ihs) $M_t^{k,\varphi}$ | | (0.77) | (0.32) | (0.24) | (0.22) | (0.21) |
| (ihs) Distance | -0.16*** | -0.08* | -0.09** | -0.10** | -0.10** | -0.10** |
| to Vienna | (0.05) | (0.05) | (0.04) | (0.04) | (0.04) | (0.04) |
| Log Likelihood $\ell_2(\alpha)$ | -324.83 | -316.08 | -317.22 | -317.31 | -317.49 | -318.04 |
| (ihs) Pop in 1500 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 67620 | 67620 | 67620 | 67620 | 67620 | 67620 |

*Notes*: *p<0.1; **p< 0.05; ***p<0.01. Robust standard errors are reported in parentheses. Similar models were run for every $\alpha$ between 0 and 1 at an interval of 0.05. Here, we only report results for 0, 0.1, 0.3, 0.5, 0.7, and 1.
Distance to Vienna is computed using the Haversine formula. All models include (ihs) population in 1500. All the variables with (ihs) are transformed into inverse hyperbolic sine.

1 at an interval of 0.05 to obtain the log likelihoods which enter the joint maximization described in Section 3.3.

### 3.3.4 Results

The joint likelihood $\ell_1(\alpha) + \ell_2(\alpha)$ achieves a maximum of $-652.568$ at $\hat{\alpha} = 0.45$, which we interpret as the estimated link activation probability. Appendix 3.E reports the full likelihood table, the maximization plot, and the Cox model results at $\hat{\alpha} = 0.45$. In the remainder of the paper, we therefore hold $\alpha$ fixed at this estimated value. An important open question is whether $\alpha$ varies systematically across fields, periods, religions, institutional types, or other dimensions. Unfortunately, the available data do not permit a reliable estimation of such variation. A related concern is that, because the auxiliary outcomes reflect not only intellectual exposure but also the capacity and willingness of institutions to act on it, $\hat{\alpha}$ may compound two distinct probabilities: that an interaction successfully transmits an idea, and that cumulatively, exposure translates into an observable institutional outcome. Disentangling the two would require an outcome that captures exposure

independently of local material conditions at any given point in time: a measure that, to our knowledge, is not systematically recoverable for early modern European academia. Alternative proxies that might offer a more direct link to, say, *individual* (rather than institutional) exposure with comparable coverage are equally unavailable. In a contemporary setting, for instance, citations to Fuchs or Regiomontanus could serve this purpose; in early modern Europe, however, not only were citation practices not standardized, but such a measure would in any case capture only the fraction of scholars who left a footprint in today's libraries (estimated, for scholars in our database, at roughly one third), not the broader academic population through which ideas circulated.

Before interpreting the results, we verify the proportional hazards assumption—which requires that the effect of each covariate on the hazard remains constant over time—using scaled Schoenfeld residuals (Schoenfeld, 1982). For Botanical Realism, the coefficients on Prevalence $P_t^{k,\varphi}$, Eligible Mass $M_t^{k,\varphi}$, and their interaction show no evidence of drift over time when tested individually. The global test (including population and distance to Tübingen) yields a p-value of 0.075, indicating borderline evidence against proportionality at the 10% level. Inspection of the Schoenfeld residual plots in Appendix 3.F reveals no systematic trend, suggesting the borderline p-value reflects noise rather than a genuine violation. For Mathematical Astronomy, both individual and joint tests are comfortably non-significant (global p-value = 0.189), and residual plots confirm no drift. We proceed on this basis for both models.

We interpret results through hazard ratios, computed by exponentiating the estimated coefficients. For Botanical Realism, the association between Prevalence $P_t^{k,\varphi}$ and the hazard of founding a botanic garden is not constant but varies with Eligible Mass $M_t^{k,\varphi}$—and vice versa. Specifically, with coefficients $\beta_p = 3.61$, $\beta_m = 0.90$, and interaction term $\beta_{pm} = -1.70$, the hazard ratio for a change $\Delta P_t^{k,\varphi}$ in Prevalence at a given level of Eligible Mass is:

$$HR(\Delta P_t^{k,\varphi} | M_t^{k,\varphi}) = \exp((\beta_p + \beta_{pm} \times M_t^{k,\varphi}) \Delta P_t^{k,\varphi}) \tag{3.9}$$

A 10 percentage point increase in Prevalence is associated with a 23% higher hazard of founding a botanic garden at average Eligible Mass ($\mu = 0.93$),[15] and a 43% higher hazard when Eligible Mass is zero. At high Eligible Mass

[15] See summary statistics of the covariates in Table 3.D.1 in the Appendix.

the association attenuates, with the $HR$ approaching 1 once Eligible Mass exceeds approximately 2.1—about one standard deviation above the mean. Turning to Eligible Mass, at median Prevalence (zero), doubling Eligible Mass raises the hazard by 86% ($HR = 1.86$); at mean Prevalence (0.084), the same doubling raises it by 69% ($HR = 1.69$). The negative interaction captures a substitutability between size and exposure: where the susceptible pool is small, a given increase in Prevalence is more decisive; where it is large, the marginal contribution of Prevalence diminishes. These patterns are robust across all values of $\alpha \geq 0.3$.

The pattern of Mathematical Astronomy results mirrors that of Botanical Realism. With $\beta_p = 1.93$ ($\mu = 0.12$, $SD = 0.28$), $\beta_m = 1.12$ ($\mu = 0.62$, $SD = 1.06$), and $\beta_{pm} = -1.13$, a 10 percentage point increase in Prevalence is associated with a 13% higher hazard of establishing an observatory at average Eligible Mass, and a 21% higher hazard when Eligible Mass is zero. As with Botanical Realism, the interaction is negative: size amplifies diffusion when exposure is low but adds little once Prevalence is already high.

## 3.4 Further empirical assessments

In the analysis of Botanical Realism and Mathematical Astronomy, we examined both the correlation between the Prevalence (Equation (3.5)) of these ideas and the Eligible Mass (Equation (3.4)) of scholars, and the time required to establish a botanic garden or an observatory at the headquarters of institution $k$.

We now extend our empirical analysis beyond the early modern period and the context of the Scientific Revolution. As in the previous sections, the scope of the analysis is shaped by the availability of outcome data. We examine three distinct cases: a backlash against an idea in the context of the relationship between scholasticism and Protestantism; an extension of the plague–pogrom nexus that incorporates exposure to anti-Judaic ideas; and the diffusion of a demonstrably false belief, namely that the Swedes are descendants of the inhabitants of Atlantis.

These experiments focus on outcomes at the level of European cities. For cities without academic scholars, Prevalence and Eligible Mass are not meaningful concepts. Nevertheless, such cities may still be exposed to intellectual influences originating from nearby institutions. For this reason, the following analyses relate city-level exposure (Equation (3.7)) to city-level

outcomes.

### 3.4.1 Scholasticism and Protestantism

Scholastic theology is an approach to theological questions that uses logical analysis and systematic reasoning, influenced by ancient Greek philosophers (see an example in Appendix 3.G). It is more a paradigm than a single idea. Petrus Lombardus is often recognized as an early proponent and influential figure in the scholastic tradition. According to Genet (2019) and Herbermann (1913), he taught at what would become the University of Paris from 1145 to his death in 1160. Mazzetti (1847) claims that he was at the University of Bologna in about 1150.

Petrus Lombardus' primary work is the *Sentences*. Completed in the mid-12th century, the *Sentences* cover key theological topics such as the nature of God, creation, the Trinity, grace, and sacraments. The *Sentences* became a foundational text for theological education in medieval universities and was the starting point for many scholastic theologians who followed, including Thomas Aquinas, who wrote extensive commentaries on it.

Martin Luther (1483–1546), the 16$^{\text{th}}$-century German monk and theologian who sparked the Protestant Reformation, was initially trained in the scholastic tradition and engaged with its methods. But as his personal spiritual crisis deepened, he became increasingly critical of many aspects of the Catholic Church's theology—including scholasticism's emphasis on human reason. Luther laid out his objections in a striking document, the *Disputatio contra scholasticam theologiam* (1517), a series of 97 theses. In it, he made provocative claims such as: "No syllogistic form is valid when applied to divine terms," and "... the whole Aristotle is to theology as darkness is to light" (theses 47 and 50, respectively).

A major grievance that fueled the rise of Protestantism was the desire to reform theological teachings and Church practices that, in the Reformers' view, were not grounded in Scripture. Scholastic theology—especially in its later form known as nominalism—was a central target. The historian Chaunu (2014) argues that this style of theology, which emphasized logic and abstraction, distanced ordinary believers from their faith and ultimately left them receptive to the message of the Reformation. From this we draw our hypothesis that the adoption of Protestantism was a backlash to (exposure to) scholasticism. This view is rarely made explicit in the scholarly literature,

but it underlies much of the Reformation's intellectual context.[16] To be clear, our hypothesis does not imply that scholastic reasoning was abandoned after the Reformation; on the contrary, Protestant theologians themselves later developed systematic theological frameworks that drew on scholastic method (Muller, 1986). The backlash we posit concerns rather the popular reception of late medieval scholastic theology as institutionally practiced.

We highlight three key ways in which Luther offered a clear and spiritually compelling response to the crisis of salvation induced by scholastic theology—each emphasizing *direct access to God* by *faith* rather than through a rational merit-based system (Chaunu, 2014).

(A) The Catholic Church taught that salvation comes through both faith and works, a position formalized in scholastic doctrines of grace and merit. Luther rejected this, arguing that Scripture teaches that salvation comes by faith alone (*sola fide*), not through human effort or achievement.

(B) Catholic theology placed Scripture and Church tradition—along with papal authority—on equal footing. This framework relied on a scholastic synthesis of Aristotelian philosophy and ecclesiastical tradition. Luther broke with this, insisting that Scripture alone (*sola scriptura*) is the final authority in matters of faith. He viewed the scholastic approach as placing human reason above divine revelation.

(C) The Catholic Church mediated grace through a complex sacramental system, including the sale of indulgences and a strict divide between clergy and laity. Luther opposed this mediation of grace and denied the special status of clergy rooted in scholastic definitions of ordination and apostolic succession. Instead, he emphasized the "priesthood of all believers" (*sola gratia*).

To test for a positive correlation between exposure to scholasticism and the rise of Protestantism, we propose the following experiment. We feed the idea of scholasticism to Petrus Lombardus. The idea would spread to his colleagues in Paris, and then beyond, thanks to the mobility of scholars. We simulate the spread of this idea in Europe, using the epidemiological approach described in Section 3.2.4, all the yearly network snapshots built from our data, and our

[16] Theologian Matthew Barrett (2023) argues that it was the degeneration of scholasticism in the later Middle Ages that was a significant catalyst for the Reformation. This view is not universally accepted. For example, Cross (2024) sees substantial continuity between Luther and late scholasticism.

$\hat{\alpha} = 0.45$.

Averaging the outcomes of many simulations, we compute the $S_t^{k,\varphi}$ exposure of each institution to scholasticism in 1508, the year Luther started teaching at the University of Wittenberg. We measure exposure by counting the number of theologians exposed to scholasticism in 1508, by institution. We then obtain the exposure to scholasticism of each institution in the network, using Equation (3.6). Figure 3.5 depicts this level of exposure. We can further compute the exposure to this paradigm for European cities even if they do not host any institution, as in Equation (3.7). We do this by computing the distance between each city in our sample and each institution in our network, and summing up institutions' exposure at the city level weighted by the inverse of distance, $w_{ck}$.

The sample of cities used to compute exposure is taken from Rubin (2014), which provides a database of over 800 European cities and classifies them as Catholic or Protestant based on the dominant religion in three different years: 1530, 1560, and 1600.[17] Figure 3.5 illustrates the sample of cities and the spread of Protestantism as of 1600. Cities that remained Catholic are shown in grey, while those that became Protestant are marked in red. The blue bubbles represent institutions' exposure to scholasticism as in 1508. The figure already suggests a possible positive correlation between exposure to scholasticism and the likelihood that a city would become Protestant.

We employ a linear probability model to estimate the correlation between exposure to scholasticism and the likelihood that a city became Protestant in 1530, 1560, and 1600. For this experiment, we cannot use a Cox proportional hazard model because in some European regions, such as England and Scotland, the shift towards Protestantism was a top-down decision, meaning hundreds of cities share the same event year. With ties of this magnitude, the partial likelihood approximation underlying the Cox model breaks down, and a linear probability model is more appropriate.

The top panel of Table 3.4 shows the results. The key variable of interest is exposure to scholasticism $S_{1508}^c$, whose estimated coefficient is consistently positive and statistically significant in 1560 and 1600. Columns (1)-(3) show the variable of interest alone. Columns (4)-(6) show that the results are robust to the addition of controls and fixed effects. We control for the presence of all universities active in 1500 as indicated in our database to show that

[17] We updated Rubin's data to better account for the fact that in France and the Low Countries (modern-day Belgium), several cities adopted Protestantism temporarily before being reconquered by Catholic forces. See Appendix 3.H.

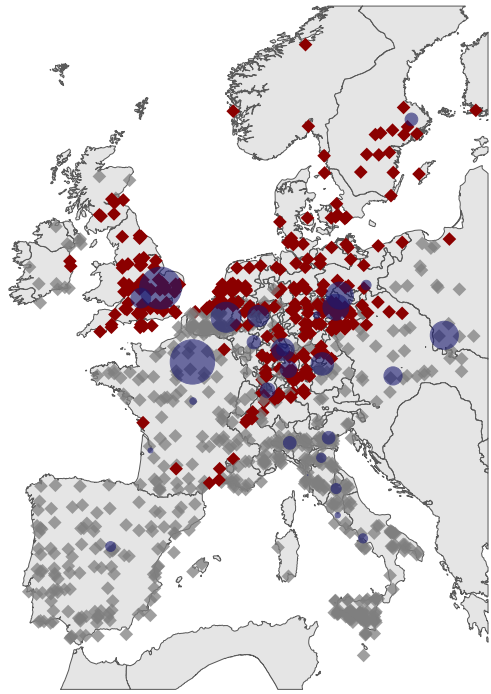


**Figure 3.5.** Blue bubbles represent the exposure to scholasticism in 1508. $\alpha = 0.45$ and $D = 5,000$. Protestant cities are the red diamonds, and Catholic cities are the grey diamonds. Data on cities' religion are taken from Rubin (2014) and updated as in Appendix 3.H; they reflect the religious status in 1600.

exposure to scholasticism $S^c_{1508}$ is not merely proxying for the presence of a university. We also include the city populations in 1500 taken from Buringh (2021), transformed in inverse hyperbolic sine (ihs) to account for cities with no recorded population estimates.[18] The remaining control variables are taken from Rubin (2014) and include factors related to the economic status of cities: the presence of a printing press by 1500, whether the city was a free city by 1517, the market potential of the city in 1500, its membership in the Hanseatic League by 1517, whether it hosted a bishop or archbishop by 1517, and whether the city had direct access to water. We also include fixed effects capturing time-invariant characteristics common to each imperial circle and historical country as of 1500.

Using the estimated coefficients in the benchmark panel of Table 3.4, Columns (1)–(3), we quantify the association between scholastic exposure and the estimated probability of adopting Protestantism. Taking Trier—the most exposed city in the sample—as a reference point, we evaluate how estimated Protestant adoption probabilities differ for other cities when exposure is set to Trier's observed level. For Barcelona, in the lowest quartile of the exposure distribution, the estimated probability is 9.4 percentage points higher in 1530, 39 percentage points higher in 1560, and 47 percentage points higher in 1600. The corresponding differences are 7.1, 28.5, and 35.6 percentage points for Copenhagen; 5.3, 21, and 26.3 percentage points for Padova; and only 0.1, 0.2, and 0.3 percentage points for Leuven, whose observed exposure is already close to Trier's.

To rule out the possibility that the result reflects intellectual centrality rather than scholasticism specifically, we repeat the analysis replacing scholastic exposure with exposure to two contemporaneous but non-theological ideas: Practical Surgery and Civil Law (see Appendix 3.I for further details and context). The placebo panels of Table 3.4 report the results alongside the benchmark regression. Neither idea predicts the adoption of Protestantism in any specification, reinforcing the interpretation that the association documented above is specific to scholasticism rather than to university culture or intellectual activity in general.

[18]The inverse hyperbolic sine is similar to the logarithmic transformation but can accommodate zero (Bellemare and Wichman, 2020), which makes it particularly useful when dealing with the large number of cities from Rubin (2014) with no available population estimates in Buringh (2021).

**Table 3.4.** Linear Probability Model - Exposure to scholasticism in 1508 and cities' probability to become protestant in 1530, 1560, and 1600

| | Protestant in | | | Protestant in | | |
|---|---|---|---|---|---|---|
| | 1530 | 1560 | 1600 | 1530 | 1560 | 1600 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Benchmark regression with exposure to scholasticism | | | | | | |
| Exposure scholasticism $S^c_{1508}$ | 0.001 | 0.004*** | 0.005*** | 0.001** | 0.004*** | 0.004*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Adjusted $R^2$ | 0.016 | 0.103 | 0.185 | 0.503 | 0.719 | 0.736 |
| Placebo regression with exposure to practical surgery | | | | | | |
| Exposure to Practical Surgery $PS^c_{1508}$ | 0.00000 | -0.001 | -0.001 | -0.0003 | -0.001 | -0.001 |
| | (0.001) | (0.002) | (0.002) | (0.0004) | (0.001) | (0.001) |
| Adjusted $R^2$ | -0.001 | 0.002 | 0.006 | 0.499 | 0.708 | 0.726 |
| Placebo regression with exposure to civil law | | | | | | |
| Exposure to Civil Law $CL^c_{1508}$ | 0.0005 | 0.0004 | 0.001 | -0.00001 | -0.0004 | 0.0001 |
| | (0.001) | (0.002) | (0.002) | (0.0005) | (0.001) | (0.001) |
| Adjusted $R^2$ | 0.001 | -0.001 | 0.0003 | 0.499 | 0.707 | 0.726 |
| Rubin's controls | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ |
| Imperial Circle FE | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ |
| 1500 Country FE | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ |
| Observations | 867 | 867 | 867 | 867 | 867 | 867 |

*Notes:* Robust SE clustered by territory in parentheses. A constant term is included in all regressions. Dependent variable 'Protestant' takes value 1 if the city is protestant in 1530, 1560, 1600, respectively. Data on cities' religion taken from Rubin (2014) and updated as in Appendix 3.H. "Exposure to scholasticism $S^c_{1508}$", "Exposure to Practical Surgery $PS^c_{1508}$", and "Exposure to Civil Law $CL^c_{1508}$" are computed as in Equation (3.7). Rubin's controls, selected from Rubin (2014), include: printing press by 1500, (ihs) city population in 1500, free imperial city by 1517, market potential in 1500, Hanseatic League membership by 1517, lay magnate, (arch)bishop by 1517, and access to water. We also include in all specifications the dummy variable"Presence of university in 1500", taking value 1 if the city had a university in 1500, as in our database (de la Croix, 2021).

### 3.4.2 Anti-Judaism and the Persecution of Jews

Our model can propagate not only good ideas, but also bad or even false ones. One such idea is anti-Judaism. The availability of data on Jewish persecutions from Anderson et al. (2017) and Jedwab et al. (2019) allows us to apply the same approach as in Section 3.4.1. The roots of anti-Judaism run deep and are widely discussed in the literature. Scholasticism, once again, played a role in rationalizing prejudice against Jews. Thomas Aquinas (1225–1274), one of the most influential scholars at the University of Paris and the intellectual heir of Lombardus, endorsed many prevailing medieval Christian views about Jews. He supported the idea that Jews should live in subjugation as a reminder of their supposed rejection of Christ. In the *Summa Theologiae*, he also discusses Jews in ways that reinforce their marginalization.

The next generation of scholastic theologians, such as John Duns Scotus (c. 1266–1308) and William of Ockham (c. 1287–1347), contributed to the broader scholastic discourse that pathologized Judaism as a theological error (see Abulafia (2011), which provides a substantial discussion of how scholastic theology and legal reasoning shaped Christian–Jewish relations between 1000 and 1300). While these theological ideas do not explicitly advocate persecution, they may have interacted with the mechanisms identified by Anderson et al. (2017) and Jedwab et al. (2019). In Anderson et al. (2017), cold temperatures are shown to increase the probability of persecution of Jewish communities. Jedwab et al. (2019) extend this insight, showing that negative shocks more broadly, particularly plagues, raise the likelihood of minority persecution.

To study if there is an association between yearly exposure to scholasticism $S_t^c$ and the probability of violent acts against Jews, we use data from Anderson et al. (2017) and Jedwab et al. (2019).[19] Building on this literature, we hypothesize that the association between negative shocks and the likelihood of Jewish persecutions is stronger in cities with greater exposure to scholasticism $S_t^c$. One possible interpretation of this pattern is that, in cities where local priests were more exposed to scholastic teachings and disseminated anti-Judaic arguments, the local intellectual and religious climate was more receptive to scapegoating Jews in times of crisis. To test this hypothesis, we augment the empirical framework of Anderson et al. (2017) by including an interaction term between our measure of scholastic exposure and the incidence of plague

[19] We thank the authors for sharing the most recent (and still unpublished) version of the pogroms and persecutions data.

outbreaks.

**Table 3.5.** Linear Probability Model - Yearly Exposure to scholasticism and cities' probability of persecution of Jews

| | Persecutions | |
|---|---|---|
| | Replication | $S_{ct}$ x Plague |
| | (1) | (2) |
| Temperature$_{c,t-1}$ | -0.467*** | -0.472*** |
| | (0.125) | (0.124) |
| Plague | 5.100* | -1.880 |
| | (2.149) | (1.172) |
| Exposure to scholasticism $S_{ct}$ | | 0.162* |
| | | (0.064) |
| Exp. to scholasticism $S_{ct}$ x Plague | | 5.158*** |
| | | (1.520) |
| Controls | ✓ | ✓ |
| City Fixed Effects | ✓ | ✓ |
| Observations | 273,879 | 273,879 |
| $R^2$ | 0.013 | 0.018 |

*Notes:* *p<0.1; **p< 0.05; ***p<0.01. Standard errors clustered at the climate grid level in parentheses. City fixed effects are always included.
Column (1) replicates Anderson et al. (2017) specification (2), Table 3, p.940. Coefficients are multiplied by 100 to represent percentage points.
Controls include a slope variable for the 10 years surrounding the Black Death and a measure of population density as in Anderson et al. (2017).

Table 3.5 presents the results of this panel regression. The dependent variable, Persecutions, takes the value 1 when either an expulsion or another violent act against Jews occurred in a given year, following the definition in Jedwab et al. (2019). As in both Anderson et al. (2017) and Jedwab et al. (2019), we restrict the analysis to cities with a documented Jewish presence.

Column (1) replicates specification (2) of Table 3 in Anderson et al. (2017, p. 940). The estimated coefficient on lagged temperature is nearly identical: we find that a one-degree decrease in temperature increases the probability of Jewish persecutions in the following year by 0.467 percentage points, compared to 0.464 percentage points in Anderson et al. (2017). Despite this similarity, our sample differs slightly for two main reasons: (i) our dependent variable is drawn from a more recent version of the dataset used by Anderson et al. (2017), and (ii) we eliminate duplicate city entries prior to estimation.

Column (2) introduces an interaction between yearly exposure to

scholasticism $S_t^c$ at the city level and a dummy variable indicating the presence of a plague, following Anderson et al. (2017). Interestingly, the plague dummy is not statistically significant on its own anymore, while exposure to scholasticism $S_t^c$ is only slightly significant at the 10% level. However, their interaction is both statistically and substantively significant. This suggests that the coexistence of a theoretical framework portraying Jews as a threat and a plague outbreak is associated with a significantly higher likelihood of violence against Jews. The coefficient on the interaction term is also sizable: during a plague, a one-unit increase in exposure to scholasticism is associated with a 5.2 percentage point higher likelihood of Jewish persecutions.

### 3.4.3 Finding Atlantis: A True Story of Genius and Madness

We now turn to Olaus Rudbeck's (1630–1702) claim that Sweden was the cradle of civilization and the site of the lost city of Atlantis. It is an interesting case to simulate using our model for two reasons. First, it illustrates that ideas–whether accurate or not–can still spread through affiliation networks, which further underscores the crucial role of institutions in shaping intellectual diffusion. Second, it highlights that individuals do not necessarily need to agree with an idea in order to help propagate it.

Rudbeck, who was professor of medicine at the University of Uppsala from 1658 to 1692 (Von Bahr, 1945), claimed that Sweden was in fact the mythical island of Atlantis, and thus the cradle of all ancient civilization. He supported this sweeping theory by drawing connections between Norse mythology, the Bible, and classical sources. Rudbeck explained his thesis in the book *Atlantica* (also known as Atland eller Manheim), first published in four volumes between 1679 and 1702 (King, 2005). It was written in Latin (vol. 1) and Swedish (vols. 2–4). *Atlantica* was not fully translated into major European languages during the 18th century. The length, complexity, and eccentricity of Rudbeck's arguments likely discouraged publishers. His ideas were seen by many contemporaries as extravagant, although some Nordic nationalist thinkers admired them. Despite its eccentricity, *Atlantica* was referenced and critiqued by various 18th-century thinkers, including two academic scholars, Denis Diderot and Ludvig Holberg. They may have been exposed to Rudbeck's ideas indirectly through the affiliation network, given

that they are unlikely to have read his books in Swedish themselves.

Denis Diderot (1713–1784) was a member of the Prussian Academy of Sciences from 1751 (Amburger, 1950). Diderot's radical thinking, conflict with French authorities, and nonconformist personality kept him outside the fold of French academies. In the article *Étymologie* of his Encyclopédie (published over the period 1751–1765), Diderot used Rudbeck's work as a cautionary example of how speculative etymology can lead to erroneous conclusions, critiquing the methodology employed in *Atlantica*. Our simulation reveals that Diderot had a 100% chance of being exposed to Rudbeck's ideas as early as 1751–the year of his election–because he hypothetically encountered numerous members who were already exposed to the idea.

Ludvig Holberg (1684–1754) was a prominent Danish-Norwegian writer and philosopher, and professor at the University of Copenhagen from 1717 to 1754 (Slottved, 1978). He satirized Rudbeck's theories, mocking the idea of Sweden as Atlantis and highlighting the speculative nature of Rudbeck's claims. According to our simulation, the chance Holberg was exposed to Rudbeck's idea is 0 until 1753, when it goes to 48.74%. It gains an additional 46.52% in the following year, before he died. This highlights two features of our approach: first, our exposure is a lower bound, and in this case Holberg might have been acquainted with Rudbeck's work through other means. In other words, even in cases where we know from historical evidence that a scholar engaged with an idea, our model still captures eventual exposure through the network alone. Second, Holberg was never affiliated with an academy, but only with a university, which in our model means it must have taken more time for ideas to reach him. As we will later show, this institutional feature plays a key role in shaping how easily ideas circulate through the network.

## 3.5 Counterfactual experiments

In this section, we identify the features of the academic network that are more conducive to spreading ideas. We perform two kinds of experiments. We assign ownership of an idea to fictitious inventors, in order to track whether the idea would spread differently in alternative realities. In the second experiment, we remove some parts of the network to assess their importance in spreading ideas.

### 3.5.1 Placebo inventors of Botanical Realism

To better understand how the structure of a network influences the speed at which ideas spread, we run counterfactual experiments using Fuchs' Botanical Realism as a case study. In these experiments, we imagine that it was not Fuchs who introduced the new paradigm of Botanical Realism, but another contemporary scientist from a different region of Europe. We simulate the diffusion of this paradigm, still originating in 1542, but emerging in various alternative locations: in Salamanca with Juan Aguilera (1507-1560), in Oxford with John Warner (c. 1500-1565), in Louvain with Jeremius Dryvere (1504-1554), in Wittenberg with Andreas Goldschmidt (1513-1559), in Cracow with Mikołaj Mleczko Wieliczki (1490-1559), in Rostock with Jacob Bording (1511-1560), in Montpellier with Antoine Saporta (1507-1573), in Padua with Girolamo Donzellini (1513-1587), from the Royal College of France with Oronce Fine (1494-1555), in Pisa with Realdo Colombo (1510-1559), and in Leipzig with Georg Joachim Porris (1514-1574). Appendix 3.J presents a short biography of each of these scholars. These counterfactual simulations allow us to explore how regional networks and academic hubs would have shaped the spread and influence of Botanical Realism across Europe.

Many of these scholars came from strong intellectual backgrounds and were part of the broader Renaissance shift toward empiricism and direct observation. Several, particularly those with medical training (e.g., Saporta, Colombo, Bording), had practical reasons to study plants carefully and could have contributed to a scientific approach to botany.[20]

Figure 3.6 illustrates the percentage of individuals in medicine and sciences who were exposed to the idea across the twelve simulated scenarios. This simulation offers valuable insights into the diffusion process. Notably, in one case, the idea fails to spread: Wieliczki in Cracow (professor from 1513 to 1552) lacked other mobile peers for meaningful intellectual exchange, and the idea does not take hold. In ten other cases, we observe that by the end of the period, nearly all relevant scholars had encountered the idea. It is important to note that these results are averaged over 5,000 simulations, meaning a consistently high rate of diffusion across all simulations. This reflects how

[20]That said, Botanical Realism required not just empirical skills but also an interest in plants themselves, which some of these figures might not have had. Fuchs' success came from a combination of his medical background, interest in plants, access to talented illustrators, and the specific intellectual environment in Germany at the time.

effectively the European intellectual network functioned to disseminate ideas. Regardless of their point of origin, ideas eventually spread throughout Europe in the long run. For instance, Warner's idea remained localized in Oxford for some time, but reached Cambridge and Gresham College by 1650, before spreading further. Similarly, Saporta's concept, originating in Montpellier, reached Basel, Lincei, and Toulouse by 1600, and continued to propagate across Europe afterward. After two centuries, both Bording's and Saporta's ideas had spread at a similar rate to various locations. This demonstrates that, despite differences in individual pathways and speed of diffusion, the overall outcome was the same: the widespread diffusion of ideas across Europe.

This confirms that the diffusion process generated by our model is non-ergodic: the success of an idea remains dependent on its initial conditions – specifically, the network position of its inventor. Appendix 3.J.1 provides a more detailed discussion about this point. This implies that the distribution of successful ideas does not converge to a single stationary form, i.e., even as time progresses, the expected success of an idea remains contingent on the initial conditions. It also implies that outcomes across different realizations of the process do not average out over time (in an ergodic process, averaging over time should yield the same result as averaging over different realizations of the process: see Peters (2019) for a history of the idea of ergodicity).

The case of Fuchs (4) is particularly interesting, as the diffusion of his herbal plateaus at 40%, unlike the other simulations where ideas either reach full exposure or fade out entirely. This reflects the fact that, in roughly 60% of the simulations, Fuchs' idea dies out quickly, keeping exposure at zero. In the other 40%, Fuchs' idea spreads successfully, reaching high levels of adoption after a century. This outcome suggests that the spread of Fuchs' herbal hinges on a fragile initial phase, where its survival occurs with a probability of about one tenth.

To understand the European academic network, we can examine the transmission of ideas through the case of Fuchs' Botanical Realism. Figure 3.7 highlights the key individuals and institutions involved. This idea originated at the University of Tübingen, where it thrived for over a century due to the steady presence of scholars in science and medicine. However, the Thirty Years' War effectively closed the university and disrupted this continuity, especially between 1628 and 1634, during its occupation by Imperial (Catholic) forces of the Holy Roman Empire.

Despite this disruption, the idea spread to other institutions via Tübingen

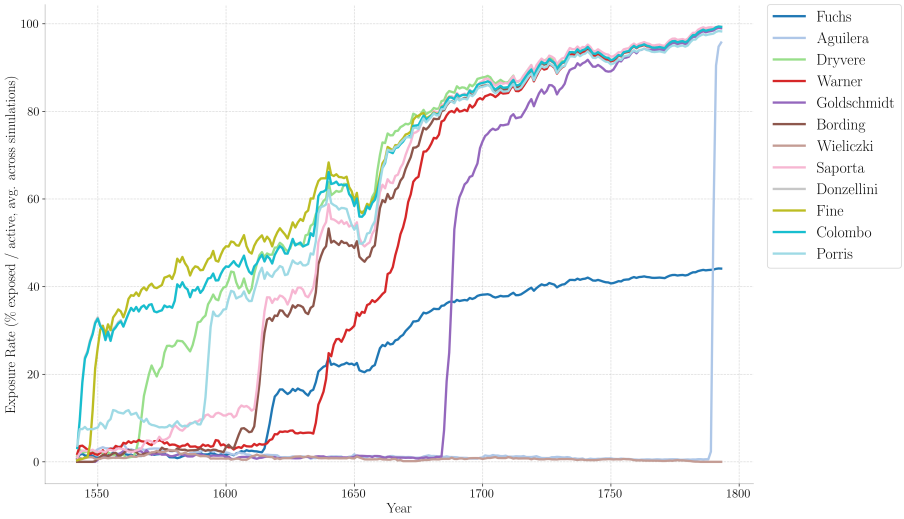



**Figure 3.6.** Exposure rates of active scholars in the network from 1084 to 1793, considering different hypothesized proponents of Botanical Realism.

scholars who secured positions elsewhere. Jakob Degen taught briefly in Strasbourg (Berger-Levrault, 1890), while Michael Mästlin held a temporary post in Heidelberg (Drüll, 2002). However, these transfers did not result in sustained knowledge transmission: Strasbourg was too small to establish permanent positions in the sciences, and Heidelberg faced the same wartime challenges as Tübingen. Nevertheless, in Strasbourg, the physician Kasper Maliński may have encountered Fuchs' ideas. His subsequent move to the University of Zamość (Kedzoria, 2021) could have carried the concept further.

At Zamość, the mathematician Adrien Van Roomen (also known as Romanus) might have engaged with the idea. Near the end of his life, Romanus became a member of the Accademia dei Lincei, where he potentially reintroduced the concept. Through the Lincei, an informal academy with prominent members such as Galileo and Kepler, the idea could have spread internationally. Thus Adrien Van Roomen, Jakob Degen, Michael Mästlin, and Kasper Maliński are necessary for the survival of the idea. Van Roomen is only the last of a series of key players according to Zenou's (2016) definition, which was developed in the context of criminal networks: "the key player who is the agent that should be targeted by the planner so that, once removed, she will generate the highest level of reduction in total activity" (p. 1403).

This hypothetical trajectory highlights three features of European academia in the sixteenth and seventeenth centuries. First, the dense network of connections ensured that ideas could survive even amid significant disruptions, such as the Thirty Years' War. Marginal institutions, like the University of Zamość, played a crucial role in this resilience. Second, early informal academies, such as the Lincei, were vital for preserving and disseminating ideas across borders. Third, the European academic network was strongly path-dependent and non-ergodic as described in (David, 1985): it was shaped by more or less random historical events "rather than systematic forces" (p.332).

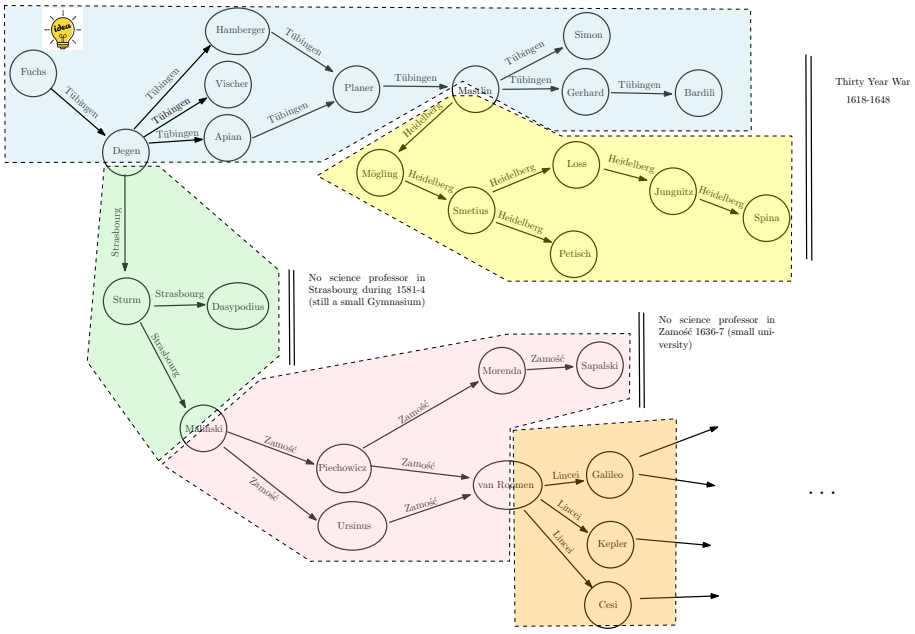



**Figure 3.7.** Botanical Realism path

## 3.5.2 Removing components of the network

As explained above, we simulate the spread of ideas in a network stripped of certain components. This gives us better insight into the role of each component. Under this approach, we rewrite Equation (3.1) as

$$I^B_{t+1} = B_t I^B_t + I^B_t \tag{3.10}$$

where the new affiliation matrix $B \leq A$ in the Hadamard order (that is, every entry of $B_t$ is less than or equal to the corresponding entry of $A_t$). Then it is obvious that $I_t^B \leq I_t \ \forall t$, assuming the same initial condition $I_0 = I_0^B$. Indeed, in the new dynamics, there will be fewer exposed persons at every time step, since reducing the number of connections reduces the opportunities for ideas to spread. This follows from the fact that matrix multiplication with a reduced adjacency matrix $B_t$ leads to a weakly lower infection count at every step.

In the stochastic version, as $\mathbb{E}[\Omega^d(B_t)] = \alpha B_t$, we have $\mathbb{E}[\Omega^d(B_t)] < \mathbb{E}[\Omega^d(A_t)]$ in the Hadamard order, and $\mathbb{E}[I_t^B] \leq \mathbb{E}[I_t]$, $\forall t$. This means that, with an infinite number of simulations $D$, the world with $B$ (fewer contacts) will still have fewer exposures than the world with $A$, even in the presence of stochastic transmission. But this statement is no longer strictly true in every realization. If we simulate the process many times, the law of large numbers implies that the average outcome of these simulations should converge to the expectation. The required number of simulations $D$ may however be very large, because of strong non linear effects coming from the topology of the network.

First, we let Botanical Realism and Mathematical Astronomy spread over the affiliation network as if academies never emerged, to assess whether academies were key for the diffusion of the ideas of the Scientific Revolution (McClellan, 1985; Pedersen, 1992). To measure diffusion, we use Equation (3.7), which computes the exposure of any European city to an idea. We use the set of cities in Buringh (2021), excluding those in the Ottoman Empire and in Russia. This leaves us with 1,916 cities. We report quartiles of the distribution of exposure across this set of cities relative to the benchmark.

In reporting the results, we distinguish between two roles of academies. First, they contribute directly to the exposure of nearby cities to ideas – for example, when Greenwich benefits from the presence of the Royal Society in London. Second, they facilitate the diffusion of ideas within the affiliation network by bridging university communities – for instance, when Greenwich benefits from scholars at Oxford and Cambridge, whose intellectual development and connectivity have been enhanced by the Royal Society. In Table 3.6, the line "No direct contribution" gives the exposure distribution when the direct contribution is shut down. Practically, we keep the vector of the individual exposures $I_t^d$ from the benchmark, but we remove academies' exposure $S_t^{k,\varphi}$ from the computation of city exposure $S_t^{c,\varphi}$. The line "No academies at all" is based on an alternative affiliation matrix where all the

edges stemming from academies are removed, and the various measures of exposures are computed with this matrix. The first line for each year "With academies" represents the baseline exposure distribution, which can be used for comparison.

The following insights can be drawn from Table 3.6: there are very few academies in 1600. The Ricovrati in Padua was just created (in 1599), and it was mostly literary at that time. The same is true of the Accademia della Crusca (founded in Florence in 1583 to preserve the purity of the Italian language). The Lincei, already mentioned above in the context of its relationship with Romanus and Botanical Realism, was founded in 1603 (Gabrieli, 1989). As a result, academies did not have a big influence on the exposure to Botanical Realism, and both lines remain the same as in the benchmark. However, academies already played a role in spreading Mathematical Astronomy, as the exposure distribution drops with respect to the benchmark: the median drops by almost 23% in both scenarios. Even at this early stage, when academies are few and primarily informal, they contribute to the dissemination of ideas.

By 1650, academies begin to influence Botanical Realism, both as part of the network (due to figures like Romanus) and directly. By this time, both Botanical Realism and Mathematical Astronomy have reached nearly all cities. After 1700, academies are increasingly significant. For Botanical Realism, academies are essential, as indicated by the third row in each scenario dropping to zero. In contrast, while Mathematical Astronomy does not depend strictly on academies for its survival, they play a crucial role in amplifying exposure. By 1793, the absence of academies would lead to a dramatic reduction in exposure to Mathematical Astronomy: in the scenario without academies at all, the median would drop by more than 95%.

We can use this same tool to analyze the role of specific regions or nations. The literature has examined the contributions of each nation to the rise of science and knowledge in Europe. Each country possessed unique characteristics that, when combined, created a fertile environment for intellectual and scientific transformations. For example, Italy laid the foundations with the Renaissance and early scientific methods (Applebaum, 2003); the British Isles drove empiricism and practical applications (Mokyr, 2011a); France spearheaded Enlightenment thinking and institutional science (Ferris et al., 2010); the Iberic Peninsula advanced economic theory and natural law, and the Holy Roman Empire advanced theoretical frameworks in

**Table 3.6.** Counterfactual experiment with and without academies.

| **Botanical Realism** | Q1 | Median | Q3 |
|---|---|---|---|
| With academies in 1600 | 0 | 1.31 | 2.73 |
| No direct contribution in 1600 | 0 | 1.31 | 2.73 |
| No academies at all in 1600 | 0 | 1.31 | 2.73 |
| With academies in 1650 | 8.86 | 25.02 | 41.80 |
| No direct contribution in 1650 | 7.13 | 20.65 | 34.48 |
| No academies at all in 1650 | 0 | 0.13 | 0.43 |
| With academies in 1700 | 31.07 | 93.63 | 153.18 |
| No direct contribution in 1700 | 20.91 | 47.95 | 74.33 |
| No academies at all in 1700 | 0 | 0 | 0 |
| With academies in 1750 | 58.77 | 190.58 | 292.93 |
| No direct contribution in 1750 | 23.86 | 57.09 | 98.47 |
| No academies at all in 1750 | 0 | 0 | 0 |
| With academies in 1793 | 94.67 | 376.86 | 616.56 |
| No direct contribution in 1793 | 21.15 | 62.04 | 109 |
| No academies at all in 1793 | 0 | 0 | 0 |
| **Mathematical Astronomy** | Q1 | Median | Q3 |
| With academies in 1600 | 0 | 6.20 | 17.10 |
| No direct contribution in 1600 | 0 | 4.79 | 13.59 |
| No academies at all in 1600 | 0 | 4.78 | 13.58 |
| With academies in 1650 | 9.16 | 27.02 | 45.72 |
| No direct contribution in 1650 | 6.60 | 20.80 | 35.66 |
| No academies at all in 1650 | 0 | 7.19 | 24.97 |
| With academies in 1700 | 26.24 | 82.13 | 131.15 |
| No direct contribution in 1700 | 16.52 | 37.51 | 62.11 |
| No academies at all in 1700 | 0.11 | 6.90 | 21.91 |
| With academies in 1750 | 67.64 | 220.03 | 334.68 |
| No direct contribution in 1750 | 25.27 | 60.47 | 105.78 |
| No academies at all in 1750 | 5.61 | 21.40 | 37.53 |
| With academies in 1793 | 121.76 | 455.02 | 744.32 |
| No direct contribution in 1793 | 19.50 | 65.41 | 105.18 |
| No academies at all in 1793 | 7.90 | 22.51 | 39.05 |

*Notes*: Quartiles of cities' exposure distributions to ideas when [0] academies are fully considered (benchmark) [1] academies have no direct contribution on cities but are present in the network [2] academies have no effect at all.

mathematics and astronomy.

**Table 3.7.** Counterfactual Experiment with and without European regions.

| **Botanical Realism** | Q1 | Median | Q3 |
|---|---|---|---|
| No Italian Peninsula | 0 | 0 | 0 |
| No British Isles | 83.77 | 285.19 | 543.02 |
| No France | 70.98 | 280.54 | 478.80 |
| No Iberic Peninsula | 92.87 | 374.70 | 614.42 |
| No Holy Roman Empire | 0 | 0 | 0 |
| Benchmark | 94.67 | 376.86 | 616.56 |
| **Mathematical Astronomy** | Q1 | Median | Q3 |
| No Italian Peninsula | 0 | 0 | 0 |
| No British Isles | 95 | 334.40 | 630 |
| No France | 83.46 | 346.68 | 581.16 |
| No Iberic Peninsula | 118.94 | 453.81 | 742.57 |
| No Holy Roman Empire | 118.13 | 433.65 | 694.37 |
| Benchmark | 121.76 | 455.02 | 744.32 |

*Notes*: Quartiles of cities' exposure distributions to ideas in five counterfactual networks without a specific European region, and the benchmark in 1793.

We apply our model to study separately the importance of each nation in spreading and keeping alive each idea. To measure how nations are key for an idea, we construct five counterfactual networks, removing institutions belonging to specific geographical areas: one without the Italian peninsula, one without the British Isles, one without France,[21] one without the Iberian peninsula, and one without the Holy Roman Empire (as defined in Stelter et al. (2021). We then simulate the spread of Botanical Realism and Mathematical Astronomy in these five networks, and in the benchmark model.

Results are presented in Table 3.7. Each number is a summary statistic of the exposure $S^{c,\varphi}_{1793}$ of the cities' distributions, together with the relative exposure distribution in the benchmark. Without the Italian Peninsula, neither Botanical Realism nor Mathematical Astronomy would have survived. For example, for the latter, the universities of Rome and Bologna were critical hubs, allowing scholars previously exposed to Mathematical Astronomy in

[21] We exclude from France cities which became French towards the end of the period: Strasbourg (1681), Molsheim (1648), Nancy (1766), Pont-a-Mousson (1766), Nice (1860), Perpignan (1659), Arras (1659), Douai (1667).

Vienna or Bratislava to continue disseminating it among their colleagues. It appears that the Holy Roman Empire is necessary for Botanical Realism, but not for Mathematical Astronomy. No other regions in this simulation were necessary for either idea to spread. For example, both Botanical Realism and Mathematical Astronomy spread throughout Europe even without the British Isles, although the median drops by 24.3% and by 26.5%, respectively. Even if the British Isles had a key role in the propagation and implementation of useful knowledge (Hallmann et al. (2022) show that British inventors worked on technologies that were more central within the innovation network), it is not the case as far as our example of propositional knowledge is concerned (Mokyr, 2011b).

The same reasoning holds for the other regions: the spread of ideas across Europe is not significantly hindered by the absence of France, and the contribution of the Iberian Peninsula appears particularly limited or even negligible.

Overall, this analysis underscores the resilience of the European network of academies and universities. Even when some parts are removed, the network remains sufficiently dense to sustain the circulation of ideas.

Finally, we focus on the role of Jesuits and present the simulation results when Jesuits are removed from the network. The Society of Jesus, founded in 1540 by Ignatius of Loyola, was a highly influential religious order in the Catholic Church. Its members underwent rigorous training, including years of spiritual exercise and intellectual formation. To counter Protestantism, Jesuits rapidly established an extensive network of schools, colleges, and universities across Europe (Grendler, 2019) and beyond. In the RETE database, we count 52 Jesuit institutions among the 213 universities and colleges of some renown, and more than 6800 scholars (Jesuit priests)–approximately 8% of all recorded scholars between 1000 and 1800, a figure that rises to 11% when the sample is limited to the period after the Jesuit order was founded. Known for their high academic standards, Jesuits taught humanities, sciences, philosophy, and theology. Their growing influence led to political tensions and, subsequently, expulsions from several countries: Portugal (1759), France (1764), Spain (1767), and Naples (1767). In 1773, under pressure from European rulers, Pope Clement XIV suppressed the order, though it survived in Russia and Prussia (and later in the USA, where Georgetown University was founded in Washington DC).

Table 3.8 presents the results for the year 1750 (a few years before the

**Table 3.8.** Counterfactual experiment with and without Jesuits

| **Botanical Realism** | Q1 | Median | Q3 |
|---|---|---|---|
| Benchmark | 58.77 | 190.58 | 292.93 |
| No direct contribution | 53.76 | 174.62 | 265.14 |
| No Jesuits at all | 52.75 | 172.09 | 261.64 |
| **Mathematical Astronomy** | Q1 | Median | Q3 |
| Benchmark | 67.64 | 220.03 | 334.68 |
| No direct contribution | 57.68 | 188.67 | 283.43 |
| No Jesuits at all | 56.98 | 186.90 | 281.26 |

*Notes*: Quartiles of cities' exposure distributions to ideas in the counterfactual network without Jesuits and the benchmark in 1750.

dissolution of the order). When we remove all Jesuit nodes from the network and simulate the spread of Mathematical Astronomy, we observe some differences, especially in peripheral regions largely neglected by non-Jesuit institutions (such as Sicily, Andalusia, and Romania). Given the widely recognized contributions of Jesuits to science—particularly astronomy—as evidenced by the many lunar craters named after Jesuit astronomers and all the observatories they built (Udías, 2003), it is unsurprising that the median drop is around 15% when analysing Mathematical Astronomy in both scenarios. On the other hand, removing the Jesuits does not really impact the spread of Botanical Realism, where the median drop is around 9% in both cases. One possible explanation for their limited role in the broader diffusion of Botanical Realism is that they operated within a relatively isolated network, separate from the rest of academia.

To test this hypothesis, we examine whether Jesuit nodes are densely connected internally while remaining weakly connected to the rest of the network. We define two groups of nodes: Jesuits and non-Jesuits. To quantify Jesuit insularity, we compute the Coleman (1958) inbreeding homophily index, which ranges from -1 (no internal Jesuit connections) to 0 (connections similar to a random network) to 1 (complete inbreeding). Following Currarini et al. (2009), we define $IH$ as follows:

$$IH_{Jesuits,t} = \frac{H_{Jesuits,t} - \omega_{Jesuits,t}}{1 - \omega_{Jesuits,t}}$$

where $H_{Jesuits,t}$ is the fraction of edges entailing only Jesuits at time $t$,

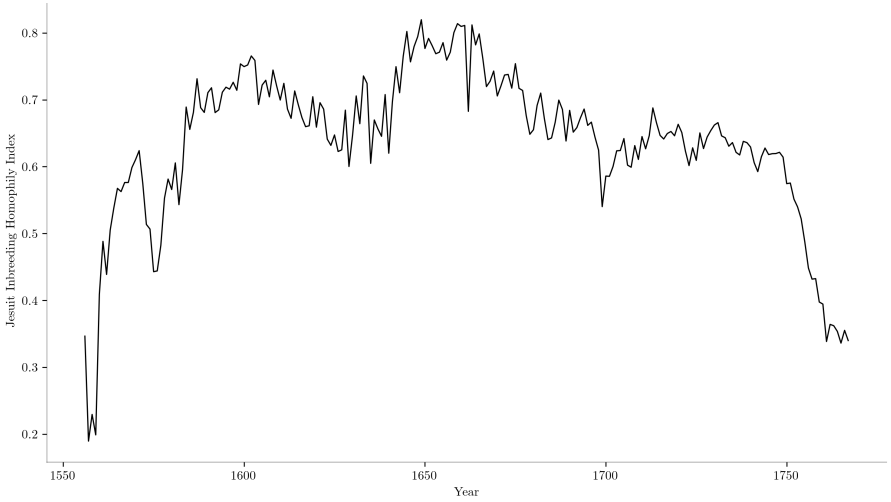



**Figure 3.8.** Jesuit inbreeding homophily over time

and $\omega_{Jesuits,t}$ is the relative fraction of Jesuits in the scholar population at any given point in time $t$.

Figure 3.8 plots $IH_{Jesuits,t}$ from 1556 to 1767, the period over which Jesuits were most active. For most of the timeframe, the index remains between 0.6 and 0.8, indicating a high degree of inbreeding. Jesuit universities were typically closed to non-Jesuit professors, and Jesuit scholars rarely taught outside them. However, the index is slightly lower at the beginning, when the Jesuits were establishing their university network, and at the end, preceding their gradual dissolution.[22]

## 3.6 Conclusions

We have studied how academic networks in the premodern era provided support to the spread of ideas. Using dynamic network models and counterfactual experiments, we showed that features like the emergence of academies and the connections they created across regions helped ideas to spread more widely. By examining the diffusion of groundbreaking ideas and paradigm shifts such

[22] Appendix 3.K provides additional statistics on the Jesuits' position and connectivity in the network, including their number, density, conductance, and decompositions by field.

as Botanical Realism, Mathematical Astronomy, and scholasticism, we have highlighted the role of higher-education institutions in European development.

The counterfactual experiments reveal the nuanced importance of academies: not only did they act as hubs of direct idea dissemination, but they also enhanced the connectivity of the broader network, bridging university communities. Without academies, the spread of ideas born in university settings would have been significantly slower. Our approach also provided insights into regional contributions to scientific progress, highlighting the resilience of the European network of academies and universities, which was dense enough to sustain the circulation of ideas even if certain parts were removed.

While our results are robust to several modelling choices, the analysis treats the observed affiliation network as given. This is a limitation of our approach. In practice, the network arises from individuals' location choices, which are shaped by agglomeration forces and positive sorting. These forces are thus integral to the broader process we study. Mobilizing network formation theory would be a challenging next step.

# Appendix

## 3.A Descriptive Statistics

### 3.A.1 Completeness of the data

Figure 3.A.1 provides an overview of missingness in the scholars dataset across three indicators: the presence of a Wikipedia page, the availability of birth and death years (and thus longevity), and the availability of a place of birth. The largest segment consists of scholars for whom none of these pieces of information is observed—no place of birth, no vital dates, and no Wikipedia link. In the plot, this group is labelled "no.info" and comprises 26,549 individuals. A second large segment (25,381), labelled "Origin.known", includes scholars for whom only the place of birth is known. The third-largest group, with 17,200 scholars, has complete information on all three indicators. Two additional groups are also sizeable. One includes scholars with complete information except for a missing Wikipedia page (13,528). Finally, the set-size bars indicate that scholars with a Wikipedia page — roughly corresponding to the samples used in studies such as Laouenan et al. (2022) — number about 20,000, i.e., approximately one quarter of the full dataset.

Figure 3.A.2 offers a similar overview for affiliations. The number of affiliations exceeds the number of scholars because many individuals held multiple affiliations over their lifetimes. The figure focuses on academic fields. Roughly 80% of scholars are associated with a single field, most commonly the humanities, followed by the sciences (including medicine), theology, and law. 6,939 scholars have an unknown field, and 3,744 are classified as "honorary" because they are not scholars in the strict sense but instead hold honorary or administrative roles (e.g., academy protectors, or students serving as rectors in Italian universities). The plot also highlights frequent combinations of fields among multi-field scholars, notably theology with humanities (6,707) and humanities with sciences (3,382).

Figure 3.A.3 shows how complete our information is for identifying scholars' periods of activity. In the vast majority of cases, the start year is known, typically the year of appointment at a university or the year of election to an academy. End dates, however, are recorded less systematically, as many sources report the year a professorship begins but not when it ends. For academies, it is generally reasonable to assume that membership lasts until death, whereas no

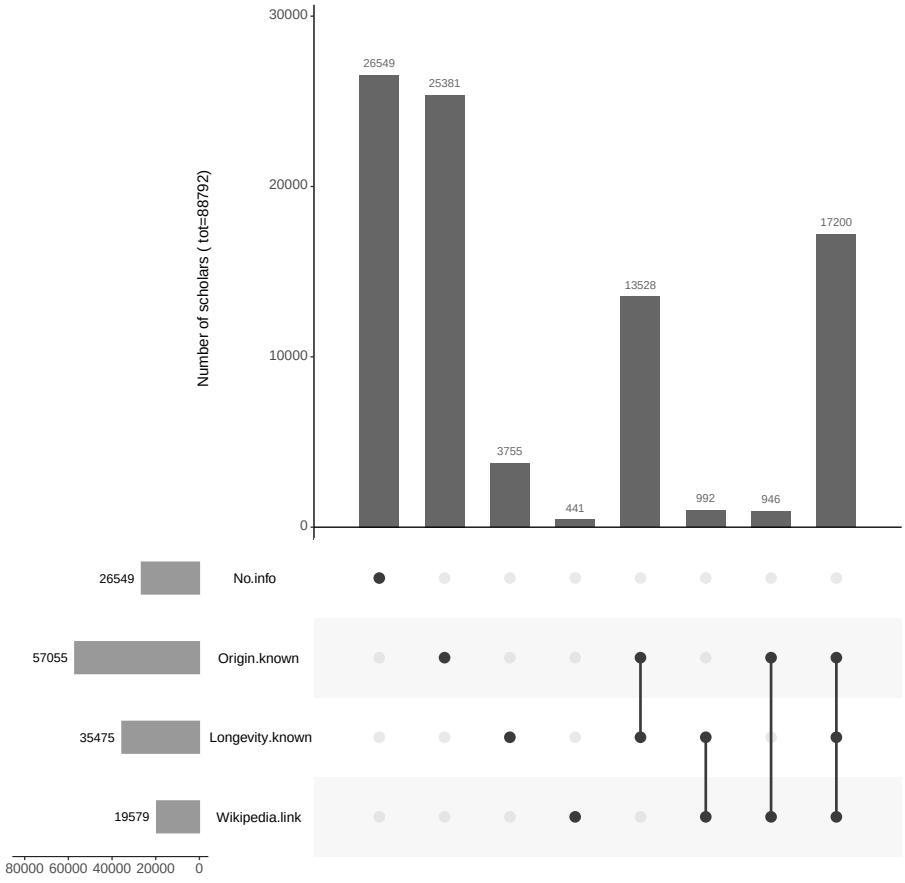



**Figure 3.A.1.** UpSet diagram for persons (matrix-based visualization of intersections across variables).

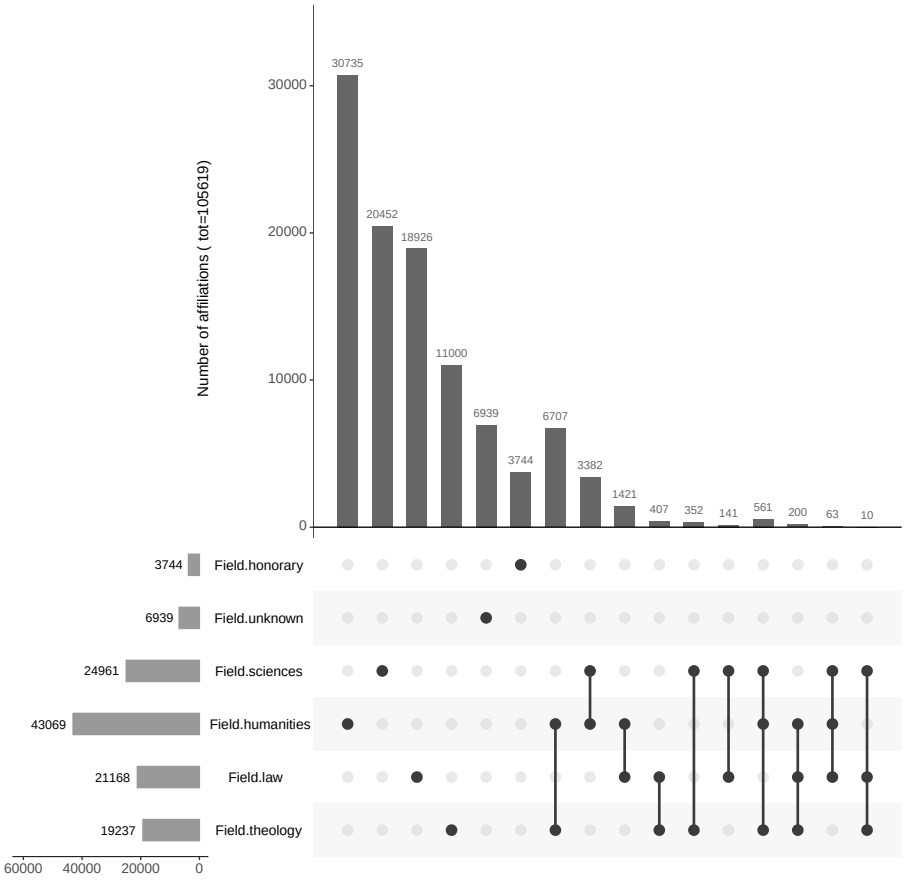



**Figure 3.A.2.** UpSet diagram for affiliations and academic fields (matrix-based visualization of intersections across variables).

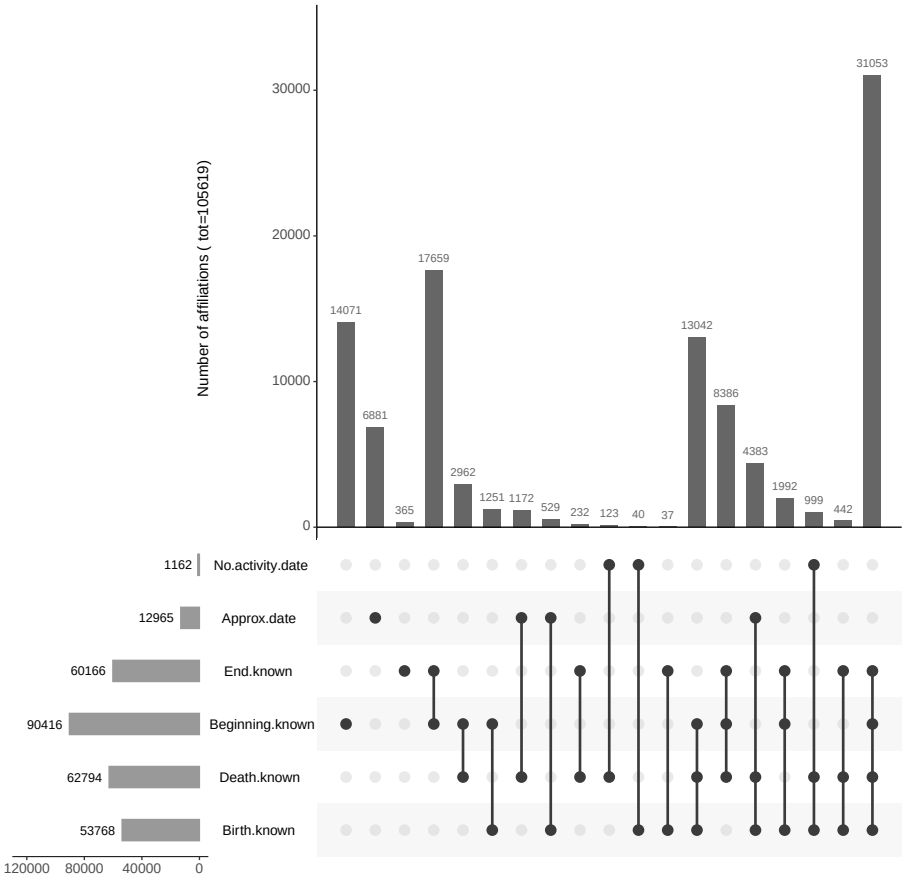



**Figure 3.A.3.** UpSet diagram for affiliations and activity periods (matrix-based visualization of intersections across variables).

comparable assumption can be made for university posts. For 12,965 scholars we can only assign an approximate activity date, and for a small minority (around 1,000) no dating information is available at all.

### 3.A.2 Universities, academies, scholars over time

Figure 3.A.4 and Figure 3.A.5 show the number of scholars over time of the top-10 universities and academies ranked by size in 1793, respectively. By the end of the period, Cambridge is the most populated university and the Royal Society the most populated academy. Tables 3.A.1 and 3.A.2 report detailed descriptive statistics for each institution.

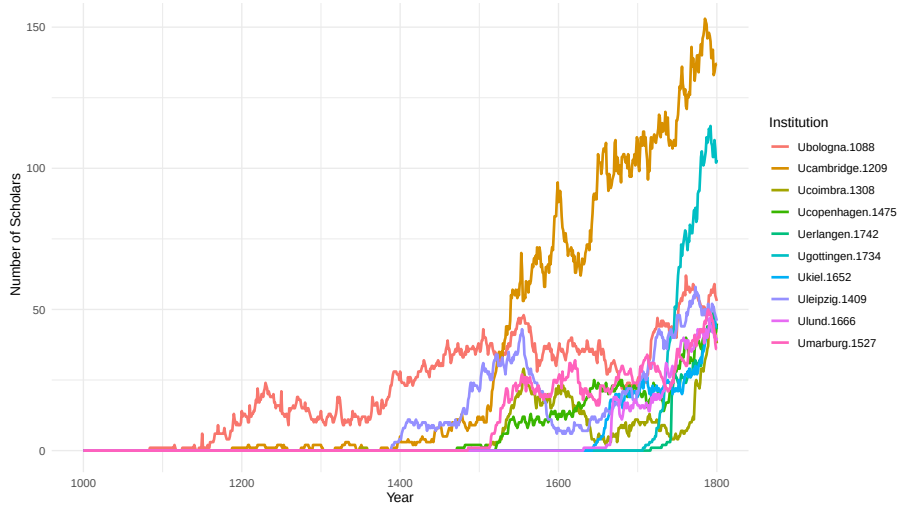



**Figure 3.A.4.** Number of university professors between 1000 and 1800 of the top 10 universities considering the number of scholars in 1793. In the legend, the name of the institution with the foundation date.

## 3.B Definition of active periods

The active period of each scholar commences at the start of their academic career and concludes at their retirement or death. Activity begins in the earliest known year of affiliation with a formal educational institution, when available. For university professors, this is the year they began teaching, while for academy members, it is the year they were elected as members of the

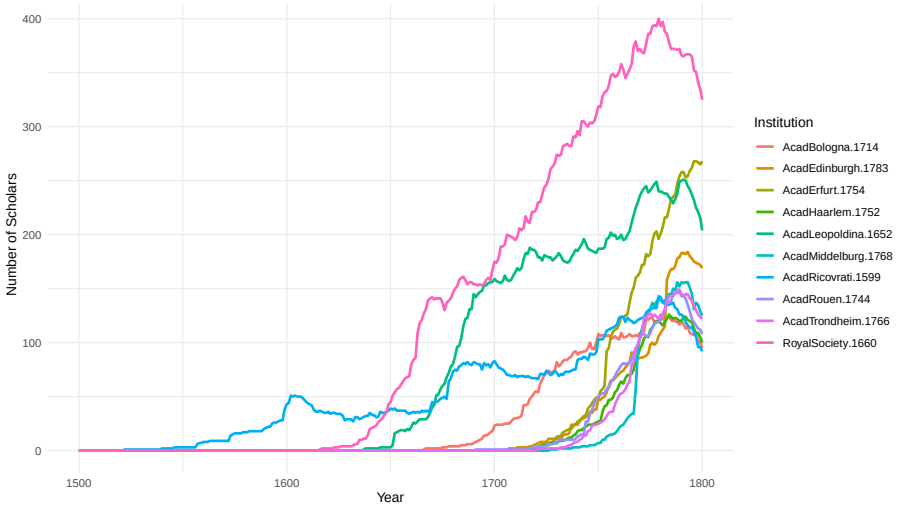



**Figure 3.A.5.** Number of academicians between 1500 and 1800 of the top 10 academies considering the number of scholars in 1793. In the legend, the name of the institution is given with the foundation date.

| Institution | Mean | Median | SD | Min | Q1 | Q3 | Max |
|---|---|---|---|---|---|---|---|
| Ubologna.1088 | 34.14 | 37 | 13.02 | 14 | 26.5 | 41 | 53 |
| Ucambridge.1209 | 69.57 | 105 | 60.38 | 1 | 8.5 | 113.5 | 137 |
| Ucoimbra.1308 | 8.29 | 4 | 13.70 | 0 | 0 | 8 | 38 |
| Uerlangen.1742 | 9.71 | 0 | 17.38 | 0 | 0 | 12.5 | 43 |
| Ugottingen.1734 | 22.86 | 0 | 41.23 | 0 | 0 | 28.5 | 103 |
| Ukiel.1652 | 12.71 | 4 | 16.66 | 0 | 0 | 20.5 | 44 |
| Ucopenhagen.1475 | 17.57 | 22 | 17.45 | 0 | 1 | 27 | 45 |
| Uleipzig.1409 | 21.29 | 20 | 19.17 | 0 | 7 | 33.5 | 48 |
| Ulund.1666 | 13.14 | 1 | 17.67 | 0 | 0 | 25.5 | 40 |
| Umarburg.1527 | 16.43 | 18 | 16.35 | 0 | 0.5 | 30 | 36 |

**Table 3.A.1.** Descriptive statistics for the top 10 universities, considering the number of scholars in 1793. In the first column, the name of the institution is given with the foundation date. "SD" stands for Standard Deviation, "Min" for minimum, "Q1" for first quartiles, "Q3" for third quartiles, and "Max" for maximum.

| Institution | Mean | Median | SD | Min | Q1 | Q3 | Max |
|---|---|---|---|---|---|---|---|
| AcadBologna.1714 | 32.29 | 0 | 48.17 | 0 | 0 | 59 | 108 |
| AcadEdinburgh.1783 | 31 | 0 | 63.43 | 0 | 0 | 24 | 169 |
| AcadErfurt.1754 | 45.29 | 0 | 99.89 | 0 | 0 | 24.5 | 268 |
| AcadHaarlem.1752 | 18.14 | 0 | 37.47 | 0 | 0 | 13.5 | 100 |
| AcadLeopoldina.1652 | 79 | 3 | 98.48 | 0 | 0 | 173 | 204 |
| RoyalSociety.1660 | 123.43 | 45 | 149.10 | 0 | 0 | 247 | 325 |
| AcadRicovrati.1599 | 45.29 | 39 | 46.78 | 0 | 0 | 87.5 | 103 |
| AcadRouen.1744 | 22.71 | 0 | 41.94 | 0 | 0 | 25.5 | 108 |
| AcadTrondheim.1766 | 21 | 0 | 45.50 | 0 | 0 | 12.5 | 122 |
| AcadMiddelburg.1768 | 18.86 | 0 | 46.88 | 0 | 0 | 3.5 | 125 |

**Table 3.A.2.** Descriptive statistics for the top 10 academies, considering the number of scholars in 1793. In the first column, the name of the institution is given with the foundation date. "SD" stands for Standard Deviation, "Min" for minimum, "Q1" for first quartiles, "Q3" for third quartiles, and "Max" for maximum.

academy. If the exact affiliation date is unavailable, we infer the first year of affiliation from approximate dates. In more extreme cases, we use the earliest available date among: 30 years after the birth year, the year of death, the institution's closing date, or 1793, which marks the end of our study period. This approach aims to provide a conservative estimate of each scholar's active period.

Scholars cease to be active when they leave the institution, if this date is available. For university professors, it is the year their teaching ends. For members of academies, it is usually the death year, or in some rare cases, the year the member was expelled. When there is no precise information about the end of their activity, we infer it in one of two ways for university professors and academicians. For university professors without a precise end affiliation date, we assume it is equal to the approximate affiliation date when available. Otherwise, if these pieces of information are unknown, we assume that a university professor will teach in that university for eight years.[23] Hence, we take the earliest date between the beginning date (after adding eight years) and the year of death. For academicians without a precise end date of their

[23]Eight years being the median of the affiliation years in the sample of university professors for whom we know the precise beginning and end affiliation dates. This data is consistent with the literature: Koschnick (2025b) finds that the median length of academic careers at Oxford and Cambridge is 9 years.

affiliations, we also assume it from more approximate dates if available. When not possible, for these scholars, we assume a lifelong affiliation,[24] so we take the year of death when there is no more precise information. Otherwise, when we do not know the year of death either, we assume they stay in the academy only one year, imposing the end affiliation date equal to the beginning date.

## 3.C Contextual information

### 3.C.1 Botanical Realism and botanic gardens

In Europe, natural history traces its roots back to ancient Greek philosophers such as Aristotle, Theophrastus, and Dioscorides. During the Scientific Revolution, botany underwent major advancements, transitioning from a primarily descriptive field into a more systematic and experimental science. By the $16^{th}$ and $17^{th}$ centuries, botany started encompassing not only the identification and classification of plant species, but also the growing field of plant physiology, investigating the properties and functions of plant life. This marked a shift in botanical practices, expanding beyond the descriptive and illustrative focus of ancient authors (Applebaum, 2003), towards a more empirical approach that we will refer to as "Botanical Realism". Before 1650, botany was considered merely a complement to medical studies, but it became an independent field of study during the Scientific Revolution. Universities renown for their medical faculties began offering innovative botany lectures, where students were taken directly to gardens to observe plant species first-hand. These universities were also the first to establish their own botanic gardens to support further research and development in botany. Following this trend, private citizens and local lords also recognized the importance of botanical studies and funded the creation of such gardens (Applebaum, 2003).

One key figure in this transformation was Leonhart Fuchs, a German physician and botanist. He is best known for his book *De historia stirpium commentarii insignes*, which translates to "Notable commentaries on the history of plants". First printed in Basel in 1542, one year before Nicolaus Copernicus' *De revolutionibus orbium coelestium* and Andreas Vesalius' *De humani corporis fabrica*, this work laid the foundation for modern botany. Fuchs not only provided ideal visual representations of 511 plant species, but

[24] Academies usually grant a lifelong affiliation.

he also included his own critical observations on their uses and characteristics, highlighting differences from ancient texts (Applebaum, 2003). Figure 3.C.1 shows a page of this book, emphasizing the realistic description of a plant.




**Figure 3.C.1.** A page of *De historia stirpium commentarii insignes*

Furthermore, we gathered information on the existence and founding dates of European botanic gardens. Our starting point was the first annual report of the Montreal Botanic Garden, (1886), which lists the botanic gardens open worldwide in 1885. From this, we selected only European gardens and determined their founding dates using AI-assisted tools, which were then manually verified through sample checks. We then matched this sample of botanic gardens with our university cities, assuming that a city without a botanic garden was not listed in the first annual report by Montreal Botanic Garden, (1886). To fix ideas, Figure 3.C.2 shows the *Hortus botanicus* (botanic garden) of the University of Leiden, which was opened in 1590. This is a case of a large garden with a dedicated building, but some other gardens were much smaller.

## 3.C.2 Mathematical Astronomy and astronomical observatories

The 15th and 16th centuries witnessed a growing interest in experimental science and increasing dissatisfaction with the explanations offered by ancient astronomical authorities, such as Claudius Ptolemy (c. 100–c.170, Alexandria). Similar to the approach seen in Botanical Realism, this paradigm shift led mathematicians and astronomers to question the accuracy of Ptolemy's models and refine them through observation and mathematical analysis. This era

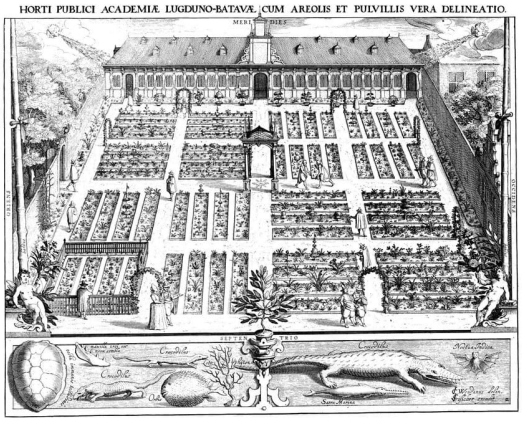



**Figure 3.C.2.** The Hortus botanicus of Leiden

marked the beginning of a mathematical reform in astronomy, characterized by advances in trigonometry, new geometric formulas, and the adoption of decimal calculations. The focus shifted from simply explaining celestial motions to understanding the physical mechanisms behind them.

A key figure in this revolution was Regiomontanus, pseudonym of Johannes Müller. His mastery of Greek and mathematics enabled him to study the original works of Ptolemy and other ancient thinkers. At the University of Vienna, around 1454, he and his mentor, Georg Peurbach (1423–1461), began collaborating on *Theoricæ novæ planetarum*. This seminal work introduced new methods for solving plane and spherical trigonometry problems, including the use of sine and tangent functions. Regiomontanus also created extensive trigonometric tables with values calculated to decimal units, which remained influential for centuries. As such, he can be considered a pioneer of Mathematical Astronomy. While trigonometry had been used in astronomy and other sciences, Regiomontanus's contributions greatly enhanced its application. His work, alongside Peurbach's, laid the groundwork for later revolutionary astronomers such as Copernicus, Kepler, and Galileo (Applebaum, 2003).

After Peurbach's death, Regiomontanus moved to Northern Italy, then to Hungary in 1467. Later, he settled in Nuremberg, drawn by its status as a free city and central location. There, he established a workshop and printing press, dedicating himself to the dissemination of scientific knowledge. In 1463, he published *Epitoma in Almagestum Ptolemaei*, which clarified, corrected,

and expanded Ptolemaic astronomy. Two pages of *Epitoma* are shown in Figure 3.C.3. In 1472, he published *Theoricæ novæ planetarum*, his collaboration with Peurbach. In 1475, Pope Sixtus IV invited him to Rome to work on calendar reform, but Regiomontanus died shortly after, at age 41.

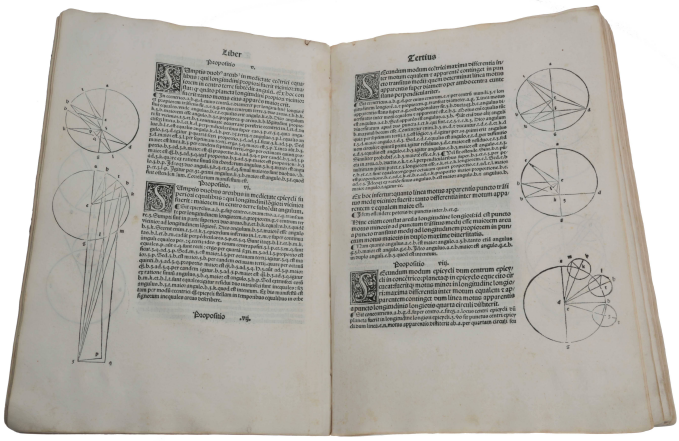


**Figure 3.C.3.** Two pages of *Epitoma in Almagestum Ptolemaei*

We examine the creation of observatories in the same 185 university cities defined in Section 3.3.2. We collected the names and foundation dates of observatories from *The Greenwich List of Observatories* compiled by Howse (1986). As with the botanic gardens, we only considered observatories in continental Europe, assuming that a location lacked one if it was not listed in our source. To fix ideas, Figure 3.C.4 shows the main building of the University of Prague, the Clementinum. There, a tower was built in 1722. Later, in 1751, instruments were installed, and the tower became an astronomical observatory.

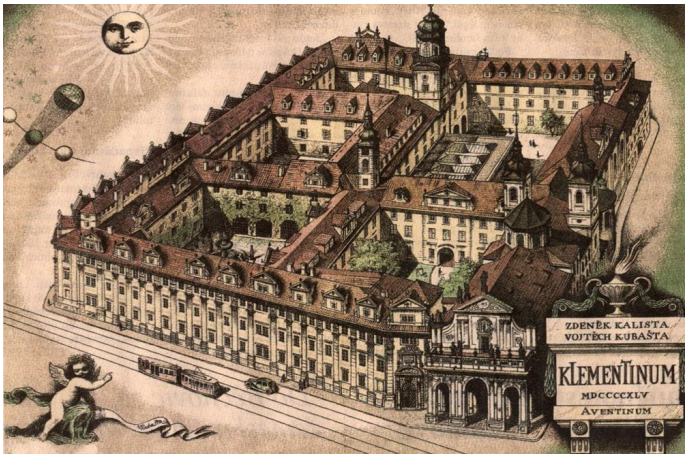



**Figure 3.C.4.** The *Clementinum* in Prague with its observatory

# 3.D Empirical Assessments

## 3.D.1 Descriptive Statistics

Table 3.D.1 presents some descriptive statistics, which supports the results in the main text, Table 3.2 and Table 3.3. Here we consider $\hat{\alpha} = 0.45$.

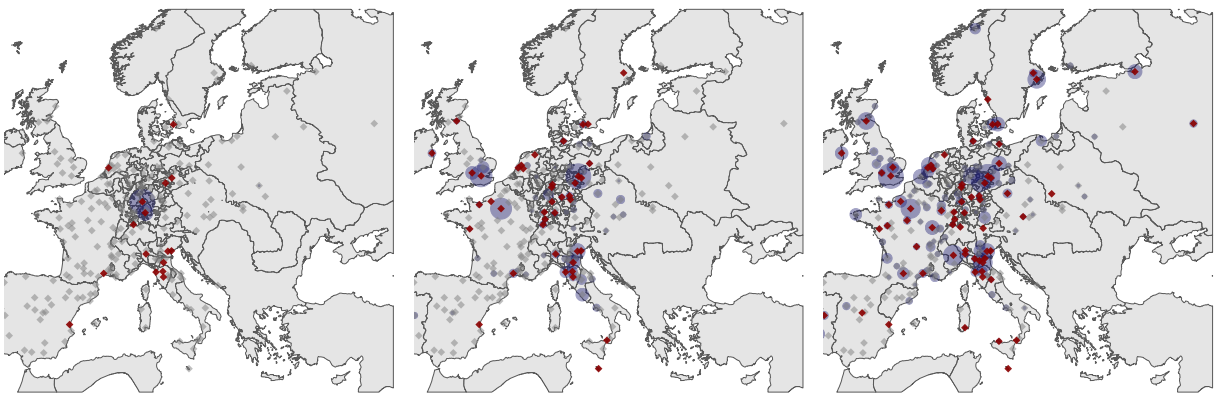
**Table 3.D.1.** Summary statistics.

| Variable | Obs | Mean | Median | SD | Min | Max |
|---|---|---|---|---|---|---|
| Prevalence of Botanical Realism | 67620 | 0.084 | 0 | 0.172 | 0 | 1 |
| (ihs) Eligible Mass of Botanical Realism | 67620 | 0.928 | 0 | 1.302 | 0 | 6.471 |
| Prevalence × Eligible Mass of Botanical Realism | 67620 | 0.226 | 0 | 0.522 | 0 | 2.856 |
| Prevalence of Mathematical Astronomy | 67620 | 0.120 | 0 | 0.288 | 0 | 0.843 |
| (ihs) Eligible Mass of Mathematical Astronomy | 67620 | 0.619 | 0 | 1.063 | 0 | 6.054 |
| Prevalence × Eligible Mass of Mathematical Astronomy | 67620 | 0.298 | 0 | 0.813 | 0 | 5.103 |
| (ihs) City population in 1500 | 67620 | 2.531 | 2.492 | 1.243 | 0 | 5.521 |
| (ihs) Distance to Tübingen | 67620 | 7.026 | 7.101 | 0.823 | 0 | 8.338 |
| (ihs) Distance to Vienna | 67620 | 7.388 | 7.466 | 0.748 | 0 | 8.434 |

*Notes*:(ihs) refers to the transformation in inverse hyperbolic sine of the relative variable.

## 3.D.2 Botanical Realism

Figure 3.D.1 shows yearly exposure to Botanical Realism at three points in time. It visualizes how exposure evolves over time and space, from a concentration around Tübingen in 1600 to a spread across Europe by 1700, and a further expansion to smaller and more distant urban centers by 1793. Blue bubbles represent exposure to Botanical Realism, while red diamonds indicate cities that had at least one botanic garden by the respective year.

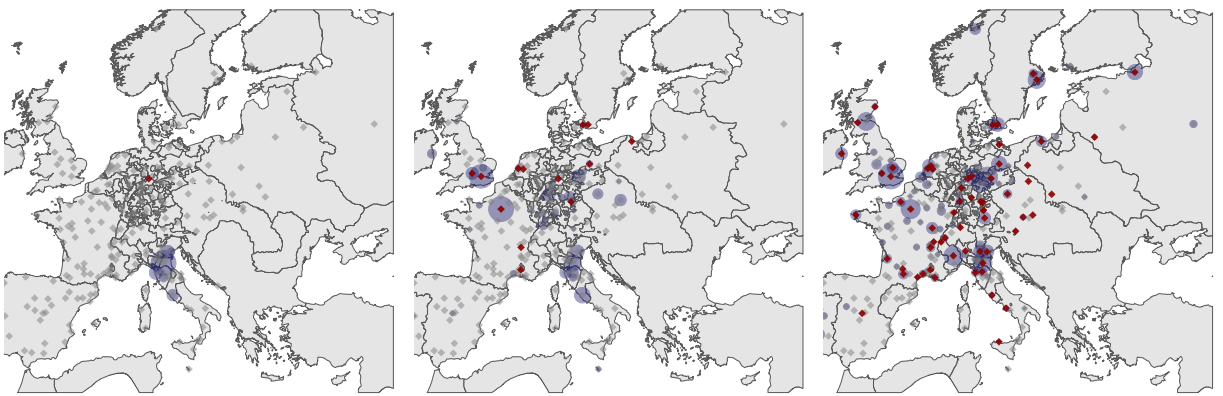


**(a)** Year 1600 **(b)** Year 1700 **(c)** Year 1800

**Figure 3.D.1.** Blue bubbles represent the yearly exposure of Botanical Realism in years 1600, 1700, and 1793, respectively. $\alpha = 0.45$ and $D = 5,000$. Cities with a botanic garden are the red diamonds, and without botanic gardens are the grey diamonds.

### 3.D.3 Mathematical Astronomy

Figure 3.D.2 visualizes the sample of university cities with their yearly exposure to Mathematical Astronomy in 1600, 1700, and 1793. This figure shows how the exposure evolves over time and its interaction with the creation of observatories. As in the first experiment, blue bubbles represent exposure to Mathematical Astronomy, while red diamonds indicate cities that had at least one observatory in the relative year.



**(a)** Year 1600 **(b)** Year 1700 **(c)** Year 1800

**Figure 3.D.2.** Blue bubbles represent the yearly exposure of Mathematical Astronomy in years 1600, 1700, and 1793, respectively. $\alpha = 0.45$ and $D = 5,000$. University cities with an astronomical observatory are the red diamonds, and without astronomical observatories are the grey diamonds.

## 3.E Structural estimation

Table 3.E.1 reports the log-likelihoods for each level of $\alpha$ relative to both Botanical Realism (Flora, column 1) and Mathematical Astronomy (Cosmos, column 2). Column 3 shows the joint likelihood $\ell_1(\alpha) + \ell_2(\alpha)$, confirming that the maximum is reached at $\hat{\alpha} = 0.45$. Figure 3.E.1 confirms it visually. Finally, Table 3.E.2 shows the results of the Cox models with $\hat{\alpha} = 0.45$ for both Flora and Cosmos.

**Table 3.E.1.** Log-likelihood values for different values of $\alpha$.

| $\alpha$ | **Flora** $\ell_1$ | **Cosmos** $\ell_2$ | **Sum** $\ell_1(\alpha) + \ell_2(\alpha)$ |
|---|---|---|---|
| 0 | -343.178 | -324.834 | -668.012 |
| 0.01 | -342.610 | -313.493 | -656.103 |
| 0.02 | -342.625 | -313.633 | -656.258 |
| 0.03 | -342.623 | -314.785 | -657.408 |
| 0.04 | -342.648 | -315.396 | -658.044 |
| 0.05 | -342.671 | -315.552 | -658.223 |
| 0.10 | -342.511 | -316.084 | -658.595 |
| 0.15 | -339.284 | -316.661 | -655.945 |
| 0.20 | -337.442 | -316.995 | -654.437 |
| 0.25 | -336.348 | -317.147 | -653.495 |
| 0.30 | -335.757 | -317.216 | -652.973 |
| 0.35 | -335.453 | -317.242 | -652.695 |
| 0.40 | -335.330 | -317.262 | -652.592 |
| **0.45** | -335.286 | -317.282 | **-652.568** |
| 0.50 | -335.280 | -317.308 | -652.588 |
| 0.55 | -335.289 | -317.342 | -652.631 |
| 0.60 | -335.300 | -317.385 | -652.685 |
| 0.65 | -335.310 | -317.432 | -652.742 |
| 0.70 | -335.315 | -317.488 | -652.803 |
| 0.75 | -335.315 | -317.562 | -652.877 |
| 0.80 | -335.312 | -317.645 | -652.957 |
| 0.85 | -335.307 | -317.735 | -653.042 |
| 0.90 | -335.299 | -317.832 | -653.131 |
| 0.95 | -335.288 | -317.939 | -653.227 |
| 1.00 | -335.276 | -318.044 | -653.320 |

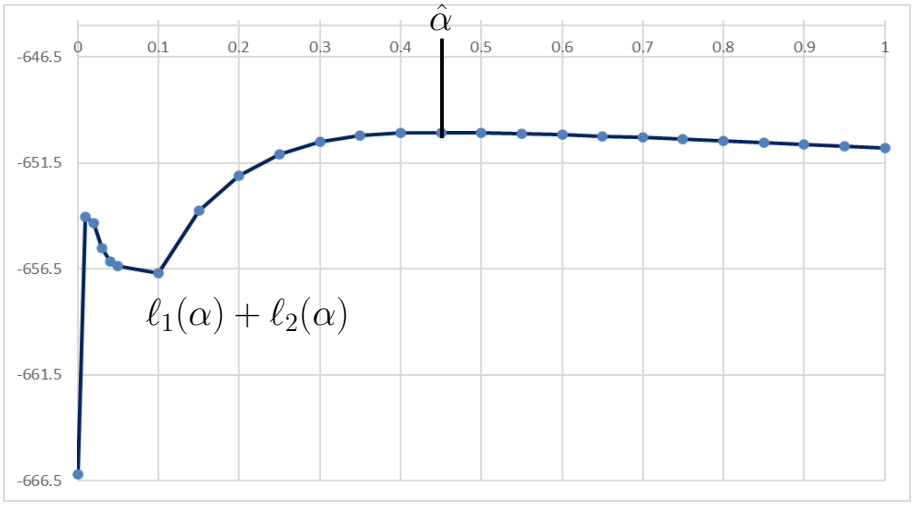


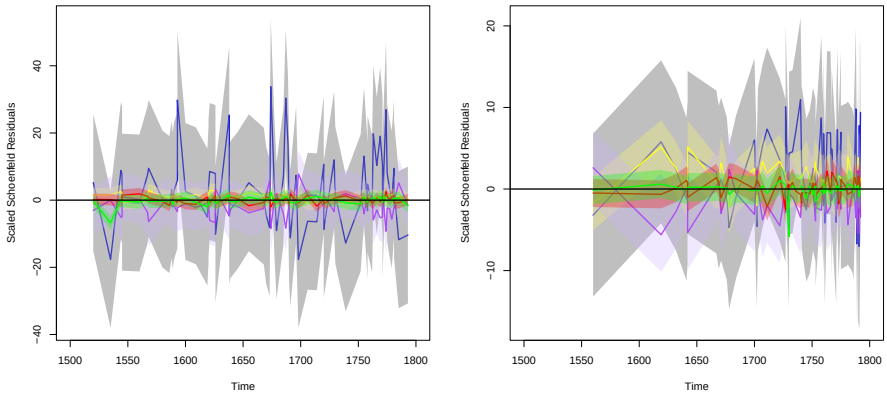

**Figure 3.E.1.** Plot of log-likelihood values for different values of $\alpha$.

**Table 3.E.2.** Cox Proportional Hazards Models for alpha = 0.45

| | Botanical Realism & botanic gardens | Mathematical Astronomy & astronomical observatories |
|---|---|---|
| $\alpha$ | 0.45 | 0.45 |
| Prevalence $P_t^{k,\varphi}$ | 3.61*** | 1.93*** |
| | (1.23) | (0.56) |
| (ihs) Eligible Mass $M_t^{k,\varphi}$ | 0.90*** | 1.12*** |
| | (0.10) | (0.19) |
| Prevalence $P_t^{k,\varphi}$ × (ihs) Eligible Mass $M_t^{k,\varphi}$ | -1.70*** | -1.13*** |
| | (0.35) | (0.25) |
| (ihs) Distance to Tübingen | -0.11 | |
| | (0.08) | |
| (ihs) Distance to Wien | | -0.10** |
| | | (0.04) |
| Log Likelihood $\ell(\alpha)$ | -335.286 | -317.282 |
| (ihs) Pop in 1500 | ✓ | ✓ |
| Observations | 67620 | 67620 |

## 3.F Tests for the proportionality of Hazard Functions

In this section we test the proportionality of the hazard functions (e.g.,scaled Schoenfeld residuals) of all the covariates of Table 3.2 and Table 3.3, respectively, and time, with $alpha = 0.45$. The global correlation is only slightly significant in the case of Botanical Realism, while it is not at all significant for Mathematical Astronomy. We can see that by looking at the confidence intervals almost always overlapping with the zero line. The blue line and grey confidence interval relate to the hazard ratios of Prevalence $P_t^{k,\varphi}$ of Botanical Realism (Fig 3.F.1a) and of Mathematical Astronomy (Fig 3.F.1b), respectively. The yellow line and confidence interval correspond to the hazard ratios of Eligible Mass $M_t^{k,\varphi}$ of Botanical Realism (Panel a) and of Mathematical Astronomy (Panel b). The purple line and confidence interval relate to the hazard ratios of the interaction term between Prevalence $P_t^{k,\varphi}$ and Eligible Mass $M_t^{k,\varphi}$. The red line and confidence interval correspond to the scaled Schoenfeld residuals of "(ihs) City population in 1500", while the green line and interval of confidence represent the hazard ratios of "(ihs) Distance to Tübingen" (Fig 3.F.1a) and "(ihs) Distance to Vienna" (Fig 3.F.1b), respectively.




**(a)** Botanical Realism **(b)** Mathematical Astronomy

**Figure 3.F.1.** Joint correlations between the scaled Schoenfeld residuals (e.g. hazard ratios) of all the covariates in column (3) of (a) Table 3.2 and time, and (b) Table 3.3 and time.

## 3.G Example of scholastic reasoning

To fix ideas, a good example of reasoning using the tools of scholastic theology is the second proof of the existence of God by Aquinas as reported by Copleston (1993). Remark that it does not rely much on the scriptures, but rather on a kind of mathematical/logical argumentation:

1. In the world, we can see that things are caused.
2. But it is not possible for something to be the cause of itself because this would entail that it exists prior to itself, which is a contradiction.
3. If that by which it is caused is itself caused, then it too must have a cause.
4. But this cannot be an infinitely long chain, so there must be a cause which is not itself caused by anything further.
5. This everyone understands to be God.

## 3.H Augmenting Rubin's data

Rubin (2014) compiled data on whether European cities were Protestant in 1530, 1560, and 1600, focusing primarily on cities within the Holy Roman Empire. His approach assumed that cities located in officially Catholic countries remained Catholic throughout the period. However, this assumption does not hold for regions such as France and the Low Countries (approximately corresponding to the modern-day Benelux area), where several cities adopted Protestantism temporarily before being reconquered by Catholic forces.

To address this limitation, we have updated the religious status of the following cities to reflect periods of Protestant control, based on information from the Catholic Encyclopedia (Herbermann, 1913).

- Die. Protestant control: 1562-1628/1629. Became a Protestant stronghold in Dauphiné during the Wars of Religion. Occupied by royal troops during Richelieu's repression of Huguenot fortresses.
- La Rochelle. Protestant period: 1550s-1628. Key Huguenot stronghold. Besieged and subdued by royal forces under Richelieu in 1628.
- Montauban. De facto Protestant Rule: 1561-1629. Montauban became one of the most fortified and independent Huguenot cities in France.

Finally capitulated (1629) to royal troops under Richelieu after the fall of La Rochelle and a renewed campaign to suppress Huguenot political autonomy.

- Montpellier. Protestant control: 1562-1622. Served briefly as a de facto capital of Huguenot political assemblies. Lost military and political autonomy after the Siege of Montpellier (1622) by royal troops under Louis XIII.

- Nîmes. Protestant control: 1561-1629. After violent iconoclasm in 1561, the city came under Huguenot control. Held by Protestants throughout the Wars of Religion; formally lost autonomy in 1629 after Richelieu's campaigns.

- Ostend. In 1572, Ostend joined the Dutch Revolt and came under the control of the rebel States-General of the Netherlands, aligning with Protestant (Reformed) forces. After the Siege of Ostend (1601-1604), the city was almost entirely destroyed, and Catholicism was restored under Spanish Habsburg rule.

- Uzès. Protestant control: ca. 1562-1629. Important center in the Languedoc region. Remained predominantly Protestant until submission to royal forces in 1629, when Richelieu dismantled Protestant strongholds.

For the sake of completeness, we list here the cities with Protestant control outside Rubin's years 1530,1560, and 1600.

- Antwerp. Protestant period: ca. 1566-1585. During the Reformation, Antwerp became a Calvinist stronghold. After the Fall of Antwerp in 1585, Protestant worship was banned, and many Protestants fled north.

- Caen. Protestant control: 1562-1572. Important Protestant stronghold in Normandy. The St. Bartholomew's Day Massacre (1572) led to widespread killings and ended Protestant rule.

- Ghent. Protestant period: ca. 1577-1584. Calvinist Republic of Ghent established during the Dutch Revolt. Ended with Spanish reconquest.

- Tournai. Protestant control: 1577-1581. Strong Calvinist presence; fell to Spanish forces in 1581.

# 3.I Additional results with alternative ideas: Practical Surgery and Civil Law

In addition to the benchmark results in Table 3.4, we present results controlling for two specific, orthogonal ideas: *Exposure to Practical Surgery* $PS^c_{1508}$ and *Exposure to Civil Law* $CL^c_{1508}$. This is interesting because it captures a distinct effect, separate from scholasticism. We compute exposure to Practical Surgery and Civil Law similarly to scholasticism, with the only difference being the starting point. The "inventor" of the former, Guy de Chauliac (c. 1300–1368), was the most prominent physician and surgeon of the Middle Ages. His most famous work *Chirurgia Magna*, published in 1363, was the first to detail surgical procedures, previously handled mostly by charlatans. It remained the main reference well into the $17^{\text{th}}$ century (The Editors of Encyclopaedia Britannica, 2024). The inventor of the latter idea, Irnerius, already mentioned in the paper's Introduction, established the school of jurisprudence which eventually grew into the University of Bologna. He was a jurist credited with reviving the scholarly study of Roman civil law, knowledge of which had endured in the Byzantine East but largely faded in Western Europe following the fall of the Western Roman Empire (Berman, 1983).

The orthogonal relationship between scholasticism and Practical Surgery is evident in Figure 3.I.1: Practical Surgery was more prominent in Southern Europe, spreading independently of scholasticism. Similarly, Figure 3.I.2 shows that exposure to Civil Law followed its own geographic pattern, radiating outward from Bologna and concentrated in regions where the revival of Roman jurisprudence had the deepest institutional roots. In the linear probability model shown in the first placebo panel of Table 3.4, *Exposure to Practical Surgery* $PS^c_{1508}$ always has a negative sign, but it is significant only in 1560 and 1600 in the specifications without controls and fixed effects, and only in 1560 when controls and fixed effects are included. The results for Civil Law are equally unambiguous: as shown in the second placebo panel of Table 3.4, *Exposure to Civil Law* $CL^c_{1508}$ is never statistically significant in any specification and at any date. We interpret this as further evidence that the relationship documented above is specific to scholasticism: exposure to other intellectual traditions rooted in medieval universities—one medical, one juridical—bears no association with a city's likelihood of becoming Protestant. This reinforces our hypothesis of a "disgust" effect triggered by scholasticism,

rather than by university culture or intellectual activity in general.

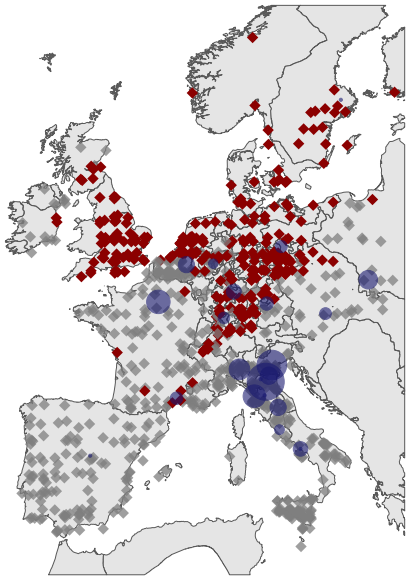


**Figure 3.I.1.** Blue bubbles represent the yearly exposure to Practical Surgery in 1508. $\alpha = 0.45$ and $D = 5,000$. Protestant cities are the red diamonds, and Catholic cities are the grey diamonds. Data on cities' religion are taken from Rubin (2014) and updated as in Appendix 3.H; they reflect the religious status in 1600.

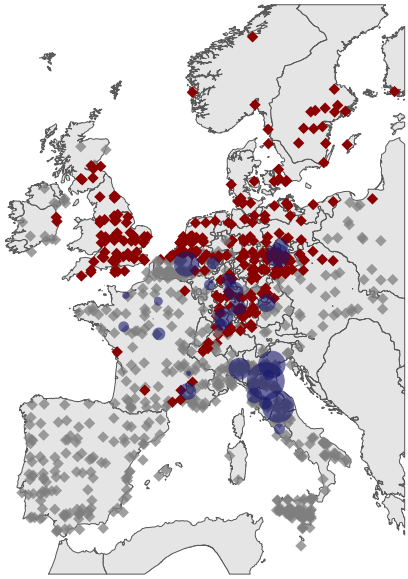


**Figure 3.I.2.** Blue bubbles represent the yearly exposure to Civil Law in 1508. $\alpha = 0.45$ and $D = 5,000$. Protestant cities are the red diamonds, and Catholic cities are the grey diamonds. Data on cities' religion are taken from Rubin (2014) and updated as in Appendix 3.H; they reflect the religious status in 1600.

## 3.J Placebo Inventors of Botanical Realism

Here, we list the twelve scholars we use as "placebo" inventors in the counterfactual analysis in Section 3.5.1. While it is speculative to say whether each of these twelve individuals could have invented Botanical Realism, many of them were indeed prominent scholars in fields that could have contributed to the development of a more empirical approach to botany. However, the emergence of a paradigm like Botanical Realism depended on a combination of factors–intellectual, cultural, and scientific–beyond the work of individual scholars. Below, a closer look at the potential of each of the individuals:

- Juan Aguilera was professor of medicine and sciences at the University of Salamanca from 1538 to 1560 (Vidal y Díaz et al., 1869; Esperabé de Arteaga et al., 1917): as he had a background in natural philosophy or medicine, he could have contributed to a more empirical study of plants,

as Salamanca was a leading university with a strong focus on scientific inquiry during the Renaissance.

- John Warner taught medicine at the University of Oxford from 1520 to 1554 (Gunther, 1937) and was a member of the Royal College of Physicians (1561): he might have had access to Renaissance humanist ideas, but Oxford was more conservative at the time, and Warner would need a strong inclination toward natural science to spearhead Botanical Realism.

- Jeremius Dryvere taught medicine at the University of Louvain from 1522 to 1554 (Lamberts and Roegiers, 1990): Louvain was a center of scientific learning, and someone like Dryvere could have contributed to botanical studies.

- Andreas Goldschmidt taught medicine at the University of Königsberg from 1550 to 1559 (Schwinges and Hesse, 2019): As a scholar trained in Wittenberg, where humanism and scientific inquiry were encouraged, Goldschmidt could have been part of the intellectual currents that led to developments like Botanical Realism.

- Mikołaj Mleczko Wieliczki was professor of medicine at the University of Cracow from 1512 to 1552 (Uniwersytet Jagielloński, 2019): Cracow had a strong tradition in astronomy and natural sciences, and a scholar like Wieliczki could have contributed to the empirical study of nature.

- Jacob Bording was professor of medicine at the University of Rostock from 1549 to 1556 and at the University of Copenhagen from 1556 to 1560 (Slottved, 1978): as a prominent physician, Bording would likely have been interested in botany as it related to medicine, which was a key motivator for many early botanists.

- Antoine Saporta was professor of medicine at the University of Montpellier from 1531 to 1573 (Dulieu, 1979): Montpellier was a leading medical school, and Saporta, as a physician, would have had a strong interest in medicinal plants. He could have been well-positioned to develop a more scientific approach to botany.

- Girolamo Donzellini taught medicine at the University of Padua from 1541 to 1543 (Facciolati, 1757): The University of Padua was a hub

of medical and scientific learning, so Donzellini, with his interest in medicine, might have had the right environment to develop Botanical Realism.

- Oronce Fine taught sciences at the Royal College in Paris from 1530 to 1555 (Collège de France, 2018) : Although primarily a mathematician and cartographer, Fine was part of a broader Renaissance movement that emphasized empirical study, and he could have contributed to a more systematic approach to botany.

- Realdo Colombo taught medicine at the universities of Padua (1538–1544), Pisa (1544–1548) and Roma (1548–1559), see Del Negro (2015): he was a noted anatomist, and his empirical methods in anatomy could have translated well into botany, particularly in the detailed study of plant structures.

- Georg Joachim Porris taught sciences at the universities of Wittenberg (1537–1542), Leipzig (1542–1551) and Vienna (1554–1555), see Schwinges and Hesse (2019) and von Aschbach (1865). Also known as Rheticus, Porris was an astronomer and mathematician. While not a botanist, his scientific mindset might have inclined him toward an empirical approach in natural studies if he had turned his attention to plants.

### 3.J.1 Centrality of the Placebo inventors

The likelihood of an idea spreading depends on how well-connected its inventor is, typically measured by degree centrality—the number of edges a node has. However, degree centrality (and other centrality measures) does not fully capture the dynamics of idea diffusion, which unfolds over time rather than at a single moment. A high degree centrality does not necessarily equate to greater reach. To illustrate this, we report the degree centrality of each placebo inventor in 1542 (i.e., the number of colleagues in the same field) or in the first available year of activity.

If Fine (2) or Colombo (19) had been the originators, the idea would have already spread to around 50% of the academic population by the second half of 1600. The next fastest spread would have occurred with Dryvere (8) and Porris (22), followed by Saporta (6) and Bording (2 in 1550). For Warner (10), the spread would occur significantly later, only accelerating after the establishment of the first major academies, with a sharp increase around 1640. If Goldschmidt (0 in 1542, but 1 in 1549) had been the inventor, the idea would have struggled to survive initially, only gaining rapid traction around 1680. Interestingly, had Aguilera (8) been the inventor, the idea would have remained confined to Salamanca, persisting without spreading elsewhere until the end of the 18th century, when we observe a sudden spike. This shift coincides with scientists from the Spanish university beginning to affiliate with more international academies. In one case (Wieliczki (3) in Cracow), the idea fails to spread altogether. These simulations demonstrate how academic institutions can play a crucial role in preserving ideas that might otherwise remain obscure due to their development in less influential locations.

## 3.K Placebo networks and the role of the Jesuits

In the paper, we explore a counterfactual scenario examining the spread of ideas after removing Jesuit scholars from the network. Here, we present key statistics on the position and connectivity of Jesuits within the affiliation network. The network metric in Figure 3.K.1d quantifies how well-connected Jesuit-affiliated nodes are to the rest of the network. In our network structure, conductance reflects the extent to which Jesuits were integrated into mixed institutions rather than remaining isolated. The initially high conductance suggests that early Jesuits were present in diverse academic environments. Over time, the decline in conductance aligns with their increasing concentration within Jesuit institutions. Figure 3.K.2b shows the Homophily Index of Jesuits by field between 1556 and 1767. Notably, the decreasing trend in inbreeding homophily among Jesuit scientists is likely a consequence of the rise of academies, which were largely centered on scientific disciplines and saw a relatively active participation from Jesuits.

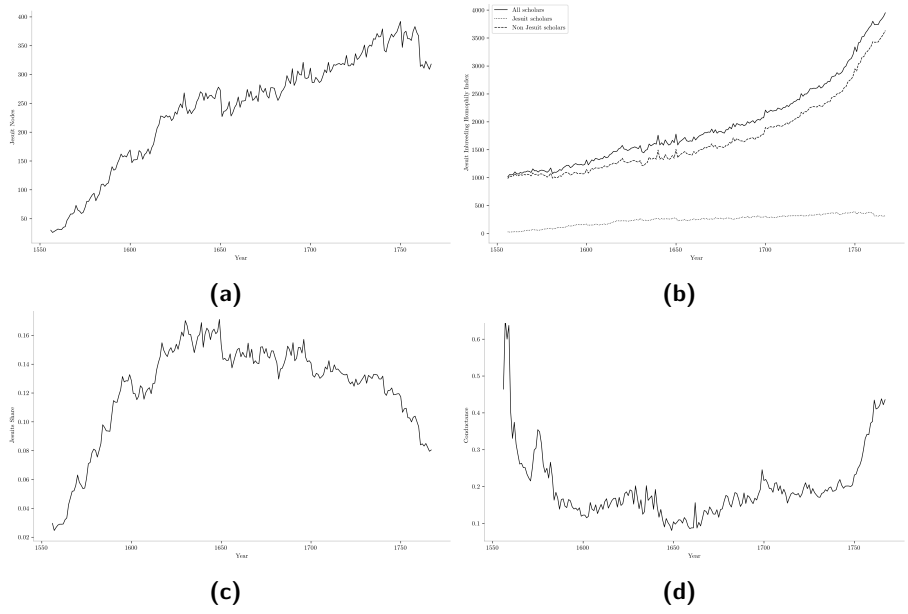



**Figure 3.K.1.** (a) Number of Jesuit scholars active in the network, 1556-1767. (b) Comparison of number of scholars by type: all (solid line), Jesuits (dotted line), and non-Jesuit scholars (dashed line),1556-1767. (c) Fraction of Jesuit scholars active in the network, 1556-1767. (d) Conductance of Jesuits in the affiliation network, 1556–1767.

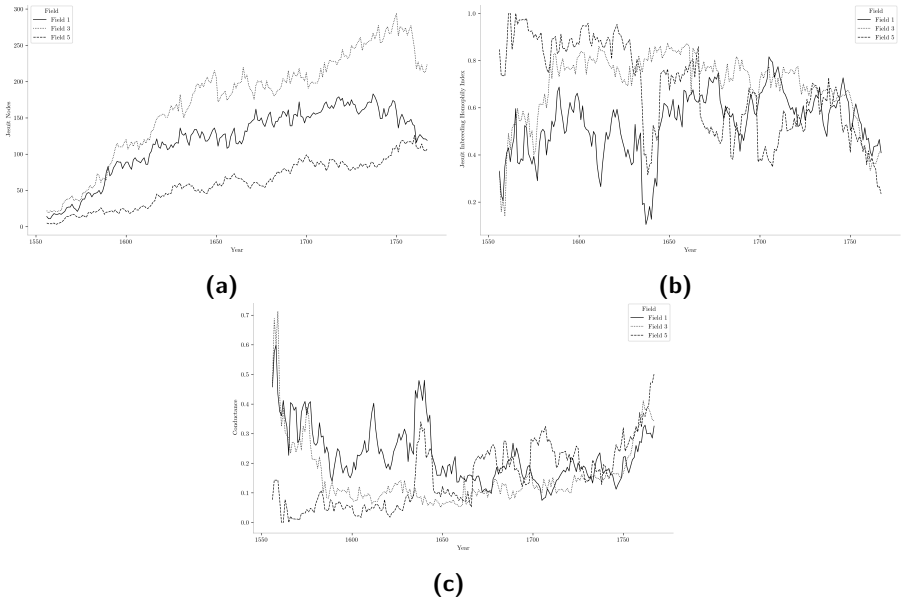



**Figure 3.K.2.** (a) Number of Jesuit scholars active in the network by field, 1556-1767. Field 1 stands for theology, field 3 for humanities, and field 5 for sciences. (b) Homophily Index of Jesuits, by field, 1556-1767. Field 1 stands for theology, field 3 for humanities, and field 5 for sciences. (c) Conductance of Jesuits in the affiliation network by field, 1556–1767. Field 1 stands for theology, field 3 for humanities, and field 5 for sciences.

# 4

# Lov'Ego: Printers and their Partnerships in the Academic Book Trade of Early Modern Louvain (1501–1797)

*This chapter is co-authored with Violet Soen.*

*In this paper, we examine the networks of five selected printers in the academic town of Louvain at pivotal moments across the three centuries of the early modern era. Our bibliographic network analysis enriched with prosopographical data combined with historical evidence reveals three distinct configurations. First, the networks around Martens and Rescius were locally anchored in academic circles, some (but not all) with pronounced humanist orientations. Second, the network of Masius was characterized by recurring, commission-based partnerships. Finally, the networks associated with Jansenist-leaning printers displayed strong spiritually driven affiliations, suggesting that religious orientations could outweigh purely economic considerations.*

## 4.1 Introduction

The ways in which printers and printing houses in university towns positioned themselves within the broader book market, and more specifically within local academic circuits, have been the subject of multiple recent contributions in the field of early modern book history.[1] This article revisits the relationships between printers, their partners and their portfolio for early modern Louvain, the first town in the delta region of Rhine, Meuse and Scheldt to host a university, established by papal charter in 1425 and eventually abolished by French decree in 1797 (Soen, 2026). Increasingly, scholarship has emphasized that university towns formed distinctive niches within the book economy, both before and after the advent of movable type in Europe: students and scholars generated sustained demand for textbooks and manuals, thesis sheets and defense booklets, humanist text editions as well learned treatises, while local printers sought to capitalize on these opportunities through specialization and close collaboration with the local academic community.[2] This new research agenda in book history has sparked significant novel contributions manuscript and print culture for early modern Louvain as well,[3] the first town in the delta region of Rhine, Meuse and Scheldt to host a university, established by papal charter in 1425 and eventually abolished by French decree in 1797 (Soen, 2026). However, the manner and extent in which printers operating within the same academic center differed in their partnerships, networks, and marketing strategies remains underexplored. At the intersection of economic history and book history, this article aims to address this gap by offering a data-driven complement to the recent studies of the academic book world in Louvain. Specifically, we examine the networks underpinning book production in early modern Louvain, and especially those between printers (whether or not also taking up a publishing role) and their partners (being authors, faculties,

[1] To name but some examples: Oosterhoff (2018); the contributions by Kirwan, Thomas, Walsby and Limbach in Pettegree, ed (2017); Blair (2010); Soen and Fantoli (forthcoming).

[2] Relating to manuscript culture and printed culture: Brinkmann et al., eds (2021); Facca, ed (2025).

[3] Feys et al., eds (2025); Van Der Biest (2025); Cammaerts (2025); Soen et al., eds (2025); Dekoninck and Soen, eds (2025); the Research Infrastructure https://www.studium-ai.org, a digital gateway to the University of Louvain (1425–1797).

colleges, etc.).[4] The advent of movable type transformed the costs and logistics of textual reproduction: early modern printshops could reuse type with unprecedented efficiency, distinguishing typographic printing from the labor-intensive practices of manuscript production and even xylography (woodblock printing, also in vogue in Europe from the second half of the fifteenth century). Yet, as Andrew Pettegree has shown, the technological advantages of movable type required substantial capital investment, with the publishers bearing significant financial risk and printers carrying out the job of printing itself, hence the reason why the two profiles gradually drifted apart (Pettegree, 2010).[5] Anthony Grafton, in turn, has emphasized that printing and publishing overall remained an artisanal craft, one in which 'the fingers got inky and dirty' mostly by printers, but sometimes by publishers, correctors and scholars: printing early books was time consuming, collaborative, and dependent on manual labor (from inking and setting type to drying sheets and correcting proofs) resulting in inevitable variations even within a single edition (Grafton, 2020). Handpress printing thus remained both capital and (manual) labor-intensive throughout the early modern era.

Even if the early modern book trade thus worked on proto-capitalist dynamics of supply, demand and investment, these were never 'free markets', be it through regulations of guild membership, or for academic printers specifically, by imposed matriculation at the university, thus by obedience to the rector and the university's rules.[6] However, the most explicit and implicit form of regulation might have been censorship. Long before the Reformation, and even more so after it, scholars operated within regulated environments that sought to control texts and ideas, both their content and their spread. After the Reformation, the Papacy and Catholic Universities contributed to curating orthodoxy by compiling 'Lists of Forbidden Books'. In Catholic regions, such as Louvain (even throughout the religious and civil wars of the sixteenth century), professors exerted preventive censorship through approbation (signaled by paratextual markers such as the *nihil obstat*, *placet* or *fiat*). These mechanisms imposed economic constraints, as printers,

[4]Constantinidou et al. (2025) brings an orienting state-of-the-art and bibliography.

[5]We acknowledge the complexity behind the figure of the 'publisher', a role often encompassing early modern printer-publishers but also a range of other individuals and institutions, which we do not pursue in the present article. See der Weduwen and Cullen (2025).

[6]Grendler (1977); for the Low Countries: Soen et al., eds (2017)

publishers and booksellers navigated shifting legal requirements that structured what could be printed, sold, or circulated (De Bujanda, 1986, 1990, 2002).[7] Despite such efforts, as we will also see further along this article, censorship regimes remained porous: legal pluralism and jurisdictional fragmentation created opportunities for negotiation, smuggling, and clandestine production, including the use of "fake imprints" to obscure a book's true origin. How much this regulatory pressure shaped the partnerships and portfolios printers build (and for which of them it mattered the most) is one of the aspects this article puts to the test.

By adopting a Social Network Analysis (SNA) approach to the question of partnerships in academic book markets, we examine not only the interplay of financial incentives, as proposed by economic historians, but also the local context and intellectual setting as proposed by historians of the book: as SNA enables to systematically observe patterns of social interactions, we can examine the structure of interpersonal contacts and intellectual affiliations at the First, or 'Old' University in Louvain, where from circa 1473 a printing press operated. Drawing on bibliographic metadata from the library collection holding the lion's share of rare books published by scholars of the university (*Collectio academica antiqua*, henceforth Caa) and the Short Title Catalog Vlaanderen (henceforth, STCV), we model patterns of specialization and collaboration within the Louvain book trade. The newly curated dataset Lov'Ego enables us to assess the interplay of commercial and confessional motives that structured the book trade of Louvain printers. As recent work in computational historical research has shown, bibliographic metadata, catalogs, and digitized full texts can serve as reliable and scalable proxies for studying long term trends in surviving book production, provided that their reconstructed character and selection biases are critically addressed(Lahti et al., 2019a; Tolonen et al., 2019; Valleriani and Ottone, eds, 2022; Scebba, 2026).

[7]In Protestant territories, censorship often came to operate post-publication, targeting rival confessional works; Tudor England, for example, adjusted the regulation of imported continental literature according to changing political and theological agendas.

## 4.2 Lov'Ego: Ego-networks of Selected Louvain Printers (1501-1797)

Introduced by Westphalian printers, the printing press reached Louvain in 1473–1474, and supported a flourishing book industry for nearly three centuries. The city developed an exceptionally dense publication network in the sixteenth century, maintaining a high output throughout the early modern period. Among its most renowned products is the 1516 *editio princeps* of Thomas More's *Utopia*, printed by Dirk Martens, whose humanist profile and program also attracted figures such as Desiderius Erasmus of Rotterdam, co-founder of the Collegium Trilingue, before his later move to Basel (François, 2020). The university's presence stimulated the local book trade, yet Louvain's scholarly and academic production also relied on the proximity of the dynamic presses of Antwerp, located within the same duchy of Brabant. Both centers operated under the shared system of Habsburg censorship, which required formal approval for all editions by appointed censors (in Louvain they were found next door at the Faculty of Theology), yet academic printers in Louvain also needed to operate under the supplementary statutes of the university, and as mentioned, obedience to the rector (Adam, 2017). Beyond Pierre Delsaerdt's seminal study published in 2001, a 2023 exhibition marking 450 years of printing in Louvain (with a focus on censorship) and two recent (yet until now unpublished) dissertations have expanded our understanding: Dieter Cammaerts on printed university handbooks in the curricula of the five faculties of the University of Louvain (with the corresponding dataset Manuale Lovaniense), and Heleen Wyffels on female agency within the printing communities of Louvain, Douai, and Antwerp, the first two being rival academic towns in the Habsburg Low Countries from 1562-3 onwards, until Douai fell to France in the seventeenth century.[8]

As documented most recently by the meticulous investigations of Dieter Cammaerts, in Louvain, too, printing and publishing for the academic textbook market required substantial investments well before any sales were made to students, scholars or their colleges, and the eventual demand for a book was

[8]Delsaerdt (2001); Cammaerts (2024b); and the corresponding dataset: Cammaerts (2024a); single records can be checked through http://www.odis.be; Wyffels (2021). The exhibition 'Kennis aan de ketting' ran at the KU Leuven Central Library from 6 October 2023 until 14 January 2024, and is now available as an online exhibition from KU Leuven Libraries: https://expo.bib.kuleuven.be/exhibits/show/kennisaandeketting.

very hard to foresee. Consequently, those involved in this trade operated under constant uncertainty about whether initial costs would be recouped, and in a second instance, whether any profit would be generated. In this contribution, therefore, we focus specifically on the partnerships between local printers (and/or publishers) and authors (including university professors yet excluding students who had their 'theses' printed for formal exercises, defenses or exams). The collaborations around these printed books emerged through economic motives and confessional (e.g. Catholic) agendas. For example, Louvain condemned Luther's works in respectively 1519 (as the Sorbonne would do in 1520) and proceeded with imperial backing to the burning of books on the Louvain Grand Square. Two decades later, the 1543 confiscation of Hieronymus Cloet's book shop in Louvain illustrates the material and financial risks faced by booksellers. Some years later only, the Faculty of Theology solicited Emperor Charles V to draw up a list of forbidden books, complemented with a list of books recommended for teaching children and adolescents. This lists of forbidden books, the *Index Prohibitorum*, appeared in 1546 and a revised one in 1550 and 1558, thus before the Papacy proceeded to its own Index in the year 1559 (De Bujanda, 1986). In turn, the restricted access to certain texts and created demand for illicit or underground editions in the Low Countries.

This paper works from the curated dataset Lov'Ego, compiled from two distinct datasets compiling (meta) data of early modern handpress books: the *Collectio academica antiqua* (Caa) and the STCV (Short Title Catalogue Vlaanderen), . The Caa dataset hosts the cataloguing metadata of an 'artificial' library collection curated by KU Leuven Libraries Special Collections, which focuses on rare books printed from 1501 onwards for the first University of Leuven (1425-1797). Assembled after the Second World War to acquire all rare books printed by or associated with scholars of the Old University of Louvain (1425-1797), the Caa was partially fragmented following the institutional division between KU Leuven and UCLouvain in the 1970s. The Caa has been assembled and reassembled, and the KU Leuven Libraries have certainly not been able to acquire all books within this scope of the collection's profile. Despite, or because of their general use, some textbooks seem to have been lost, while sometimes earlier 'lost books' can be identified through meticulous book historical research.[9] Despite these

[9]See on the topic Bruni and Pettegree, eds (2016) and Cammaerts (2024b) for the in-depth reconstruction of the universe of printed handbooks used at the Old University of Louvain between 1474 and 1650.

observations of specific curatorial decisions at KU Leuven, on the one hand, and general survival bias in the historical record, on the other, an earlier data-driven analysis on the Caa-dataset showed that it provides a more nuanced and representative view of Louvain's early modern intellectual network than might be anticipated, with notable figures like Erasmus not overly prominent in terms of centrality in the network drawn from the collection.(Scebba and Fantoli, 2026). With some exceptions,[10] the Caa does not contain the 'ephemeral' thesis sheets or booklets for rehearsal and/or exam disputations printed in Louvain, and as outlined above, we also excluded students as 'author' category.

In order to identify other possible books printed in Louvain but not present in the Caa at KU Leuven Libraries' Special Collections, we have supplemented the Caa with metadata retrieved from the large-scale retrospective bibliographies of handpress books (the 'Short Title Catalogues'). The Short Title Catalogue Vlaanderen (henceforth STCV) positions itself as the bibliographic repertory for the handpress book in Flanders in its current boundaries, thus excluding Walloon Flanders being part of the historical County of Flanders until the seventeenth century. The STCV, managed today by Bibliothecair Erfgoed.be, contains the descriptions of books printed in the historic Southern Low Countries and more specifically of the surviving copies from libraries in its heritage network (excluding Walloon libraries); the KU Leuven University Libraries (henceforth, KU Libraries) being one of them, housing the Caa in their Special Collections' Division, alongside the Maurits Sabbe Library attached to the Faculty of Theology and Religious Sciences (included in STCV, yet not in the Caa) (*STCV: Bibliography of the Hand Press Book*, n.d.). Covering the period from the advent of handpress printing to 1801, the STCV undoubtedly remains incomplete with respect to the full historical print output of the Southern Low Countries. Notably, the rare book holdings from the *Koninklijke Bibliotheek van België* (KBR) have not yet been fully integrated in the STCV, and holdings from non-Belgian or Walloon institutions are unevenly represented. Nevertheless, within the limits of

[10] An estimation of the number of theses (sheets) having ended up in the Collectio academica antiqua is difficult, as this information has not been standardized from a cataloguing point of view. Based on querying the titles and additional titles for the words 'defendet' and "præsid*", which seemed fitting common denominators and avoided noisy date, in the Caa (not including publications on the Lovaniensia platform; 2056 publications in total): about 80 theses (sheets) were found (circa 3.9%). On the Lovaniensia platform (2057 publications in total): about 35 thesis (sheets) were found (circa 1.7%).

survival and retrospective cataloging, the Caa and STCV combined remain the most exhaustive bibliographic resources currently available to systematically reconstruct Louvain's early modern print output.[11]

As an important methodological remark, the comparative design of the present study does not require an exhaustive bibliographical census of each printer's output; rather, it requires a corpus that is large enough and internally consistent enough to reveal structural differences between printers' networks. An incomplete but uniform dataset serves this purpose more reliably than a more complete but unevenly assembled one. In principle, Caa and STCV could be supplemented with records from union catalogs aggregating holdings from non-participating libraries, such as the Heritage of the Printed Book database (HPB) maintained by the Consortium of European Research Libraries (CERL). However, such integration poses significant methodological challenges. Union catalogues like the HPB aggregate copy-level catalogue entries from contributing libraries, each following its own cataloguing conventions and varying considerably in the depth of metadata recorded. A single edition may therefore appear as multiple separate records, while the metadata reported for each (particularly regarding contributors appearing on the title page) is inconsistent across sources. By contrast, the Caa and STCV apply uniform cataloging standards, record at the level of the edition rather than the individual copy, and systematically capture all contributors appearing on the title page – the very metadata on which our analysis is based. Moreover, Caa and STCV have a high degree of interoperability, which allows us to rely on the catalogers' expertise to identify the overlapping holdings properly, and are available in exports for bulk processing, ensuring the reproducibility for a large-scale network analysis like the one undertaken here.[12]

Today, the dataset Lov'Ego seeks to offer the most systematic curated corpus of available bibliographic records associated with five key printers from

[11] We have not made any attempts to identify other editions beyond the STCV and the Caa dataset, as these would in terms of SNA not significantly alter the results; for the portfolio of Maesius, however, we did manually correct the assignations made by the Caa to enhance historical correctness of the dataset, see Appendix 4.A

[12] A preliminary search of the HPB for Hieronymus Nempaeus, one of Lov'Ego's selected printers, illustrates these challenges: the same 1702 edition of Jacques-Hyacinthe Serry's *Histoire des congregations de auxiliis* appeared as five separate records from libraries in Switzerland, Italy, Germany, Spain, and the STCV itself, confirming the copy-level redundancy.

early modern Louvain. We selected these figures on the basis of their established prominence in the historiography of Louvain's print culture, a prominence confirmed by their high centrality scores in the whole co-occurrence network constructed from the bibliographic metadata of Caa (in almost its entirety), which has been the subject of an earlier chapter by one of the authors and Margherita Fantoli. Figure 1 supports this argument by showing the evolution over time of betweenness centrality among the fifteen highest-ranking printers active in Louvain. This implies that the configurations described in the present article primarily characterize printers who occupied bridging positions within Louvain's academic book trade, not the full range of printing activity in the city – a complete reconstruction of which is not a priority of the present study and would in any case be impaired by the limited scope of the selected bibliographic repertories, as well as by the irreducible losses of the historical books itself. We selected across three centuries 'caesuras' in the context of early modern Louvain: the turn from the fifteenth to the sixteenth century (Dirk Martens and Rutgerus Rescius), the late sixteenth century to the seventeenth century (Johannes Masius), and the later seventeenth to early eighteenth century (Hieronymus Nempaeus and Aegidius Denique).

The five selected printers/bridging figures across the three centuries of the early modern era operated in the same university city of Louvain but within starkly different economic and confessional environments. The paper approaches their centrality through an explanatory framework built around three ideal types of business models based on distinct partnerships' configurations: (1) printers operating with a program gearing towards (Christian) humanism and the *studia humanitatis* (Martens and Rescius); (2) printers acting on demand, the so-called 'jobbing printers' (Masius); and (3) the printer-propagandists (Nempaeus and Denique) engaged in doctrinally charged at times semi-clandestine networks around Jansenism, theological ideas from the Louvain professor Cornelius Jansenius (1585-1638) which eventually proved to bring a rupture within the broader Catholic church, as well provoking heavy censorship by several archbishops of the ecclesiastical province of Malines in which Louvain was situated.[13] These three configurations are illustrative profiles from the book-history literature, each anchored to a different historical period. Our goal is not to present them in

[13] A. Smets and W. François, 'Cornelius Jansenius', in: ODIS, http://www.odis.be/lnk/en/PS_126898, last modified 16 July 2024; Vanden Bosch (2017a).

chronological progression as successive stages of development, nor to establish them as "examples" of their cohorts, but to show how each printer's surviving output instantiates one of these distinct ideal types, through descriptive analysis of their networks.

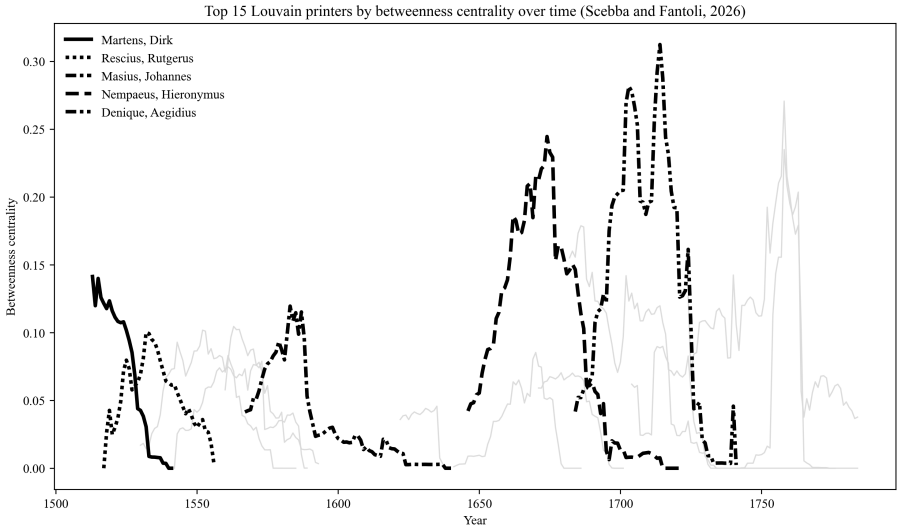



**Figure 4.1.** Top 15 Louvain printers by betweenness centrality over time (from Scebba and Fantoli (2026)) The case studies were selected from among the printers with the highest betweenness centrality in the network constructed for the Scebba/Fantoli paper, a co-occurrence network drawn from the Collectio academica antiqua (Caa), representing the intellectual community at and around the Old University of Louvain. We chose to focus on individual printers or joint printing ventures that mark transitional moments across three centuries: at the turn of the fifteenth to the sixteenth century (Dirk Martens and Rutgerus Rescius), the sixteenth to the seventeenth century (Johannes Masius), and the seventeenth to the eighteenth century (Hieronymus Nempaeus and Aegidius Denique). More specifically, the figure plots the betweenness centrality over time for the top 15 Louvain-based printers with the highest average scores. Betweenness centrality measures the extent to which a node occupies an intermediary position in the network, that is, how frequently it lies on the shortest paths connecting other nodes. Printers with high values therefore act as structural brokers. In the figure, we highlight the trajectories of the five focal printers, while the remaining printers appear in lighter gray.

To assemble the Lov'Ego dataset, we first selected Caa and STCV bibliographic records whose title pages attest the names of the selected five printers. In several instances, detailed in the Appendix, we revised the original attributions following human-in-the-loop validation. Next, to avoid double counting of KU Libraries holdings that are represented in both the

Caa and the STCV, we compared the metadata of both the Caa and STCV and identified their overlap by tracing duplicates through shared call numbers and electronic location references. STCV records carrying a KU Libraries call number that also appears in the Caa (or, in fewer cases, explicitly linked in the Caa metadata under the note: '[T]his edition in STCV') are treated as duplicates. The Lov'Ego 'universe' was thus constructed by retaining all Caa records and adding only those STCV records that do not correspond to any KU Libraries holding already represented in the Caa. We then applied the following inclusion criteria: (i) the item had to be printed by one of the selected five printers; (ii) it had to be published in Louvain after 1501 (in order to ensure comparison between Caa/STCV and thus exclude incunables); and (iii) specific categories of material were excluded from the corpus. These exclusions entail academic thesis sheets and dissertation booklets (materials that have not, to date, been incorporated into a systematic catalog or dataset[14]—and to broadsheets and rare books, whose bibliographic units differ too markedly to allow for uniform comparison[15]), legislation issued for non-academic authorities, as well as prints produced for civic or clerical entities (for example, those commissioned by the city of Louvain or by the Order of the Annunciates in close-by Heverlee). The composition of the final corpus is illustrated in Figure 2. The Caa subset comprises 239 records, while the corresponding STCV subset contains 363 records. Of these, 163 records represent KU Libraries holdings appearing in both datasets and are therefore treated as overlap. The final corpus thus consists of 439 bibliographic records or, equivalently, editions of rare books: 239 Caa items complemented by 200 additional STCV records corresponding to copies preserved outside the KU Libraries.

Table 1 shifts the focus to the internal structure of Lov'Ego by detailing the contribution of each printer to the corpus of the single printers and situating them in their biographical and professional timelines. For instance, Dirk Martens started in Louvain from 1498 onwards but interrupted his activity there between 1502 and 1512 in order to be active in Antwerp, where

[14]In KU Leuven Libraries, except for some oddities, rare books have been cataloged under Caa, and Thesis Sheets under CaaTH. The digitized collection of Thesis Sheets in the KU Leuven Libraries has now been integrated under https://studium-ai.org/thesissheets, and will soon be updated with thesis sheets kept in the Maurits Sabbe Library.

[15]A methodological reflection on the impact of integrating thesis sheets into SNA of academic book towns is to be found in Soen and Fantoli (forthcoming).

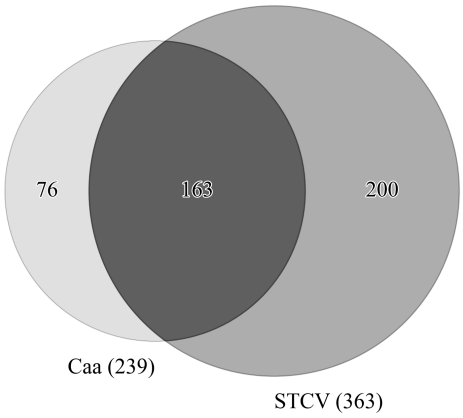



**Figure 4.2.** Source composition of the Lov'Ego dataset. The Venn diagram represents the composition of the Lov'Ego dataset in terms of its sources, namely the bibliographic metadata drawn from the Collectio academica antiqua (Caa) and the Short Title Catalog Flanders (STCV). In the diagram, each item corresponds to a single bibliographic record provided by one of the datasets. We manually inspected all records and corrected cataloging inconsistencies and attribution errors. As a result, the Lov'Ego dataset comprises 439 bibliographical records, 200 of which come exclusively from the STCV, and with an intersection of 163 records present in both sources.

he had been operating presses in the four years arriving in Louvain. He transferred his business to Rescius some five years before his death. Moreover, we note that the Lov'Ego corpus includes a single co-imprint shared by both Nempaeus and Denique, the former in the capacity of printer and the latter in that of bookseller and publisher. The item is the 1689 edition of *Commentaria in primam partem divi Thomæ de Deo trino et uno, de angelis, et operibus sex dierum*, a late Louvain reprint of a standard Thomist university textbook (USTC No. 6169520, KU Libraries control number 998871040101471). Since this co-imprint enters the count for both printers, the arithmetic sum of imprints reported per printer would exceed the number of unique bibliographic records by one.

**Table 4.1.** The *Lov'Ego* dataset: printers' chronology, activity span, and number of bibliographic records. The table reports the biographical and bibliographic footprint of the focal printers based on the *Lov'Ego* dataset. For each printer, it lists the available birth and death dates, the beginning of documented activity, and the first and last item identified in the *Lov'Ego* corpus. The table also reports the total number of bibliographic records attributed to each printer, together with their authority identifiers in ODIS, STCV, and VIAF.

| | Birth | Death | Start Louvain | First Item in *Lov'Ego* | Last Item in *Lov'Ego* | # *Lov'Ego* records |
|---|---|---|---|---|---|---|
| *Martens* | 1467–9 | 1534 | c. 1498 | 1501 | 1529 | 37 |
| *Rescius* | 1495 | 1545 | 1529 | 1529 | 1544 | 35 |
| *Masius* | ? | 1616 | 1567 | 1570 | 1625 | 113 |
| *Nempaeus* | ? | c. 1702 | ? | 1645 | 1702 | 132* |
| *Denique* | ? | c. 1728 | c. 1685 | 1667 | 1702 | 123* |
| TOTAL* | | | | | | 439* |

*Notes*: *The arithmetic sum of individual *Lov'Ego* bibliographic records equals 440, but the corpus total, as reported in the table, is 439 because one co-imprint jointly produced by Nempaeus and Denique is counted once in the aggregate but appears in both individual printer ego counts.

On the basis of this final corpus of bibliographic records in Lov'Ego, we construct a network representation in which personal named entities (authors, printers, editors, and other actors) constitute the nodes, and co-occurrence on the same title page defines the edges. We interpret such co-occurrence as material evidence of formal association within a specific edition, and, when contemporaneity permits, of collaborative endeavors. Figure 3 provides a simplified representation of the notion of nodes and (weighted) edges in the Lov'Ego network. The basic elements of this representation are nodes (i.e., the dots) and edges (that is, the segments connecting the dots). Nodes

correspond to individual actors involved in the bibliographic records of the corpus, including printers, authors, editors, and other contributors. We draw an edge between two nodes when they appear together on at least one bibliographic record in our corpus. Edge thickness is proportional to the number of shared records, so that thicker lines indicate more frequent collaboration. The figure offers a simplified illustration of this weighting principle. The upper example shows the nodes of Louvain 'on-demand' printer Johannes Masius and Louvain professor Henricus Cuyckius (1546-1609, also known as Henri or Hendrik van Cuyk[16]), connected by an edge of weight 13, indicating that they co-appear on thirteen records—one of the strongest ties in the Lov'Ego network. The lower example shows the link between the nodes of Dirk Martens and Thomas More, with weight 1, corresponding to their single yet renowned collaboration in the earlier mentioned first-ever publication of the Utopia in Louvain in 1516.




**Figure 4.3.** Example of nodes and (weighted) edges. Nodes represent individual actors; edges connect actors who co-appear on at least one title page. Edge thickness is proportional to the number of shared records. The upper example shows a high-weight tie (Johannes I Masius–Henricus Cuyckius, weight 13); the lower example shows a single shared edition (Dirk Martens–Thomas More, weight 1).

[16] A. Smets, G. Bertrand and V. Soen, 'Henricus Cuyckius', ODIS, http://www.odis.be/lnk/en/PS_126852, record last modified 24 July 2024.

While Figure 3 isolates two examples of connected node pairs, Figure 4 displays the full network constructed from the Lov'Ego dataset. The selected printers are highlighted in black and stand out visually as they are, structurally, the network hubs reflecting the criteria by which the corpus was defined.[17] The positions of the remaining actors (white nodes) are not random but result from a force-directed layout algorithm: nodes repel one another while edges exert attractive forces, so that more strongly connected actors tend to cluster toward the center. Moreover, because some individuals co-occur in the bibliographic records of more than one selected printer, the corresponding ego-networks partially overlap, generating visible intersections within the overall graph. Later in the article, we segment the aggregate network depicted in Figure 4 to isolate the various subnetworks stemming from each of the focal printers, also known as ego-networks. Unlike a so-called 'whole' network, which represents all ties within an entire community, an ego-network centers on a single individual, the 'ego' (in our case, the five focal printers), and maps only ties involving the focal node, plus any reciprocal ties among the nodes connected to it. Such networks typically have a star-shaped form, showing a community as it radiates outward from one node rather than a snapshot of the whole. Since each network is built around its focal node, that node is always the hub of the star, the most connected point by design. We thus focus on relational and compositional metrics rather than standard centrality measures, which reveal little here. By examining each printer's immediate connections, we capture the specific configuration of partnerships and editorial strategies associated with each one of our case studies.[18]

To systematically characterize the printers' associates and make emerging categories immediately visible, we complement our bibliographic corpus with

[17]The network representation of the Lov'Ego corpus is subject to an additional constraint: only co-occurrences between imprint actors and paratextual attributions are retained. All other role types are excluded *a priori*. In practice, roles associated with dedicatory functions, censorship and approbatory interventions, illustration and engravings, and other forms of ambiguous involvement are systematically filtered out whenever they appear as either source or target in the edge list.

[18]Although in Chapters 1 and 2 of the present dissertation I advocate a multilayer representation of bibliographic co-occurrence networks, I argue that the multilayer design adds little here: the corpus is restricted to two tie types (imprint and paratextual attributions), which removes most of the role heterogeneity that justifies layering, and each printer's ego-network is by construction a star centered on the focal node, so the analytically relevant information lies in the *composition* of the neighborhood and in dyadic edge weights rather than in inter-layer topology.

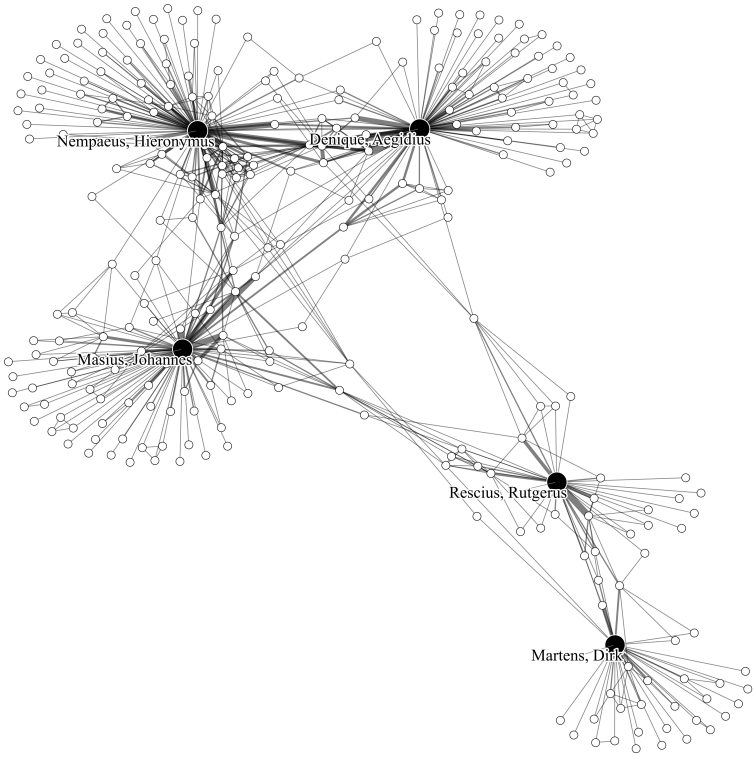



**Figure 4.4.** The Lov'Ego full weighted co-occurrence network. The figure represents the network constructed from the full Lov'Ego dataset. Nodes correspond to individual actors, and edges connect actors who co-appear on at least one bibliographic record. Edge thickness is proportional to the number of shared records. Node positions derive from a force-directed (spring) layout in which nodes repel one another while edges exert attractive forces proportional to their weight. The five focal printers appear as black nodes and are labeled, whereas all other actors appear as white nodes. In the following figures, we segment this full network to isolate the ego-network of each focal printer.

external prosopographical datasets. In particular, we rely on the *Repertorium Eruditorum Totius Europae* (RETE henceforth) to identify, among the printers' partners, university scholars, as well as affiliates to the Society of Jesus. Hosted at UCLouvain and scheduled for full release at the end of 2026, the RETE project is a prosopographical database of professors and *eruditi* affiliated with European academic institutions from their inception to the first Industrial Revolution.[19] The database is compiled from a wide range of secondary sources such as university roasters, biographical dictionaries, and institutional histories. In Lov'Ego, for Louvain specifically, affiliations are reconstructed primarily from the dataset Scholars@OldLouvain, which underpins the identification of scholars within the Research Infrastructure STUDIUM.AI, a linked open data resource launched in November 2025.[20] For each printer, we compute the rate of repeated collaboration, the share of scholars (with an affiliation to Louvain's Theology and Artes faculties, to the Collegium Trilingue, as well as other European universities), the proportion of collaborators belonging to the Jesuit order, and the presence of classical and medieval authorities within the printer's corpus. We also assess the structural configuration of the network, distinguishing between highly centralized or star-like structures and more cohesive collaborative circles beyond the printer's involvement. Our aim is to compare how these metrics vary within and across each printer category to identify systematic similarities and differences. We report these figures for each printer in Table 2 and discuss them throughout the article.

## 4.3 The Printer-Scholar: Academic networks of Martens and Rescius

In the earlier cited frameworks proposed by Pettegree and Grafton, humanist-printers or 'scholar-printers' are identified as notable exceptions to the prevailing paradigm of entrepreneurial and proto-capitalist craftsmanship

[19] D. de la Croix (principal investigator), M. Vitale, C. Zanardello, M. Curtis, F. Manfredini (contributors), *Repertorium Eruditorum Totius Europae*, UCLouvain, v. 1.0, 2026: https://perso.uclouvain.be/david.delacroix/menushiny.html.

[20] A. Smets (author), R. Scebba and V. Soen (contr.), *Scholars@OldLouvain*, KU Leuven RDR, v. 1, 2025: https://doi.org/10.48804/I3CETK. V. Soen et al., *STUDIUM.AI: A digital gateway to the University of Leuven (1425–1797)*, launched November 2025.

**Table 4.2.** Summary metrics per focal printer. Each row corresponds to one focal printer. Each row reports a network metric capturing structural and compositional characteristics of the printer's ego network. From top to bottom, we report: the total number of collaborators (authors, editors, contributors, and other printers or publishers appearing in books issued by the focal printer); the ratio of repeated partnerships, measured as the share of collaborators who appear in more than one book published by the printer; the percentage of academic scholars among collaborators (source: RETE scholars affiliated to at least one European university); the percentage of Jesuit collaborators (source: RETE attributes and bibliographic metadata); the percentage of classical authors represented in the printer's editions; the percentage of medieval authors; the percentage of collaborators who were themselves printers or publishers; the percentage affiliated with the Collegium Trilingue, as well as with Louvain's Theology and Artes Faculties; the entropy of role distribution (capturing diversity of functional roles within the network); and the density of the collaborator subnetwork after removing the focal printer.

| Metric | Martens | Rescius | Masius | Nempaeus | Denique |
|---|---|---|---|---|---|
| *Partners (N)* | 34 | 32 | 76 | 94 | 78 |
| *Repeated collaboration (%)* | 20.6 | 34.4 | 35.5 | 35.1 | 35.9 |
| *Ego density (no printer)* | 0.03 | 0.08 | 0.02 | 0.03 | 0.02 |
| *University scholars (%)* | 55.9 | 40.6 | 32.9 | 48.9 | 42.3 |
| *Jesuits (%)* | 0 | 0 | 6.6 | 1.1 | 3.8 |
| *Classical authors (%)* | 17.6 | 43.8 | 3.9 | 1.1 | 1.3 |
| *Medieval authors (%)* | 5.9 | 0 | 1.3 | 3.2 | 0 |
| *Printer partners (%)* | 0 | 6.2 | 11.8 | 6.4 | 12.8 |
| *Collegium Trilingue (%)* | 17.6 | 15.6 | 1.3 | 4.3 | 5.1 |
| *Theology faculty (%)* | 8.8 | 6.2 | 10.5 | 21.3 | 20.5 |
| *Artes faculty (%)* | 17.6 | 9.4 | 10.5 | 13.8 | 11.5 |

that characterized early modern printing on the continent. This particular category of printers aligned more closely with the profile of humanist scholars, whose engagement with printing was motivated not by commercial interests but by a commitment to the dissemination of scholarship and the promotion of the *bonae literae*. Their primary objective was to advance the *studia humanitatis* (the comprehensive study of humanity and literature) and to contribute to what they perceived as a necessary reform of the medieval scholastic tradition and text edition (and/or theology). During the Renaissance, a distinct subgroup of printers emerged, many of whom had originally pursued scholarly careers. Book historians commonly refer to these individuals as "humanist printers" or "scholar-printers". These learned figures were distinguished by their mastery of Latin and, in some cases, Greek and Hebrew. Beyond their technical role in establishing printing presses, they also authored original works and undertook editorial responsibilities for classical texts. In this capacity, they mirrored the broader humanist ethos by striving to produce accurate editions of texts, informed by the possibilities afforded by new media technologies. Prominent representatives of this tradition include Aldus Manutius (1450–1515) in Venice, and Josse Bade (1461/2–1535), Robert Estienne (1503–1559), and Chrestien Wechel (ca. 1494–1554) in Paris.[21]

In Louvain, two printers have long been recognized as exemplifying this distinctive profile. The first is Dirk (or Thierry) Martens (1446/7–1534), the only figure in this article referred to primarily by his vernacular name rather than the less familiar Latin form 'Theodoricus Martinus'.[22] Martens is widely acknowledged as one of the earliest printers active in the Low Countries, perhaps even the very first. He began his career in Aalst (which explains the aliases van Aelst or Alostensis), where he partnered with Johannes of Westphalia (fl. 1473-1503)[23] ) , a printer from Paderborn newly arrived in Louvain, then he worked briefly in Antwerp, and then established himself in Louvain in 1492. Martens later returned to Antwerp between 1502 and 1512 before ultimately settling once more in Louvain. In Louvain, in particular, Dirk

[21]Constantinidou et al. (2025, pp. 100–101) for an introduction to this type of printer; Bénévent et al., eds (2012).

[22]D. Cammaerts and H. Wyffels, 'Theodoricus Martinus', ODIS, http://www.odis.be/lnk/en/PS_147826, record modified 27 November 2020; see also Adam (2018, p. 175).

[23]D. Cammaerts, 'Joannes de Westfalia Paderbornensis', ODIS, http://www.odis.be/lnk/en/PS_172022, record last modified 1 March 2021.

Martens conducted a business of significance, turning out numerous editions of the classics and modern authors. He was the first to introduce italic type in the Netherlands and also experimented with Greek and Hebrew characters and typesets; he also enjoyed the patronage of Erasmus for editions of his works published in the Netherlands, and through him, secured the edition of the very first edition of Thomas More's *Utopia*.

The other clear profile of a scholar-printer would be Rutgerus Rescius (c. 1495-1545), [24] one of the correspondents of Erasmus, first Chair of Greek at the newly founded humanist Collegium Trilingue. He needed to move out of the college after marrying Anna Moons[25] and engaged in a quite remarkable life as printer-publisher, teacher, spouse and father. Both men knew each other well, as the talented Latin linguist Rescius started as corrector in Martens' printing offices, and maybe it was through this connection that he was called to serve at the Collegium Trilingue in Louvain. With more certainty, it is documented that he eventually took over the Martens' printing office in Louvain in 1529, years before Martens' death, and became one of the chief printers of Greek texts in the Low Countries, most often in partnership with the Louvain printer Gravius. It has been argued that Rutgerus Rescius was not only a commercial partner for Martens, but also his 'spiritual heir'.

Figures 5 and 6 present the ego networks of Martens and Rescius, constructed according to the methodology outlined above. It is important to note that, because the Caa uses 1501 as its cutoff start date (and Lov'Ego followed in this respect), early incunable production from 1473 to 1500 is not represented in the networks. Although the STCV does contain incunabula attributed to the focal printers, we excluded them in order to maintain comparability between the two sources of the Lov'Ego corpus. Moreover, as already mentioned, these early editions were for the most part not printed in Louvain, since Martens was primarily active in Aalst and Antwerp during those years—so they would fall outside the scope of Lov'Ego, which is restricted to imprints issued in Louvain. In addition, the production capacity of presses in the early sixteenth century remained substantially more limited than it would become from the mid-sixteenth century onwards. This helps

[24] D. Cammaerts et al., 'Rutgerus Rescius', ODIS, http://www.odis.be/lnk/en/PS_142106, record last modified 7 August 2024, with a more introductory bibliography; also see Van Rooy (2022).

[25] H. Wyffels, 'Anna Moons', ODIS, http://www.odis.be/lnk/en/PS_147641, last modified 27 January 2020.

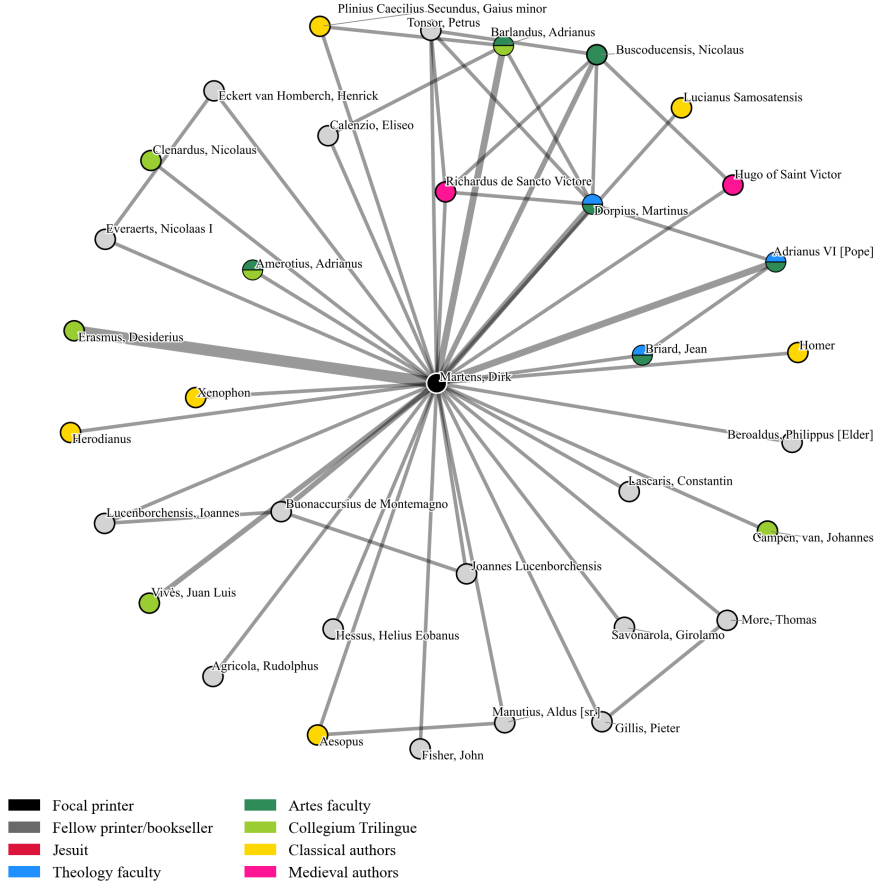



**Figure 4.5.** Ego network of Dirk Martens. Nodes represent individuals appearing in at least one book published by the focal printer; edges indicate co-occurrence in the same publication. Edge thickness is proportional to the number of shared books. The focal printer is shown in black. Nodes are color-coded according to categorical attributes: fellow printers or booksellers (dim gray); Jesuits (crimson), affiliation with Louvain's Theology Faculty (blue), Artes Faculty (green), and Collegium Trilingue (yellow green); classical authors (gold); and medieval authors (pink). When an individual possesses up to two of the listed attributes, the node is split into two colored semicircles.

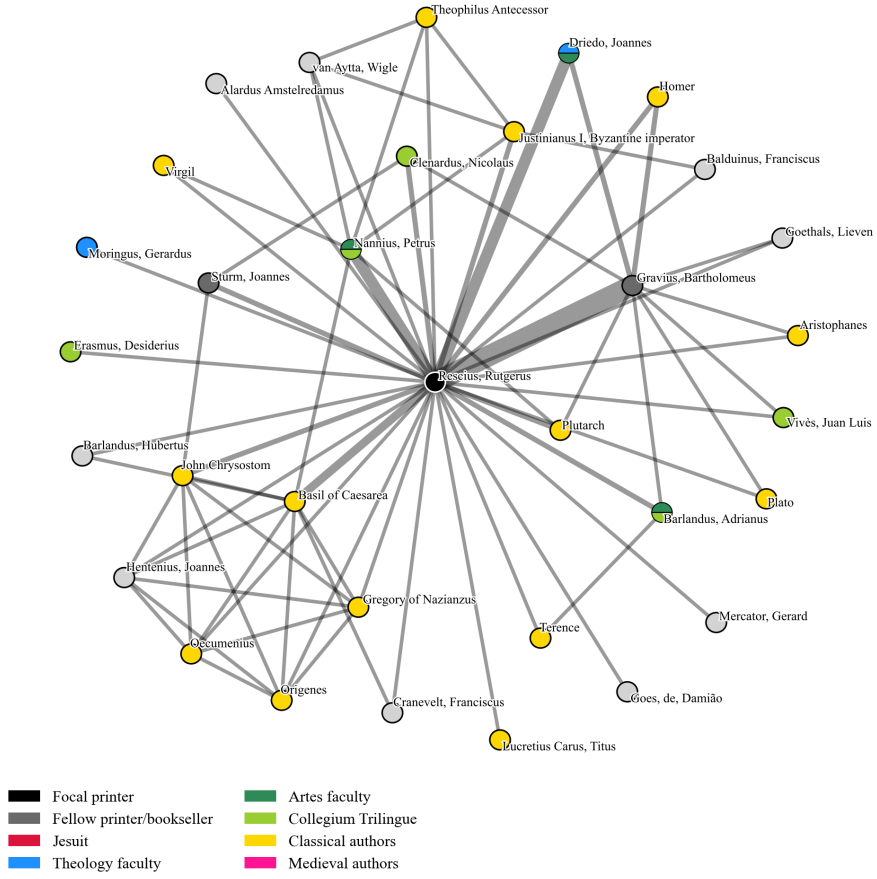



**Figure 4.6.** Ego network of Rutgerus Rescius. Nodes represent individuals appearing in at least one book published by the focal printer; edges indicate co-occurrence in the same publication. Edge thickness is proportional to the number of shared books. The focal printer is shown in black. Nodes are color-coded according to categorical attributes: fellow printers or booksellers (dim gray); Jesuits (crimson), affiliation with Louvain's Theology Faculty (blue), Artes Faculty (green), and Collegium Trilingue (yellow green); classical authors (gold); and medieval authors (pink). When an individual possesses up to two of the listed attributes, the node is split into two colored semicircles.

explain the comparatively smaller size of the networks reconstructed for the two humanist printers active in Louvain, as compared to the printers in the next two sections.

In particular, Martens shows the second smallest number of partners (34 associated individuals) and the lowest rate of repeated partnerships (20.6 percent). The low ego density (0.03) and null tendency to partner with other printers indicate a star-like configuration centered strongly on Martens rather than a cohesive circle among associates. In other words, Martens appears structurally central yet not embedded in a tightly interconnected intellectual milieu. Similarly, Rescius's network is the smallest in size out of all this article's case studies, counting 32 individuals only. But in contrast to Martens's, Rescius's network has a significant rate of repeated partnerships (34.4 percent) and the highest ego-density score in the sample (0.08). Taken together, these figures suggest a cohesive network among associates, in opposition to a purely star-shaped structure. Also, in Rescius's network, the proportion of classical authors is exceptionally high (43.8 percent), indicating a strong editorial orientation toward classical scholarship, which is consistent with Rescius's dual position as both printer and humanist scholar, notably being a lecturer of Greek at the Collegium Trilingue. This is also embodied by the highest fraction of partners affiliated with the Collegium, 15.6 percent.

The question, then, remains how their partnership network was composed. Overall, both Martens and Rescius overwhelmingly partnered with scholars of the local university (55.9 and 40.6 percent, respectively, of the total nodes in their networks). The presence of the itinerant scholars with Louvain links, Desiderius Erasmus and Juan Luis Vives as well as Nicolaas Clenardus and the less mobile Hadrianus Barlandus both networks indicates a clear overlap in the classical-humanist segment of the book market, as well as with the teaching staff of the Collegium Trilingue. Given that Rescius took over Martens's printing house, this pattern plausibly reflects the inheritance of editorial lines and scholarly contacts, with both printers acting as publishers as well. However, it is important to note that scholars affiliated with the Faculty of Arts of the Old University of Louvain represent a non-negligible fraction of partners, with 17.6 and 9.4 percent respectively, this is more remarkable given the fact that scholars of the Collegium Trilingue often discredited both the linguistic skills as well as the scholastic teaching at that Faculty. Hence, the link with humanism is clear, but should also not be exaggerated in the first decades. In fact, Martens' first produced book in Aalst in 1474 and again 1475, commissioned by Johannes

of Westphalia, was a text edition of Aristoteles' *Organon* (a companion to the so-called *Logica Vetus* prepared by Pomponius Porphyrius) in a handsome format 4°, which would sustain rather than refute the scholastic curriculum in Louvain. Martens would also secure further editions of Aristotle for the Faculty of Arts. He was eager to bring classical authors to the market in Louvain (and later the Antwerp market), whether Aesopus, Xenophon, Homer or Herodianus. Classical authors (17.6 percent) were never as dominant in the Martens' portfolio as they were for Rescius, however. More than humanist, thus, Martens and Rescius' networks remained strongly academic, embedded within the Louvain learned world, with Rescius leaning more towards the figure of the humanist printer, or Printer-Scholar. In both cases, what sets these earlier presses apart was above all partnerships clustered overwhelmingly within the local scholarly community, from the university and Collegium Trilingue, and editions of classical authors.

## 4.4 The Jobbing Printer: University Commissions for Johannes Masius (d. 1616)

At the junction of the sixteenth and seventeenth centuries, we selected Johannes Masius (Jean/Jan Maes, d. 1616) because he exemplifies a category of printers in early modern Louvain whose economic and social position has remained understudied: the *jobbing printer*. Unlike the proto-capital or the humanist publishers who shaped the university's intellectual profile through large-scale editorial ventures, jobbing printers survived on a patchwork of small, local commissions and rarely developed a sustained publishing program of their own. Their output was dominated by ephemeral prints, the kind of documents intended for immediate use. Within university towns across Europe, such printers filled a crucial niche: they could be hired quickly to produce thesis sheets for public defenses, occasional prints for academic ceremonies, and other small-format works that larger houses considered unprofitable or too routine. This work offered no prestige, but it provided steady income within the volatile economies of university book production. Itinerancy was common among jobbing printers: they often sent out sons or grandsons to scout for work across regional book markets (Walsby, 2017). The multibranched Masius' printing house fit this profile almost paradigmatically:

members of the family oscillated between Louvain and Ath in Hainaut (home to an important Latin School) and sought opportunities wherever educational or governmental institutions required cheap and rapid print work.

Despite the meticulous biographical and book historical research of Christopher Warner, published in 2018,[26] substantial uncertainty remains concerning the genealogy of the Masius family, as well as its geographic spread. Between roughly 1560 and 1650, four homonymous individuals bearing the name Johannes or Joannes Masius operated in Louvain's book world, whether as printers, bibliophiles, or patrons. Their overlapping lifespans and partially intersecting activities have generated confusion since at least the middle of the seventeenth century, obscuring the identification of specific imprints and complicating bibliographical attribution. This article focuses exclusively on the earliest of these figures, the 'founding father'Johannes Masius, who died in 1616.[27] Appendix 4.A clarifies how we disentangled the four men and corrected the raw data accordingly.

As a general pattern, Johannes Masius printed for a wider range of commissioners than the university alone. His workshop produced materials for the Habsburg government, the city of Louvain, and multiple religious orders (including female ones), indicating that he operated within a broad and diversified network of small-scale patrons. These commissions fall outside the Lov'Ego project's focus on academic book production and were therefore excluded from our sample. Likewise, although Masius printed numerous thesis sheets (some preserved in the Caa holdings and others recently rediscovered in the Abbey of Park and the *Bibliothèque nationale de France*), we excluded them for reasons of comparability. Yet these ephemeral university prints must have constituted a significant share of his output and income; however, so few survive in the Caa that their inclusion would not have materially altered the analysis presented here. Book historians have described the Masius press as producing work of modest or even poor quality: faulty paper, textual omissions, weak ornamental framing, and the absence of Greek or Hebrew type suggest a workshop operating with limited resources (Walsby, 2017; Warner, 2018a). Typographic modesty and cheapness were two faces of the same coin, common to academic ephemeral print. Even the

[26] Warner (2018b), correcting information from Walsby cited above; Dewert (1926); as well as Rouzet (1975).

[27] A. Soetaert and D. Cammaerts, 'Jean Maes', ODIS, http://www.odis.be/lnk/en/PS_127109, record last modified 8 January 2021.

most celebrated university printers, like Abraham Elzevier in Leiden, produced cheaply made dissertations alongside scholarly, higher quality works. Precisely this combination of economic precarity, technical limitations yet wide-ranging patrons makes Johannes Masius an illuminating case study.

From his ego networks drawn from the Lov'Ego corpus, Masius operated with a relatively large set of partners (76) and showed one of the highest rates of repeated collaboration in the sample of our printers (35.5 percent), supporting the notion of his stable base of recurring commissions. Yet his ego density remained low (0.02), which seems to imply that his associates engaged only weakly with one another. This combination of high recurrence and limited lateral connections among partners suggested a structure in which production flowed abundantly through the printer himself, but only marginally within a tightly knit intellectual community. A second observation reinforces this interpretation of Masius as a jobbing printer: he also displayed the highest ratio of fellow printer partners (11.8 percent), a substantial collaboration with Theology faculty presence (10.5 percent) and the highest faction of joint ventures with affiliates to the Jesuit Order (6.6 percent)—who, in Louvain, were not associated to the university directly but mostly to the stand-along Jesuit college. Rather than constructing a selective prestige circle, he rather tied in within the local printing economy and a tendency to engage with broad sets of partners. Consistently, his share of academic scholars was not insignificant but nevertheless represented the lowest among the focal printers (32.9 percent). Engagement with canonical authors, whether classical or medieval (3.9 and 1.3 percent, respectively), declined sharply compared with the network structures of humanist printers. Set beside the scholarly embedding of Martens and Rescius, Masius marks a different answer to the same question: with the lowest share of academic scholars (32.9 percent), classical authors all but absent, and the highest proportion of fellow-printer ties, his partnerships were organized by the breadth of the local commission economy rather than by any shared intellectual program.

## 4.5 The Printer-Propagandist: The Jansenist Cause of Nempaeus and Denique

The final section reconsiders the operational and intellectual networks surrounding the printing house of Hieronymus Nempaeus (Jeroen Nempe, fl.

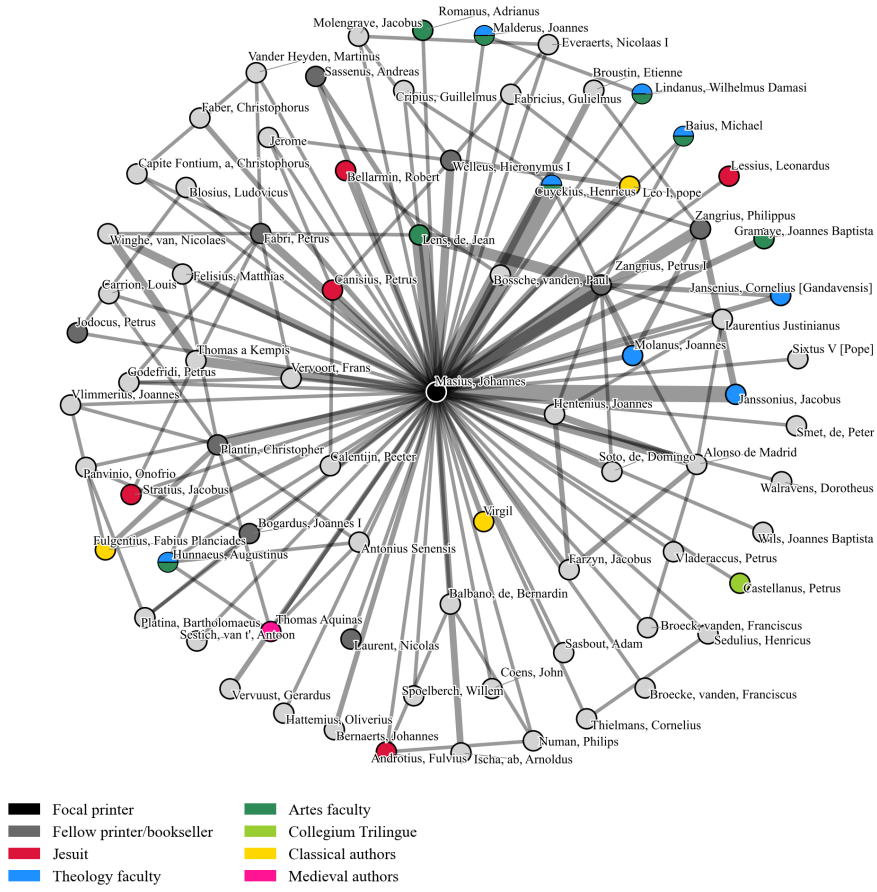



**Figure 4.7.** Ego network of Johannes Masius. Nodes represent individuals appearing in at least one book published by the focal printer; edges indicate co-occurrence in the same publication. Edge thickness is proportional to the number of shared books. The focal printer is shown in black. Nodes are color-coded according to categorical attributes: fellow printers or booksellers (dim gray); Jesuits (crimson), affiliation with Louvain's Theology Faculty (blue), Artes Faculty (green), and Collegium Trilingue (yellow green); classical authors (gold); and medieval authors (pink). When an individual possesses up to two of the listed attributes, the node is split into two colored semicircles.

1646–1702),[28] which by the early eighteenth century was absorbed into the printing workshop of Aegidius (I) Denique.[29] Nempaeus first entered the historical record upon his matriculation at the University of Louvain as a minor on 24 June 1641, and within less than a decade he had transformed himself into one of the city's central printers. Between 1646 and 1702 he operated a vigorous workshop under his own name—a development made possible not simply by entrepreneurial talent but by his strategically advantageous marriage around 1645 to a female printer, Paschasia de Zangere, heiress to (at least part of) her late husband Zegers's printing house and born in the Zangrius' printing dynasty.[30] Because she (most probably) enjoyed the legal right to continue her late husband's business as a widow, Nempaeus effectively gained access to an established printing enterprise, its clientele, and its university connections. The location of the printing house—on the corner of Naamsestraat and Muntstraat, directly facing the University Hall and in sight of Holy Spirit College—was equally significant. Nempaeus's printer's device, which depicted the Holy Spirit inspiring a Catholic author, visually affirmed his desired ties to Louvain's university and its most prominent college.

Nempaeus quickly built a flourishing career by specializing in substantial legal commentaries indispensable to students at the faculties of Civil and Canon Law. Crucially, he secured the privilege to publish the posthumous works of Petrus Gudulinus (Pierre Goudelin), whose writings represented some of the earliest attempts to conceptualize *public law* within Louvain scholarship. His output, however, was not limited to jurisprudence: he became a prolific publisher of astronomical (and astrological) handbooks, theological treatises, and biblical commentaries, positioning his shop as a major supplier of academic handbooks across the university's faculties. Control over these reliable bestsellers solidified Nempaeus's financial base and intellectual prestige, likely financing his capacity to sustain more controversial theological production without immediate economic risk. Unlike many of his

[28] D. Cammaerts and H. Wyffels, 'Hieronymus Nempaeus', ODIS, http://www.odis.be/lnk/en/PS_147582, record last modified 24 February 2022.

[29] H. Wyffels, 'Aegidius I Denique', ODIS, http://www.odis.be/lnk/en/PS_145665, record last modified 13 June 2018.

[30] D. Cammaerts and H. Wyffels, 'Paschasia de Zangere', ODIS, http://www.odis.be/lnk/en/PS_147597, record last modified 24 February 2022. She was first married to Jacobus Zegers, and from 1644 until 1645 appears under the imprint 'widow Zegers'.

contemporaries, he seems to have invested less in the ephemeral market of printed thesis sheets and more in durable academic treatises, thereby cultivating a stable, long-term niche.

The childless marriage between Nempaeus and Paschasia Zangrius set the stage for the consolidation of the printing house into the Denique dynasty at the beginning of the eighteenth century. Yet it is during Nempaeus's own lifetime that the workshop achieved its most decisive impact. His press played an instrumental (and often overlooked) role in sustaining and normalizing Louvain's participation in the production of Jansenist-associated literature, the controversial theological movement drawing its ideas from the writings of Cornelius Jansenius, bishop of Ypres and former Louvain professor. The risks of such engagement were well known: in 1640, the university printer and librarian Jacobus Zegers (Paschasia's first husband) had nearly bankrupted himself and died shortly after the publication and subsequent condemnation of Jansenius's *Augustinus*.[31] Nempaeus's later stability, by contrast, suggests not open defiance but operation within (and at times testing the boundaries of) censorship. Nempaeus positioned his printing shop as a conduit for the continued circulation of Augustinian theology, even under pervasive ecclesiastical scrutiny.

By re-examining the networks that sustained Nempaeus's career, we gain clearer insight into how one Louvain printing house became a critical node in the diffusion of a controversial (but in Louvain circles cherished) theology. Nempaeus operated the largest network in the sample (94 partners), combined with a high rate of repeated collaboration (35.1 percent). As with the previous network from Masius, ego density remains low (0.03), indicating a centralized configuration relying primarily on Nempaeus himself rather than on strong lateral ties among associates. What sets Nempaeus apart is the composition of his network: he has the highest share of affiliates to Louvain's Faculty of Theology (21.3 percent) and Artes (13.8 percent) combined. In contrast, classical (1.1 percent) and medieval authors (3.2 percent) are marginal, marking a decisive shift away from the humanist reprint model. Nempaeus's configuration must be read against the confessional climate of the later sixteenth and early seventeenth centuries. His strong embedding in theology suggests engagement with current doctrinal debates. Among his strongest ties, we report those with Libertus Fromondus, Gummarus

[31] D. Cammaerts and H. Wyffels, 'Jacobus Zegers', ODIS, http://www.odis.be/lnk/en/PS_148192, record last modified 24 February 2022.

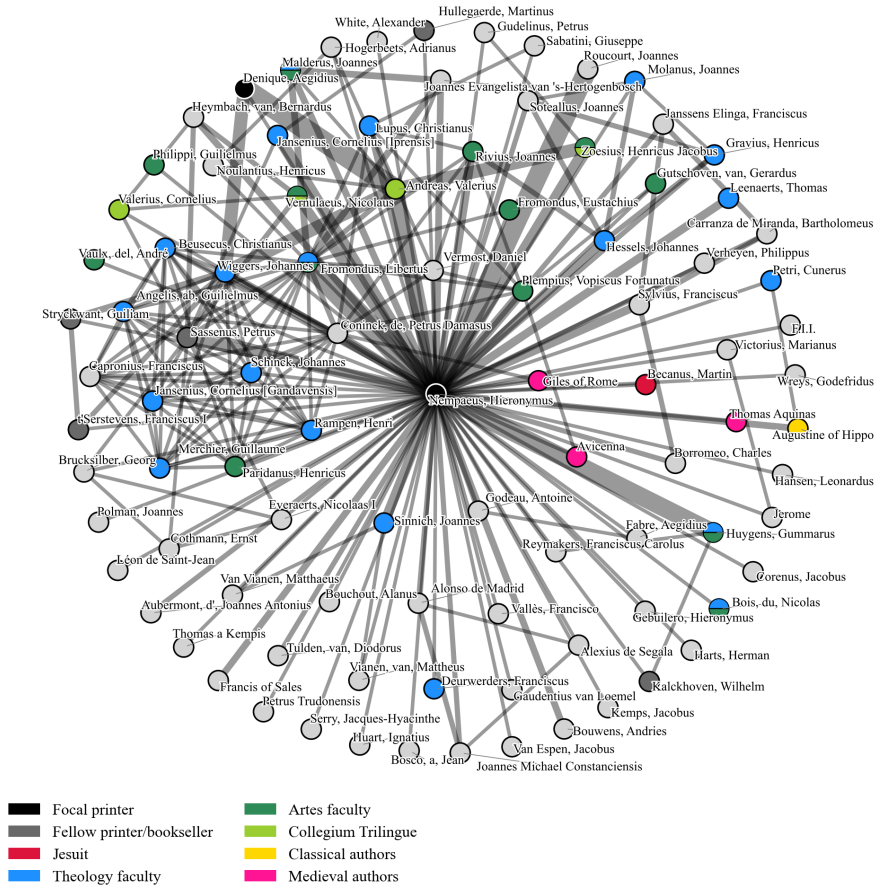



**Figure 4.8.** Ego network of Hieronymus Nempaeus. Nodes represent individuals appearing in at least one book published by the focal printer; edges indicate co-occurrence in the same publication. Edge thickness is proportional to the number of shared books. The focal printer is shown in black. Nodes are color-coded according to categorical attributes: fellow printers or booksellers (dim gray); Jesuits (crimson), affiliation with Louvain's Theology Faculty (blue), Artes Faculty (green), and Collegium Trilingue (yellow green); classical authors (gold); and medieval authors (pink). When an individual possesses up to two of the listed attributes, the node is split into two colored semicircles.

Huygens, Johannes Wiggers, Thomas Leenarts, Christianus Lupus, and Johannes Hessels, all embedded in Louvain's theological Faculty during or adjacent to the Jansenist controversy. Moreover, Nempaeus's negligible involvement with Jesuits (1.1 percent) is consistent with Louvain's broader anti-Jesuit stand during phases of the Jansenist controversy.

Also, the chronological dimension is decisive. Nempaeus prints Jansenius in 1659 and again in 1685, continues publishing Fromondus long after his death, from 1646 through 1678, and issues Lupus in 1654 and 1682, which fall within phases of intensified grace controversy. This is relevant because these editions post-date the papal bull Cum occasione (1653), which had condemned key propositions drawn from Augustinus. Their appearance in 1654, 1659, and again in the 1680s demonstrates sustained post-condemnation circulation. From 1675 onwards, Catholic authorities increasingly steered towards a more systematic form of book censorship, both preventive and curative, often under the influence of anti-Jansenist (arch)Bishops. Lists of bookshops and printers' portfolios were registered in Malines as well as in Brussels with the Council of Brabant. Yet this tightening of oversight is not visibly reflected in Nempaeus's production profile within the Lov'Ego corpus, which does not include information on censorship directly. Still, the continued publication of these authors situates Nempaeus within the infrastructure of post-condemnation Augustinian discourse. At the same time, the publication of figures such as Joannes Malderus and Joannes Molanus demonstrates continuity with Tridentine reform and episcopal orthodoxy. Taken together, these pieces of evidence indicate that Nempaeus's press helped preserve and circulate theological works that, even if not openly rebellious, were directly implicated in the central doctrinal conflicts of mid- and late-seventeenth-century Catholicism in the Southern Netherlands.

The transition to Nempaeus's successor, Aegidius Denique, reinforces this continuity. Denique already appears within Nempaeus' network before assuming control of his print shop. Structurally, he is more embedded in the wider cluster of Nempaeus's associates than directly tied to Nempaeus himself, with a single explicit co-occurrence: the 1689 reprint of *Commentaria in primam partem divi Thomæ de Deo trino et uno, de angelis, et operibus sex dierum*, where Nempaeus's name appears in the role of printer and Denique's in that of bookseller and publisher (sumptibus), indicating the growing differentiation between typographical production and financial responsibility characteristic of later seventeenth-century print culture (Pettegree, 2010).

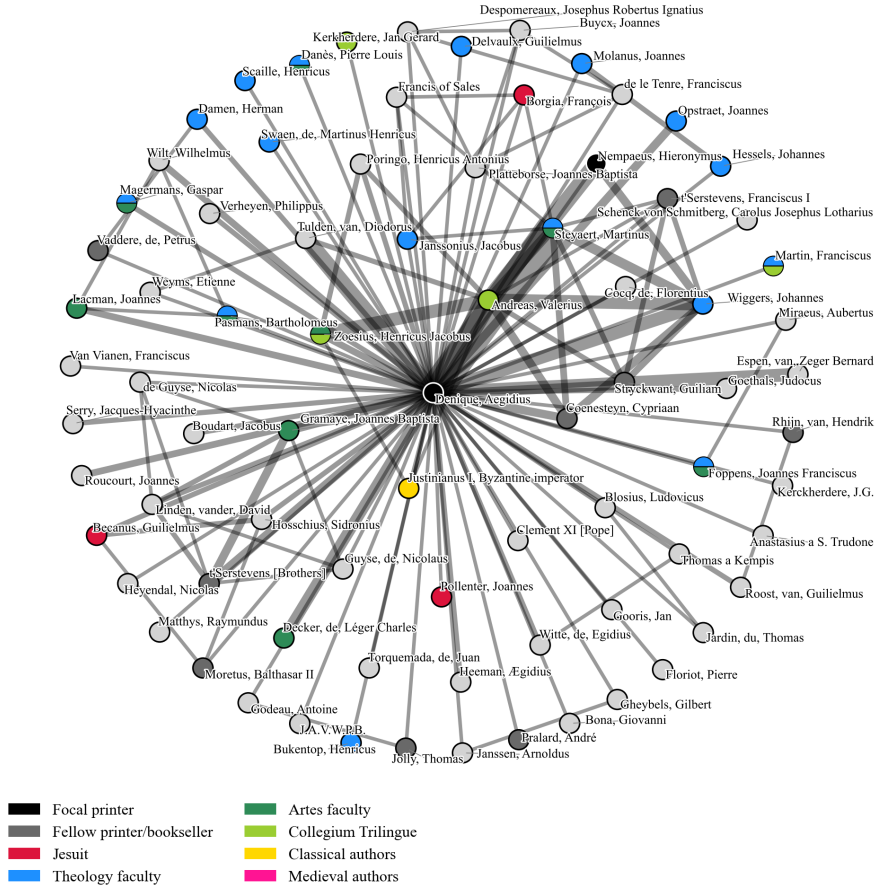



**Figure 4.9.** Ego network of Aegidius I Denique. Nodes represent individuals appearing in at least one book published by the focal printer; edges indicate co-occurrence in the same publication. Edge thickness is proportional to the number of shared books. The focal printer is shown in black. Nodes are color-coded according to categorical attributes: fellow printers or booksellers (dim gray); Jesuits (crimson), affiliation with Louvain's Theology Faculty (blue), Artes Faculty (green), and Collegium Trilingue (yellow green); classical authors (gold); and medieval authors (pink). When an individual possesses up to two of the listed attributes, the node is split into two colored semicircles.

Overall, the configuration of the network of Denique does not diverge sharply from Nempaeus's. Comparable in size (78 partners). It displays a similarly high rate of repeated collaboration (35.9 percent), combined with low ego density (0.02), again pointing to a centralized structure. Classical and medieval authors remain marginal (1.3 and 0 percent, respectively). The share of university scholars remains substantial (41.8 percent), with affiliates to Louvain's faculties of Artes (11.5 percent) and especially Theology (20.5 percent) continuing to dominate. At the core of Denique's network, in fact, stands a group of Louvain theologians: Johannes Hessels, Johannes Wiggers, Henricus Jacobus Zoesius, and Johannes Molanus—all associates of Nempaeus as well. In most cases, these associations reflect posthumous editions and the continued circulation of established academic authorities. Other prominent actors in both Denique's and Nempaeus's networks are Valerius Andreas (Wouter Dries), professor at the university and Collegium Trilingue and first librarian of the university, and Louvain jurist Diodorus Tuldenus (Diodorus van Tulden). Andreas was a contemporary of Nempaeus and co-occurred with him often in the role of editor and commentator, including in the 1689 *Commentaria*. Tuldenus likewise aligns more with Nempaeus's period of activity. Yet it was Denique who reprinted several of his works posthumously, often corrected and enlarged, between 1701 and 1715. This period, the beginning of the eighteenth century, coincides with Nempaeus's decreased activity and death and Denique's gradual consolidation as a key figure in the Louvain scene. Appointed as Beadle of the Faculty of Arts in 1684, a position he held until 1719, Denique occupied a stable institutional role within the university. His activity can therefore be understood as the institutional consolidation of a network originally forged amid doctrinal controversy, transforming a structure shaped by seventeenth-century theological conflict into a more stabilized system of academic canon preservation.

## 4.6 Conclusions

In this paper, we have examined the ego-networks of five selected Louvain printers from the fifteenth to the eighteenth century, based on our curated dataset Lov'Ego. From this diachronic network analysis, it emerges that working in the same university town did not translate to similar academic networks or partnerships. Rather, the visualizations of ego-networks showed three distinct configurations. First, the networks around Martens and Rescius

were locally anchored in academic circles, with pronounced (but not exclusive) humanist orientations. Second, Masius's network was marked by recurring partnerships, yet these did not unfold within a tightly knit intellectual milieu, reinforcing the characterization of his activity as that of a jobbing printer. Johannes Masius ultimately relied on commissioned and second-tier work in the Louvain book market. Finally, the networks associated with Jansenist-leaning printers such as Nempaeus and Denique displayed strong spiritually driven affiliations, suggesting that spiritual affiliation could outweigh purely economic considerations. If we consider all five printers to be "academic printers" based on their print portfolio with academic authors (rather than based the early modern qualification of academic printer), than it seems that the category is highly diverse: some printers were closely aligned with the ongoing scholarly debates (e.g. humanist-printers) or theological debates (e.g. Jansenist printers) while others just took on commissions from scholars and academic authorities (e.g. the jobbing-printer). This is an important observation for the field of book history, which generally focuses on the aggregated print production in academic towns. Our observations encourage to better discern the individual strategies within the same academic town, and within time-bounded scholarly debates. Some methodological observations remain. First, gender was not part of the initial analytical framework. Because of the selection procedures, only widow-printer Paschasia Zangrius (widow Zegers) came into view, but other contributions have made clear that women could attain some degrees of centrality in university towns (Soen/Fantoli forthcoming). Second, had thesis sheets been included in the scope of data collection, jobbing printers such as Masius would have registered substantially higher centrality scores. Finally, it is worth mentioning that our initial working hypothesis - that Catholic censorship (directed not only at Protestant texts but also at Catholic works with Jansenist affiliations) would visibly shape these networks - could not be confirmed; rather, on the contrary. This may partly stem from the absence of censors from the initial dataset, but the evidence also shows that printers such as Nempaeus and Denique operated within tightly knit partnership networks willing to challenge censorship: these connections enabled them to publish Jansenist material even when such works were under close scrutiny or explicitly prohibited by the Archbishop.

The implications of our analysis extend beyond the specifics of Louvain's book market. By demonstrating that the composition of a printer's network

varies systematically with a printer's business model, the article contributes to broader historiographical debates about the organization of early modern print culture. What we offer here is a data-driven framework that in principle could be tested against more case studies and other university towns, provided comparable bibliographic and prosopographical data exists or is assembled. Fundamentally, we illustrated how computational approaches like SNA can serve as diagnostic tools for book historians, complementing close reading by making patterns of collaboration and specialization systematically visible at a larger scale.

# Appendix

## 4.A Detecting Johannes Masius (d. 1616) in Caa/STCV

For the 145 records originally associated with 'a' Jean Masius in the raw data of Caa and STCV, we have retained only 113 items, based on the steps outlined in this appendix, and after verification and disambiguation with the data present in ODIS (compiled by Alexander Soetaert, Heleen Wyffels and Dieter Cammaerts for their respective dissertations). In preparation for the publication of this article, Violet Soen has, however, recalibrated identifications and links in ODIS as outlined here. First, we have readjusted and corrected false identifications (both in STCV and Caa) and disambiguated the three homonymous men behind the attestation of 'printer' Jean Maes/Johannes Masius (Jan Maes, apud Iannem Masium, excudebat Ioannes Maes/Maesius...), based on what we can accept as certain after the detailed 2018 study of Christopher Warner (the exact nature of the kin relationships between these three men however remain uncertain). His 'conjectures' seem to be validated by the fact that there is a serious time gap between Johannes Masius (d. 1616), with a last registered book in 1611 (in ODIS), and the printer Johannes Masius (?-1647/8) with publications from 1624-1625 onwards until 1654.

- We are concerned with Jean Masius, Johannes Masius (d. 1616), "Lovaniensis" active in Antwerp and Louvain, first as bookseller (from 1557 onwards), consequently also as printer (from 1567 onwards): http://www.odis.be/lnk/en/PS_127109
- We excluded attributions to the relative Jean Masius (c. 1568-1622), "Athensis", active possibly as apprentice in Louvain, but certainly in Ath, active before and after 1616: http://www.odis.be/lnk/en/PS_130516, except for those items where they collaborated in a joint venture:

| Place | Year | STCV ID | Title | Author ID | Author |
|---|---|---|---|---|---|
| Athi,Lovanii | 1601 | c:stcv:12925263 | Opera | a::06.10078 | Masius, Joannes II |
| Athi,Lovanii | 1601 | c:stcv:12925263 | Opera | a::06.3153 | Masius, Joannes I |
| Athi,Lovanii | 1601 | c:stcv:12925264 | Aeneidos | a::06.10078 | Masius, Joannes II |
| Athi,Lovanii | 1601 | c:stcv:12925264 | Aeneidos | a::06.3153 | Masius, Joannes I |

- We excluded attributions to the relative Jean Masius (?-1647/8, buried 2 April 1648),[32] printing between 1624-1654 (not yet listed in ODIS now, a unique ID will be created for the publication)
- We excluded abbot Johannes Masius (1592-1647, abbot of the Norbertine Abby of Park near to Louvain between 1635-1647) who never was a printer himself but he was an active book patron ("mecenas") and thus appears as dedicatee: http://www.odis.be/lnk/en/PS_178850. His exact relationship to the printing family remains uncertain.

Secondly, we have excluded some additional publications from the dataset to ensure harmony. We have included texts in the vernacular when relating to religious purposes or when translated from Latin texts by Louvain scholars; however, we excluded:

- Publications for the Spanish Habsburg dynasty/Governor-General (e.g. Placcaet ons heeren des conincx: aengaende de voldoeninghe, moderatie ende quijtscheldinghe vande chijnsen, renten, pachten [...] op landen [...] duer dese tegenwoirdige troubelen geruineert [...] gebleuen zijnde, 1583, ... c:stcv:12926634)
- Print jobs for the archbishop (Clare. Belijdinghe des christen gheloofts [...] deur ordinantie ende beuel van [...] Ian Hauchin eerts-bisschop van Mechelen wtghegheuen, 1585, c:stcv:12925152)
- Print jobs for the city (and the variations) (e.g. Ordinantien ende statvten, opt beleyden vanden processen, administreren vander iusticien, salarisen vanden officiers, ende tgene des daer aen cleeft, gemaeckt [...], inden iaere xvc. xcijtich. c:stcv:12913155)
- For the Annuntiaten (close to Leuven) (Corte devote oeffeningen op de veertich wercken der maghet Marie, begrepen inde [...] reghels vande orden der annvntiaten, 1614 c:stcv:3140551, or Het beghinsel oft oorspronck vande orden onser L. Vrovwen, ghenoempt annvntiaten, 1614 c:stcv:3140552)
- Print jobs of news events, e.g. siege of Lier 1596 (Joh. Bernartii De Lirani oppidi ab Hollandis occupati, per Mechlinianos et Antverpianos admirabili liberatione commentariolus ... c:stcv:7038560)

[32] *Lavdatio fvnebris ... D. Ioann. Masii ... Dicta a Iacobo Pontano in academia Lovaniensi*, Anno 1648, Die 2 April.

- thesis sheets (Infra-scriptas ingenvi ac probi adolescentes svbscripti, artivm baccalavrei, pædagogii Falconensis alumni, publicè tueri conabuntur, Louanij in scholis artium, 25. septembris, anno 1602, c:stcv:12911845Theses or Dispvtatio clementinæ: si furiosus. De homicidio [...] præside [...] Andrea Kemmerio [...] defendet Melchior Moretvs [...] xv. febrvarii. Lovanii, anno 1597 c:stcv:12919367)

# 5

# Network-mediated screening under uncertainty: Evidence from early modern academic publishing

*In markets where quality is difficult to observe, decision-makers may rely on social ties to screen unfamiliar counterparts. I study this mechanism in the early modern academic book market of the Low Countries, where printers selected scholarly manuscripts for publication under commercial risk and incomplete information. The Dutch Revolt (1568–1648), which split the region into the Dutch Republic and the Spanish Netherlands, intensified these conditions by disrupting the institutional environment in which printers operated. Combining bibliographic records from the Short Title Catalogs of Flanders and the Netherlands with academic affiliation data from the Repertorium Eruditorum Totius Europae, I construct a novel database of 5,275 printer–author dyads spanning 1476–1800. Using a conditional logit model of the printer's choice among potential authors, I find that shared affiliation between a printer's prior collaborators and a new candidate strongly predicts selection, but only during periods of institutional disruption. The network channel is near zero before the 1560s, peaks around the Fall of Antwerp (1585), when the split of the Low Countries first emerged, and fades as the conflict approached its end.*

## 5.1 Introduction

In markets where information about individual quality is limited, economic agents may rely on mutual ties to evaluate unfamiliar counterparts. Networks are information channels (Montgomery, 1991), but even in absence of information, they allow members to inherit the collective reputation of the group (Tirole, 1996), and they sustain the relational trust that accumulates between parties through repeated interaction (Greif, 1993; Granovetter, 1985). Reliance on such mechanisms becomes especially salient in two contexts: (i) during periods of institutional disruption, when established systems of assessment are destabilized (North, 1990), and (ii) in knowledge and innovation markets, where quality is inherently difficult to observe ex ante (Nelson, 1959). Under these circumstances, the weight placed on mutual ties may shape market coordination and opportunity allocation, with consequences for the broader trajectory of knowledge production. The same informational logic operates in contemporary entrepreneurial finance, where investors rely on network referrals to overcome the severe information asymmetry of early-stage funding, and where an investor's network centrality predicts investment performance (Hochberg et al., 2007).

The early modern academic book market of the Low Countries offers a setting in which to examine this prediction. Academic publishing produced scholarly works written by the 'upper tail of human capital' and played a key role in Europe's knowledge economy (Pettegree, 2008; Squicciarini and Voigtländer, 2015; Mokyr, 2016). Printers were the gatekeepers of this market, selecting manuscripts and committing capital and reputation to their publication. They operated under persistent commercial risk, since the success of a book was uncertain. During periods of instability, they also faced rising political and ideological risk. The Dutch Revolt, beginning in the late 1560s, disrupted this environment, introducing instability into a previously integrated market and altering the conditions under which printers made publication decisions. The question this paper asks is whether printers turned to the networks they had built with scholars, through prior collaborations, as a way of screening unfamiliar candidates. Historical scholarship anecdotally documents the circulation trust within scholarly communities (Maclean, 2012; Mauelshagen, 2003), but there is little systematic evidence on the extent to which these ties shaped who was being published, or whether their role intensified during periods of heightened uncertainty.

By combining bibliographic data from the Short Title Catalogs of Flanders and the Netherlands with academic affiliation records from the Repertorium Eruditorum Totius Europae, I examine whether shared affiliation between printers' prior collaborators and new candidate authors became a stronger predictor of selection as institutional disruption intensified. Three channels could produce this pattern. Information transmitted by previously published authors about candidates they know through shared affiliation. Trust and mutual recognition accumulated within scholarly communities, which extend to new members even without direct information about them. And the printer's own experience with scholars from a given community, which provides evidence about what publishing such scholars tends to yield. I find that the association between printers' prior ties to a scholarly community and their selection of new authors from that community strengthened during the decades surrounding the onset of the Dutch Revolt and weakened as stability returned.

The paper proceeds as follows. Section 5.2 discusses the related literature. Section 5.3 describes the institutional context of the academic book market of the Low Countries. Section 5.4 presents the data used. Section 5.5 introduces the empirical strategy. Section 5.6 reports the results. Section 5.7 concludes.

## 5.2 Related literature

This paper speaks to several strands of the literature, including information and labor economics, quantitative economic history, and the study of knowledge markets. Classical models in information economics show how incomplete information affects market selection and how observable characteristics can reveal hidden quality (Akerlof, 1970; Spence, 1973; Stiglitz and Weiss, 1981). Printers in the early modern book market faced an analogous pre-contractual problem: committing capital and reputation to a manuscript required inferring an author's quality, reliability, and ideological conformity. This setting also resembles delegated investment under imperfect signals, where agents must allocate resources despite limited verifiability (Bhattacharya and Pfleiderer, 1985; Palomino and Prat, 2003). I contribute to this literature by studying author selection in a knowledge market where an institutional change alters the cost of selecting the wrong candidate.

A distinct body of work in labor economics studies how decision-makers form beliefs about agent types under uncertainty, and identifies three informational channels that are relevant to the printer's problem in my

setting. The first channel concerns information transmitted through social ties. Tie-holders who know a candidate can convey private information about their reliability and quality, reducing the informational frictions that complicate selection (Montgomery, 1991; Topa, 2001; Dustmann et al., 2016). The second channel concerns group-level reputation. Tirole (1996) demonstrates that the reputation of an institution spills over onto its individual members, so that decision-makers use group membership as a basis for inference when direct information about an individual is scarce. The third channel concerns accumulated experience. Jovanovic (1979) models employer learning about worker productivity, and Farber and Gibbons (1996) and Altonji and Pierret (2001) show empirically that employers update beliefs about worker types as direct experience accumulates. Each of these channels is likely to be operating in the early modern academic book market. Printers' past collaborations with scholars of a given affiliation provide direct evidence about what publishing such scholars tends to yield, repeated interaction with a scholarly community generates mutual recognition that extends to colleagues, and previously published authors may themselves transmit information about candidates they know through shared affiliation. All three channels predict that candidates from communities the printer has stronger ties to are more likely to be selected, and that this reliance intensifies when uncertainty rises.

This paper also speaks to the notions of embeddedness and social capital in sociology, which holds that economic transactions are shaped by the social structures in which they occur rather than taking place between isolated agents (Granovetter, 1985), and that interpersonal networks create the trust that makes productive activity possible (Coleman, 1988). A parallel literature in economics, from Greif (1993) to Guiso et al. (2004), shows that the value of such networks rises when formal institutions are weak or disrupted. In early modern academic publishing, printers and scholars operated within tight networks of trust, on which they likely relied as institutional, political, and confessional change swept the Low Countries.

A central difficulty in studying how social structure shapes economic outcomes is that ties tend to concentrate within homogenous groups (Manski, 1993; Golub and Jackson, 2012; Gentzkow and Shapiro, 2011). Early modern academic and print networks were no exception, especially after confessionalization became a more prominent feature of European academia. de la Croix and Morault (2026) document how the Reformation reconfigured

the European academic network, with heterogeneous effects on the scientific output of university professors. These institutional shifts also reshaped the environment in which printers operated, as the same period saw the rise of censorship as an institutional force affecting knowledge production. Blasutto and de la Croix (2023) quantify the impact of Catholic censorship on scientific output, while Comino et al. (2024) document how inquisitorial oversight altered incentives for creativity and publication. I contribute to this literature by shifting the unit of analysis from aggregate output to the matching decision inside the market: the printer's choice about whom to publish, conditional on their position in networks of past partnerships and on the academic ties of the pool of potential new authorial counterparts.

## 5.3 Historical context

### 5.3.1 The printing industry and academic authors

The introduction of movable type to Europe in the mid-fifteenth century created a new industry with high fixed costs and uncertain returns.[1] Producing a book required substantial upfront investment in labor and quality paper (Dittmar, 2011), and costs were difficult to recoup quickly: "[a]n edition of a single title required, on the basis of the above figures, more money than the initial equipping of a printing shop" (Febvre and Martin, 1997, p. 111).

As profit-seeking entrepreneurs, printers sought manuscripts that promised commercial success.[2] The printer's primary risk was economic: was the author good, would the manuscript sell, was the scholar reliable enough to justify committing capital and reputation to a print run? This risk was present at all times, in peace as in war, but periods of instability layered additional ones

[1] Gutenberg's innovation, which spread across the continent within fifty years of its introduction around 1440–50, combined movable metal type, ink, paper, and the press into a single system, as opposed to the woodblock printing techniques used in previous decades in Europe and elsewhere in the world, like China and Korea, where movable type — including metal type — predated Gutenberg by centuries, though it never supplanted woodblock printing in East Asia.

[2] Early modern printers usually combined editorial judgment, financial risk-taking, and manufacturing under one roof. For larger or riskier projects, they occasionally formed partnerships with other printers and booksellers or sought patronage from wealthy individuals or institutions to share the costs of a print run.

on top of the commercial risk, such that networks of trust, possibly already used to gauge quality, became correspondingly more valuable when the stakes of misjudgment rose. In the early decades of print, printers were notably cautious in their selection of texts. One common strategy to manage risk was to develop a reputation in a particular segment of the market (e.g. theology, law, classical editions), since high initial investment in specialized typefaces and expertise discouraged competitors from entering the same niche (Pettegree, 2010). Yet over-specialization carried its own dangers: shifts in taste or political circumstances could devastate a house that had invested too narrowly. The most successful printers balanced specialization with diversification, reinvesting the returns from a strong field into broader ventures—sometimes beyond the printing industry itself (Maclean, 2012, p. 130).

Once a printer identified a promising manuscript, the next step was to secure a *privilege*, a primitive form of intellectual property right granted by a local territorial authority, typically conferring an exclusive monopoly to print and sell a given title for a fixed number of years within a specific jurisdiction. In the Low Countries, these privileges were issued by institutions such as the Council of Brabant and the Privy Council. In the Northern provinces, after the *de facto* secession from the Spanish Netherlands, the institutional framework differed: after the Revolt, it was the States of Holland and the other provincial States that assumed this regulatory function. Regardless of the specific jurisdiction, the system of privileges served as the principal mechanism through which regulatory authorities policed unauthorized reprinting; without a privilege, a printer had no legal recourse against competitors who copied the text (Pettegree, 2010; Johns, 1998).

The majority of privileges were held by printers and publishers, not by authors. This asymmetry reflected the economic structure of the trade: authors typically sold or donated their manuscript to a printer, thereby transferring all intellectual property rights unless a contract specified otherwise (Munck, 2019, p. 59). The few authors who obtained privileges in their own name were usually either exceptionally well-connected at court or independently wealthy (Maclean, 2012, p. 167). For most authors, publication brought reputational rather than financial returns, while the economic rights often remained with the printer.[3] Indeed, printers frequently used the intellectual

[3]Erasmus was a rare and celebrated exception: almost certainly the first living author to earn a substantial income from writing. Even after his death, few sixteenth-century authors managed to replicate his commercial success (Pettegree, 2010).

prestige of their authors as a marketing device, promoting their credentials and reputation (including their prestigious academic affiliations) to attract buyers (Cammaerts, 2024b).

This asymmetry in economic power also shaped who initiated the publishing relationship. In the majority of cases, it was the printer who sought out manuscripts and selected authors, rather than authors approaching printers with unsolicited work. There were, of course, some exceptions: Louvain professors Johannes Driedo and Michael Baius, for instance, are documented as having taken the initiative to submit manuscripts to printers (Cammaerts, 2024b, p. 155). My data are consistent with this characterization: in the sample, the number of distinct printers is consistently smaller than the number of distinct authors, suggesting that most often it was the printer who selected from a pool of candidate scholars, and not the other way around. This pattern motivates a key modeling choice in this paper: I treat the printer as the decision-maker who chooses among candidate authors, while acknowledging that this is a simplifying assumption of the historical complexity and that in some cases the initiative ran in the opposite direction.

### 5.3.2 The Low Countries

The region was a patchwork of semi-independent counties and duchies, each with their own laws and privileges. The Dukes of Burgundy gradually assembled most of them through inheritance and purchase, creating what is often called the "Burgundian Netherlands", a loose personal union rather than a unified state (Arblaster, 2018; De Ridder and Soen, 2020). When the last Burgundian duke Charles the Bold died in 1477, the territories passed to the Habsburgs through his daughter Mary, who married the Archduke Maximilian I of Habsburg. Their son Philip the Handsome married Joanna of Castile, uniting the Burgundian and Spanish inheritances in their son Charles V, who became both Holy Roman Emperor and King of Spain.

By 1556, when Charles V abdicated, he divided his domains: the Holy Roman Empire passed to his brother Ferdinand I, while the seventeen provinces of the Low Countries[4] went to his son Philip II as part of the Spanish Crown.

[4]The seventeen provinces included: the Duchies of Brabant, Limburg, Luxembourg, and Gelderland; the Counties of Flanders, Artois, Hainaut, Holland, Zeeland, and Namur; the Lordships of Friesland, Mechelen, Overijssel, Groningen, and Drenthe; the Bishopric turned Lordship of Utrecht; and Tournai with the Tournaisis. The exact enumeration varies between sources, as the nature of historical boundaries before the emergence of national states is

Roughly covering modern Belgium, the Netherlands, Luxembourg, and parts of northern France (chiefly Artois), the Low Countries now fell under Spanish rule, and with it, under the religious authority of the Roman Catholic Church.

Meanwhile, the ideas of the Protestant Reformation, and particularly Calvinist movements, had been gaining ground across the region, clashing with Charles V's and Philip II's determined enforcement of Catholic orthodoxy. The resulting religious tensions, compounded by resentment of centralizing Spanish rule and heavy taxation, set the stage for open revolt. Tensions boiled over in the summer of 1566 with the Iconoclasm or Iconoclastic Fury (*Beeldenstorm*), a wave of Protestant-led destruction of Catholic church interiors that swept across the provinces. Philip II responded by dispatching the Duke of Alba with a large army and a special tribunal, the Council of Troubles, popularly known as the Blood Council (*Bloedraad*), to crush dissent. Alba's repression, far from restoring order, pushed the provinces toward outright rebellion: the so-called Dutch Revolt against Spanish Habsburg rule, also known as the Eighty Years' War (1568–1648).

The seven northern provinces (Holland, Zeeland, Utrecht, Gelderland, Friesland, Groningen, and Overijssel) formed the Union of Utrecht in 1579 and declared independence as the Dutch Republic.[5] The Fall of Antwerp to the Duke of Parma in 1585 proved decisive: the front line where Spanish forces halted their reconquest effectively became the *de facto* border between North and South. The southern provinces (Flanders, Brabant, Hainaut, Namur, Artois, Luxembourg, Tournai, and the lordship of Mechelen) remained under Spanish control as the Spanish Netherlands. This divide ran roughly where the modern Dutch-Belgian border sits today, though it was imperfectly aligned: a strip of territories known as the Generality Lands (*Generaliteitslanden*), including States-Brabant (modern North Brabant) and parts of Limburg, were conquered by the Dutch Republic but never granted full provincial status.

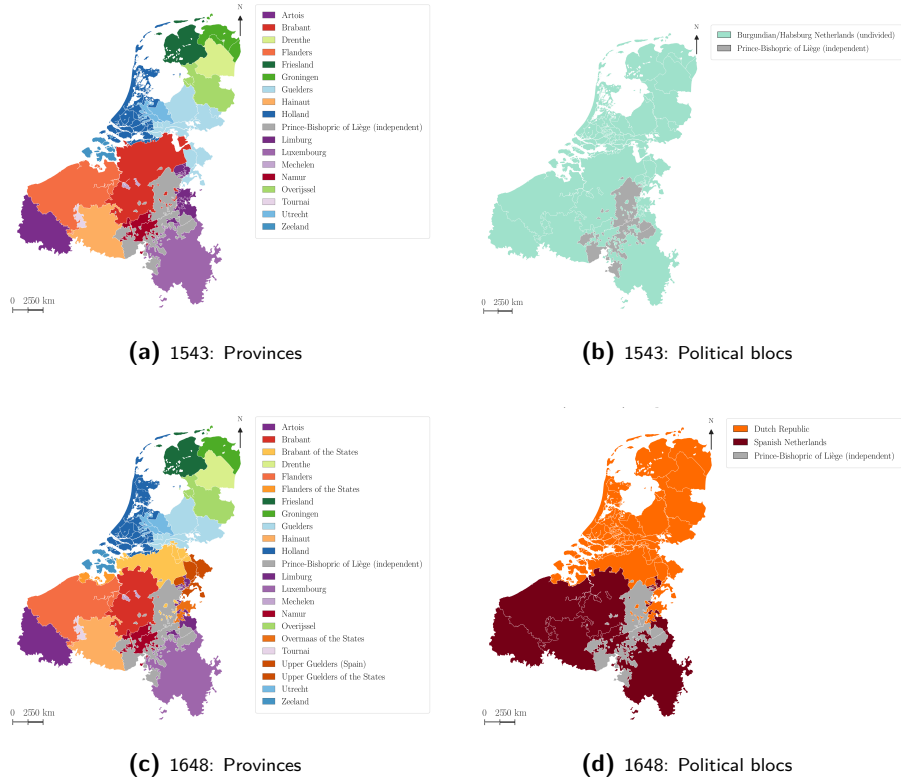
Importantly, the border between North and South was not the product of a pre-existing confessional or cultural divide: before the revolt, Antwerp had a substantial Calvinist community, and along Brussels and Ghent, it was under Calvinist control from 1579 and 1581 (Arblaster, 2018, p. 126), while

inherently complex; the number "seventeen" was itself partly a political construct formalized by the Pragmatic Sanction rather than a precise administrative description.

[5]Drenthe participated in the revolt but lacked full provincial status and held no vote in the States-General.

Amsterdam remained largely Catholic well into the 1570s (Arblaster, 2018, p. 133). The line was drawn where Parma's army stopped.

Figure 5.1 illustrates the territorial evolution of the Low Countries at two key moments: the completion of the seventeen provinces under Habsburg rule (1543), and the crystallization of the North–South divide after the end of the Dutch revolt (1648). Each moment is shown by administrative province (left) and by political bloc (right).




**(a)** 1543: Provinces **(b)** 1543: Political blocs

**(c)** 1648: Provinces **(d)** 1648: Political blocs

**Figure 5.1.** Administrative divisions and political blocs of the Low Countries before and after the split. The top row shows the unified Seventeen Provinces in 1543; the bottom row shows the North–South division at the end of the Dutch Revolt. The polygons come from the Mint Authorities of the Low Countries (Stapel, 2025).

The Low Countries occupied a central position in the book trade, supplying printed books well beyond their borders despite the political turmoil that engulfed the region throughout the sixteenth century (Pettegree,

2008). According to historical accounts, in this period, the southern provinces of the Low Countries (specifically, the County of Flanders and the Duchy of Brabant) were among the most urbanized region of Europe, along with the Po Valley and the Naples area in the Italian peninsula (Arblaster, 2005). They also dominated the book trade, with Antwerp as the "epicenter of Nederlandish print" (Buning et al., 2024, p. 2). The city's position on the Scheldt river made it a hub of international commerce (Portuguese merchants, among others, based their trading operations there) and its proximity to the University of Louvain, the first university in the region (founded 1425), had nurtured a printing trade since the late fifteenth century. It was in Louvain that Dirk Martens, a printer in the circle of Erasmus of Rotterdam, published the first edition of Thomas More's *Utopia* in 1516 (François, 2020). The South's commercial vitality attracted migrants from across Europe; for instance, Christophe Plantin, originally from France, who established in Antwerp one of the most important printing houses of the period. Northern cities such as Utrecht and Deventer also sustained active presses, but they remained secondary to Antwerp (Pettegree, 2008).

The balance shifted decisively in the last quarter of the century. As the Spanish reconquered the southern provinces, waves of migrants (printers, scholars, merchants, and skilled artisans) fled north to the newly independent Dutch Republic. The founding of Leiden University in 1575 provided a gravitational pull for intellectual life, and as Europe's commercial center of gravity moved from Antwerp to Amsterdam, the northern provinces inherited the conditions for a flourishing book trade of their own.

Throughout this period, the production and circulation of printed books was subject to an evolving censorship apparatus that preceded, and in important respects anticipated, the one later adopted by the Holy See.[6] From the 1520s onward, a series of imperial edicts issued under Charles V imposed escalating penalties on the printing, distribution, and possession of heretical material (Agten, 2017; Vanden Bosch, 2017b). The centerpiece of the regime was a system of preventive licensing: every manuscript was to be submitted to a censor, typically a theologian, who certified that the work contained nothing contrary to the faith. The risks for printers were severe and well documented—fines, confiscation of stock, revocation of licenses, and in extreme cases imprisonment (Comino et al., 2024)—but the formal system

[6]This was not unique to the Low Countries; as Pettegree (2010) writes: "[f]rom the 1520s onwards most states with a local book industry had taken measures to regulate production."

was neither uniform nor consistently enforced (Pettegree, 2010). Enforcement depended on local magistrates and ecclesiastical authorities whose priorities varied widely across provinces and cities and who frequently resisted or delayed the publication of anti-heresy edicts to protect trade and commerce (Arblaster, 2018). The proliferation of placards itself, with each new edict implicitly acknowledging the failure of its predecessor, testifies to the difficulty of implementation. Censorship is best understood here as one additional source of the heightened uncertainty printers faced, layered on top of the ever-present commercial risk inherent to the printer's activity, raising the penalty for a misjudged selection: a manuscript could fail in the market because of low demand or poor quality, but it could now also be prohibited outright.

The Fall of Antwerp in 1585, drawing the *de facto* political border between the Dutch Republic and the Spanish Netherlands, also triggered a broader reorganization of the institutional environment in which printers operated. Established universities lost cross-border reach, new institutions consolidated their position (like the universities of Leiden, mentioned above, and of Franeker, founded in 1585, among others), waves of scholars and printers migrated across the new frontier, and the unified market that had previously linked the Low Countries fragmented into two distinct commercial zones (Arblaster, 2018; Israel, 1995). Within this broader reorganization, censorship regimes also crystallized on both sides: in the South, Catholic authorities consolidated preventive licensing and ex-post oversight under a clearer division of institutional responsibilities; in the North, the Dutch Republic maintained censorship law on paper but enforcement was dispersed across provincial and municipal jurisdictions, producing a more permissive but jurisdictionally uneven regime (Van Eijnatten, 2004; der Weduwen, 2021).

## 5.4 Data

Data for the empirical analysis come from distinct sources that I combine into a novel database. Academic affiliation records, from which I track scholarly careers and infer co-presence at the same institution, come from the Repertorium Eruditorum Totius Europae (henceforth, RETE). RETE records the affiliations of European university professors and academy members from the inception of the first medieval universities until the First Industrial Revolution, as well as a wide set of scholar attributes, including biographical

information, fields of teaching and research, and affiliation to religious orders such as the Jesuits.

Data on publications derive from the Short Title Catalog Netherlands (STCN) and the Short Title Catalog Vlaanderen (STCV). Both are retrospective bibliographies covering today's Netherlands and the region of Flanders (Belgium), which combined partly correspond to the historic Low Countries. From the bibliographic metadata of catalog records, I extract information on published items including the names of the author and the printer/publisher, the year and place of publication, and the genre and topic.

My goal is to segment the academic book market of the region by isolating publications involving academically affiliated authors. To this end, I focus on the subset of books (i) involving RETE scholars as authors, (ii) printed in cities located in the historic Low Countries,[7] (iii) published by non-fictitious printers; from the earliest recorded year (around the 1470s) to 1800. In Table 5.A.1 in Appendix 5.A, I present an overview of how these restrictions affect the sample base provided by the two catalogs.[8]

For each RETE scholar found as an author in the STCN and/or STCV, I identify their publications in the catalogs, provided they occurred before their death, as I exclude posthumous publications.[9] This yields 1,804 academic authors, whose affiliations to universities and academies are summarized in Table 5.A.3 for the top 15 institutions, in the Appendix 5.A, published by 1,398 distinct printers. The distribution of academic titles across printers is concentrated: 39% of printers appear with a single publication, while 35% of

[7]The STCN and STCV also contain imprints with place of publication beyond the boundaries of the historic Low Countries.

[8]To what extent this subset was financially important for printers cannot be reconstructed from the available data. What can be approximated, however, is the share of each printer's surviving output that these academic titles represent. At present, this exercise is feasible only for the printers extracted from the STCV; for the STCN, the comparison is in principle possible but has not yet been carried out, as, given the catalog's size, I directly extract and process publications involving a RETE-affiliated author only. On average, academic titles account for roughly a third of recorded output (mean: 33%, median: 17%, interquartile range: 7–50%).

[9]I exclude posthumous publications because the matching decision behind such works often differed from that of ordinary trade publishing: the author cannot negotiate, and the initiative frequently lies with the deceased scholar's family and colleagues, who may commission or subsidize the edition as an act of commemoration (Maclean, 2012; Cammaerts, 2024b). The printer's selection in those cases would therefore not be a clean observation of a first-time printer–candidate author dyad in the sense my model assumes.

authors likewise publish only once in the sample. Table 5.A.2 in the Appendix 5.A details the full distribution. More summary statistics on the number of printers, authors, and dyads by decade are available in Table 5.A.4 in the Appendix 5.A.

The empirical analysis focuses on the first time each printer publishes with a given author: every dyad enters the sample as a single event, the year in which the printer and the author first appear together in print, and excludes any repeated collaboration between the same printer-author pair. This restriction isolates the decision to initiate a new publishing relationship, which is the margin on which the network channel is expected to matter.

Next, I track the collegial ties of these academic authors through institutional co-presence. Two scholars are considered colleagues if they held overlapping affiliations at the same institution, that is, if their tenure spells at a shared university or academy overlap temporally and if they worked in the same field, broadly defined (i.e., theology, law, humanities and sciences). From these co-presence records, I construct the key network variables, described in detail in Section 5.5.

To assign each publication to a historical jurisdiction, I use the place of publication recorded in the STCN/STCV. I match each city to its geographic coordinates and determine the administrative unit it belonged to at the time of publication using the Mint Authorities of the Low Countries (Stapel, 2025), a GIS dataset providing time-varying polygon boundaries for the major sovereign authorities in the region from the twelfth century onwards.

At any point in time, I can also assign each publication to a higher-level political entity: the Burgundian/Habsburg Low Countries, and following the Union of Utrecht (1579) and the Fall of Antwerp (1585) either the Dutch Republic in the north or the Spanish Netherlands in the south. In the analysis, I split the sample into North and South based on the jurisdiction each city falls into after 1585. Cities in provinces that become part of the Dutch Republic are assigned to the North; cities in provinces that remain under Spanish control are assigned to the South.

## 5.5 Empirical strategy

### 5.5.1 A random utility model of the printer's choice

The setting described in Section 5.3 motivates a discrete choice framework. The printer's publication decision can be viewed as a selection from a pool of candidate authors, where the printer's assessment of each candidate depends on observable characteristics of the candidate and of the printer–candidate author dyad. A conditional logit model provides a natural way to estimate which of these characteristics predict selection.[10] That said, what I observe in the data is simply a realized match between a printer and an author, which may reflect initiative on either side, but the conditional logit frames it entirely as the printer's choice. A direct consequence of this framing is that authors in a realized match are treated as willing to publish with whichever printer selects them, which is the same as assuming an elastic supply on the authors' side. I acknowledge that this abstracts from the two-sided nature of the publishing market, in which authors also held preferences, searched for printers, and at times initiated the relationship themselves. My assumption is thus not innocuous, but, as previously discussed, it is backed by historical evidence that the printer was the more plausible party to move first: printers were fewer than authors, held the economic rights, and typically sought out manuscripts.

The model rests on a random utility framework: each candidate generates a payoff for the printer that consists of an observable component and a random shock, and the printer selects the candidate yielding the highest realized payoff. Each publication decision defines a *choice occasion*: a single event in which printer $i$ selects one author from the pool of candidates active in year $t$. The corresponding *choice set* $\mathcal{C}_{it}$ is the collection of candidate authors available at that occasion, and a printer–candidate pair $(i, j)$ within the choice set is a *dyad*, the unit of observation in the econometric model. Because a printer may publish more than one author in a given year, each $(i, t)$ can contain several choice occasions, each entering the sample as a separate observation. The choice set itself is printer- and year-specific: each $(i, t)$ corresponds to a distinct set of

[10] The network reliance I aim to study lives in the variation across the candidates a printer faces at a single decision: holding the printer and the year fixed, my hypothesis is that the more connected candidate is the more likely to be chosen. This within-decision variation is what the conditional logit exploits, and a more aggregate design (say, a design built on aggregate counts per city) would discard it. Still, I discuss an alternative design exploiting the geography of the reconquest in Section 5.7.

scholars at risk of being published in that year, so the 'alternatives' the printer faces vary across choice occasions in both size and composition.

I focus on choice occasions in which the selected author is publishing for the first time with that specific printer, excluding repeated dyads that reflect the continuation of an existing relationship. For each choice occasion, the choice set consists of all academic scholars who publish at least once in the Low Countries during their career and whose academic career is active in year $t$. This restriction excludes academics who never published at all: in this period publishing was not the norm for every scholar, and one who never appeared in print with any printer over an entire career was plausibly not a participant in this market, so including them would pad the choice sets with alternatives no printer would realistically have weighed. The restriction does condition on an outcome, and I cannot rule out that the reasons some scholars never reached print are correlated with the network ties I study. If anything, the estimate is then a likely lower bound: the never-published are presumably the least connected, so adding them back would raise the weight of $P^{\text{same}}$, not lower it.

A scholar enters the risk set for year $t$ if their earliest recorded affiliation begins at or before $t$ and their career end (year of death, or latest recorded affiliation) falls at or after $t$. The resulting choice sets have a mean size of 250 and a median of 245. Every scholar who appears as a non-chosen alternative in at least one choice set is also observed as the chosen alternative in at least one other, confirming that the candidate pool consists of genuinely active participants in the market.

The utility printer $i$ obtains from publishing candidate $j$ in year $t$ is:

$$U_{ijt} \;=\; \beta\, x_{ijt} \;+\; \varepsilon_{ijt}, \tag{5.1}$$

where $x_{ijt}$ is a vector of dyad-specific attributes that vary across printer–candidate dyads within a choice occasion, $\beta$ is the associated parameter vector, and $\varepsilon_{ijt}$ is a random term capturing unobserved determinants of the printer's payoff.

Assuming $\varepsilon_{ijt}$ is independently and identically distributed Type-I extreme value across candidates within a choice occasion, and letting $y_{ijt} = 1$ if printer $i$ selects candidate $j$ in year $t$ and $y_{ijt} = 0$ otherwise, the probability of selection takes the conditional logit form:

$$P(y_{ijt} = 1) \;=\; \frac{\exp(\beta\, x_{ijt})}{\sum_{j' \in \mathcal{C}_{it}} \exp(\beta\, x_{ij't})}. \tag{5.2}$$

The probability of choosing a given candidate depends on the attributes of that printer–candidate dyad (the numerator) relative to the attributes of all other printer–candidate dyads in the choice set (the denominator).

The parameter vector $\beta$ is estimated by maximum likelihood. Because any component of utility that is constant across candidates within a choice occasion cancels from both numerator and denominator, the conditional logit absorbs all printer-level and year-level characteristics by construction, and estimation is driven only by variation in $x_{ijt}$ across alternatives within the same occasion. This same logic explains why the model does not need printer-by-year fixed effects: such an effect is identical for every candidate in a given occasion, so it cannot explain why one was chosen over another, and is already netted out together with every other characteristic of the printer and the year, whether observed or not.

The baseline specification includes a field-match control that captures whether a candidate's discipline aligns with the printer's predominant area of publication, but this addresses only the level of the printer's preference for a given field, and it does not correct the substitution pattern among candidates. Two theologians remain closer substitutes for each other than either is for a botanist, regardless of whether the printer specializes in theology. The conditional logit, by construction, cannot accommodate this asymmetry. To relax this restriction, in Section 5.6 I re-estimate the model as a nested logit, grouping candidates into nests defined by academic field. This structure allows candidates within the same field to be correlated in their unobserved attributes, so that substitution occurs primarily within fields rather than uniformly across them, while preserving the conditional logit structure for substitution across fields (McFadden, 1978).

### 5.5.2 Variables construction and specifications

The main explanatory variable, $P_{ijt}^{\text{same}}$, measures the number of candidate $j$'s university colleagues in year $t$ who have previously published with printer $i$ before year $t$ and who work in the same field as $j$. Formally, it counts the size of the intersection between $j$'s set of co-present colleagues at their current institution and the set of scholars appearing in printer $i$'s publication record prior to year $t$, restricted to those sharing at least one field of activity with $j$. The variable is transformed using the inverse hyperbolic sine to accommodate

zeros and reduce the influence of extreme values.[11]

By counting only colleagues who published with printer $i$ at least once before year $t$, I ensure that the variable reflects information available to the printer at the time of the decision. This rules out mechanical reverse causality from the decision itself feeding back into the network measure, though it does not eliminate other forms of endogeneity.

Two related variables decompose the network channel further. $P_{ijt}^{\text{diff}}$ counts colleagues who published with the printer but work in a different field from $j$, capturing cross-disciplinary information spillovers. $P_{ijt}^{\text{pred}}$ counts scholars at the same institution and in the same field whose tenure ended before $j$ arrived, that is, institutional predecessors who could not have been personal contacts but whose prior publication with the printer signals an established institutional connection. Together, the three variables distinguish personal same-field ties ($P^{\text{same}}$), personal cross-field ties ($P^{\text{diff}}$), and institutional reputation ($P^{\text{pred}}$). They are also meant to give some empirical traction to the channels discussed in Section 5.2. In particular, $P_{ijt}^{\text{pred}}$ counts colleagues whose tenure ended before the candidate arrived, so by construction they never overlapped with him and cannot have passed on private information about him. A positive coefficient of $P^{\text{pred}}$ would therefore reflect group-level reputation and/or the printer's accumulated experience, but not the personal-information channel, which only co-present colleagues ($P^{\text{same}}$) can carry. I return to this in Section 5.6.

The baseline specification introduces four controls that vary across candidates or across printer–candidate dyads within a choice set. *Birthplace distance* is the haversine distance between the printer's place of activity (which in many cases coincided with their own birthplace) and the candidate's birthplace, transformed using the inverse hyperbolic sine.[12] *Total colleagues* is the number of co-present colleagues at the candidate's institution in year $t-1$, capturing university size as a mechanical scale factor for the network variable. *Field match* is an indicator equal to one if the candidate's field matches the printer's modal field of publication, computed using an expanding window of the printer's past output up to but not including year $t$. *Experience* is the number of years since the candidate's first recorded university affiliation, serving as a proxy for academic seniority.

[11] The inverse hyperbolic sine, $\text{ihs}(x) = \ln\left(x + \sqrt{x^2+1}\right)$, approximates the natural logarithm for large values while being defined at zero.

[12] Birthplace is unknown for approximately 5% of scholars in the sample; for these observations, distance is imputed at the sample median.

To assess whether the network channel strengthens during periods of political uncertainty, I interact the main regressor $P^{\text{same}}$ with a binary indicator that splits the sample at a given year. Rather than imposing a single cutoff year, I re-estimate the interaction across a sequence of possible cutoff years from the 1540s onward, plotting the coefficient and its confidence interval against each cutoff to trace when the network channel emerges. In other specifications, I will split the sample at the year 1585, the Fall of Antwerp, which marked the *de facto* territorial separation between the Dutch Republic and the Spanish Netherlands. This date also provides a natural boundary for regional comparison, as the Northern printing market remained too small before the mid-1580s to support separate estimation. Table 5.A.5 in Appendix 5.A presents descriptive statistics of the variables defined above.

## 5.6 Results

**Baseline.** Table 5.1 reports conditional logit estimates on the pooled sample of choice occasions. Specification (1) includes only $P^{\text{same}}_{ijt}$, the number of candidate $j$'s same-field colleagues who have previously published with printer $i$. Specifications (2) through (5) progressively add the full set of controls: birthplace distance, total colleagues at the candidate's institution, the field-match indicator, and years of the candidate's academic experience. The coefficient on $P^{\text{same}}$ is stable across specifications, indicating that the association does not shift when progressively controlling by the dyad's cultural proximity, institutional size, the printer's field of specialization, or the candidate's seniority. The controls behave sensibly. Birthplace distance, which proxies for cultural and linguistic similarity between printer and author, enters negatively. The estimated coefficient of total colleagues is negative too because it rescales the network measure: holding $P^{\text{same}}$ fixed, the same number of connected colleagues signals a weaker tie when the candidate's faculty is large than when it is small, so a bigger faculty lowers the odds of selection. Field match enters positively, in line with printers being more likely to select within their own area of specialization. The one coefficient that may look counterintuitive is that of experience, which is negative; this is consistent with the empirical design, as each event is a *first* collaboration between a given printer–author pair. More senior scholars have likely already established their publishing relationships, so they are less likely to form a first-time partnership with a printer *later* in their career.

**Table 5.1.** Baseline specifications.

| | Dependent variable: $P(y_{ijt} = 1)$ | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| $P_{ijt}^{\text{same}}$ | 1.008*** | 0.866*** | 1.013*** | 1.014*** | 1.008*** |
| | (0.026) | (0.027) | (0.028) | (0.028) | (0.028) |
| Birthplace dist.$_{ij}$ | | −0.221*** | −0.211*** | −0.214*** | −0.207*** |
| | | (0.007) | (0.007) | (0.007) | (0.008) |
| Total n. colleagues$_{jkt}$ | | | −0.133*** | −0.133*** | −0.135*** |
| | | | (0.007) | (0.007) | (0.007) |
| Field match$_{ij}$ | | | | 0.222*** | 0.233*** |
| | | | | (0.034) | (0.034) |
| Experience (yrs)$_{jt}$ | | | | | −0.010*** |
| | | | | | (0.001) |
| Log-likelihood | -28,053 | -27,683 | -27,503 | -27,482 | -27,436 |
| Observations | 5,275 | 5,275 | 5,275 | 5,275 | 5,275 |

*Notes:* *p<0.1; **p<0.05; ***p<0.01
Conditional logit. Standard errors in parentheses.

To give a sense of the magnitude of the coefficient, I run a counterfactual exercise. I re-compute the predicted choice probabilities from Specification (5) after setting $P_{ijt}^{\text{same}}$ to zero for all dyads, holding the other covariates at their observed values. In each choice occasion, a printer faces on average 250 candidate authors, so a uniform guess would assign each candidate a probability of roughly $1/250 \approx 0.004$. Under the baseline model, the candidate ranked first by the printer receives on average a probability 7.1 times the uniform benchmark (0.028). In the counterfactual where $P^{\text{same}}$ is set to zero, this ratio falls to 4.5 (0.018), a 37% relative decline. The network channel therefore accounts for roughly a third of the model's ability to concentrate probability on the printer's most likely candidate, above and beyond what is captured by the geographic distance between the printer's place of activity and the author's birthplace, the size of the author's academic peer group, whether the author's field matches the printer's specialization, and the scholar's years of academic experience.

**Distance thresholds.** As currently designed, each choice set averages around 250 active candidates (see Table 5.A.3), that is, around 250 academics who eventually publish in the Low Countries, but it is agnostic about how far each candidate is from the printer's place of activity. This implies that, for a printer based in, say, Louvain in the 1610s, far-away scholars—e.g., a

professor at Leiden, Cambridge, or Halle—count as options just as plausible as scholars affiliated with a closer, domestic institution such as the University of Louvain. If most of these distant candidates were never realistic choices in the first place, their inclusion could inflate the estimated $P_{ijt}^{\text{same}}$ coefficient by padding the choice set with alternatives that were never truly in play. That said, out of the 1,804 academic authors in the sample, around half are affiliated with a 'foreign' institution, that is, one outside the Low Countries, which underscores the transregional nature of the region's book market and suggests that the relevant choice was not always a local one.

In any case, to check this, I re-estimate the baseline on choice sets that keep only candidates within a fixed distance of the place of activity of the printer—specifically, 50, 100, 200, and 300 km—and I drop the few decisions whose chosen author is itself farther than the cutoff. Table 5.2 shows the results. The average choice set falls from about 250 candidates to 39 at the tightest cutoff, and only 44% to 63% of decisions survive. Even so, $P_{ijt}^{\text{same}}$ stays large, positive, and highly significant in every column, between 0.77 and 1.15. If anything it gets bigger under moderate restriction (1.15 and 1.09 at the 300 and 200 km cutoffs, against 1.01 in the full sample) which is the opposite of what we would see if distant, implausible candidates were inflating it. At the two tightest cutoffs the coefficient is a bit smaller (0.92 at 100 km, 0.77 at 50 km) but still highly significant; these samples also keep fewer than half of all decisions, so part of the drop simply reflects the smaller, more local set of cases rather than a weaker effect. The reliance on the network is therefore not driven by implausibly far candidates populating the choice sets.

**Nested logit specifications.** The conditional logit implies the Independence of Irrelevant Alternatives (IIA) assumption: the relative probability of selecting one candidate over another depends on the attributes of those two candidates, not on who else is available. Adding or removing a candidate from the choice set rescales every other candidate's selection probability by the same proportion. This is implausible when some candidates are closer substitutes than others. In my setting, there are several dimensions along which this might directly matter, including disciplinary specialization and geographical proximity. On the first, printers often developed an expertise in particular areas of scholarship, making candidates within the same field plausibly more interchangeable from the printer's perspective than candidates working in unrelated disciplines. On the second,

**Table 5.2.** Subsamples with distance-restricted choice sets.

| | Dependent variable: $P(y_{ijt} = 1)$ | | | | |
|---|---|---|---|---|---|
| | Baseline | $\leq$ 300 km | $\leq$ 200 km | $\leq$ 100 km | $\leq$ 50 km |
| | (1) | (2) | (3) | (4) | (5) |
| $P_{ijt}^{\text{same}}$ | 1.008*** | 1.148*** | 1.093*** | 0.919*** | 0.769*** |
| | (0.028) | (0.036) | (0.037) | (0.042) | (0.046) |
| Controls | ✓ | ✓ | ✓ | ✓ | ✓ |
| Log-likelihood | -27,436 | -14,678 | -12,989 | -9,328 | -7,679 |
| Observations | 5,275 | 3,321 | 3,064 | 2,588 | 2,326 |
| % of full sample | 100 | 63 | 58 | 49 | 44 |
| Avg. # candidates/set | 250 | 122 | 101 | 53 | 39 |

*Notes*: *p<0.1; **p<0.05; ***p<0.01. Baseline conditional logit re-estimated on choice sets restricted to candidates within the stated distance of the printer; events whose chosen author lies beyond the cutoff are dropped. Standard errors in parentheses.

because all printers in the sample operated in the Low Countries, scholars located within the region may have constituted a distinct pool of candidates relative to scholars based elsewhere.

To check whether these dimensions structure substitution patterns, I compare the conditional logit with nested logit specifications that group candidates by field and by geographic proximity. Table 5.3 presents the results. In each nested logit, the inclusive value or dissimilarity parameter $\lambda$ for a given nest measures the degree of correlation in unobserved utility among candidates within that nest. When $\lambda = 1$, the nested logit collapses to the conditional logit; values below one indicate stronger within-nest substitution, which is the pattern consistent with random utility maximization (McFadden, 1978; Börsch-Supan, 1990). Values of $\lambda$ above one, instead, imply greater substitution across nests than within them, which is inconsistent with global utility maximization, though local consistency may still hold for the observed range of the data (Kling and Herriges, 1995; Herriges and Kling, 1996; Zanardello, 2023).

Specification (2) in Table 5.3 groups candidates into four nests by field of activity (theology, law, humanities, and sciences—drawn from the academic affiliation attributes present in the RETE database). The coefficient on $P^{\text{same}}$ rises from 1.008 in the baseline conditional logit to 1.138 in this nested logit specification. The estimated $\lambda$ parameters are 1.138 (s.e. 0.052) for theology,

1.129 (0.067) for law, 1.091 (0.052) for sciences, and 1.169 (0.055) for humanities. Each lies above one, and a Wald test rejects $\lambda = 1$ for theology and humanities at the 5% level. Although the point estimates fall outside the globally consistent range, formal tests of local consistency following Herriges and Kling (1996), reported in Table 5.A.6 in the Appendix 5.A, show that neither the first- nor the second-order condition is rejected at the 5% level for any nest. The model is therefore not badly behaved at the observed values of the data, but the $\lambda > 1$ estimates indicate that the available four-field classification does not capture a substitution pattern in which candidates within a field are closer substitutes than candidates across fields. This may reflect the coarseness of the classification rather than the absence of any disciplinary structure in substitution, and warrants a more fine-grained investigation in future iterations of the analysis, possibly relying on the genre and topic attributes of the books available from the STCV/STCN catalogs themselves.

Specification (3) instead groups candidates by geographic proximity: scholars affiliated with an institution in the Low Countries form one nest, and scholars based outside the region form another. Here, both $\lambda$'s fall below one—0.792 (0.057) for the Low Countries nest and 0.752 (0.061) for the outside nest—and both are significantly below one at the 1% level, consistent with random utility maximization. The nested logit significantly improves on the conditional logit (likelihood ratio statistic 28, $p < 0.001$ on 2 degrees of freedom). The coefficient on $P^{\text{same}}$ attenuates from 1.008 to 0.803 but remains large and highly significant, and the other coefficients shift in the same direction without changing sign or significance.

The contrast between the two nesting structures suggests that printers treated local and foreign scholars as distinct pools but did not substitute along disciplinary lines in a way the current four-field classification captures. Still, the conditional logit provides a tractable baseline whose qualitative conclusions are stable across both: the key coefficient retains its signs, significance, and approximate magnitude. The remainder of the analysis thus proceeds with the conditional logit.

**Tracking institutional disruption.** I now turn to the cutoff-scan exercise described in Section 5.5.1. A central prediction I want to test is whether printers were more likely to rely on their networks during periods of institutional disruption. For each candidate cutoff $t^*$ between 1540 and 1650,

**Table 5.3.** Conditional logit vs. nested logit specifications with alternative nest groupings.

| | Dependent variable: $P(y_{ijt} = 1)$ | | |
|---|---|---|---|
| | Conditional logit | Nested logit | |
| | | Field nests | Geographic nests |
| | (1) | (2) | (3) |
| $P_{ijt}^{\text{same}}$ | 1.008*** | 1.138*** | 0.803*** |
| | (0.028) | (0.058) | (0.056) |
| $\lambda_{\text{theology}}$ | | 1.138*** | |
| | | (0.052) | |
| $\lambda_{\text{law}}$ | | 1.129*** | |
| | | (0.067) | |
| $\lambda_{\text{sciences}}$ | | 1.091*** | |
| | | (0.052) | |
| $\lambda_{\text{humanities}}$ | | 1.169*** | |
| | | (0.055) | |
| $\lambda_{\text{Low Countries}}$ | | | 0.792*** |
| | | | (0.057) |
| $\lambda_{\text{Outside Low Countries}}$ | | | 0.752*** |
| | | | (0.061) |
| Log-likelihood | -27,436 | -27,402 | -27,422 |
| Observations | 5,275 | 5,275 | 5,275 |
| Controls | ✓ | ✓ | ✓ |

*Notes:* $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$
Standard errors in parentheses.

the utility that printer $i$ derives from publishing candidate $j$ in year $t$ is:

$$U_{ijt} = \beta_1 P_{ijt}^{\text{same}} + \beta_2 \left( P_{ijt}^{\text{same}} \times \mathbb{I}(\text{year} \geq t^*) \right) + \gamma' \mathbf{z}_{ijt} + \varepsilon_{ijt} \tag{5.3}$$

where $\mathbf{z}_{ijt}$ collects the baseline controls. The coefficient of $P_{ijt}^{\text{same}}$ is $\beta_1$ before $t^*$ and $\beta_1 + \beta_2$ after $t^*$, so parameter constancy holds when $\beta_2 = 0$.

Figure 5.2 plots $\hat{\beta}_2$ across candidate cutoffs. The estimated coefficient is close to zero for cutoffs in the 1540s and early 1550s, rises through the 1560s and 1570s, reaches its maximum around the Fall of Antwerp, and declines gradually thereafter. This pattern indicates that the differential weight printers placed on their prior same-field ties strengthened during the decades surrounding the onset of the Dutch Revolt and weakened as stability returned. $\hat{\beta}_2$ is largest at cutoffs near 1583–1585, the years immediately surrounding the Fall of Antwerp, which motivates the choice of 1585 as the cutoff year in the specifications that follow.

However, searching over many cutoffs increases the chance of finding a significant $\hat{\beta}_2$ even when $\beta_2$ is in fact zero throughout. Thus, I run a formal parameter constancy test, specifically the sup-Wald test of Andrews (1993).[13] The null hypothesis is that $\beta_2 = 0$ for every cutoff in the search range—that is, that no candidate break date divides the sample into periods with different reliance on networks. To implement the test, I compute the Wald statistic at each candidate cutoff $t^*$ as the squared ratio of $\hat{\beta}_2$ to its standard error, and take the supremum across all cutoffs in a feasible range. The feasible range is determined by a trimming proportion $\pi_0 = 0.05$, which requires that at least 5% of observations fall on each side of any candidate cutoff, yielding a search window of 1569–1650 in my case. The sup-Wald statistic is 27.80, attained at $t^* = 1583$, and as it exceeds the 5% critical value of 9.84 (Andrews, 1993, Table I, $p = 1$, $\pi_0 = 0.05$), I reject the null of parameter constancy.

Since the coefficient on $P_{ijt}^{\text{same}}$ appears to shift around the time of the Dutch Revolt, I re-estimate the baseline specification separately before and after 1585,

[13] The sup-Wald test of Andrews (1993) is a standard structural break test with unknown break date, originating in the time-series literature but applicable to any model estimated by maximum likelihood, including conditional logit. Andrews (1993) considers three asymptotically equivalent variants—Wald, likelihood ratio, and Lagrange multiplier—of which the Wald is the most convenient here, as it requires only the coefficient and standard error of the interaction term at each cutoff, both of which are already available from the cutoff scan.

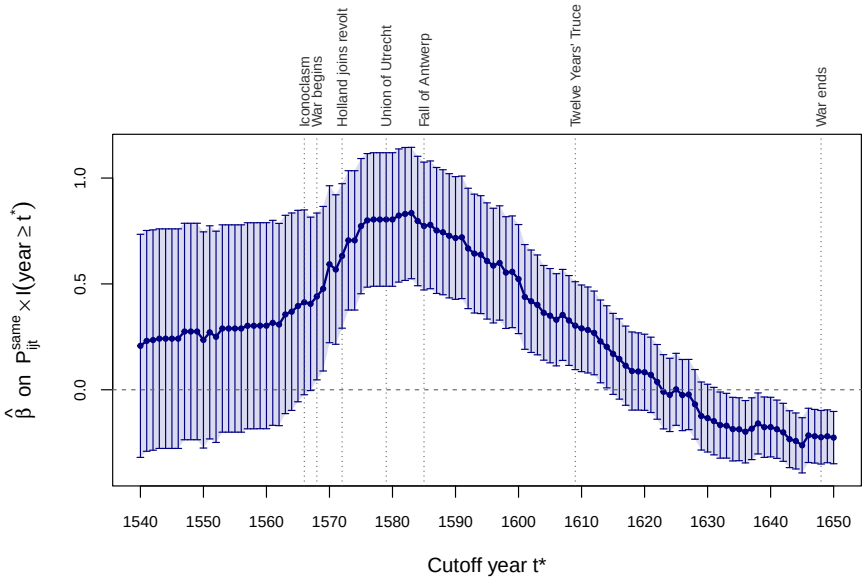



**Figure 5.2.** Estimated coefficient on $P_{ijt}^{\text{same}} \times \mathbb{I}(\text{year} \geq t^*)$ across potential cutoff years $t^* \in [1540, 1650]$. Each point is a separate conditional logit on the pooled sample of choice occasions, with the full set of baseline controls. Vertical bars show 95% confidence intervals; dotted lines mark key events around and during the Dutch Revolt. The coefficient captures the differential weight placed by printer $i$ on mutual ties connecting him to candidate author $j$ at time $t^*$ relative to before.

and on the North and South subsamples after 1585.[14] Table 5.4 reports the estimates. After 1585, the estimated coefficient of $P^{\text{same}}$ is 1.653 in the South and 1.001 in the North: the association between a printer's prior ties to a scholarly community and the selection of new authors from that community is quite pronounced in the post-1585 period and present on both sides of the new political border, with a larger magnitude in the South. Before 1585, the coefficient on $P^{\text{same}}$ is 0.021 and is not distinguishable from zero. However, this pre-1585 subsample is small, with only 359 dyads against 942 in the post-1585 South and 3,974 in the post-1585 North subsamples, so the null result should be read with some caution: it is consistent with the absence of a network channel before the *de facto* split, but the limited number of observations leaves it relatively underpowered.

**Table 5.4.** Subsample analysis.

| | Dependent variable: $P(y_{ijt} = 1)$ | | |
|---|---|---|---|
| | Pre-1585 | South Post-1585 | North Post-1585 |
| | (1) | (2) | (3) |
| $P_{ijt}^{\text{same}}$ | 0.021 | 1.653*** | 1.001*** |
| | (0.164) | (0.107) | (0.029) |
| Controls | ✓ | ✓ | ✓ |
| Log-likelihood | -1,481 | -4,885 | -20,908 |
| Observations | 359 | 942 | 3,974 |

*Notes:* *p<0.1; **p<0.05; ***p<0.01
Conditional logit. Standard errors in parentheses.

**Mechanism decomposition.** Next, I implement the decomposition of the network channel introduced in Section 5.5.1, distinguishing $P^{\text{same}}$ from $P^{\text{diff}}$ and $P^{\text{pred}}$ in Table 5.5. Specifications (1) and (2), respectively, include $P^{\text{diff}}$ and $P^{\text{pred}}$ alongside $P^{\text{same}}$, while Specification (3) includes all three variables simultaneously. The estimated coefficients suggest that for a candidate author, having co-present colleagues in unrelated fields who published with the printer is not, on its own, associated with selection. The opposite is true for co-present

[14] I use 1585 because the Fall of Antwerp that same year marked the *de facto* secession of the Northern provinces. The regional partition corresponds to the post-revolt configuration of provinces lying in the Dutch Republic (North) and the Spanish Netherlands (South).

colleagues in the candidate's own field and for same-field predecessors at the candidate's institution, whose tenure ended before the candidate arrived: both enter with positive and significant coefficients, with the coefficient on $P^{\text{pred}}$ stable across specifications (2)–(5), and concentrated in the post-1585 period in the interacted specification. The informational content of the network channel therefore concentrates in ties that share the candidate's field, whether those ties are personal contemporaries or institutional predecessors. This separation between predecessors and contemporaries lets me say something about the channels discussed in Section 5.2: personal information, group reputation, and the printer's own experience. Because predecessors never overlapped with the candidate, $P^{\text{pred}}$ cannot carry private information about him; its positive coefficient therefore shows that an effect survives once the personal-information channel is excluded by construction. The personal channel is thus not what drives the result on its own: group reputation, the printer's own experience with the community, or both, also move selection. However, I cannot push the decomposition further. Group reputation and accumulated experience operate through predecessors and contemporaries alike, so neither $P^{\text{pred}}$ nor $P^{\text{same}}$ (nor their residual) can tell them apart. The decomposition thus rules out the personal channel as the sole driver, but leaves group reputation and experience indistinguishable.

Specifications (4) through (6) interact each of the three network variables with the Post-1585 indicator. Once the interaction on $P^{\text{same}}$ is included in Specification (4), the main coefficient falls from 0.881 to 0.110 and the interaction term absorbs the effect at 0.796. The same-field network channel is therefore concentrated in the post-1585 period, consistent with the subsample analysis in Table 5.4. Specification (6) adds the interaction on $P^{\text{pred}}$, which is positive, significant, and equal to 0.943. The post-1585 strengthening therefore appears in both the personal and the institutional components of the same-field network channel. The interaction on $P^{\text{diff}}$, by contrast, is large and negative ($-1.71$) but not statistically significant, in line with the absence of a stable cross-field effect in the pooled specifications.

### 5.6.1 Robustness checks

**Heterogeneity analysis by discipline.** So far, the analysis pools across fields. I now examine whether the post-1585 intensification is uniform across

**Table 5.5.** Mechanisms: same field, different field, predecessors.

| | Dependent variable: $P(y_{ijt} = 1)$ | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $P_{ijt}^{\text{same}}$ | 1.010*** | 0.879*** | 0.881*** | 0.110 | 0.112 | 0.284* |
| | (0.028) | (0.028) | (0.028) | (0.151) | (0.152) | (0.157) |
| $P_{ijt}^{\text{diff}}$ | −0.650* | | −0.756** | −0.772** | 1.119 | 0.840 |
| | (0.356) | | (0.353) | (0.354) | (1.149) | (1.150) |
| $P_{ijt}^{\text{pred}}$ | | 0.707*** | 0.709*** | 0.711*** | 0.711*** | −0.192 |
| | | (0.039) | (0.039) | (0.039) | (0.039) | (0.210) |
| $P_{ijt}^{\text{same}} \times \text{Post}_{1585}$ | | | | 0.796*** | 0.794*** | 0.612*** |
| | | | | (0.153) | (0.153) | (0.159) |
| $P_{ijt}^{\text{diff}} \times \text{Post}_{1585}$ | | | | | −1.987 | −1.714 |
| | | | | | (1.209) | (1.210) |
| $P_{ijt}^{\text{pred}} \times \text{Post}_{1585}$ | | | | | | 0.943*** |
| | | | | | | (0.213) |
| Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Log-likelihood | -27,434 | -27,284 | -27,281 | -27,266 | -27,265 | -27,253 |
| Observations | 5,275 | 5,275 | 5,275 | 5,275 | 5,275 | 5,275 |

*Notes:* *p<0.1; **p<0.05; ***p<0.01
Conditional logit. Standard errors in parentheses.

academic disciplines: theology, humanities, and sciences.[15] Table 5.6 reports the interacted specification separately by field. The pre-1585 coefficient on $P^{\text{same}}$ is positive and significant only in theology, suggesting that there, network-based selection was already operating before the Revolt. The Post-1585 interaction is positive and significant across all fields but most pronounced in theology (0.940), followed by sciences (0.727) and humanities (0.646). Theology, after all, was the field most exposed to doctrinal scrutiny, at least leading up to the 1570s, as suggested by the descriptive evidence presented in Figure 5.3: theological works made up the dominant share of titles prohibited in the sixteenth-century Habsburg Netherlands indexes, suggesting that the heightened censorship risk borne by theological authorship underlies the stronger network effect estimated for that field.[16]

[15] I exclude the field of law because the sample is too small to deliver statistical power.

[16] Lists of prohibited books were promulgated through edicts, placards, and indexes issued at discrete dates rather than continuously, so prohibition is observed only at these fixed editions and cannot be measured as an annual flow. The indexes used here also do not exhaust censorship in the Low Countries: after the 1570s, prohibition in the Southern provinces was increasingly governed by the centralized Roman *Index librorum prohibitorum*—beyond the scope of this figure. That later material is catalogued in volume XI of de Bujanda's series




**(a)** Prohibitions in force




**(b)** New prohibitions

**Figure 5.3.** Subject composition of prohibited works across Habsburg Netherlands placards and indexes, 1540–1570. Panels show the categorical composition of individually-listed titles (excluding general author bans) in the Louvain (1546, 1550, 1558) and Antwerp (1569, 1570) indexes, as edited in De Bujanda (1986) and De Bujanda (1988), respectively. The former volume additionally documents works condemned in the imperial placard of 1540 issued by Charles V (included here), while the latter also contains a 1571 *Index Expurgatorius*, which lists specific passages to be corrected rather than entire works to be suppressed (and is therefore excluded from this figure). The assignment of titles to subject categories is my own elaboration. Panel (a) reports the stock of prohibitions in force in each edition; panel (b) reports only the newly listed titles relative to the previous edition within the same index.

**Table 5.6.** Heterogeneity by field.

| | Dependent variable: $P(y_{ijt} = 1)$ | | |
|---|---|---|---|
| | Theology | Humanities | Sciences |
| | (1) | (2) | (3) |
| $P_{ijt}^{\text{same}}$ | 0.413** | 0.270 | 0.083 |
| | (0.205) | (0.181) | (0.431) |
| $P_{ijt}^{\text{same}} \times \text{Post}_{1585}$ | 0.940*** | 0.646*** | 0.727* |
| | (0.210) | (0.184) | (0.433) |
| Controls | ✓ | ✓ | ✓ |
| Log-likelihood | -9,176 | -14,681 | -7,643 |
| Observations | 1,819 | 2,827 | 1,447 |

*Notes*: *p<0.1; **p<0.05; ***p<0.01. Conditional logit. Standard errors in parentheses.

**Disentangling behavioral versus compositional change.** The post-1585 intensification could mechanically reflect composition rather than disruption, with the coefficient on $P_{ijt}^{\text{same}} \times \text{Post}_{1585}$ potentially comparing choice occasions made by largely different printers. In that perspective, a higher coefficient might capture a change in *who* publishes rather than a change in how a given printer screens candidates. Ideally, restricting the sample to printers active in both periods would separate the two channels, but only 14 printers straddle the cutoff (179 dyads, 3.4% of the sample), too few to estimate the interaction reliably. This scarcity in fact is not incidental: many presses closed or relocated as a direct consequence of the disruption.

I thus proceed as follows. I classify a printer as an *incumbent* if their first publication in the sample predates 1585, and as an entrant otherwise. I then add the triple interaction $P_{ijt}^{\text{same}} \times \text{Post}_{1585} \times \text{Incumbent}_i$ to a specification that already includes $P_{ijt}^{\text{same}}$, its interaction with $\text{Post}_{1585}$, and the baseline controls, so that the post-1585 change in network reliance is free to differ between incumbents and entrants. If composition alone drove the result, the post-1585 strengthening should appear only among entrants, not among incumbents. The triple interaction captures precisely this difference: a large negative triple would mean incumbents barely changed, consistent with the composition story, while a triple near zero would mean incumbents and entrants changed alike. This

(De Bujanda, 2002), but its scale places a comprehensive coding beyond the scope of this paper.

comparison rests on relatively few observations: 101 post-1585 dyads from 14 incumbent printers—so it should be read with that limitation in mind.

Table 5.7 reports the estimates. The triple interaction is small and statistically insignificant (−0.094), so incumbents and entrants cannot be distinguished. The linear combination of this term and the post-1585 coefficient gives the post-1585 effect for incumbents alone—the "incumbent total" reported at the foot of the table—which is 0.68 and significant under standard errors clustered by printer. Incumbents therefore raised their reliance on shared ties after 1585 by an amount comparable to entrants.[17] This is inconsistent with the composition story alone, which requires the incumbent effect to be absent, although the small number of incumbent printers (14) warrants caution.

**Table 5.7.** Composition vs. behavior: incumbent printers.

| | Dependent variable: $P(y_{ijt} = 1)$ |
|---|---|
| $P^{\text{same}}_{ijt}$ | 0.259* |
| | (0.149) |
| $P^{\text{same}}_{ijt} \times \text{Post}_1585$ | 0.776*** |
| | (0.154) |
| $P^{\text{same}}_{ijt} \times \text{Post}_{1585} \times \text{Incumbent}_i$ | −0.094 |
| | (0.201) |
| Controls | ✓ |
| Log-likelihood | -27,422 |
| Incumbent total | 0.682*** (0.248) |
| Observations | 5,275 |

*Notes*: *p<0.1; **p<0.05; ***p<0.01. Conditional logit; SEs clustered by printer in parentheses. Incumbent: first academic publication before 1585. Incumbent total: post-1585 network effect for incumbent printers, with clustered SEs.

## 5.7 Discussion and conclusions

This paper asks whether printers in the early modern Low Countries leaned on their networks to screen unfamiliar authors, and whether they did so more when

[17] This holds when each of the five largest incumbent presses is excluded in turn (the incumbent total ranges from 0.63 to 0.93— leave-one-out exercise not reported).

the institutional environment became unstable. I find that a printer was more likely to select a candidate whose colleagues had published with him before, and this reliance on shared ties was not constant over time: it was close to zero before the 1560s, grew through the onset of the Dutch Revolt, peaked around the Fall of Antwerp and secession of the Northern provinces, and faded as the conflict settled. The same pattern holds on both sides of the new border, being stronger in the South. To the best of my knowledge, this is the first systematic evidence that network-based screening in a knowledge market responds to the institutional environment: the same ties that carry little weight under stable conditions become central to selection when uncertainty rises.

The effect runs through ties in the candidate's own field, whether colleagues he overlapped with or predecessors who left before he arrived, and not through ties in other fields. It is also strongest in theology, the field most exposed to doctrinal scrutiny at least before the seventeenth century. Both patterns support the reading that printers relied on shared affiliation where the cost of publishing the wrong author was highest.

Three features of my findings push the interpretation beyond a simple correlation. First, the reliance on ties follows the timing of the disruption rather than a smooth long-run trend; second, it holds when I keep the printer fixed, so the same printer screens differently before and after the shock; and third, it is confined to ties of a precise and informative kind—those within the candidate's own field. At the same time, I acknowledge that network ties are not randomly assigned, so unobserved traits of a candidate linked to both shared affiliation and selection could still explain part of the result. Moreover, the within-printer evidence, which does the most to move the results beyond correlation, rests on few incumbents. A deeper limitation concerns how I capture the uncertainty triggered by the institutional disruption. The post-1585 split bundles the Dutch Revolt, the institutional reorganization that followed, the migration of printers and scholars across the new border, the intensification of censorship, as well as ordinary time trends. I do not measure uncertainty directly, and so cannot say which of these components drives the network response. While the cutoff scan shows that reliance on shared ties closely tracks the timing of the disruption, and the effect is strongest in theology, the field most exposed to doctrinal scrutiny, admittedly a continuous measure of uncertainty would let the data speak to what specifically about the disruption mattered. The counts of prohibited titles and the timing of the edicts banning them are one direction I am pursuing, already started with the digitization of the Louvain

and Antwerp indexes (De Bujanda, 1986, 1988) presented in Figure 5.3.

Another set of caveats concerns the data used. My sample is thinner before 1585 and in the North. This partly reflects the real evolution of the market, but it also follows from the matching procedure that links RETE scholars to authors in the STCN and STCV, which relies on shared Wikidata and VIAF identifiers; I recover only the authors for whom this link exists, and expect more matches as the procedure is refined. The catalogs themselves cover *surviving* imprints rather than full output. The current four-field classification is coarse, and could be sharpened using the genre and topic metadata in the catalogs.

A further extension which would complement the present analysis is to exploit the geography of the Spanish reconquest more directly. Drawing on the shift in jurisdiction over time, I could classify cities into those retaken by Spain, those that remained Dutch, and those on the moving frontier, and compare how network reliance evolved across these groups. Such a design would shift the unit of analysis from printer–author dyads to the city-year level and add a spatial dimension: in a difference-in-differences logic, cities that never changed regime would serve as a control for the trends affecting the whole market, so that what remains is more plausibly the effect of the reconquest itself rather than of everything else moving around a single year. Its feasibility, however, depends on first reconstructing detailed time-varying jurisdiction at the city level[18] and hinges on whether enough printing centres actually changed regime to identify such an effect.

[18]This is not trivial: the Stapel (2025) polygons record minting authorities over time, but not their jurisdiction assignment, which I impose myself following the consolidated post-split configuration.

# Appendix

## 5.A Supplementary tables and figures

**Table 5.A.1.** Sample restrictions applied to the STCN and STCV catalog

| Restriction | STCN | STCV |
|---|---|---|
| Starting sample | 876,613 | 29,530 |
| RETE scholar as author | 33,637 | 4,364 |
| Published in the LC by non-fictitious printers | 18,612 | 4,011 |
| Published during author's academic career | 7,123 | 1,895 |

*Notes:* The unit of observation is the bibliographic record, corresponding in both catalogs to a distinct edition of a printed work. The starting sample sizes consider exports of the catalogs downloaded in August 2025. Ephemeral print (such as broadsheets, periodicals, state publications, and the like) is excluded prior to any further restriction from the second row onwards. The final row reports the approximate base sample; the effective sample of publication events used in the analysis (5,275 observations) is however smaller than the sum of the two reported totals, because (i) there is a small overlap across catalogs, and (ii) only first-time collaborations between a unique printer–author pair are retained.

**Table 5.A.2.** Distribution of academic titles per printer

| Titles | Printers | % |
|---|---|---|
| 1 | 545 | 39.0 |
| 2 | 212 | 15.2 |
| 3 | 137 | 9.8 |
| 4 | 74 | 5.3 |
| 5–9 | 191 | 13.7 |
| 10–19 | 113 | 8.1 |
| 20+ | 126 | 9.0 |
| Total | 1,398 | |
| Mean | 8.0 | |
| Median | 2 | |

**Table 5.A.3.** Top 15 institutions by number of published members in the Low Countries as recorded in the STCN/STCV, 1476–1800.

| Institution | Founded | Closed | N. members | Published in LC |
|---|---|---|---|---|
| Leiden University | 1575 | — | 268 | 200 |
| University of Louvain | 1425 | 1797 | 1051 | 157 |
| Royal Society of London | 1660 | — | 1654 | 149 |
| Royal Zeeland Scientific Society | 1769 | — | 312 | 128 |
| University of Franeker | 1585 | 1811 | 143 | 110 |
| Holland Society of Sciences | 1752 | 1804 | 223 | 110 |
| Utrecht University | 1636 | — | 117 | 98 |
| French Academy | 1635 | 1793 | 254 | 85 |
| University of Harderwijk | 1647 | 1811 | 103 | 78 |
| University of Groningen | 1612 | — | 96 | 71 |
| Academy of Sciences Leopoldina | 1652 | — | 922 | 67 |
| Athenaeum Illustre of Amsterdam | 1632 | 1877 | 72 | 55 |
| University of Cambridge | 1209 | — | 7508 | 52 |
| Batavian Society for Experimental Philosophy | 1769 | — | 82 | 50 |
| University of Oxford | 1200 | — | 1579 | 43 |

*Notes:* A dash in the "Closed" column indicates the institution is still active. "N. members" reports the number of affiliated scholars active between 1476 and 1800. "Published in LC" counts scholars with at least one publication recorded in the STCN/STCV in cities that were part of the Low Countries.

**Table 5.A.4.** Summary statistics by decade.

| | | | | Choice-set size | |
|---|---|---|---|---|---|
| Decade | Printers | Authors | Dyads | Mean | Median |
| 1470s | 4 | 14 | 5 | 12.6 | 13 |
| 1480s | 6 | 14 | 11 | 12.7 | 13 |
| 1490s | 1 | 11 | 1 | 11.0 | 11 |
| 1500s | 2 | 21 | 3 | 15.0 | 13 |
| 1510s | 14 | 37 | 25 | 32.2 | 32 |
| 1520s | 20 | 54 | 30 | 42.2 | 44 |
| 1530s | 20 | 81 | 48 | 58.2 | 55 |
| 1540s | 17 | 87 | 28 | 68.2 | 67 |
| 1550s | 20 | 96 | 36 | 82.7 | 83 |
| 1560s | 40 | 108 | 91 | 86.9 | 87 |
| 1570s | 23 | 134 | 43 | 104.1 | 103 |
| 1580s | 26 | 146 | 70 | 119.5 | 119 |
| 1590s | 39 | 179 | 88 | 143.3 | 143 |
| 1600s | 66 | 221 | 152 | 178.6 | 182 |
| 1610s | 110 | 259 | 252 | 204.6 | 202 |
| 1620s | 82 | 277 | 221 | 231.7 | 229 |
| 1630s | 105 | 300 | 255 | 234.2 | 232 |
| 1640s | 111 | 311 | 294 | 248.1 | 248 |
| 1650s | 106 | 318 | 254 | 251.1 | 251 |
| 1660s | 105 | 320 | 247 | 256.1 | 256 |
| 1670s | 98 | 318 | 206 | 250.3 | 250 |
| 1680s | 121 | 302 | 248 | 245.0 | 245 |
| 1690s | 91 | 293 | 188 | 234.6 | 232 |
| 1700s | 87 | 297 | 163 | 247.8 | 249 |
| 1710s | 73 | 289 | 182 | 227.7 | 230 |
| 1720s | 66 | 259 | 152 | 203.4 | 199 |
| 1730s | 64 | 237 | 146 | 192.2 | 193 |
| 1740s | 94 | 279 | 239 | 207.4 | 203 |
| 1750s | 77 | 327 | 185 | 252.2 | 248 |
| 1760s | 104 | 403 | 301 | 304.3 | 303 |
| 1770s | 121 | 469 | 365 | 373.4 | 373 |
| 1780s | 150 | 462 | 409 | 375.2 | 373 |
| 1790s | 108 | 386 | 301 | 344.0 | 359 |
| Total | 1,398 | 1,804 | 5,275 | 250.4 | 245 |

*Notes:* "Printers" is the number of distinct printers selecting at least one new author in the decade. "Authors" is the number of distinct authors appearing in the choice sets. "Dyads" is the number of unique printer–author pairs.

**Table 5.A.5.** Descriptive statistics: main variables.

| | Choice set | | | Selected | | |
|---|---|---|---|---|---|---|
| | Mean | SD | Median | Mean | SD | Median |
| $P_{ijt}^{\text{same}}$ | 0.390 | 1.671 | 0.000 | 0.858 | 2.261 | 0.000 |
| $P_{ijt}^{\text{diff}}$ | 0.002 | 0.074 | 0.000 | 0.002 | 0.041 | 0.000 |
| $P_{ijt}^{\text{pred}}$ | 0.068 | 0.425 | 0.000 | 0.285 | 0.989 | 0.000 |
| Birthplace dist.$_{ij}$ | 659.373 | 1,136.459 | 387.944 | 476.143 | 928.691 | 226.530 |
| Total n. colleagues$_{jkt}$ | 67.299 | 115.048 | 18.000 | 47.376 | 95.696 | 10.000 |
| Field match$_{ij}$ | 0.296 | 0.457 | 0.000 | 0.323 | 0.468 | 0.000 |
| Experience (yrs)$_{jt}$ | 18.343 | 13.736 | 16.000 | 16.101 | 15.112 | 14.000 |
| Observations | | 1,320,772 | | | 5,275 | |

*Notes:* Each observation is a printer–candidate dyad within a choice occasion. "Choice set" includes every candidate author available for selection; "Selected" is the author the printer publishes. $P_{ijt}^{\text{same}}$, $P_{ijt}^{\text{diff}}$, $P_{ijt}^{\text{pred}}$, birthplace distance, and total n. colleagues enter the regression under the inverse hyperbolic sine (ihs) transformation.

**Table 5.A.6.** Consistency tests of nested logit model with RUM.

| | | First-order condition | | Second-order condition | |
|---|---|---|---|---|---|
| Nest | $\hat{\lambda}_s$ | $\hat{U}_{1s}(\bar{v})$ | $t$ ratio | $\hat{U}_{2s}(\bar{v})$ | $t$ ratio |
| Theology | 1.138 | 1.526 | -7.47 | 1.156 | -0.34 |
| Law | 1.129 | 1.132 | -0.05 | 1.021 | 1.61 |
| Sciences | 1.091 | 1.334 | -4.68 | 1.086 | 0.11 |
| Humanities | 1.169 | 1.404 | -4.27 | 1.110 | 1.07 |

# 6

# General conclusion

This dissertation set out to bridge two research traditions that have long drawn on the same empirical materials (bibliographic metadata, prosopographical records, geospatial data) while developing largely in parallel: the network-visualization tradition of digital humanities and computational history on the one hand, and the structural, model- and inference-based tradition of quantitative economic history on the other. The Old University of Louvain (1425–1797) served both as the anchoring case for a dissertation aiming to integrate the two traditions, and as the point of departure for a wider regional and European inquiry.

The first two chapters laid the methodological groundwork. Chapter 1 converted the metadata of the *Collectio academica antiqua* into machine-readable, relational data and modeled it as a time-evolving multilayer network, showing among other things the structurally central position of censors in a publication-based network, the comparatively modest bias of a curated collection, and the long-run intellectual shifts materialized in the changing centrality of individual scholars.

Chapter 2 generalized this experience into a cross-disciplinary methodological reflection, arguing that a multilayer representation preserves the role-differentiated forms of participation encoded in catalog records rather than flattening them into undifferentiated co-occurrence, while also establishing the limits of bibliographic metadata used in isolation: their analytical potential is realized only when relational representations are linked to complementary institutional, social, and biographical sources.

The remaining chapters put these principles to work at contrasting scales. Chapter 3 moved to the continental level, reconstructing the European academic affiliation network and proposing an epidemiological model of idea diffusion through institutional co-presence. In this chapter we adopt a full-fledged structural and quantitative approach to a highly complex phenomenon, and we acknowledge that this may at times be unsettling to humanist scholars. With many caveats in mind, it can be read as something close to a thought experiment brought to life, addressing—at least collaterally, through the lens of the distinctively European network of academic institutions—the long-standing question of the rise of Europe and the Great Divergence.

Chapter 4 returned to Louvain to study in greater depth the strategies and partnerships of five notable printers operating in the city across very different periods, based on a segment of the surviving bibliographic record of their production, as summarized in the *Lov'Ego* dataset and network. The comparative analysis of compositional metrics for each ego-network revealed three contrasting configurations of printer–scholar partnership (humanist, commission-based, and ideologically charged), showing that sharing a university town did not mean sharing the same kind of network.

Chapter 5 settles on a middle-ground scale, extending from the city of Louvain to the academic book market of the historic Low Countries, but stopping short of a Europe-wide scope while taking into account the highly transregional nature of this market. It tackles a classical question: how trust is built, how it functions, and under what circumstances it comes to matter most. This is formalized under a classical economic lens, examining whether printers leaned on accumulated social ties to screen unfamiliar authors, and whether this reliance intensified during the institutional disruption that marked the region's history from the first stirrings of the Reformation onward. Preliminary findings reveal that reliance on shared ties grew strongest around the eruption of the Dutch Revolt, the Fall of Antwerp, and the ensuing split of the region.

As the title of the present dissertation suggests, the intended contribution of these chapters is to integrate and bridge two traditions to enrich each by supplying what the other lacks. This inter- and cross-disciplinary approach attempts to break silos and open a dialogue, at the cost of some discomfort for readers on both sides, who are confronted with methods, approaches, and even formulations of research questions that may sound puzzling and unfamiliar. Overall, the records left by universities, printers, and scholars are rich enough

to reward close-reading research methods *and* formal empirical modeling, and the most interesting questions about how knowledge was made, circulated, and survived sit precisely on the border where book history and economics meet.

# Bibliography

**Abramitzky, Ran and Isabelle Sin**, "Book translations as idea flows: The effects of the collapse of Communism on the diffusion of knowledge," *Journal of the European Economic Association*, 2014, *12* (6), 1453–1520.

**Abulafia, Anna Sapir**, *Christian–Jewish Relations 1000–1300: Jews in the Service of Medieval Christendom*, Harlow: Pearson, 2011.

**Adam, Renaud**, *Jean de Westphalie et Thierry Martens: la découverte de la "Logica Vetus (1474)" et les débuts de l'imprimerie dans les Pays-Bas méridionaux (avec un fac-similé)* number 8. In 'Nugae humanisticae: sub signo Erasmi Series.', Turnhout: Brepols, 2009.

_ , "The Profession of Printer in the Southern Netherlands before the Reformation," in Violet Soen, Dries Vanysacker, and Wim François, eds., *Church, Censorship and Reform in the Early Modern Habsburg Netherlands*, Turnhout: Brepols, 2017, pp. 13–25.

_ , *Vivre et imprimer dans les Pays-Bas méridionaux (des origines à la Réforme)*, Vol. 2, Turnhout: Brepols, 2018.

**Agten, Els**, "Ottavio Mirto Frangipani, First Papal Nuncio to Flanders (1596–1606) and His Thoughts on Book Censorship," in Violet Soen, Dries Vanysacker, and Wim François, eds., *Church, Censorship and Reform in the Early Modern Habsburg Netherlands*, number 101. In 'Bibliothèque de la Revue d'histoire ecclésiastique.', Turnhout: Brepols, 2017, pp. 73–96.

**Ahmadpoor, Mohammad and Benjamin F. Jones**, "The dual frontier: Patented inventions and prior scientific advance," *Science*, 2017, *357* (6351), 583–587.

**Akcigit, Ufuk, Santiago Caicedo, Ernest Miguelez, Stefanie Stantcheva, and Valerio Sterzi**, "Dancing with the Stars: Innovation

Through Interactions," Working Paper 24466, National Bureau of Economic Research 2018.

**Akerlof, George A.**, "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), 488–500.

**Almelhem, Ali, Murat Iyigun, Austin Kennedy, and Jared Rubin**, "Enlightenment ideals and belief in progress in the run-up to the industrial revolution: A textual analysis," 2023. Working paper.

**Altonji, Joseph G and Charles R Pierret**, "Employer learning and statistical discrimination," *The Quarterly Journal of Economics*, 2001, *116* (1), 313–350.

**Amburger, Erik**, *Die Mitglieder der Deutschen Akademie der Wissenschaften zu Berlin 1700–1950*, Berlin: Akademie-Verlag, 1950.

**Anderson, Robert Warren, Noel D Johnson, and Mark Koyama**, "Jewish persecutions and weather shocks: 1100–1800," *The Economic Journal*, 2017, *127* (602), 924–958.

**André, Catherine and Jean-Philippe Platteau**, "Land Relations Under Unbearable Stress: Rwanda Caught in the Malthusian Trap," *Journal of Economic Behavior and Organization*, 1998, *34* (1), 1–47.

**Andrews, Donald**, "Tests for parameter instability and structural change with unknown change point," *Econometrica*, 1993, pp. 821–856.

**Angrist, Joshua D. and Jörn-Steffen Pischke**, "The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics," *The Journal of Economic Perspectives*, 2010, *24* (2), 3–30.

**Applebaum, Wilbur**, *Encyclopedia of the Scientific Revolution: from Copernicus to Newton*, New York: Routledge, 2003.

**Arblaster, Paul**, "Policy and Publishing in the Habsburg Netherlands, 1585–1690," in Brendan Dooley and Sabrina A. Baron, eds., *The Politics of Information in Early Modern Europe*, London: Routledge, 2005, pp. 187–206.

_ , *A history of the Low Countries*, Bloomsbury Publishing, 2018.

**Arora, Abhishek, Emily Silcock, Leander Heldring, and Melissa Dell**, "Contrastive Entity Coreference and Disambiguation for Historical Texts," 2024.

**Ashraf, Quamrul and Oded Galor**, "Dynamics and Stagnation in the Malthusian Epoch," *American Economic Review*, 2011, *101* (5), 2003–41.

**Atkin, David, M Keith Chen, and Anton Popov**, "The returns to face–to–face interactions: Knowledge spillovers in Silicon Valley," Technical Report, National Bureau of Economic Research 2022.

**Avram, Henriette**, "MARC: Its History and Implications," Technical Report LC-1.2:M18/16, Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 1975.

**Azoulay, Pierre, Christian Fons-Rosen, and Joshua S Graff Zivin**, "Does science advance one funeral at a time?," *American Economic Review*, 2019, *109* (8), 2889–2920.

**_ , Joshua S Graff Zivin, and Jialan Wang**, "Superstar extinction," *The Quarterly Journal of Economics*, 2010, *125* (2), 549–589.

**Bahr, Gunnar Von**, *Medicinska fakulteten i Uppsala*, Stockholm: Almqvist & Wiksell International, 1945.

**Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson**, "The diffusion of microfinance," *Science*, 2013, *341* (6144), 1236498.

**Barrett, Matthew**, *The Reformation as Renewal: Retrieving the One, Holy, Catholic, and Apostolic Church*, Zondervan Academic, 2023.

**Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy**, "Gephi: An Open Source Software for Exploring and Manipulating Networks," in "Proceedings of the International AAAI Conference on Web and Social Media," Vol. 3 2009, pp. 361–362.

**Becker, Sascha and Luigi Pascali**, "Religion, Division of Labor, and Conflict: Anti-semitism in Germany over 600 Years," *American Economic Review*, 2019, *109* (5), 1764–1804.

_ , **Volker Lindenthal, Sharun W. Mukand, and Fabian Waldinger**, "Persecution and Escape: Professional Networks and High-Skilled Emigration from Nazi Germany," *American Economic Journal: Applied Economics*, 2024, *16* (3), 1–43.

_ , **Yuan Hsiao, Steven Pfaff, and Jared Rubin**, "Multiplex Network Ties and the Spatial Diffusion of Radical Innovations: Martin Luther's Leadership in the Early Reformation," *American Sociological Review*, 2020, *85* (5), 857–894.

_ , _ , _ , **and** _ , "Competing social influence in contested diffusion: contention and the spread of the early Reformation," *Social Forces*, 06 2025, p. soaf079.

**Bellemare, Marc F and Casey J Wichman**, "Elasticities and the inverse hyperbolic sine transformation," *Oxford Bulletin of Economics and Statistics*, 2020, *82* (1), 50–61.

**Ben-David, Joseph**, *The scientist's role in society: A comparative study*, Foundations of Modern Sociology Series, 1971.

**Bénévent, Christine, Annie Charon, Isabelle Diu, and Magali Vène, eds**, *Passeurs de textes: Imprimeurs et libraires à l'âge de l'humanisme*, Vol. 37 of *Études et rencontres de l'École des chartes*, Paris: École nationale des chartes, 2012.

**Bennett, James S., Erin Mutch, Andrew Tollefson, Ed Chalstrey, Majid Benam, Enrico Cioni, Jenny Reddish et al.**, "Cliopatria: A Geospatial Database of World-Wide Political Entities from 3400 BCE to 2024 CE," *Scientific Data*, 2025, *12* (247).

**Berger-Levrault, Oscar**, *Catalogus professorum Academiarum et Universitatum alsaticarum XVI-XVIII seculi*, Nancy: Impr. de Berger-Levrault, 1890.

**Berman, Harold J.**, *Law and Revolution: The Formation of the Western Legal Tradition*, Cambridge, MA: Harvard University Press, 1983.

**Bhattacharya, Sudipto and Paul Pfleiderer**, "Delegated portfolio management," *Journal of Economic Theory*, 1985, *36* (1), 1–25.

**Blair, Ann M.**, *Too Much to Know: Managing Scholarly Information before the Modern Age*, New Haven and London: Yale University Press, 2010.

**Blasutto, Fabio and David de la Croix**, "Catholic censorship and the demise of knowledge production in early modern Italy," *The Economic Journal*, 2023, *133*, 2899–2924.

**Blei, David M., Andrew Y. Ng, and Michael I. Jordan**, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 2003, *3*, 993–1022.

**Blonder, Benjamin, Tina W Wey, Anna Dornhaus, Richard James, and Andrew Sih**, "Temporal Dynamics and Network Analysis," *Methods in Ecology and Evolution*, 2021, *3* (6), 958–972.

**Borowiecki, Karol Jan**, "Geographic clustering and productivity: An instrumental variable approach for classical composers," *Journal of Urban Economics*, 2013, *73* (1), 94–110.

**Börsch-Supan, Axel**, "On the compatibility of nested logit models with utility maximization," *Journal of Econometrics*, 1990, *43* (3), 373–388.

**Bosch, Gerrit Vanden**, "Censorship in the Archdiocese of Malines (1559–1795)," in Violet Soen, Wim François, and Dries Vanysacker, eds., *Church, Censorship and Reform in the Early Modern Habsburg Netherlands*, Turnhout: Brepols, 2017, pp. 97–106.

**Bramoullé, Yann and Garance Genicot**, "Diffusion and targeting centrality," *Journal of Economic Theory*, 2024, *222*, 105920.

**Brinkmann, Stefanie, Giovanni Ciotti, Stefano Valente, and Eva Maria Wilden, eds**, *Education Materialised: Reconstructing Teaching and Learning Contexts through Manuscripts*, Vol. 23 of *Studies in Manuscript Cultures*, Berlin and Boston: De Gruyter, 2021.

**Bruni, Flavia and Andrew Pettegree, eds**, *Lost Books: Reconstructing the Print World of Pre-Industrial Europe*, Vol. 46 of *Library of the Written Word*, Leiden: Brill, 2016.

**Brunt, Liam and Cecilia García-Peñalosa**, "Urbanisation and the onset of modern economic growth," *The Economic Journal*, 2022, *132* (642), 512–545.

**Bujanda, Jesús Martínez De**, *Index de l'Université de Louvain 1546, 1550, 1558*, Vol. 2 of *Index des livres interdits*, Sherbrooke: Éditions de l'Université de Sherbrooke, 1986.

_ , *Index d'Anvers 1569, 1570, 1571*, Vol. 7 of *Index des livres interdits*, Sherbrooke: Éditions de l'Université de Sherbrooke, 1988.

_ , *Index de Rome 1557, 1559, 1564*, Vol. 8 of *Index des livres interdits*, Sherbrooke: Éditions de l'Université de Sherbrooke, 1990.

_ , *Index Librorum Prohibitorum 1600–1966*, Vol. 11 of *Index des livres interdits*, Genève and Sherbrooke: Droz / Éditions de l'Université de Sherbrooke, 2002.

**Buning, Marius, Marlise Rijks et al.**, "Introduction: Printing privileges in the early modern Low Countries," *Early Modern Low Countries*, 2024, *8* (2), 135–144.

**Buringh, Eltjo**, "The population of European cities from 700 to 2000: Social and economic history," *Research Data Journal for the Humanities and Social Sciences*, 2021, *6* (1), 1–18.

**Cammaerts, Dieter**, "Receiving Approbations from Jansenist Censors: A Strategy for Circumventing the Anti-Jansenist Pre-Publication Censorship of Malines Archbishop Humbertus de Precipiano (1690–1711)," *In Monte Artium. Journal of the Royal Library of Belgium (KBR)*, 2023, *15.*

_ , "Manuale Lovaniense," KU Leuven RDR, V1 2024. Dataset; single records also consultable via http://www.odis.be.

_ , "Manuale Lovaniense: Een sociaaleconomische en typografische studie van het gedrukte handboek aan de vroegmoderne Leuvense universiteit (1474–1650)." PhD dissertation, KU Leuven 2024.

_ , "A Typographical Evolution in the Louvain Lectures on Logic? How Students Used Printed Textbooks between 1474 and 1532," in Xander Feys, Maxime Maleux, Andy Peetermans, and Raf Van Rooy, eds., *Student Notes from Latin Europe (1400–1750): A Research Companion*, Leuven: Leuven University Press, 2025, pp. 47–68.

**Cantoni, Davide and Noam Yuchtman**, "Historical Natural Experiments: Bridging Economics and Economic History," in "The Handbook of Historical Economics," Elsevier, 2021, pp. 213–241.

**Catoire, Guillaume, Valentine Debois, David de la Croix, Julie Duchêne, Maximilian Ganterer, and Mara Vitale**, "Scholars and Literati at the Universitas Lovaniensis (1425–1797)," *Repertorium Eruditorum Totius Europae*, 2021, *4*, 53–66.

**Cervellati, Matteo, Sara Lazzaroni, Gianni Marciante, and Paolo Masella**, "The Rise of the Knowledge Economy: Republic of Letters and Communication Infrastructures in Early Modern England," 2025. Working paper.

**Chaney, Eric**, "Religion and the Rise and Fall of Islamic Science," 2023. Working paper.

_ , "Modern Library Holdings and Historic City Growth," 2024. Working paper.

**Chaunu, Pierre**, *Le Temps des réformes: La crise de la chrétienté, l'éclatement (1250-1550)*, Fayard, 2014.

**Chiopris, Caterina**, "The Diffusion of Ideas," 2024. Working paper.

**Coleman, James S**, "Relational analysis: The study of social organizations with survey methods," *Human organization*, 1958, *17* (4), 28–36.

_ , "Social capital in the creation of human capital," *American Journal of Sociology*, 1988, *94*, S95–S120.

**Collège de France**, "Liste des professeurs du Collège de France depuis 1530," 2018. PDF document.

**Comino, Stefano, Alberto Galasso, and Clara Graziano**, "Censorship, industry structure, and creativity: Evidence from the Catholic Inquisition in Renaissance Venice," *The Journal of Law, Economics, and Organization*, 2024, pp. 1045—-1074.

**Conrad, Ernst**, *Die Lehrstühle der Universität Tübingen und ihre Inhaber (1477-1927)* 1960.

**Constantinidou, Natasha, Dieter Cammaerts, and Violet Soen**, "Book History: The Basics with Two Case Studies," in Xander Feys, Maxime Maleux, Andy Peetermans, and Raf Van Rooy, eds., *Student Notes from Latin Europe (1400–1750): A Research Companion*, Leuven: Leuven University Press, 2025, pp. 95–124.

**Copleston, Frederick Charles**, *Medieval Philosophy: From Augustine to Duns Scotus*, Image Books, 1993.

**Coppens, Chris, Mark Derez, and Jan Roegiers**, *Leuven University Library, 1425–2000*, Leuven: Leuven University Press, 2005.

**Cross, Richard**, "An accidental Reformation?," in "The Hanover Review: The Journal of the London Lyceum: 3.1: The Reformation as Renewal Symposium," Hanover Press, 2024, pp. 5–13.

**Currarini, Sergio, Matthew O Jackson, and Paolo Pin**, "An economic model of friendship: Homophily, minorities, and segregation," *Econometrica*, 2009, *77* (4), 1003–1045.

**Curtis, Matthew and David de la Croix**, "Seeds of Knowledge: Premodern Scholarship, Academic Fields, and European Growth," 2025. SSRN Working Paper.

**David, Paul A**, "Clio and the Economics of QWERTY," *The American economic review*, 1985, *75* (2), 332–337.

**Davies, John**, "Retrospect and Prospect," in Claire Taylor and Kostas Vlassopoulos, eds., *Communities and Networks in the Ancient Greek World*, Oxford: Oxford University Press, 2015.

**de Arteaga, Enrique Esperabé et al.**, *Historia pragmática e interna de la Universidad de Salamanca*, Salamanca: Imp. y lib. de Francisco Núñez Izquierdo, 1917.

**de la Croix, David**, "Scholars and Literati in European Academia before 1800," *Repertorium Eruditorum Totius Europae*, 2021, *5*, 35–41.

_ **and Marc Goñi**, "Nepotism vs. Intergenerational Transmission of Human Capital in Academia (1088–1800)," *Journal of Economic Growth*, 2024.

_ **and Pauline Morault**, "Winners and Losers from the Protestant Reformation: An Analysis of the Network of European Universities," *Journal of Economic History*, 2026, *86* (2), 460–500.

_ **and Rossana Scebba**, "Geolocalization and the Birth-to-Death Distance," *Repertorium Eruditorum Totius Europae*, 2024, *14*, 37–42.

_ **, Frédéric Docquier, Alice Fabre, and Robert Stelter**, "The Academic Market and the Rise of Universities in Medieval and Early Modern Europe (1000-1800)," *Journal of the European Economic Association*, 2024, *22*, 1541–1589.

_ **, Matthias Doepke, and Joel Mokyr**, "Clans, guilds, and markets: Apprenticeship institutions and growth in the preindustrial economy," *The Quarterly Journal of Economics*, 2018, *133* (1), 1–70.

**de Pleijt, Alexandra and Julius Koschnick**, "Alienated Intellectuals? Exploring the Political Consequences of the Educational Revolution in Early Modern England," 2025. Working paper.

**De Ridder, Bram and Violet Soen**, "Transregional History: New Perspectives on Early Modern Borders and Borderlands in the Low Countries and the Habsburg Worlds," in "Transregional Territories," Turnhout: Brepols, 2020, pp. 13–19.

**Dekoninck, Ralph and Violet Soen, eds**, *Leçons du passé. Quatre siècles de notes d'étudiants à l'Université de Louvain (1425–1797)*, Louvain–la–Neuve: Presses universitaires de Louvain, 2025.

**Delsaerdt, Pierre**, *Suam quisque bibliothecam: boekhandel en particulier boekenbezit aan de oude Leuvense universiteit, 16de–18de eeuw*, Leuven: Universitaire Pers Leuven, 2001.

**der Weduwen, Arthur**, "Sold in a Closed Room. Auctioning Libri Prohibiti in the Dutch Republic, 1670–1720," in Arthur der Weduwen, Andrew Pettegree, and Graeme Kemp, eds., *Book Trade Catalogues in Early Modern Europe*, Vol. 93 of *Library of the Written Word*, Leiden: Brill, 2021, pp. 319–360.

_ **and Barnaby Cullen**, "King, Bishop, Professor and Postmaster: The Early Modern Swedish Publisher," in Barnaby Cullen, Ian Maclean, and

Arthur der Weduwen, eds., *Early Modern Publishers: Identities and Strategies in the Book Trade*, number 138. In 'Library of the Written Word.', Leiden: Brill, 2025, pp. 9–28.

**Dewert, Jules**, "Jean Maes père et Jean Maes fils (1610–1638). Les deux premiers imprimeurs athois," *Annales du Cercle archéologique d'Ath et de la région*, 1926, *13*, 67–96.

**Dittmar, Jeremiah E**, "Information technology and economic change: the impact of the printing press," *The Quarterly Journal of Economics*, 2011, *126* (3), 1133–1172.

**Domonkos, Leslie S**, "The Polish Astronomer Martinus Bylica de Ilkusz in Hungary," *The Polish Review*, 1968, pp. 71–79.

**Drüll, Dagmar**, *Heidelberger Gelehrtenlexikon: 1386–1651*, Berlin: Springer, 2002.

**Dulieu, Louis**, *La médecine à Montpellier, vol II: La Renaissance*, Avignon: Les presses universelles, 1979.

**Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker**, "Referral-based job search networks," *The Review of Economic Studies*, 2016, *83* (2), 514–546.

**Ehrmann, Maud, Ahmed Hamdi, Elvys Linhares Pontes, Matteo Romanello, and Antoine Doucet**, "Named Entity Recognition and Classification in Historical Documents: A Survey," *ACM Computing Surveys*, 2023, *56* (2), 1–47.

**Eijnatten, Joris Van**, "Between practice and principle: Dutch ideas on censorship and press freedom, 1579–1795," *Redescriptions: Political Thought, Conceptual History and Feminist Theory*, 2004, *8* (1).

**English, Edward D.**, *Encyclopedia of the Medieval World*, New York: Facts On File, Inc., 2005.

**Facca, Danilo, ed.**, *The Students and their Books: Early Modern Practices of Teaching and Learning*, Vol. 2 of *Renaissance Mind*, Berlin and Boston: De Gruyter, 2025.

**Facciolati, Jacopo**, *Fasti gymnasii Patavini. Jacobi Facciolati opera collecti ab anno MDXVII quo restitutae scholae sunt ad MDCCLVI*, Padova: typis Seminarii, 1757.

**Fantoli, Margherita, Jukka Suomela, Toon Van Hal, Mark Depauw, Lari Virkki, and Mikko Tolonen**, "Quantifying the Presence of Ancient Greek and Latin Classics in Early Modern Britain," *Journal of Cultural Analytics*, 2025, *10* (1).

**Farber, Henry S and Robert Gibbons**, "Learning and wage dynamics," *The Quarterly Journal of Economics*, 1996, *111* (4), 1007–1047.

**Febvre, Lucien and Henri-Jean Martin**, *The coming of the book: the impact of printing 1450-1800*, Vol. 10, Verso, 1997.

**Ferris, Timothy, Fred Stella, and Kevin Yon**, *The science of liberty: Democracy, reason, and the laws of nature*, Harper New York, 2010.

**Feys, Xander, Maxime Maleux, Andy Peetermans, and Raf Van Rooy, eds**, *Student Notes from Latin Europe (1400–1750): A Research Companion*, Leuven: Leuven University Press, 2025.

**Fogel, Robert William**, "Railroads and American Economic Growth: Essays in Econometric History," 1964.

**Fogli, Alessandra and Laura Veldkamp**, "Germs, social networks, and growth," *The Review of Economic Studies*, 2021, *88* (3), 1074–1100.

**François, Wim**, "Authority Revisited: Leuven's 500th Year Commemoration of Erasmus' Novum Instrumentum and More's Utopia," in Wim François, Violet Soen, Anthony Dupont, and Andrea. Aldo Robiglio, eds., *Authority Revisited. Towards Thomas More and Erasmus in 1516*, number 10. In 'Lectio.', Turnhout: Brepols, 2020.

**Frijhoff, Willem**, "Patterns," in Hilde de Ridder-Symoens, ed., *A History of the University in Europe. Vol. II: Universities in Early Modern Europe (1500–1800)*, Cambridge University Press, 1996, chapter 2.

**Gabriel, Astrik L**, *The Medieval Universities of Pécs and Pozsony*, Frankfurt-am-Main: Joseph Knecht, 1969.

**Gabrieli, Giuseppe**, *Contributi alla storia della Accademia dei Lincei*, Accademia nazionale dei Lincei, 1989.

**Gavin, Michael**, "Historical Text Networks: The Sociology of Early English Criticism," *Eighteenth-Century Studies*, 2016, *50* (1), 53–80.

**Genet, Jean-Philippe**, "Projet Studium Parisiense," 2019. Université Paris 1.

**Gentzkow, Matthew and Jesse M Shapiro**, "Ideological segregation online and offline," *The Quarterly Journal of Economics*, 2011, *126* (4), 1799–1839.

**Geraerts, Jaap and Demival Vasques Filho**, "Networks of Confessional Affiliation: Religious Choice and the Schism of Utrecht," *Journal of Historical Network Research*, 2024, *10* (1).

**Geudens, Christophe and Serena Masolini**, "Teaching Aristotle at the Louvain Faculty of Arts, 1425–1500. General Regulations and Handwritten Testimonies," *Rivista di Filosofia Neoscolastica*, 2016, *108* (4), 813–844.

**Giorcelli, Michela, Nicola Lacetera, and Astrid Marinoni**, "How does scientific progress affect cultural changes? A digital text analysis," *Journal of Economic Growth*, 2022, *27*, 415–452.

**Gittel, Benjamin**, "An Institutional Perspective on Genres: Generic Subtitles in German Literature from 1500–2020," *Journal of Cultural Analytics*, 2021, *6* (1), 1–38.

**Goeing, Anja-Silvia, Glyn Parry, and Mordechai Feingold, eds**, *Early Modern Universities: Networks of Higher Learning* number 31. In 'Scientific and Learned Cultures and Their Institutions Series.', Leiden: Brill, 2021.

**Golub, Benjamin and Matthew O Jackson**, "How homophily affects the speed of learning and best-response dynamics," *The Quarterly Journal of Economics*, 2012, *127* (3), 1287–1338.

**Goyal, Sanjeev**, *Networks: An economics approach*, MIT Press, 2023.

**_ , Marco J Van Der Leij, and José Luis Moraga-González**, "Economics: An emerging small world," *Journal of political economy*, 2006, *114* (2), 403–412.

**Gráda, Cormac Ó**, "Did science cause the industrial revolution?," *Journal of Economic Literature*, 2016, *54* (1), 224–239.

**Grafton, Anthony**, *Inky Fingers: The Making of Books in Early Modern Europe*, Cambridge, MA and London: The Belknap Press of Harvard University Press, 2020.

**Grandjean, Martin**, "Introduction to Social Network Analysis: Basics and Historical Specificities," in "Proceedings of the HNR+ResHist Conference 2021" 2021.

**Granovetter, Mark**, "Economic action and social structure: The problem of embeddedness," *American Journal of Sociology*, 1985, *91* (3), 481–510.

**Gregory, Ian**, "Challenges and Opportunities for Digital History," *Frontiers in Digital Humanities*, 2014, *1*, 1–2.

**Greif, Avner**, "Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition," *American Economic Review*, 1993, *83* (3), 525–548.

**_ , Joel Mokyr, and Guido Tabellini**, *Two Paths to Prosperity*, Princeton University Press, 2025.

**Grendler, Paul F.**, *The Roman Inquisition and the Venetian Press, 1540–1605*, Princeton: Princeton University Press, 1977.

**Grendler, Paul F**, *Jesuit Schools and Universities in Europe 1548–1773* Brill Research Perspectives in Jesuit Studies, Brill, 2019.

**Greteman, Blaine**, *Networking Print in Shakespeare's England: Influence, Agency, and Revolutionary Change*, Stanford University Press, 2021.

**Guiso, Luigi, Paola Sapienza, and Luigi Zingales**, "The role of social capital in financial development," *American Economic Review*, 2004, *94* (3), 526–556.

**Gunther, Robert T.**, *Early Science in Oxford, Vol. xi Oxford colleges and their men of science*, Oxford, 1937.

**Hallmann, Carl, W Walker Hanlon, and Lukas Rosenberger**, "Why Britain? The Right Place (in the Technology Space) at the Right Time," 2022. Working paper.

**Hanlon, W Walker**, "Necessity is the mother of invention: Input supplies and Directed Technical Change," *Econometrica*, 2015, *83* (1), 67–100.

_ , "The rise of the engineer: Inventing the professional inventor during the Industrial Revolution," *The Economic Journal*, 2025, p. ueaf023.

**Hanneman, Robert. and Mark Riddle**, *Introduction to Social Network Methods*, Riverside: University of California, 2005.

**Herbermann, Charles George**, *The Catholic Encyclopedia*, New York: Encyclopedia Press, Incorporated, 1913.

**Herriges, Joseph A and Catherine L Kling**, "Testing the consistency of nested logit models with utility maximization," *Economics Letters*, 1996, *50* (1), 33–39.

**Heßbrüggen-Walter, Stefan**, "Interdisciplinarity in the Seventeenth Century? A Co-occurrence Analysis of Early Modern German Dissertation Titles," *Synthese*, 2024, *203* (2), 67.

**Hill, Mark J., Ville Vaara, and Mikko Tolonen**, "Communication and Idea Transmission Across Historical Communities: A Quantitative Analysis of Early Modern Nonconformist Networks," *Huntington Library Quarterly*, 2023, *86* (2), 377–407.

**_ , _ , Tanja Säily, Leo Lahti, and Mikko Tolonen**, "Reconstructing Intellectual Networks: From the ESTC's Bibliographic Metadata to Historical Material," in "Proceedings of the Digital Humanities in the Nordic Countries" 2019.

**Hochberg, Yael, Alexander Ljungqvist, and Yang Lu**, "Whom you know matters: Venture capital networks and investment performance," *Journal of Finance*, 2007, *62* (1), 251–301.

**Hotson, Howard and Thomas Wallnig, eds**, *Reassembling the Republic of Letters in the Digital Age*, Göttingen University Press, 2019.

**Howse, Derek**, "The Greenwich List of Observatories: A World List of Astronomical Observatories, Instruments and Clocks, 1670–1850," *Journal for the History of Astronomy*, 1986, *17* (4), i–89.

**Israel, Jonathan I.**, *The Dutch Republic: Its Rise, Greatness, and Fall, 1477–1806* Oxford History of Early Modern Europe, Oxford: Clarendon Press, 1995.

**Jackson, Matthew O.**, *Social and Economic Networks*, Princeton, NJ, USA: Princeton University Press, 2008.

**Jedwab, Remi, Noel D Johnson, and Mark Koyama**, "Negative shocks and mass persecutions: evidence from the Black Death," *Journal of Economic Growth*, 2019, *24* (4), 345–395.

**Johns, Adrian**, *The nature of the book: Print and knowledge in the making*, University of Chicago Press, 1998.

**Jovanovic, Boyan**, "Job matching and the theory of turnover," *Journal of Political Economy*, 1979, *87* (5, Part 1), 972–990.

**Judge, Jane**, *The United States of Belgium: The Story of the First Belgian Revolution*, Leuven: Leuven University Press, 2018.

**Kedzoria, Andrzej**, "Zamiowopedia," 2021. Accessed Jan 10 2025.

**King, David**, *Finding Atlantis: A True Story of Genius, Madness, and an Extraordinary Quest for a Lost World*, Harmony Books, 2005.

**Kirwan, Richard**, "Function in Form: Single-Sheet Items and the Utility of Cheap Print in the Early Modern German University," in Andrew Pettegree, ed., *Broadsheets: Single-Sheet Publishing in the First Age of Print*, Leiden and Boston: Brill, 2017, pp. 337–354.

**Kivelä, Mikko, Alex Arenas, Marc Barthelemy, James P. Gleeson, Yamir Moreno, and Mason A. Porter**, "Multilayer Networks," *Journal of Complex Networks*, 2014, *2* (3), 203–271.

**Kling, Catherine L and Joseph A Herriges**, "An empirical investigation of the consistency of nested logit models with utility maximization," *American Journal of Agricultural Economics*, 1995, *77* (4), 875–884.

**Koher, Andreas, Hartmut HK Lentz, Philipp Hövel, and Igor M Sokolov**, "Infections on temporal networks—a matrix-based approach," *PloS one*, 2016, *11* (4), e0151209.

**Koschnick, Julius**, "Did a feedback mechanism between propositional and prescriptive knowledge create modern growth?," 2025.

_ , "Teacher-Directed Scientific Change: The Case of the English Scientific Revolution," Technical Report 274, EHES Working Paper Series 2025.

**Ladd, John R.**, "Imaginative Networks: Tracing Connections Among Early Modern Book Dedications," *Journal of Cultural Analytics*, 2021, *6* (1), 180–216.

**Lahti, Leo, Jani Marjanen, Hege Roivainen, and Mikko Tolonen**, "Bibliographic Data Science and the History of the Book (c. 1500–1800)," *Cataloging & Classification Quarterly*, 2019, *57* (1), 5–23.

_ **, Niko Ilomäki, and Mikko Tolonen**, "A Quantitative Study of History in the English Short-Title Catalogue (ESTC), 1470–1800," *LIBER Quarterly*, 2015, *25* (2), 87–116.

_ **, Ville Vaara, Jani Marjanen, and Mikko Tolonen**, "Best Practices in Bibliographic Data Science," in Juhani Harri Jantunen, S. Brunni, N. Kunnas, S. Palviainen, and K. Västi, eds., *Proceedings of the Research Data and Humanities (RDHum) 2019 Conference: Data, Methods and Tools*, Studia Humaniora Ouluensia University of Oulu 2019, pp. 57–65.

**Lamberts, Emiel and Jan Roegiers**, *Leuven University: 1425–1985*, Leuven: Leuven University Press, 1990.

**Laouenan, Morgane, Palaash Bhargava, Jean-Benoît Eyméoud, Olivier Gergaud, Guillaume Plique, and Etienne Wasmer**, "A cross-verified database of notable people, 3500BC-2018AD," *Scientific Data*, 2022, *9* (290).

**Lemercier, Claire**, "Formal network methods in history: Why and how?," in G. Fertig, ed., *Social Networks, Political Institutions, and Rural Societies*, Turnhout: Brepols, 2015, pp. 281–310.

_ , "A History Without the Social Sciences?," *Annales: Histoire, Sciences Sociales*, 2015, *70* (2), 271–283. Translated by Angela Krieger.

_ **and Claire Zalc**, *Quantitative Methods in the Humanities: An Introduction*, University of Virginia Press, 2019. Translated by Arthur Goldhammer.

**Limbach, Saskia**, "Advertising Medical Studies in Sixteenth-Century Basel: Function and Use of Academic Disputations," in Andrew Pettegree, ed., *Broadsheets: Single-Sheet Publishing in the First Age of Print*, Leiden and Boston: Brill, 2017, pp. 376–400. Chapter title published as above; the Brill

preliminary table of contents listed it as "Printing Medical Disputations in Basel: A Compelling but Competitive Business".

**Luther, Martin**, "Disputatio contra scholasticam theologiam," 1517. English version "Disputation Against Scholastic Theology" accessed October 14 2024.

**Maclean, Ian**, *Scholarship, Commerce, Religion*, Harvard University Press, 2012.

**Maloney, William F and Felipe Valencia Caicedo**, "Engineering growth," *Journal of the European Economic Association*, 2022, *20* (4), 1554–1594.

**Malthus, Thomas**, *An Essay on the Principle of Population*, fourth edition ed., London: T. Bensley, 1807.

**Manski, Charles F**, "Identification of endogenous social effects: The reflection problem," *The Review of Economic Studies*, 1993, *60* (3), 531–542.

**Mauelshagen, Franz**, "Networks of trust: scholarly correspondence and scientific exchange in early modern Europe," *The Medieval History Journal*, 2003, *6* (1), 1–32.

**Mazzetti, Serafino**, *Repertorio di tutti i Professori antichi e moderni della famosa Università, e del celebre Istituto delle Scienze di Bologna*, Bologna: tipografia di San Tommaso d'Aquino, 1847.

**McClellan, James E**, *Science Reorganized: Scientific Societies in the Eighteenth Century*, New York: Columbia University Press, 1985.

**McFadden, Daniel**, "Modelling the Choice of Residential Location," in Anders Karlquist, Lars Lundqvist, Folke Snickars, and Jörgen W. Weibull, eds., *Spatial Interaction Theory and Planning Models*, Amsterdam: North-Holland, 1978, pp. 75–96.

**Meisenzahl, Ralf R and Joel Mokyr**, "The rate and direction of invention in the British industrial revolution," *The rate and direction of inventive activity revisited*, 2012, pp. 443–482.

**Miert, Dirk Van**, "Afterword: Looking Forward to Networking the Early Modern World of Epistolary Contact," *Huntington Library Quarterly*, 2023, *86* (2), 409–425.

**Mokyr, Joel**, "Long-term economic growth and the history of technology," in "Handbook of economic growth," Vol. 1, Elsevier, 2005, pp. 1113–1180.

_ , *The enlightened economy: Britain and the industrial revolution, 1700-1850*, Penguin UK, 2011.

_ , *The gifts of Athena: Historical origins of the knowledge economy*, Princeton University Press, 2011.

_ , *A Culture of Growth*, Princeton: Princeton University Press, 2016.

**Montgomery, James D**, "Social networks and labor-market outcomes: Toward an economic analysis," *American Economic Review*, 1991, *81* (5), 1408–1418.

**Montreal Botanic Garden,**, *First Annual Report 1885*, Gazette Printing Company, 1886.

**Moser, Petra**, "How Do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World's Fairs," *The American Economic Review*, 2005, *95* (4), 1214–1236.

**Muller, Richard A**, "Scholasticism Protestant and Catholic: Francis Turretin on the Object and Principles of Theology," *Church history*, 1986, *55* (2), 193–205.

**Munck, Thomas**, *Conflict and Enlightenment: print and political culture in Europe, 1635–1795*, Cambridge University Press, 2019.

**Needham, Joseph**, "Science and Society in East and West," *Science and Society*, 1964, *28* (4), 385–408.

**Negro, Piero Del**, *Clariores: dizionario biografico dei docenti e degli studenti dell'Università di Padova*, Padova: Padova University Press, 2015.

**Nelson, Richard R**, "The simple economics of basic scientific research," *Journal of political economy*, 1959, *67* (3), 297–306.

**North, Douglass C**, *Institutions, institutional change and economic performance*, Cambridge University Press, 1990.

**Nurmi, Tarmo, Arash Badie-Modiri, Corinna Coupette, and Mikko Kivelä**, "pymnet: A Python Library for Multilayer Networks," *Journal of Open Source Software*, 2024, *9* (99), 6930.

**Oosterhoff, Richard J.**, *Making Mathematical Culture: University and Print in the Circle of Lefèvre d'Étaples*, Oxford: Oxford University Press, 2018.

**Padilla, Thomas**, "Foreword," in Paul Gooding, Melissa Terras, and Sarah Ames, eds., *Library Catalogues as Data: Research, Practice and Usage*, Facet Publishing, 2025, pp. 20–21.

**Palomino, Frédéric and Andrea Prat**, "Risk taking and optimal contracts for money managers," *RAND Journal of Economics*, 2003, pp. 113–137.

**Pedersen, Olaf**, *Lovers of Learning: A History of the Royal Danish Academy of Sciences and Letters, 1742-1992*, Copenhagen: Munksgaard, 1992.

**Peters, Ole**, "The ergodicity problem in economics," *Nature Physics*, 2019, *15* (12), 1216–1221.

**Petras, Vivien, Ray R. Larson, and Michael Buckland**, "Time Period Directories: A Metadata Infrastructure for Placing Events in Temporal and Geographic Context," in "Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '06)" ACM 2006, pp. 151–160.

**Pettegree, Andrew**, "Centre and periphery in the European book world," *Transactions of the Royal Historical Society*, 2008, *18*, 101–128.

_ , *The Book in the Renaissance*, New Haven and London: Yale University Press, 2010.

_ **, ed.**, *Broadsheets: Single-Sheet Publishing in the First Age of Print*, Vol. 60 of *Library of the Written Word; The Handpress World*, Leiden and Boston: Brill, 2017.

**Preiser-Kapeller, Johannes**, "Letters and Network Analysis," in Alexander Riehle, ed., *A Companion to Byzantine Epistolography*, Leiden: Brill, 2020, pp. 431–463.

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

**Reed, Frank**, "The Centennia Historical Atlas: Academic Research Edition," 2016.

**Renazzi, Filippo Maria**, *Storia dell'Università degli studi di Roma: detta comunemente la Sapienza, che contiene anche un saggio storico della letteratura romana, dal principio del secolo XIII sino al declinare del secolo XVIII*, Rome: Nella Stamperia Pagliarini, 1803.

**Roegiers, Jan**, "Justus Lipsius, academicus," in G. Tournoy, J. Papy, and J. De Landtsheer, eds., *Lipsius en Leuven. Catalogus van de tentoonstelling in de Centrale Bibliotheek te Leuven, 18 september – 17 oktober 1997*, Louvain: Leuven University Press, 1997.

**Roller, Ramona**, "Theory-Driven Statistics for the Digital Humanities: Presenting Pitfalls and a Practical Guide by the Example of the Reformation," *Journal of Cultural Analytics*, 2023, *7* (4).

_ , "Tracing the Footsteps of Ideas: Time-Respecting Paths Reveal Key Reformers and Communication Pathways in Protestant Letter Networks," *SocArXiv*, 2023.

_ , "Tracing the Footsteps of Ideas: Time-respecting Paths Reveal Key Reformers and Communication Pathways in Protestant Letter Networks," *SocArXiv*, 2023.

**Rollinger, Christian**, "Prolegomena. Problems and Perspectives of Historical Network Research and Ancient History," *Journal of Historical Network Research*, 2020, *4*, 1–35.

**Rooy, Raf Van**, "In Rutger Rescius' Classroom at the Leuven Collegium Trilingue (1543–1544): His Study Program and Didactic Method," in Raf Van Rooy, Pierre Van Hecke, and Toon Van Hal, eds., *Trilingual Learning: The Study of Greek and Hebrew in a Latin World (1000–1700)*, Turnhout: Brepols, 2022, pp. 179–205.

**Rouzet, Anne**, *Dictionnaire des imprimeurs, libraires et éditeurs des XVe et XVIe siècles dans les limites géographiques de la Belgique actuelle*, Nieuwkoop: B. de Graaf, 1975.

**Rubin, Jared**, "Printing and Protestants: an empirical test of the role of printing in the Reformation," *Review of Economics and Statistics*, 2014, *96* (2), 270–286.

**Ruffini, Giovanni**, *Social Networks in Byzantine Egypt*, Cambridge: Cambridge University Press, 2008.

**Ryan, Yann Ciarán and Mikko Tolonen**, "The Evolution of Scottish Enlightenment Publishing," *The Historical Journal*, 2024, *67* (2), 223–255.

_ **and** _ , "Networks of Influence in Scottish Enlightenment Publishing," *Connections*, 2024, *44* (1).

**Sanh, Victor, Lysandre Debut, Julien Chaumond, and Thomas Wolf**, "DistilBERT: A Distilled Version of BERT—Smaller, Faster, Cheaper and Lighter," 2019.

**Scebba, Rossana**, "Bibliographic Metadata as Relational Data: An Interdisciplinary Methodological Guide," *Journal of Cultural Analytics*, 2026, *11* (1).

_ **and Margherita Fantoli**, "Integrating Library and Prosopographical Data in the Publication Network of the Old University of Louvain," in Violet Soen, Wouter Druwé, Wim François, and Ralph Dekoninck, eds., *Students, Scholars and Their Books at the University of Louvain (1425–1797)*, Vol. 17 of *Lectio Series Studium Lovaniense*, Brepols, 2026, pp. 395–424.

**Schich, Maximilian, Chaoming Song, Yong-Yeol Ahn, Alexander Mirsky, Mauro Martino, Albert-László Barabási, and Dirk Helbing**, "A Network Framework of Cultural History," *Science*, 2014, *345* (6196), 558–562.

**Schoenfeld, David**, "Partial residuals for the proportional hazards regression model," *Biometrika*, 1982, *69* (1), 239–241.

**Scholtes, Ingo**, "When is a Network a Network? Multi-Order Graphical Model Selection in Pathways and Temporal Networks," in "KDD '17: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining" 2017, pp. 1037–1046.

**Schwinges, Rainer Christoph and Christian Hesse**, "Repertorium Academicum Germanicum," 2019. Consulted October 2, 2024.

**Serafinelli, Michel and Guido Tabellini**, "Creativity over time and space: A historical analysis of European cities," *Journal of Economic Growth*, 2022, *27* (1), 1–43.

**Silberman, Leo**, "Hung Liang-Chi: a Chinese Malthus," *Population Studies*, 1960, *13* (3), 257–265.

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

**Smets, An**, "War, division and digital reunification, or what do the Omont collection and the lecture notes kept at the Leuven University Library have in common?," in Mel Collier, ed., *What do we lose when we lose a library? Proceedings of the conference held at the KU Leuven 9–11 September 2015*, Leuven: KU Leuven University Library, 2016, pp. 168–176.

**Smith, Adam**, *An Inquiry into the Nature and Causes of the Wealth of Nations*, London: Strahan and Cadell, 1776.

**Smith, Anthony**, "Indirect Inference," in Steven N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics*, London: Palgrave Macmillan, 2008.

**Soen, Violet**, "From Magister Dixit to STUDIUM.AI: What is New about the Old University of Louvain (1425–1797)," in Violet Soen, Wouter Druwé, Wim François, and Ralph Dekoninck, eds., *Students, Scholars and their Books at the University of Louvain (1425–1797)*, Turnhout: Brepols, 2026, pp. 21–57.

_ **and Margherita Fantoli**, "Computing Women's Centrality in an Academic Book World: Widow-Printers in a Spanish Habsburg University Town of Douai (1559–1659)," in Alicia C. Montoya and Lieke van Deinsen, eds., *Digital Approaches to Historical Women's Book Culture*, Leiden and Boston: Brill, forthcoming.

_ **, Dries Vanysacker, and Wim François, eds**, *Church, Censorship and Reform in the Early Modern Habsburg Netherlands*, Vol. 101 of *Bibliothèque de la Revue d'histoire ecclésiastique*, Turnhout: Brepols, 2017.

_ **, Wouter Druwé, Wim François, and Ralph Dekoninck, eds**, *Students, Scholars and their Books at the University of Louvain (1425–1797)* LECTIO Series, Turnhout: Brepols, 2025.

**Spence, Michael**, "Job Market Signaling," *The Quarterly Journal of Economics*, 1973, *87* (3), 355–374.

**Squicciarini, Mara P and Nico Voigtländer**, "Human capital and industrialization: Evidence from the age of enlightenment," *The Quarterly Journal of Economics*, 2015, *130* (4), 1825–1883.

**Stapel, Rombert J.**, "Mint Authorities of the Low Countries," https://hdl.handle.net/10622/HPIC74 2025. hdl:10622/HPIC74.
*STCV: Bibliography of the Hand Press Book*

***STCV: Bibliography of the Hand Press Book***, *http://www.stcv.be. Online bibliographic database.*

**Stelter, Robert, David de la Croix, and Mikko Myrskylä**, *"Leaders and Laggards in Life Expectancy among European Scholars from the Sixteenth to the Early Twentieth Century," Demography, 2021, 58, 111–135.*

**Stiglitz, Joseph E. and Andrew Weiss**, *"Credit Rationing in Markets with Imperfect Information," The American Economic Review, 1981, 71 (3), 393–410.*

**Tennant, Roy**, *"MARC Must Die," Library Journal, 2002, 127 (17), 26–27.*

**The British Library**, *"Database of Italian Academies," 2021. Consulted on June 6, 2023.*

**The Editors of Encyclopaedia Britannica**, *"Guy de Chauliac," 2024. Accessed: 27 January 2025.*

**Thomas, Drew B.**, *"Reconstructing Broadsheet Production in Reformation Wittenberg," in Andrew Pettegree, ed., Broadsheets: Single-Sheet Publishing in the First Age of Print, Leiden and Boston: Brill, 2017, pp. 114–137.*

**Tiihonen, Iiro, Leo Lahti, and Mikko Tolonen**, *"Print Culture and Economic Constraints: A Quantitative Analysis of Book Prices in Eighteenth-Century Britain," Explorations in Economic History, 2024, 94.*

**Tirole, Jean**, *"A theory of collective reputations (with applications to the persistence of corruption and to firm quality)," The Quarterly Journal of Economics, 1996, 111 (4), 1199–1235.*

**Tolonen, Mikko, Leo Lahti, Hege Roivainen, and Jani Marjanen**, *"A Quantitative Approach to Book-Printing in Sweden and Finland, 1640–1828," Historical Methods: A Journal of Quantitative and Interdisciplinary History, 2019, 52 (1), 57–78.*

**_ , Mark J. Hill, Ali Zeeshan Ijaz, Ville Vaara, and Leo Lahti**, *"Examining the Early Modern Canon: The English Short Title Catalogue*

*and Large-Scale Patterns of Cultural Production," in "Data Visualization in Enlightenment Literature and Culture" 2021, pp. 63–119.*

**Topa, Giorgio**, *"Social interactions, local spillovers and unemployment," The Review of Economic Studies, 2001, 68 (2), 261–295.*

**Tournoy, Gilbert, Jan Roegiers, and Chris Coppens**, *Vives Te Leuven. Exhibition Catalogue, Leuven: Leuven University Press, 1993.*

**Udías, Agustín**, *Searching the heavens and the earth: the history of Jesuit observatories, Vol. 286, Springer Science & Business Media, 2003.*

**Universität Wien**, *"Geschichte der Universität Wien - Logo Universität Wien 650 plus – History of the University of Vienna," 2024. accessed 11-November-2025.*

**Uniwersytet Jagielloński**, *"Corpus Academicum Cracoviense," 2019. Accessed: January 20, 2021.*

**Valleriani, Matteo and Andrea Ottone, eds**, *Publishing Sacrobosco's De sphaera in Early Modern Europe: Modes of Material and Scientific Exchange, Cham: Springer, 2022.*

**_ , Malte Vogl, Hassan el Hajj, and Kim Pham**, *"The Network of Early Modern Printers and Its Impact on the Evolution of Scientific Knowledge: Automatic Detection of Awareness Relationships," Histories, 2022, 2 (4), 466–503.*

**Van Der Biest, Jarrik**, *"From the Lecture Hall to the Confessional Frontier: Student Notebook Production and the Transmission of Biblical Knowledge from Louvain," in Xander Feys, Maxime Maleux, Andy Peetermans, and Raf Van Rooy, eds., Student Notes from Latin Europe (1400–1750): A Research Companion, Leuven: Leuven University Press, 2025, pp. 11–46.*

**van Vugt, Ingeborg**, *"Using Multi-Layered Networks to Disclose Books in the Republic of Letters," Journal of Historical Network Research, 2017, 1, 2–51.*

**Vanden Bosch, Gerrit**, *"Censorship in the Archdiocese of Malines (1559–1795)," in Violet Soen, Dries Vanysacker, and Wim François, eds., Church, Censorship and Reform in the Early Modern Habsburg Netherlands, number*

*101. In 'Bibliothèque de la Revue d'histoire ecclésiastique.', Turnhout: Brepols, 2017, pp. 97–106.*

**Voigtländer, Nico and Hans-Joachim Voth**, *"Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany," The Quarterly Journal of Economics, 07 2012, 127 (3), 1339–1392.*

**von Aschbach, Joseph Ritter**, *Geschichte der Wiener Universität und Ihre Gelehrten, Vienna: Verlag der Universität, 1865.*

**Walsby, Malcolm**, *"Cheap Print and the Academic Market: The Printing of Dissertations in Sixteenth-Century Louvain," in Andrew Pettegree, ed., Broadsheets: Single-Sheet Publishing in the First Age of Print, Leiden and Boston: Brill, 2017, pp. 355–375.*

**Warner, J. Christopher**, *"The Printing of Dissertations in Sixteenth-Century Louvain: A Reconsideration in Light of Other Samples and Factors," De Gulden Passer / The Golden Compasses: Journal for Book History, 2018, 96 (2), 183–244.*

_ , *"Quaestiones theologicae, theses sacrae, and Some Conjectures about the Masius Family," in Goran Proot, David McKitterick, Angela Nuovo, and Paul F. Gehl, eds., Lux Librorum: Essays on Books and History for Chris Coppens, Mechelen: Flanders Book Historical Society, 2018, pp. 123–140.*

**Warren, Christopher N, Daniel Shore, Jessica Otis, Lawrence Wang, Mike Finegold, and Cosma Shalizi**, *"Six Degrees of Francis Bacon: A Statistical Method for Reconstructing Large Historical Social Networks," Digital Humanities Quarterly, 2016, 10 (3), 1–16.*

**Wasserman, Stanley and Katherine Faust**, *Social Network Analysis: Methods and Applications, Cambridge University Press, 1994.*

**Wetherell, Charles**, *"Historical Social Network Analysis," International Review of Social History, 1998, 43 (S6), 125–144.*

**Wilcox, Rand**, *Modern Statistics for the Social and Behavioral Sciences. A Practical Introduction, second ed., Boca Raton: CRC Press, 2017.*

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

**Wyffels, Heleen**, *"Women and Work in Early Modern Printing Houses: Family Firms in Antwerp, Douai, and Leuven (1500–1700)." Phd dissertation, KU Leuven, Leuven 2021.*

**Xue, Melanie**, *"Enlightenment Under Autocracy: The Origins of Liberalism in China," 2025. LSE.*

**y Díaz, Alejandro Vidal et al.**, *Memoria histórica de la Universidad de Salamanca, Salamanca: Imprenta de Olivo y Hermano, 1869.*

**Zanardello, Chiara**, *"Market forces in Italian academia today (and yesterday)," Scientometrics: an international journal for all quantitative aspects of the science of science, communication in science and science policy, 2023, 128.*

_ , *"Early Modern Academies and Growth," 2024. LIDAM Discussion Paper - 2024/12.*

**Zenou, Yves**, *"Key players," The Oxford Handbook of the Economics of Networks, 2016, 4, 244–274.*