

Did a feedback mechanism between propositional and prescriptive knowledge create modern growth?*

Julius Koschnick (University of Southern Denmark)[†]

December 19, 2025

Abstract

What was the origin of modern economic growth? Joel Mokyr has argued that self-sustained modern economic growth originated from a feedback loop between propositional (theoretical) and prescriptive (applied) knowledge, which turned positive in the eighteenth century during the “Industrial Enlightenment”. While influential, this thesis has never been directly tested. This paper provides the first quantitative evidence by estimating the impact of knowledge spillovers between propositional and prescriptive knowledge on innovation in England, 1600–1800. For this, it introduces two new text-based measures for 1) the innovativeness of publications and 2) knowledge spillovers. The paper finds strong evidence that a feedback loop between propositional and prescriptive knowledge became positive during the second half of the eighteenth century. It also documents that this process had positive effects on the real economy as measured through patents. Overall, the findings provide empirical support for Mokyr’s original hypothesis.

Keywords: ECONOMIC GROWTH, INNOVATION, FEEDBACK LOOP PROCESSES, KNOWLEDGE SPILLOVERS, NATURAL LANGUAGE PROCESSING

JEL Classification: O33, O31, O14, O47, O49, N13, N33

*I am indebted to Joel Mokyr for kindling my interest in the historical knowledge economy through his academic writings and his personal encouragement and invaluable input that has made this project possible. I would further like to express my gratitude to my PhD supervisors Max-Stephan Schulze and Jeremiah Dittmar at the LSE for their invaluable guidance and support. Moreover, I am grateful to everyone at Northwestern University who supported me during my stay in the fall of 2022. Further thanks go to seminar participants at the LSE graduate economic history seminar, the World Economic History Congress in Lund, and the workshops on “Growth, History and Development” at the University of Southern Denmark, “Technology and Long-Term Development” at the University of Valencia, “Economic History in the Age of AI” at Stellenbosch University, and “History LLMs” in Zurich. Additionally, I want to thank Bruno Caprettini, Lukas Rosenberger, Erik Hornung, Caterina Chiopris, Christian Vedel, Alexandra de Pleijt, and David de la Croix for their detailed and most helpful feedback. Finally, I am obliged to Alessandro Nuvolari and Valentina Tartari for kindly sharing their data on English patents.

[†]University of Southern Denmark, Economics Department, Campusvej 55, Odense, Denmark

1 Introduction

What explains the onset of modern economic growth at the end of the eighteenth century? Joel Mokyr (2002) has argued that self-sustained modern economic growth originated from structural change in the knowledge economy following the Scientific Revolution and Enlightenment. In the centuries before, waves of invention were rare and short-lived. After 1760, however, England entered a regime of steady growth in knowledge production, patents, and economic output (Mokyr, 2002, 2016; Clark, 2007; Baten and Van Zanden, 2008; Allen, 2009; Nuvolari and Tartari, 2011). According to Mokyr (2002), the transition to modern economic growth can be explained by a feedback loop between propositional (theoretical) and prescriptive (applied) knowledge. He argues that this feedback loop only became positive in the eighteenth century following changes in the production and sharing of ideas during the Enlightenment. While the feedback loop thesis has been widely influential (Jacob, 2014; Squicciarini and Voigtländer, 2015; Kelly and Ó Gráda, 2022; Prize Committee, 2025), we still lack quantitative estimates of feedback loop processes that could either support or falsify the theory.

This paper introduces the first estimates of feedback loop processes between propositional and prescriptive knowledge for seventeenth and eighteenth century England by utilizing recent advances in natural language processing. In doing so, the paper is the first to quantitatively test Mokyr (2002)'s feedback loop hypothesis. It is also the first paper to trace long-run patterns of innovation and spillovers in Britain's knowledge economy before and during the Industrial Revolution.

To do so, the paper introduces two new measures of 1) the innovativeness of publications and 2) spillovers between subject fields that are entirely based on information from text data. The measures are based on comparisons between the past and future of different fields while operationalizing a semantically deep and historically fine-tuned BERT model. With these two measures, it becomes possible to derive measures of innovation and knowledge spillovers i) in the absence of citation data and ii) in the presence of historical and context-sensitive text. The paper successfully validates the new innovation measure with historical patent citations, showing that the NLP-based approach introduced in this paper can recover structural information from the knowledge economy. It then applies the innovation and spillover measures to the universe of ca. 300,000 publication titles (historically equivalent to short abstracts) and patent descriptions in England, 1600–1800.¹

This toolset makes it possible to directly test Mokyr's feedback loop thesis. According to Mokyr (2002), throughout most of humanity's history, feedback loop processes were either non-existent or negative. Only with the "Industrial Enlightenment" of the eighteenth century did

¹During this time period, English publishers used the full space of a book's cover to publish short abstracts of the book with an average count of 55 words.

the feedback loop process turn positive, as propositional and prescriptive knowledge began to reinforce one another, “eventually tipping the balance of the feedback mechanism from negative to positive.” (Mokyr, 2002, p. 33). To test this prediction, the paper estimates whether, over time, spillovers from one type of knowledge increased the innovativeness of the other. This follows Mokyr’s definition of a positive feedback loop where “growth in one increases the marginal product of the other” (Mokyr, 2002, p. 21).

First, the paper produces associational evidence on the development of the relationship between received knowledge spillovers and innovativeness at the title level.² Accounting for subject- and publication-year fixed effects, it finds evidence of neutral or negative knowledge spillovers at the beginning of the seventeenth century. Then, over the next 140 years, negative spillovers phased out until around 1740, the relationship became positive for the first time. These findings directly correspond to the predictions from Mokyr (2002). The paper further links these results to the real economy by estimating the association between spillovers from propositional knowledge and patent citations in 1700–1800. As before, we find evidence of a positive relationship for the second half of the eighteenth century.

To rule out that the observed relationship is driven by shifts in language or stylistic conventions, the paper conducts a series of placebo tests. First, the paper presents evidence that spillovers from 14 plausibly unrelated fields, incl. e.g. *poetry, drama, foreign languages, biographies*, and *travel descriptions* did not exhibit any positive pattern. The paper further presents evidence from estimating spillovers from all 44 subject classes in the ESTC that were not part of propositional or prescriptive knowledge. Even here, spillovers from propositional and prescriptive knowledge at the end of the eighteenth century remain larger than any of the other spillover coefficients, highlighting the importance of spillovers between propositional and prescriptive knowledge within the knowledge economy.

Overall, these findings constitute strong evidence of the emergence of a structural break in Britain’s knowledge economy in the seventeenth and eighteenth century. Growth (spillovers) in one type of knowledge indeed seems to have increased the marginal product in the other type of knowledge (innovation). Moreover, the timing of the flipping of the relationship from negative to positive spillovers seems to follow exactly the narrative account of Mokyr (2002).

To further understand the mechanism underlying the positive knowledge spillovers emerging at the end of the eighteenth century, the paper tests whether upper-tail human capital facilitated knowledge spillovers, as predicted by Mokyr (2002, 2016). Extracting information on authors’

²To classify titles as propositional or prescriptive knowledge, the paper follows the classification system from Koschnick (2025). The set of propositional knowledge includes the subject fields of *applied physics, astronomy, mathematics, chemistry, and encyclopedias* and the set of prescriptive knowledge includes the subject fields of *technical publications, navigation, scientific instruments, and patents*. This closely follows Mokyr (2002)’s original definition of propositional knowledge being broader than just science, but also comprising simple facts and explanations. The approach of classifying titles into propositional and prescriptive knowledge is explained in detail in section 3.4.

careers from the ESTC, we find that Royal Society fellows and engineers were (a) more innovative and (b) together with other indicators of upper-tail human capital, explain up to 8% of the estimated positive spillovers. We interpret this as evidence that, consistent with recent literature (Squicciarini and Voigtländer, 2015; Kelly and Ó Gráda, 2020; Hanlon, 2025), upper-tail knowledge elites were instrumental in bridging the gap between propositional and prescriptive knowledge.

Moreover, the paper goes beyond Mokyr (2002)'s theory and investigates how revolutions in methods that occurred during the seventeenth and eighteenth century could have acted as complements to feedback loop processes. The importance of revolutions in methods has been stressed by scholars such as Landes, Wootton, Kelly, and Ó Gráda. Landes (1969, 1998), and Wootton (2015) have placed special emphasis on the new scientific method that consisted of observation, formalization, and experiment. Jacob (1997, 2014) have stressed the importance of the Newtonian Revolution and Newtonian mechanics for the British Industrial Revolution. Moreover, Ó Gráda (2016) and Kelly and Ó Gráda (2020) have highlighted the importance of precise measurement as a spillover from scientific instruments that enabled a new degree of complexity in machinery.

The paper tests these channels by calculating embedding similarities to a set of descriptive terms of these methods (similar to Garg et al., 2018; Ash, Chen and Naidu, 2025). It then interacts the spillover measure with the similarity to methods measure to estimate complementarities between the two. The paper finds that for the adoption of propositional knowledge in new prescriptive knowledge, precise measurement acted as a key catalyst. Vice versa, for the adoption of prescriptive knowledge in new propositional knowledge, Newtonian mechanics and the scientific method were key catalysts.

Lastly, the paper provides direct causal evidence of the effect of knowledge spillovers on innovation by exploiting a unique shock to the cost of accessing propositional and prescriptive knowledge: The release of John Harris's *Lexicon technicum* in 1704, the first modern scientific and technical encyclopedia in Britain. Its strong focus on Newtonian science constituted a shock to propositional knowledge. At the same time, its contributions to technical knowledge were limited (Harris was a practising mathematician with hardly any experience with the arts and trades). This setting makes it possible to cleanly identify the spillovers from propositional to prescriptive knowledge. Exploiting variation from the strength of spillovers between sub-topics, the paper adopts a difference-in-differences framework for prescriptive topics that were not covered in the *Lexicon technicum*. We find that knowledge spillovers from propositional knowledge from the *Lexicon technicum* led to increased innovation over the next 20 years. This provides a) direct causal evidence of positive spillovers from propositional to prescriptive knowledge and b) highlights the importance of codified knowledge and encyclopedias as a channel.

Altogether, the evidence found in this paper provides strong support for Mokyr (2002)'s

feedback loop hypothesis. The paper finds that by the mid-eighteenth century, the relationship between knowledge spillovers and innovation had turned positive: After ca. 1740, spillovers across knowledge domains were associated with increases in marginal productivity, consistent with Mokyr's predictions. Patenting evidence confirms that these spillovers translated into greater innovation in the real economy. Investigating the mechanism, the paper shows that upper-tail human capital played a key role in facilitating knowledge spillovers. The paper has further shown that new revolutions in methods (Landes, 1998; Jacob, 1997, 2014; Wootton, 2015; Ó Gráda, 2016; Kelly and Ó Gráda, 2020) were also complementary to the innovation process. Additionally, the paper presents causal evidence of innovation-enhancing knowledge spillovers from propositional to prescriptive knowledge from the introduction of one encyclopedia at the beginning of the eighteenth century.³ Taken together, the evidence lends overwhelming quantitative support to the Mokyrian feedback loop hypothesis and thereby contributes to our understanding of the origin of modern growth.

Overall, the paper contributes to the literature across four key dimensions. First, it provides a new approach to estimate structural change in the knowledge economy and provides new evidence on the origin of modern economic growth (see e.g. Romer, 1986, 1990; Pomeranz, 2000; Mokyr, 2002; Clark, 2007; Allen, 2009). Here, it provides the first quantitative test of Mokyr (2002)'s theory of a feedback loop process between propositional and prescriptive knowledge. The evidence found stands as strong support of Mokyr (2002)'s account of ideas and knowledge production as first-order drivers of the Industrial Revolution. Altogether, the paper provides new quantitative evidence that places additional weight on the importance of the knowledge economy for the Industrial Revolution, in contrast to e.g. natural resources (Wrigley, 2010) or factor prices (Allen, 2009), although these accounts can clearly be seen as complementary.⁴

Second, the paper moves beyond Mokyr (2002)'s theory and investigates the mechanisms behind the feedback loop mechanism. The paper places new emphasis on the role of revolutions in methods that occurred during the seventeenth and eighteenth century, such as the scientific method, the Newtonian Revolution, and precise measurement. These channels have been highlighted by Landes (1969, 1998), Wootton (2015), Jacob (1997, 2014), Ó Gráda (2016), and Kelly and Ó Gráda (2020) as potential first-order determinants of changes in the knowledge economy before the Industrial Revolution. This paper provides first quantitative evidence that supports these theories. Additionally, the paper shows that upper-tail human capital and networks through knowledge sharing societies contributed to the creation of knowledge spillovers. With this, the paper complements a long-standing literature on upper-tail human capital (Meisen-

³Here, more work on specific knowledge shocks such as other encyclopedias or publications of economic societies could be a fruitful route for further research.

⁴Here the printing press (Dittmar, 2011, 2019) and institutions (Mokyr, 2025) feature as an important precondition for the European divergence in knowledge production and science.

zahl and Mokyr, 2011; Squicciarini and Voigtländer, 2015; Hanlon, 2025; Maloney and Valencia Caicedo, 2022; Kelly and Ó Gráda, 2022; Kelly, Mokyr and Ó Gráda, 2023; Cinnirella, Hornung and Koschnick, 2025).

Third, the paper relates to the long-standing debate on the influence of science on the Industrial Revolution. Here, scholars have been divided on whether science was a key driver of the Industrial Revolution (Schofield, 1957, 1963; Musson and Robinson, 1969; Stewart, 1986; Stewart and Weindling, 1995; Jacob, 1997, 2014) or too underdeveloped to have had a direct effect (Mathias, 1972; Hall, 1974; Ó Gráda, 2016). The paper contributes to the literature by providing quantitative evidence of innovation-enhancing spillovers from science. However, these occur in a setting that allows for a multitude of indirect effects, such as the adoption of the scientific method or a quantitative mindset of precise measurement. Thereby, it reconciles both positions and contributes to a literature that has studied the diffusion of technical knowledge (Hanlon et al., 2022; Rosenberger, Hanlon and Hallmann, 2024; Chiopris, 2024) and scientific knowledge (Dittmar and Meisenzahl, 2021; Zanardello, 2024; Curtis and de la Croix, 2023; de la Croix, Scebbia and Zanardello, 2025; Koschnick, 2025) during the early modern period and Industrial Revolution.

Fourth, the paper contributes to the use of natural language processing in economics and economic history to quantify the knowledge economy. Recently, Almelhem et al. (2023) have provided evidence of an increasing separation of religion, economy, and science and an increasing culture of progress throughout the seventeenth and eighteenth century. Their findings complement the findings of this paper by also providing evidence of cultural attitudes on science and technology shifting. Additionally, Grajzl and Murrell (2024) provide evidence of quiet institutional and cultural revolutions in the late sixteenth and early seventeenth century as apparent in large text corpora.⁵

Here, the paper contributes through introducing and validating two new measures for i) innovation and ii) spillovers based on historical text data. The measures are based on a semantically deep BERT model that has been fine-tuned for seventeenth and eighteenth century scientific and technical text. However, the measures can easily be adopted to other BERT models or time periods. We are therefore confident that these measures will be useful to other researchers who are in need of text-based innovation and spillover measures in the absence of citation data. Moreover, because these measures avoid some of the biases inherent in citation data, such as peer and network effects, they also offer a useful substitute for traditional citation-based indicators.

The paper proceeds in the following way. Section 2 provides a conceptual overview of knowledge spillovers between propositional and prescriptive knowledge and illustrates them with the historical case of the improvement of water power. Section 3 introduces the text data on the

⁵Moreover, Grajzl and Murrell (2025) document a positive relationship between censorship and innovation in seventeenth century England.

universe of printed titles in Britain between 1600 and 1800, patents, and the *Lexicon technicum*. Section 4 introduces the innovation and spillover measure. The measure uses the embedding space of a historically trained and fine-tuned *SteamBERTh* model introduced in section 5.1. Section 5.2 validates the innovation measure with historical patent citations. Next, section 6 introduces the empirical framework for estimating the relationship between knowledge spillovers between propositional and prescriptive knowledge and innovation over time. Section 7 presents results. Lastly, section 8 exploits the introduction of the *Lexicon technicum* as a shock to the availability of propositional knowledge in a difference-in-differences framework. Section 9 concludes.

2 Conceptual and historical background

Conceptual background

How did propositional (theoretical) and prescriptive (applied) knowledge influence each other? This section discusses the channels of innovation-enhancing spillovers from propositional to prescriptive knowledge ($\Omega \rightarrow \lambda$) and from prescriptive back to propositional knowledge ($\lambda \rightarrow \Omega$). Thereby, it illustrates the operation of both positive and negative spillovers. An important insight here is that spillovers often operated indirectly or through a multitude of channels. At the end of the section, we illustrate the functioning of these channels with a historical case study of eighteenth century improvements in water wheels.

The operation of the feedback loop process is illustrated in figure 1. Here, spillovers between both propositional (Ω) and prescriptive knowledge (λ) increase innovation in the receiving field. New knowledge in the receiving field in turn creates new spillovers back. Spillovers themselves could operate through a range of channels denoted as A_1 – A_4 and B_1 – B_3 .

First, spillovers from propositional to prescriptive knowledge ($\Omega \rightarrow \lambda$) could have operated through channels A_1 – A_4 . Most directly, a theoretical understanding of the laws of nature could lead to exact predictions of optimal engineering techniques (A_1). For example, early eighteenth century engineers could use Hooke’s law to design springs, or model how bridges or tunnels would deform under load. Moreover, even without exact predictions, theories and observational knowledge could improve the search function over the technological space (A_2) (Kauffman, Lobo and Macready, 2000; Acemoglu, Akcigit and Kerr, 2016). Improvers could also profit from the adoption of the scientific toolset, such as mathematics or precise measurement (A_3) (Kelly and Ó Gráda, 2016). Lastly, knowledge about limits or inconsistencies in existing knowledge could also help to set the agenda for practical inquiry (A_4). As discussed later for the case of water wheels, conflicts between theories often served as a starting point for successful experimentation (Reynolds, 1983).

Notably, the presence of false predictions from false or underdeveloped scientific theories can

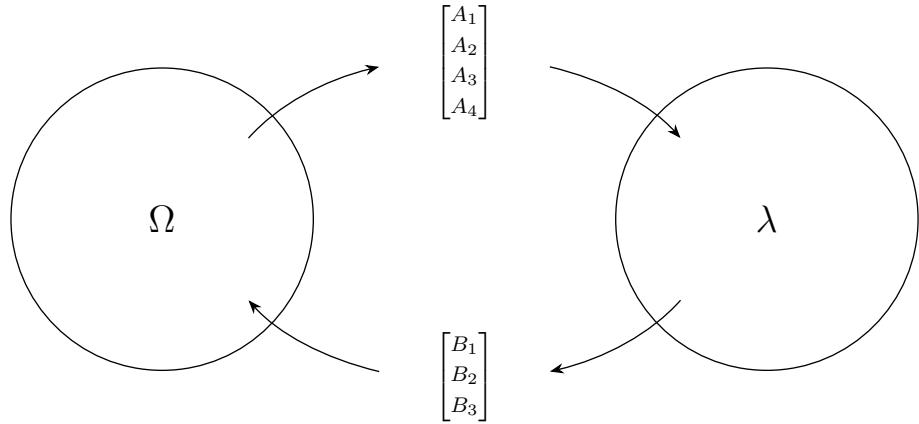


FIGURE 1: A feedback loop between propositional (Ω) and prescriptive knowledge (λ)

Notes: The figure illustrates a feedback loop between propositional (Ω) and prescriptive knowledge (λ). Innovation-enhancing spillovers from Ω to λ operate through channels A_1 – A_4 . Vice versa, innovation-enhancing spillovers from λ to Ω operate through channels B_1 – B_3 . See also (Mokyr, 2002, p. 17, 22).

also lead to negative spillovers into the technological space. Examples of this in the seventeenth century are manifold, such as the use of linear scaling laws before Galileo’s formulation of the square-cube law in 1638. The use of linear scaling in the extension of bridge constructions or ship design often led to structural failure (Valleriani, 2009). Moreover, the Aristotelian belief that nature abhorred a vacuum led to the prediction that water could be raised indefinitely by suction pumps. In fact, it failed above 10 meters, surprising contemporaries and early scientists.

Second, spillovers from prescriptive to propositional knowledge ($\lambda \rightarrow \Omega$) could likewise operate through multiple channels (B_1 – B_3). First, the practical implementation of theoretical predictions could help to reject false scientific theories (B_1). Second, research could be guided by puzzling facts (B_2). For example, Pasteur’s theoretical research was prompted by remarkable observations and puzzles, such as anomalies in fermentation that finally led to his discovery of the germ theory of disease (Stokes, 1997). Lastly, unsolved technological problems in commerce and trade could likewise help to set research agendas (B_3) (Stokes, 1997).

We should also note here, that false or incomplete empirical observations could lead to the development of incorrect and unsuccessful scientific theories. For example, early telescopes produced small discs around planets as optical artifacts. As these were mistaken for the true size of stars, this was used as an argument against heliocentrism, as star sizes appeared implausibly large (Graney, 2010; Graney and Grayson, 2011). Likewise, optical inspection of early dissections seemed to suggest that blood in arteries ebbed and flowed supporting the Galenic theory that blood is produced in the liver and therefrom flows outwards into the body to be consumed by tissues. Only later would venous valves be discovered, clearing the ground for William Harvey’s discovery of pulmonary and systemic circulation of blood (Scultetus, Villavicencio and Rich, 2001).

Historical illustration

Historians of innovation have presented evidence of an increasing interdependence of scientific theory and practical work in the eighteenth century (Reynolds, 1983; Stewart, 1986; Mokyr, 2002; Stewart, 2007). Yet, this research has also shown that innovation-inducing knowledge spillovers were often more complex than the implementation of successful predictions from science. The study of water wheels, a major source of power in the eighteenth century, serves to illustrate the complex interactions.

By the beginning of the eighteenth century, the importance of water power as a source of energy had attracted many of the most famous scientists of the age, such as Parent, Bernoulli, and Euler, to model features of an optimal water wheel (B_3). At the same time engineers like Polhem, de Parcieux, and Smeaton were reading these theories and practically experimenting on new designs. Yet, from the beginning it turned out that many of the theoretical predictions, such as the optimal number of wheels, conflicted and found little confirmation in practical experiments (Reynolds, 1983).⁶ On the one hand, this led to a continuous development and improvement of theories (B_1). On the other hand, practical experimentation, with theory as a starting point (A_2, A_4), also led to new and more effective designs such as the over-shot wheel. Reynolds (1983) has argued that the engineers' work was characterized by "(1) systematic methods of experimentation, (2) the use of working models, and (3) the application of quantitative measurements to key variables" (Reynolds, 1983, p. 232).⁷ All of this was new in comparison to the traditional approach that circled around the questions of "Will it work if I build it this way? or, If I change this element, will it work any better?" (ibid.). Engineers like Polhem, de Parcieux, and Smeaton now started from specific questions on the efficacy of designs that came from unsettled debates in physics (A_4). Thus, it were often the unsolved questions rather than successful predictions that gave engineers a specific starting place for experimentation. Furthermore, they applied a mathematical, and mechanical mindset to their experimentation where they sought to quantify effects and gather data (ibid.) (A_3).

Hence, at least for the study of water wheels, it seems that both propositional and prescriptive knowledge led to innovation-inducing spillovers, even if the channels were often not direct. Overall, innovation in the design of water wheels led to highly relevant increases in productivity. The efficiency of traditional waterwheel lay between ca. 30–40% (Viollet, 2017). Yet, Smeaton's breastshot wheel, as well as overshot and Poncelet undershot wheels⁸ delivered efficiencies of 60–80% (ibid.). In return, the scientific theory of hydrodynamics improved with each new practical insight (Reynolds, 1983). Likewise, early navigation started with practical problems encountered

⁶There is some controversy on whether the inadequacy of these theories was due to theoretical errors or idealized assumptions (Capecchi, 2013). This should not matter for the present context, their immaturity in yielding practical predictions is clear.

⁷A key input to quantitative experimentation was also the new design of testing devices (Constant, 1983).

⁸Based on theoretical work by Borda in the 1760s, but only designed by Poncelet in the 1820s.

by seamen which would then be discussed by court astronomers and university scholars (Howse, 1986; Taylor, 1957).⁹

We would expect to see the same dynamic in other scientific and technological fields. One can think of Josiah Wedgwood and the influence of his early chemistry studies on his pottery products. Here, early chemistry inspired the development of new materials and production methods even when the exact workings of the science were not yet clear. Similarly, dyeing at the end of the eighteenth century seems to have increasingly relied on chemical theory and controlled experimentation (Musson and Robinson, 1969, pp. 338–351). Schofield (1963), Stewart (1992, 2007), Stewart and Weindling (1995) and Jacob (2014) have further argued that knowledge flows between science and technological innovators in industries such as steam, cotton, weaving, metallurgy, and pottery steadily increased throughout the eighteenth century.

Yet, despite this narrative evidence of increasingly positive spillovers between propositional and prescriptive knowledge at the end of the eighteenth century, it remains hard to assess how representative these cases were for overall knowledge production (Ó Gráda, 2016). This is where quantitative evidence can significantly contribute to our understanding of the emergence of feedback loops. Therefore, the next sections will introduce data on the universe of published titles in Britain and present an empirical framework to estimate knowledge spillovers between prescriptive and propositional knowledge.

3 Data

3.1 Publication titles 1600–1800

To capture the content of the British stock of knowledge, the paper uses the universe of all unique 285,985 printed titles from the English Short Title Catalogue (ESTC) between 1600 and 1800 from Koschnick (2025). The ESTC contains all works printed in England as well as all works in English. With an average length of ca. 55 words, these titles had the form of short abstracts, usually covering the full front page of a book.¹⁰ Hence, they contain information on both the broad topic covered by a title as well as a short abbreviated summary of its content.

To classify the ESTC, we use subject classes from Koschnick (2025). Koschnick (2025) assigned higher-order categories to librarian-assigned subject labels and then used a BERT model trained on historical titles to assign subject fields to unclassified titles. The dataset was

⁹See e.g. John Flamsteed writing to Samuel Pepys in 1694 that: “All our great attainments in science and in the mechanic part also of Navigation have come out of the Chambers and from the fire-sides of thinking men within doors that were schollars and mechanics, and not from Tarpawlins, tho’ of never so great experience” (quoted from Lincoln, 1983, p. 83).

¹⁰The format was unique to England and was not adopted on the continent, e.g. in Germany or France. The format was furthermore consistent over time, with static word count levels between 1600 and 1800, see appendix figure 15.

further cleaned to remove duplicates. Appendix table 7 provides descriptive statistics of the ESTC catalogue. A typical example of an ESTC title reads:

“A compendium of algebra. Containing the principles of the analytick art, with rules for solving simple, quadratick, and cubick, &c. equations; together with the method of converging series’s; after so plain a method, that any one who understands numbers, may learn the solution of the said equations without a master. By George Gordon, assistant to the Reverend Dr. Desaguliers.” (George Gordon, 1728)

Moreover, this paper extracts information on authors’ occupations from the ESTC. Seventeenth and eighteenth century authors usually added prestigious occupations and affiliation with societies to their names on the title page. The paper extracts information occupational categories of *engineer*, *medical career*, and *academic career*, as well as memberships in the *the Royal Society* or other *enlightenment societies*. Appendix section A.1 describes this procedure in further detail.

Notably, the ESTC also covers a large amount of foreign works that were reprinted in England, thereby also accounting for knowledge inflows into the British knowledge economy. This is important, since it has been argued that the continent possessed an advantage in codifying knowledge (Mokyr, 2021). Figure 17 reports the number of all titles printed in foreign languages. For the NLP analysis, titles in other languages were translated using Google Translate (see Koschnick, 2025).

Lastly, to further assign sub-classes for the difference-in-differences approach in section 8, we adopt an unsupervised clustering approach. The key priority here is to create stable clusters that are defined by the characteristics of the embedding space and are stable under different parameter specifications.¹¹ The paper therefore uses an approach where clusters are assigned using *HDBSCAN*,¹² a density-based algorithm that identifies coherent semantic groups in an embedding space. The embedding space is generated from the historical *SteamBERT* model introduced in section 5. Appendix section F describes the approach in further detail. Using this approach, the paper identifies an average of 11 sub-clusters per topic. Appendix table 17 reports the number of clusters with tf-idf-labels assigned to ease the interpretation of the cluster groups.

3.2 Patent data

Beyond technical publications, a significant amount of societal technical knowledge is also stored in patents. To provide short summaries of patents that parallel the titles from the ESTC, the

¹¹In contrast to e.g. *kmeans* which assumes spherical and evenly sized clusters and yields highly different categories when specifying different numbers of clusters.

¹²Hierarchical Density-Based Spatial Clustering of Applications with Noise.

paper introduces a novel data source, the *Chronological Index of Patents Applied for and Patents Granted* compiled by Bennett Woodcroft in 1854a. It covers the full time period of 1617–1852.¹³ The paper follows the literature (Billington and Hanna, 2020; Billington, 2021) by excluding patents before 1700 as some of these are closer to monopoly grants than modern “patents for invention”. Patents have an average word count of ca. 25 words (see appendix figure 16). To extract the technical content from the *Chronological Index*, the paper uses an LLM-based approach to exclude legal or procedural text passages.

A typical example of a patent short description title reads:

“Machine which performs its operacions either by fire or fall of water, or both together, & the friction is thereby reduced SO as to have no solid bodys to rub but the injecting vapour & water cocks or sluices” (William Blakey, 1766)

Note that this paper necessarily has to omit valuable inventions that were kept secret instead of being patented (Moser, 2005).

3.3 The *Lexicon technicum*

John Harris’s *Lexicon technicum* (1704) was the first scientific and technological encyclopedia in Britain (Kafker, 1981a; Bradshaw, 1981b). John Harris was an early exponent of Newtonian physics with a special interest in mathematics that led him to be elected a fellow to the Royal Society (Bradshaw, 1981b). His scientific interest is reflected in the topics covered in the *Lexicon technicum*. Out of all 1201 entries, 687 were on mathematics, 242 on astronomy, and 60 each on applied physics and chemistry. In contrast, his coverage of the practical arts was relatively short, with only 152 topics on prescriptive knowledge (the trades, scientific instruments, navigation, and agriculture, see appendix table 11). The *Lexicon technicum* would later become one of the inspirations for Diderot’s and d’Alembert’s *Encyclopédie* (Kafker, 1981b; Gaukroger, 2020) (see also Squicciarini and Voigtländer, 2015). Like Diderot’s and d’Alembert’s *Encyclopédie*, it significantly lowered access costs to knowledge; in the case of *Lexicon technicum* mainly for propositional knowledge (see also Bradshaw, 1981b).

The *Lexicon technicum* was published in multiple editions. To capture the original shock to knowledge, this paper uses the original 1704 edition. It includes 13,024 individual entries with an average length of 251 words. A typical (but interesting) entry from the *Lexicon technicum* reads:

¹³The description of patents is similar to Woodcroft’s *Subject-matter index of patents of invention* (1854a) used by Billington (2019) and Billington and Hanna (2020), but has the advantage of containing the original full summaries, while the *Subject-matter index* only includes extracts and summaries from the original source. Compare e.g. patent number 1812, listed as “engine for lessening the consumption of steam and fuel in steam or fire engines, and gaining a considerable effect in time & force” in the *Chronological Index* (Woodcroft, 1854a) and listed as “Engine for saving fuel in steam-engines” in the *Subject-matter index* (Woodcroft, 1854b).

“PTOLEMAICK System of the Heavens, was that invented by Ptolemy; in which he supposes the Earth immovable any way in the Centre of the Universe, round about which the Moon first moves in a Circle; next her Mercury, then Venus above whom moves the Sun, then Mars; above him Jupiter, and last of all Saturn, all in the Zodiac from West to East. Above Saturn he places the Sphere of the fixed Stars, which he supposes to move slowly also from East to West, on the Poles of the Ecliptick. While the fixed Stars themselves, and all the Planets, move from East to West on the Poles of the Equator, in the space of a Natural Day or 24 Hours. This Vulgar System of Astronomy, (in which I omit to mention the Epicycles and Deferents, &c. with which they endeavoured to solve the Phenomena which did almost all of them contradict this Scheme) was plainly overturned and refuted as soon as ever the use of the Telescope acquainted us with the Phases of Venus and Mercury ; for from thence it was apparent, that their Orbits included the Sun, and therefore by degrees it came to be quite diffused, and consequently I shall say no more of it.(. . .)”

To assign the entries from the *Lexicon technicum* to the same subject classes as in the ESTC, the paper uses the same BERT model from Koschnick (2025). Next, to assign the same sub-classes, the paper first calculates embeddings for all *Lexicon technicum* entries using the same *SteamBERT* model as for the ESTC. Then *Lexicon technicum* entries are assigned to the ESTC sub-topics whose centroid embedding exhibits the highest cosine similarity. Hence, both the ESTC and *Lexicon technicum* follow the same classification system for subject and sub-classes, and thereby allow for direct comparisons in the difference-in-differences approach in section 8.

3.4 Defining propositional and prescriptive knowledge

Conceptual classification Mokyr (2002) defines *propositional knowledge* (Ω) as the set of knowledge that describes *how nature functions*. This includes laws of nature as formulated in science but also a collection of all empirical data about nature.¹⁴ Next, *prescriptive knowledge* (λ) is the set of knowledge that describes *how to change nature* (Mokyr, 2002). It therefore encompasses all knowledge about technologies, whether in industry, trades, or agriculture.

The distinction is similar to science vs. technology. Indeed, propositional knowledge contains all knowledge from science.¹⁵ Yet, it also includes other activities aimed at collecting, describing, and classifying natural phenomena (see Mokyr, 2002, pp. 4–6). Propositional knowledge is

¹⁴Note that according to Mokyr (2002) propositional knowledge excludes knowledge that is about interactions between humans, e.g. the arts, literature, or the social sciences.

¹⁵We should note that *scientific practice* relies on a significant amount of prescriptive knowledge, e.g. how to operate technical instruments or how to set up experiments.

therefore wider than science. Likewise, prescriptive knowledge contains the widest array of possibilities of how to interact or change nature, e.g. building machines or instruments, useful ways of navigating a ship, as well as new ways of breeding cattle.

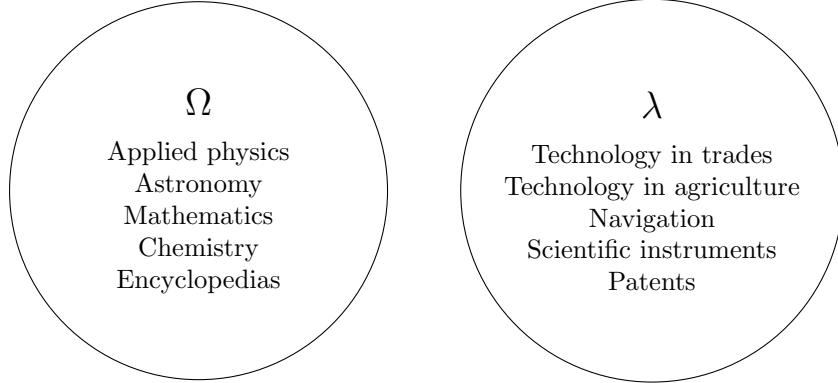


FIGURE 2: Assigning subject classes to propositional (Ω) and prescriptive knowledge (λ)

Notes: The figure reports the ESTC subject classes that form the set of propositional and prescriptive knowledge. Patents are added [Nuvolari and Tartari \(2011\)](#) and only available post 1700.

For the purpose of this paper, we aim to identify the set of propositional and prescriptive knowledge that was relevant for the Industrial Enlightenment and Industrial Revolution. As we use these categories in the downstream analysis of estimating feedback loops between related fields, we add two additional inclusion criteria: A) We require that spillovers should have been relevant to all receiving fields, and B) that the fields themselves should have been clearly distinct. For example, we plausibly expect that e.g. discoveries in *applied physics* $\in \Omega$ to have been relevant to *technology in trades* $\in \lambda$ while also being distinct fields at the same time. On the other hand, we would not expect knowledge in *medicine* $\in \Omega$ to have been relevant to *technology in trades*. Hence, one implication of this criteria is that, for this paper, we focus on the hard sciences and exclude the life sciences.¹⁶

Overall, figure 2 shows the selected subject classes of propositional and prescriptive knowledge. They include the core fields of hard science and technical subjects. Given that encyclopedias were an important means for accessing society's stock of knowledge ([Squicciarini and Voigtländer, 2015](#)), the paper further includes encyclopedic works within the group of propositional knowledge.¹⁷ Additionally, we add patents to λ as another important source for technological instructions. Here we caution that eighteenth century patents were public in principle but access to patent specifications was costly and uneven in practice ([MacLeod, 1989](#)).¹⁸ Therefore,

¹⁶This decision is purely practical and follows a narrow interpretation of the Industrial Revolution with mechanical innovation at its heart. A study of the existence of feedback loops in the live-sciences would be a promising route for future research.

¹⁷Encyclopedias were furthermore an important source for what ([Mokyr, 2002](#), p. 5) terms as technological science that includes e.g. knowledge about the properties of materials or knowledge about simple chemical reactions.

¹⁸First, before 1734 patent specifications were not requested for the registration of a patent, although the

in the following baseline specifications, we exclude patents from the construction of the spillover index. Here we assume that the access cost to patents was much higher than to printed works from the ESTC. Appendix sections G.8 and G.9 report robustness to either including patents in the spillover measure or excluding patents altogether.¹⁹ Appendix table 7 presents summary statistics for the subject classes of Ω and λ as shown in figure 2. Overall, the group of Ω includes 6,266 titles and the group of λ includes 4,784 titles between 1600 and 1800.

Time trends

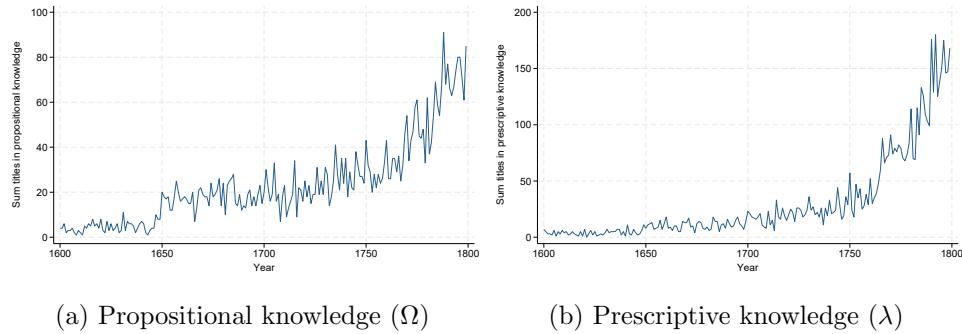


FIGURE 3: Propositional (Ω) and prescriptive knowledge (λ) over time

Notes: Propositional knowledge is defined as the set of titles in the fields of *applied physics, astronomy, mathematics, chemistry, and encyclopedias*. Prescriptive knowledge is defined as the set of titles in the fields of *technical publications, navigation, scientific instruments, and patents*.

Figure 3 plots trends in prescriptive and propositional knowledge over time. Notably, both measures closely track the timing of the Industrial Revolution (Clark, 2007; Allen, 2009; Bouscasse, Nakamura and Steinsson, 2025). We can see that the middle of the eighteenth century witnessed the transition from linear to exponential growth in both Ω and λ . Note that the exponential growth is especially pronounced for λ . However, even before that, propositional and prescriptive knowledge started to slowly grow post 1700. For comparison, figure 4 plots the number of all English publications over time. Appendix figure 18 further reports the share of propositional and prescriptive knowledge over time. It shows that the share of prescriptive knowledge in Britain's knowledge production began to continuously increase from the early eighteenth century.

Lastly, there are some salient patterns in the seventeenth century that need to be further

practice became more common during the early eighteenth century (MacLeod, 1989, p. 49). After 1734, it became standard practice to require written patent specifications for registration (ibid), after which specifications were randomly enrolled at any of three separate offices (MacLeod, 1991). Access was only granted if one knew the name of the patentee, but not provided for general technology classes (ibid). Important specifications were only released in print as late as 1794 through *The Repertory of Arts and Manufactures* (ibid). Also, note that the British patent system operated purely through registration. Therefore, neither the functioning of the invention nor the completeness of patent specifications for replicability were extensively checked. Altogether, access to patents was significantly more costly and uneven than access to books that could circulate in the country.

¹⁹Also note that since patents are only available after 1700, using patents as part of the spillover measure would introduce a discontinuity.

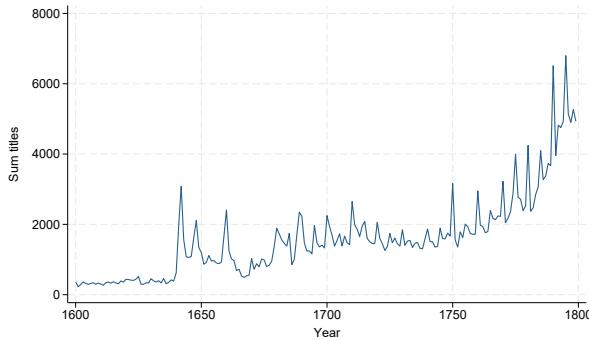


FIGURE 4: All English publications over time

Notes: The figure plots the number of all English publications over time. Publications are composed of all printed titles listed in the ESTC as well as patents.

discussed. First, we see a spike in all English publications during the period of the English Civil War, 1642–1649, and the interregnum 1649–1660.²⁰ We can also notice a jump in propositional publications post 1649. The jump likely reflects the beginning of the English Scientific Revolution in the 1650s that witnessed the invisible college as the origin of the Royal Society as well as the experimental philosophical club at the University of Oxford.²¹ Additionally, the Civil War, 1642–1649 might have either suppressed or delayed scientific publications.

Therefore, given a) the distortionary effect of the Civil War, and b) the very low number of publications pre 1650, we decided to omit all pre-1650 publications from the main analyses in section 7. Section 7 reports robustness to including this period.

4 Measuring knowledge spillovers and innovation

4.1 Conceptual framework

Innovation

How can we measure innovation and knowledge spillovers in historical data? Generally, citations are not available for the seventeenth and eighteenth century. Therefore, we have to rely on the information contained in the text data itself that can be quantified using recent advances in natural language processing.

The innovation and knowledge spillover measures introduced in this paper exploit the temporal dimension of innovation. They make use of the fundamental insight from Kelly et al. (2021) that states:

A title is innovative if:

²⁰On this, see Raymond (2003, pp. 163–165).

²¹For further studies on science at the universities, see Koschnick (2025). Merton (1938) argued that the Parliamentarian victory and Puritanism lay at the origin of the beginning of the Scientific Revolution in England. However, this has been controversially discussed in the historical literature Cohen (1994).

1. It is similar to publications in the future in its field
2. It is dissimilar to publications in the past in its field

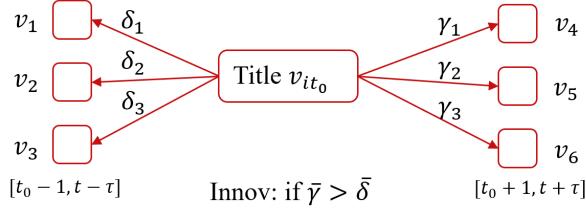


FIGURE 5: Calculating the innovation index

Notes: A title, v_{it} is innovative if its average similarity to titles in the future, $\bar{\gamma}$, is larger than its average similarity to titles in the past, $\bar{\delta}$.

Altogether, a publication is innovative if it is both *impactful* (1) and *novel* (2). We can formalize this by dividing a title's forward similarity (1) by its backward similarity (2). Concretely, for a set of titles S , and title $v_{it_0} \in S$, we calculate:

$$\text{Forward similarity}_{it_0} = f^k(v_{it_0}, \mathcal{F}_{t_0, \tau}) \quad (1)$$

$$\text{Backward similarity}_{it_0} = f^k(v_{it_0}, \mathcal{B}_{t_0, \tau}) \quad (2)$$

$$\text{Innovation}_{it_0} = \frac{f^k(v_{it_0}, \mathcal{F}_{t_0, \tau})}{f^k(v_{it_0}, \mathcal{B}_{t_0, \tau})}, \quad (3)$$

where $f^k(\cdot)$ is a function of similarity to be defined in section 4.2. Intuitively, it measures cosine similarity in the embedding space between the k -nearest titles. \mathcal{F} is a set of titles in the future, $\mathcal{F}_{t_0, \tau} := \{v_{jt} \in S \mid t_0 + 1 \leq t \leq t_0 + \tau\}$, and \mathcal{B} is a set of titles in the past, $\mathcal{B}_{t_0, \tau} := \{v_{jt} \in S \mid t_0 - \tau \leq t \leq t_0 - 1\}$. Figure 5 illustrates this logic. Innovation is greater than 1 if the average of all γ similarities is greater than the average of all δ similarities. Overall, by calculating title v_i 's similarity to the future and past, we can compare its *impact* (1) and its *novelty* (2). Novelty is a title's inverse backward similarity and impact is a title's forward similarity. Multiplying it yields equation 3.²²

Creating knowledge spillovers

We can easily extend the innovation concept to knowledge spillovers: Let $A, B \subset \mathcal{S}$ be two subject classes with $A \neq B$. Now, a title in field A creates a knowledge spillover into field B if:

1. It is similar to publications in the future in field B
2. It is dissimilar to publications in the past in field B

²²See Koschnick (2025) for a similar BERT-based innovation index. However, the Koschnick (2025) model differs in the use of the similarity function. It further uses a BERT model trained on contemporary data to derive similarities.

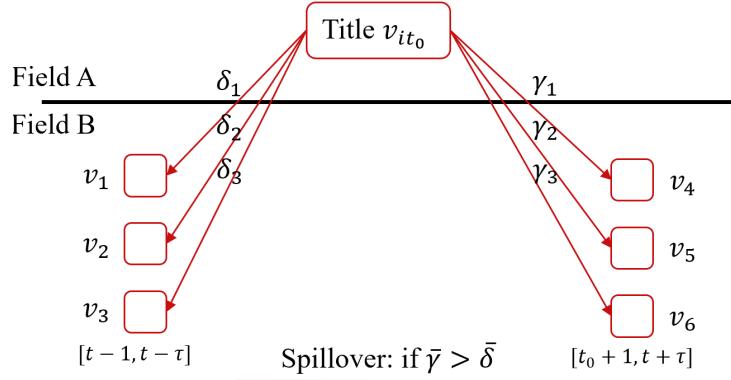


FIGURE 6: Calculating spillovers between fields

Notes: A title, v_{it} , in field A creates a spillover into field B if its average similarity to titles in the future of B , $\bar{\gamma}$, is larger than its average similarity to titles in the past of B , $\bar{\delta}$.

Figure 6 illustrates this logic. In the same way that title v_{it_0} shifted its own field in figure 5, title $v_{it_0} \in A$ now shifts field B . Thus, title $v_{it_0} \in A$ creates a spillover if it is closer to v_4, v_5, v_6 in the future of field B than to v_1, v_2, v_3 in the past of field B . We can formalize this as:

$$\text{Spillover}_{i,A \rightarrow B} = \frac{f^k(v_{it_0}, \mathcal{F}_{t_0,\tau}^B)}{f^k(v_{it_0}, \mathcal{B}_{t_0,\tau}^B)}, \quad (4)$$

where $\mathcal{F}_{t,\tau}^B$ is the forward pool in field B , $\mathcal{F}_{t_0,\tau}^B := \{v_{jt} \in B \mid t_0 + 1 \leq t \leq t_0 + \tau\}$, and $\mathcal{B}_{t,\tau}^B$ is the backwards pool in field B , $\mathcal{B}_{t_0,\tau}^B := \{v_{jt} \in B \mid t_0 - \tau \leq t \leq t_0 - 1\}$. Figure 6 illustrates this logic.

Receiving knowledge spillovers

Next, we want to define when a title in field B received a spillover from field A . The *received spillover measure* follows the same logic as the previous spillover measure from equation 4, which measures whether a title in field A created a spillover into field B . For this, we compared a title in field A to its backward and forward similarity in field B . Now, by parallel construction, to measure whether a title in field B received a spillover, we need to compare the receiving title to the past of field A and then compare the past of field B to the past of field A .

Figure 7 illustrates this logic. The figure is constructed parallel to the previous figure 6. However, instead of measuring whether v_{it_0} created a spillover, we measure whether title v_4 received a spillover. For this, we first calculate the similarity of v_4 to title v_{it_0} , γ_1 . Then we calculate the similarity of the past of field B to the title potentially creating a spillover in field A , v_{it_0} . This yields similarities $\delta_1, \delta_2, \delta_3$. Note that the signal in the measure becomes stronger when using the ρ -closest titles to v_4 for calculating the similarity between the past of field B and v_{it_0} .²³ Here we assume that the ρ -closest titles are v_1, v_2, v_3 . Now, we can calculate the

²³Note that for the baseline specification, we conveniently set $k = \rho$ where f^k is identical to f^ρ , thereby creating symmetry throughout all measures. However, in appendix G.2 we relax this assumption and show robustness to

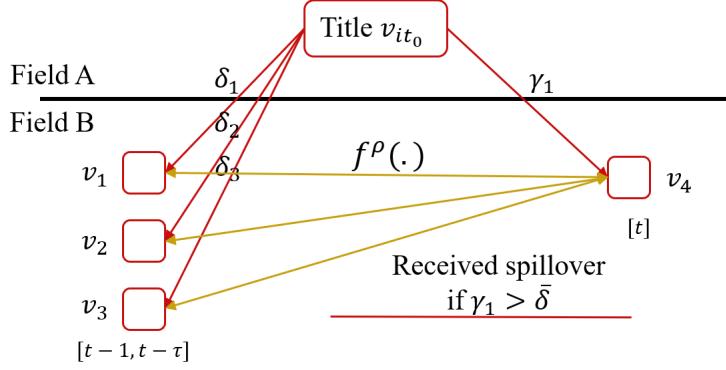


FIGURE 7: Calculating receiving knowledge spillovers (1)

Notes: The figure illustrates the logic behind title v_4 in field B receiving a spillover from title v_{it_0} in field A . It shows that it received a spillover if the similarity between v_4 and v_{it_0} was larger than the similarity between v_1, v_2, v_3 and title v_{it_0} . Similarity function f^ρ performs the similarity comparison for the ρ most similar titles to v_4 .

received spillover measure: If γ_1 is greater than $\bar{\delta}$, it indicates that title v_4 is closer to the spillover-creating title v_{it_0} than a close counterfactual in the past of its field. This implies that title v_4 *received a spillover*.

Figure 8 illustrates this logic when extending it to all titles in field A . Now, by parallel construction, title v_4 received a spillover if it was more similar to $\mathcal{B}_{t\tau}^A$ than $\mathcal{B}_{t\tau}^B$ was similar to $\mathcal{B}_{t\tau}^A$.

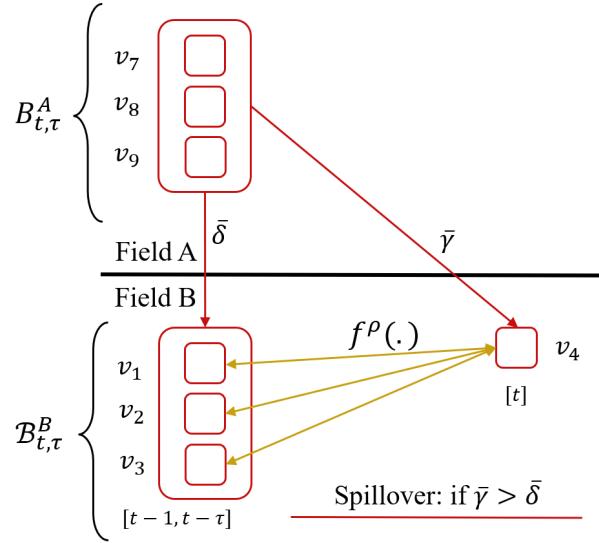


FIGURE 8: Calculating receiving knowledge spillovers (2)

Notes: The figure illustrates the calculation of the *received spillover* measure from equation 5. It extends the illustration from figure 7 by extending the previous intuition to the full field of A. It shows that title v_4 received a spillover if the average similarity $\bar{\gamma}$ of v_4 to $\mathcal{B}_{t\tau}^A$ is larger than the average similarity $\bar{\delta}$ between $\mathcal{B}_{t\tau}^A$ and $\mathcal{B}_{t\tau}^B$. This similarity is calculated by using the ρ closest titles in $\mathcal{B}_{t\tau}^B$ to v_4 ($f^\rho(\cdot)$).

different values of k and ρ .

We can formalize the received spillover measure for title v_{it} in field B as:

$$\text{Received spillover}_{A \rightarrow B}(v_{it}) := \frac{f^k(v_{it}, \mathcal{B}_{t,\tau}^A)}{f^\rho(\mathcal{B}_{t,\tau}^B, \mathcal{B}_{t,\tau}^A)}. \quad (5)$$

where $\mathcal{B}_{t,\tau}^B$ captures the backward pool of titles in field B , defined as $\mathcal{B}_{t,\tau}^B := \{v_{jt} \in A \mid t_0 - \tau \leq t \leq t_0 - 1\}$. Intuitively, the *received spillover* measure from equation 5 captures whether title v_{it} in B is closer to the past of spillover field A than counterfactual titles in the past of B were to the past of field A .

After having defined the concepts of a title i) creating innovation, ii) creating spillovers, and iii) receiving spillovers, the next section will define the *similarity* function, $f(\cdot)$, between texts.

4.2 Defining the similarity function

For calculating similarity, we rely on the semantic content of a historically trained BERT model, introduced in section 5. This model yields semantic text-vector representations in a 512×768 -dimensional embedding space. Let v_{it} be a single title embedding. Let $\text{sim}(x, y) = \frac{x \cdot y}{\|x\| \|y\|}$ be the cosine similarity between two titles in the embedding space $x, y \in \mathbb{R}^d$. Lastly, let $Y = \{y_1, \dots, y_m\} \in \mathbb{R}^d$ be a set of embeddings. Then, we define similarity as the average cosine similarity between title i and the top- k most similar titles:

$$f(v_{it}, Y; k) = \frac{1}{k} \sum_{y \in \text{Top}_k \{\text{sim}(v_{it}, y') : y' \in Y\}} \text{sim}(v_{it}, y), \quad k := \min\{k, |Y|\} \quad (6)$$

Calculating similarity only to the most similar k titles in the comparison group directly addresses two key biases in the calculation of spillovers:

First, the spillover and innovation measure can be sensitive to the inclusion of irrelevant titles. Adding irrelevant titles (e.g. imagine the emergence of a new sub-field) decreases average forward similarity. Yet, by only calculating similarity to the k -most similar titles, we ensure that irrelevant titles are not part of the comparison group as long as k is sufficiently small.

Second, the spillover and innovation measure can also be sensitive to titles assigned to the wrong subject class. By definition, titles from another subject class are highly dissimilar to other titles in this class and therefore could bias the spillover and innovation measures. By only calculating similarity to the k -most similar titles, we can exclude wrongly assigned titles from the comparison group.

For this paper, we choose a top- k value of 20. Appendix section G.2 shows robustness to different values of k . Appendix section E further shows that the k-top similarity measure (equation 6) outperforms the average similarity measure in predicting patent citations using the innovation measure from equation 3.

The next section introduces a historical BERT model for the calculation of the embedding space.

5 A BERT similarity model for historical science and patent data

To measure text similarity, the paper draws on context-sensitive BERT models (Devlin et al., 2019) that yield a multidimensional embedding space representation of the semantic content of a title. However, a key concern is that models trained on modern text data may induce two forms of anachronistic bias:

1. Bias from shifts in word-object representations in language. For example, a model might miss that historically a “fire engine” would refer to a steam engine or that “philosophical instruments” would refer to any scientific instruments.
2. Bias from a modern representation of concepts within the embedding space. For example, models trained on modern data would position “phlogiston” or “ether” as semantically distant from other scientific terms, thereby failing to reproduce the internal logic of historical scientific texts, in which “phlogiston” or “ether” were treated as equally valid scientific concepts.

To directly address these biases, the paper draws on a *MacBERTh* model that was pre-trained on a historical corpus (Manjavacas Arévalo and Fonteyn, 2021). The main weight of the corpus lies on the early modern period, 1600–1700, with *Early English Books Online* (EEBO) as one of its main sources. Yet, given a general lack of large-scale full-text data for the eighteenth century, the authors decided to extend their corpus until 1950. Nonetheless, the pre-1800 period still captures ca. 70% of the entire dataset (Manjavacas Arévalo and Fonteyn, 2021, p. 25), making *MacBERTh* a good candidate model for embedding early modern text data.

To further align the *MacBERTh* model with seventeenth and eighteenth century publications, this paper further fine-tunes the *MacBERTh* model on the ESTC and patent corpus. This step also addresses problems in using base models, such as *MacBERTh* for semantic comparison that arise from a) CLS layers trained for classification tasks and b) embedding space anisotropy. The result is a new fine-tuned *MacBERTh* model that we name *SteamBERTh*.

5.1 SteamBERTh

The paper employs a *simple contrastive learning of sentence embeddings* approach (SimCSE) (Gao, Yao and Chen, 2021) to fine-tune the *MacBERTh* model on a set of exemplary texts. The training data covers the universe of short patent descriptions from Woodcroft (1854a) and

the universe of titles within *applied physics, astronomy, chemistry, mathematics and technical publications in trades and agriculture, navigation, and scientific instruments* from the ESTC. Appendix table 6–7 provide an overview of the distribution of titles within these subsets. Appendix section B describes the concrete fine-tuning procedure and hyperparameters.

To evaluate the fine-tuning operation, appendix section C compares and discusses the key embedding space indicators between the base and fine-tuned models, such as *mean nearest neighbor cosine similarity, pc1-variance, rank-overlap, and UMAP embedding space projections* by subject classes and industries. We can see that the embedding space of the base model shows severe signs of anisotropy, while the fine-tune model is significantly more isotropic. Additionally, we see that the fine-tuned *SteamBERTh* model shows higher semantic coherence when grouping subject classes and industry groups in the embedding space.

This first set of descriptive statistics already indicates that the fine-tuned model significantly improves the performance of the base MacBERTh model. However, to assess the quality of the model as an input for historical innovation and spillover measures, we still need to validate the model with real economic data: In the next section, we assess the relationship between the innovation measure from equation 3 with similarity measures from the fine-tuned *SteamBERTh* model and historical patent citations.

5.2 Validation

Next, we validate the usefulness of *SteamBERTh* embeddings for historical innovation and knowledge spillover measures. Since we do not have historical quantitative measures of knowledge spillovers, we focus on validating the innovation measure that is conceptually closely aligned to knowledge spillovers. For this, we compare historical patent citations from Nuvolari and Tartari (2011) with the *SteamBERTh* based innovation measure from equation 3.

Figure 9 presents the relationship between the *SteamBERTh*-based innovation measure and historical patent citation counts residualized for year fixed effects. Both measures are transformed using the natural logarithm. We observe a within-year elasticity of 1.82 for the time period 1720–1799 and 1.74 for 1800–1841. The association is reassuring and of a similar magnitude to the association between different modern measures of patent quality, e.g. patent citations and patent value (Kelly et al., 2021). Therefore, the *SteamBERTh*-based innovation measure appears to successfully capture a relevant dimension of the innovativeness of historical patents.

Finally, appendix section D explores whether *SteamBERTh* constitutes an improvement over standard BERT models trained on modern data. It runs a set of regressions with different BERT models as an input into the innovation measure from equation 3 as competing predictors for patent citations. We find that the *SteamBERTh* based innovation measure has the highest R^2 and strongly outperforms the other BERT models in a horse race specification for the time

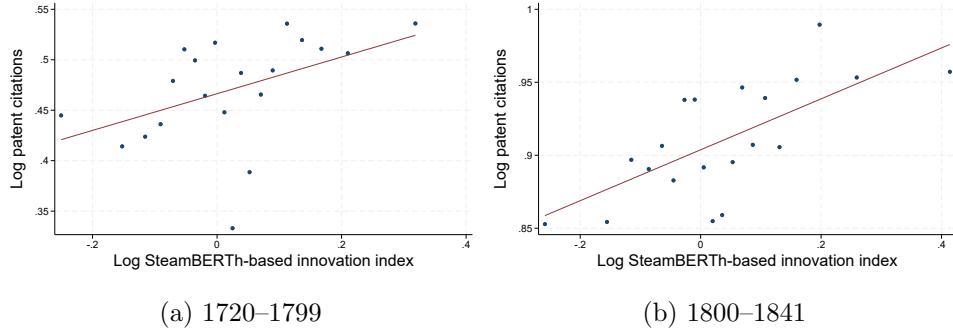


FIGURE 9: SteamBERTh-based innovation index and patent citations — binscatter plot, residualized for year fixed effects

Notes: The figure presents a binscatter plot for the SteamBERTh-based innovation measure and patent citations residualized for year fixed effects. Patent citations are taken from Nuvolari and Tartari (2011). The innovation measure is calculated for the full sample of patents, 1700–1851. Both measures are transformed using the natural logarithm. Because the innovation measure mechanically needs a sufficiently large pre- and post-period for comparison, it is calculated for the period 1720–1841. The slope is 1.82 for the time period 1720–1799 and 1.74 for 1800–1841. $N = 1,922$ for 1720–1799 and $N = 6,684$ for 1800–1841. Please refer to appendix section D for formal regression results with additional control variables, such as the length of text.

period 1720–1800. We further document that the other BERT models are performing slightly better on nineteenth century patent citation data, 1800–1849. This is expected, the fine-tuning procedure offers time-period specific performance gains at the loss of generality for more recent time periods.

As a final validation exercise for the usefulness of *SteamBERTh* for scientific publications, appendix table 15 lists the 20 most innovative ESTC titles in *applied physics*. Reassuringly, the list is composed of the most famous scientists of the day, including Robert Boyle, Thomas Hobbes, Johannes Kepler, René Descartes, Gassendi, and Robert Hooke. It further includes a large number of seminal and groundbreaking works in the history of physics, including Francis Bacon’s *Great Instauration* as well as Robert Boyle’s experiments with the air pump (see e.g. Shapin and Schaffer, 1985).²⁴

Overall, we have shown that the *SteamBERTh*-based innovation measure from equation 3 successfully captures key dimensions of historical patents’ innovativeness pre-1800. We argue that, by extension, the *SteamBERTh*-based spillover measures are also likely to perform well.

²⁴When going through this list note the following. By construction, the innovation measure identifies prominent works whose topics were novel and influential. This does not necessarily mean that the positions argued in these works were either correct or influential. The current approach would not be able to distinguish between writers arguing on the same subject, topic, and content but having differential views, nor to identify which writer won the debate. E.g. appendix table 15 includes a title by Alexander Ross where he disapprovingly comments on Sir Kenelm Digby’s natural philosophy. The impact of this title is most likely due to Sir Kenelm Digby’s influence on natural science rather than Alexander Ross’s position. Also note that the ESTC sometimes contains different versions of the same publication with differing titles. Reassuringly, these are treated consistently in the innovation measure.

6 Empirical framework

The previous section has established a new method to measure innovation and knowledge spillovers from historical text data. It has further validated the innovation measure using historical patent citations. Next, this section develops a framework to estimate the relationship between knowledge spillovers and innovation at the title level. Seeing a positive relationship between knowledge spillovers and innovation would be a key characteristic of feedback loop processes between propositional and prescriptive knowledge:

According to Mokyr, feedback loop processes are characterized by the presence of *innovation-inducing* knowledge spillovers between the two types of knowledge (Mokyr, 2002). We would expect that “growth in one increases the marginal product of the other” (Mokyr, 2002, p. 21).²⁵ Here, an increased marginal product corresponds to innovation as captured through the innovation measure from equation 3. Likewise, we can directly capture spillovers from new knowledge through the received spillover measure from equation 5. Hence, with these two measures, it becomes possible to estimate whether titles that received knowledge spillovers were more likely to be innovative. If this was the case, growth in either Ω or λ would have increased the marginal product of the other, thereby creating the foundations for a positive feedback loop.

We formally test the association between a title’s innovativeness and the strength of knowledge spillovers from $\Omega \rightarrow \lambda$ and $\lambda \rightarrow \Omega$ in the following model:

$$\text{Innovation}_{ijt}^A = \sum_{p=1600-1619}^{1760-1789} (\beta_p \cdot \text{Received spillover}_{B \rightarrow A}(v_{ijt}) \times \eta_p) + \mathbf{X}'_{ijt} \zeta + \gamma_j + \alpha_t + \varepsilon_{ijt} \quad (7)$$

Here, the dependent variable, $\text{Innovation}_{ijt}^A$, is the innovation index from equation 3 for title i in field j at time t in knowledge field A . The main explanatory variable, $\text{Received spillover}_{B \rightarrow A}(v_{ijt})$, captures the strength of knowledge spillovers that title i in time t received from knowledge field B . The measure is defined in equation 5. We use it to capture spillovers between λ and Ω as well as Ω and λ . The fields contained in Ω are *applied physics, astronomy, mathematics, chemistry, and encyclopedias*. The fields contained in λ are *technical publications, navigation, scientific instruments, and patents*. To estimate the strength of the association of these spillovers and innovation over time, the spillover coefficient is interacted with twenty-year period dummy η_p . Both the dependent and independent variables are transformed using the natural logarithm. \mathbf{X}'_{ijt} is a vector of control variables, incl. a title’s level and quadratic word count. Lastly, γ_j and α_t capture subject and year fixed effects.

Note that mechanically the calculation of the innovation index from equation 3 for title i in year t requires a pre- and post-period of length τ . Therefore, the ESTC sample is mechanically

²⁵(see also Milgrom, Qian and Roberts, 1991)

limited to $[1600 + \tau, 1800 - \tau]$. For this paper, we choose $\tau = 20$ years. Additionally, as argued in section 3.4, the pre 1660 period yields an insufficient amount of observations for reliable inference and is further contaminated by Civil War shocks. Hence, all main specifications start in 1660. Robustness to alternative τ and to including the full 1620–1780 period are shown in appendix section G.4 and G.7.

Next, we consider the prediction from Mokyr (2002) on the coefficients estimated in equation 7. According to Mokyr, it was the “Industrial Enlightenment” that changed how both types of knowledge interacted with each other, thereby “tipping the balance of the feedback mechanism from negative to positive” (Mokyr, 2002, p. 33). So, when did the “Industrial Enlightenment” begin? A concrete answer evades us, as it was a continuous process. Still, we can identify a set of important demarcation points. First, lecture series on Newtonian science had already started by the early 1700s (Stewart, 1992). Next, the publication of Voltaire’s *Letters Concerning the English Nation* in 1733 is often used to demarcate the beginning of the enlightenment (Wootton, 2015). Moreover, the 1750s and 1760s seem to have been a point of increased acceleration, marked by the publication of Diderot and d’Alembert’s *Encyclopédie* between 1751 and 1772 (Squicciarini and Voigtländer, 2015), the foundation of the Society for the Encouragement of Arts, Manufactures and Commerce in 1754 (Howes, 2020), the Lunar Society of Birmingham in 1765 (Schofield, 1963), and a series of economic societies across the continent (Stapelbroek and Marjanen, 2012; Cinnirella, Hornung and Koschnick, 2025). Taken together, we expect that the “Industrial Enlightenment” would have created innovation-inducing knowledge spillovers around the middle of the eighteenth century.

Overall, we expect to find:

1. A negative or neutral relationship between knowledge spillovers and innovation before the eighteenth century
2. A gradual “tipping of the balance” toward a positive relationship during the early eighteenth century
3. A positive relationship between knowledge spillovers and innovation by the middle of the eighteenth century

7 Empirical results

7.1 Knowledge spillovers and innovation

Figure 10 reports the main results from equation 7. Panel a) reports estimated coefficients for the relationship between spillovers from propositional to prescriptive knowledge, $\Omega \rightarrow \lambda$,

and innovation over time. Panel b) reports estimated coefficients for the relationship between spillovers from prescriptive to propositional knowledge, $\lambda \rightarrow \Omega$, and innovation over time.

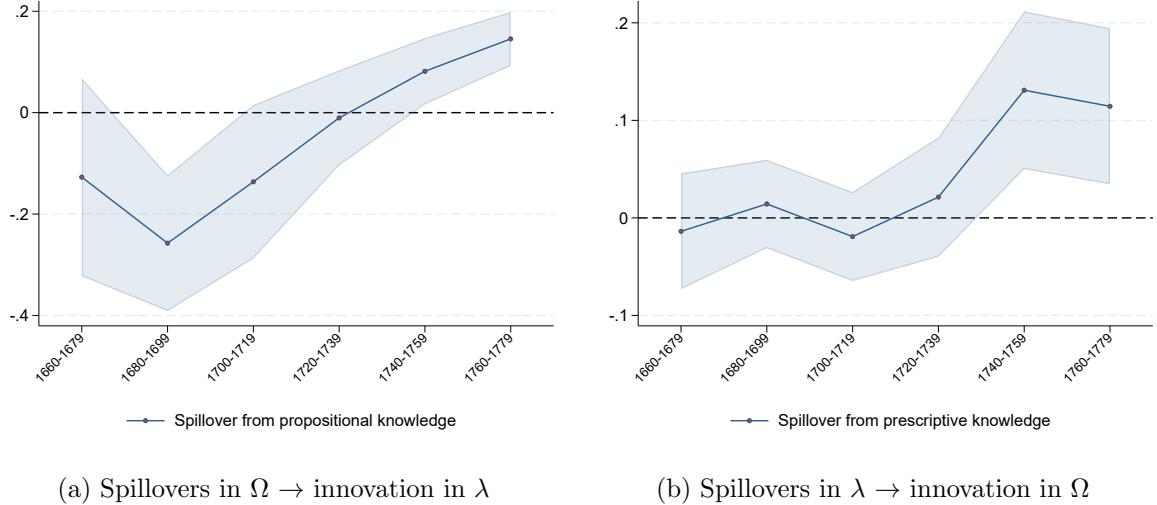


FIGURE 10: Spillovers from prescriptive (λ) and propositional knowledge (Ω) \rightarrow innovation

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The graph plots the coefficients from interacting the received spillover index from equation 5 with twenty-year time periods. The model further controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Standard errors clustered at the publication year level.

In panel a), we first find a significant negative relationship between knowledge spillovers from propositional (Ω) to prescriptive knowledge (λ) and innovation for the seventeenth century and 1700–1719. Next, we find that the relationship “flipped” in 1720–1739 and became positive in 1740–1759 and 1760–1779. For 1760–1779, we find that an increase in knowledge spillovers from Ω to λ by 100% was associated with a 10.4% increase in innovation.

In panel b), we first find a neutral or negative relationship between knowledge spillovers from prescriptive (λ) and propositional knowledge (Ω) in 1660–1679 and 1720–1739. Then in 1740–1759 and 1760–1779, the relationship turned positive. In 1760–1779, a 100% increase in knowledge spillovers from λ to Ω was associated with a 7.7% increase in innovation.

Altogether, this constitutes strong evidence of a negative or non-positive relationship between knowledge spillovers and innovation before the eighteenth century. As argued before in section 2, negative spillovers could have arisen from relying on the predictions of immature theories (A_1) or an underdeveloped and imprecise toolset (A_3). Then, as predicted from Mokyr (2002)'s theory, the relationship started to flip in the eighteenth century and became positive by 1740–1759 and 1760–1779. As argued in section 2, this could have been driven by a range of indirect channels (A_1-A_4 and B_1-B_3). A positive relationship between knowledge spillovers between both, propositional and prescriptive knowledge, would have initiated a feedback loop as growth in both types of knowledge would have increased the marginal product of the other (see Mokyr,

2002, p. 21). We further find that the transition towards a positive feedback loop took place shortly before or at the time of the beginning of the Industrial Revolution. Hence, the estimates from figure 10 correspond to the predictions from Mokyr (2002)'s thesis.

Patents and the real economy

Next, the paper specifically focuses on patents as a subset of prescriptive knowledge. Applying the model from equation 7 to patents, the paper estimates whether knowledge spillovers from propositional knowledge were associated with increased innovation in patents, and hence with direct changes in the real economy. The paper further presents results for spillovers from applied physics, capturing the applied mechanics channel most commonly associated with knowledge spillovers during the Industrial Revolution.

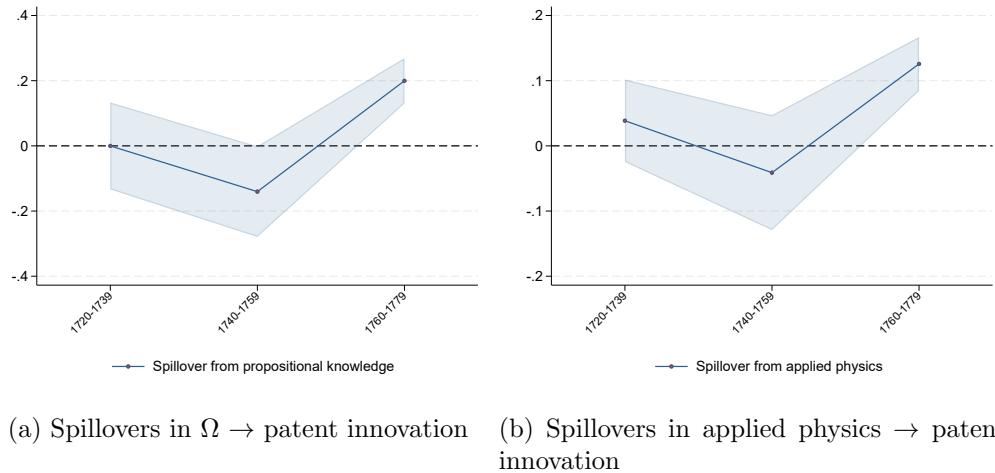


FIGURE 11: Spillovers from propositional knowledge (Ω) and applied physics \rightarrow patent innovation

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The graph plots the coefficients from interacting the received spillover index from equation 4 with twenty-year time periods. Panel a) shows results for spillovers from the full set of Ω . Panel b) shows results for spillovers from applied physics. Dependent and independent variables are transformed using the natural logarithm. The model controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Given that the patent data only starts in 1700 and the necessary comparison period of $[-\tau, \tau]$, $\tau = 20$ for the innovation index from equation 3, the model is only estimated on the post-1719 sample. Standard errors clustered at the publication year level.

Figure 11 presents the results. Panel a) presents results for spillovers from propositional knowledge (Ω) and panel b) presents results for spillovers from applied physics. Since the patent data only starts in 1700, we can only estimate effects on innovation post-1719.²⁶ In panel a), we find that the relationship between knowledge spillovers from Ω to patents and innovativeness was neutral in 1720–1739 and 1740–1759 and then became positive and significant in 1760–1779. In 1760–1779, an increase in spillovers from Ω by 100% was associated with an increase in innovativeness of 14.84%. For spillovers from applied physics in panel b), we find a similar, if

²⁶This is due to the necessary comparison period of $[-\tau, \tau]$, $\tau = 20$ for the innovation index from equation 3.

slightly smaller effect. Here, an increase in spillovers from applied physics was associated with an increase in innovativeness of 9.10%. Hence, we find that knowledge spillovers from propositional knowledge were also associated with productivity increases in the real economy, as captured through patents.

Next, we conduct a range of robustness tests. First, appendix figure 41 presents results for 10-year periods. Using patents further allows us to use patent citations as an alternative indicator for innovation. Appendix figure 42 shows results with patent citations from Nuvolari and Tartari (2011) as the dependent variable. Here, results for spillovers from propositional knowledge are insignificant, however, we find significant spillover effects post 1760 for spillovers from applied physics. Here, an increase in spillovers from applied physics by 100% is associated with an 8.54% increase in patent citations.²⁷ Note that the innovation index from equation 3 captures more variance, since 42% of all patents have never been cited. Therefore, regressions using patent citations might be underpowered. Finally, appendix section G.12 presents estimated spillover coefficients interacted by industry. We find that the relationship between spillovers from both Ω and applied physics were strongest for *engines, food, glass, instruments, mining, pottery, and ship* industries. Notably, we do not find a significant effect for *textiles*, potentially reflecting the importance of tweaking and tinkering in textile innovations (Mokyr, 1992; Ó Gráda, 2016).

Overall, this exercise shows that we find a similar relationship between knowledge spillovers from propositional knowledge and patent innovation as for the full set of prescriptive knowledge. Therefore, the patenting results indicate that changes in the knowledge economy were also closely associated with changes in the real economy.

Placebo spillovers from plausibly unrelated fields

The previous results have shown strong support for the Mokyrian hypothesis. Yet, we might be concerned that results might be confounded by general linguistic trends that were related to the innovativeness of a title. These include, for example, changes in linguistic style, sophisticated language, or references to places or events. To account for this challenge, we report results for placebo spillovers from plausibly unrelated fields. Placebo fields should have contained similar trends in language but should not have been practically relevant for knowledge production in the receiving fields.

In a first step, we define the set of plausibly unrelated fields as *drama, poetry, antiquities, amusements, foreign languages, printing, prophecies, stories, superstition, architecture, art, biographies, moral tales, state affairs, and travel descriptions*. We argue that these fields cover both a broad range of potentially confounding topics and writing styles, while also being irrelevant to knowledge production in propositional and prescriptive knowledge.²⁸ Figure 12 reports

²⁷Also note that patent citations also allow us to cover the full 1700–1800 period.

²⁸Note that we do not include religion in the list of plausibly unrelated fields given the findings from Almelhem et al. (2023) that demonstrate a structural change in relationship between science and religion throughout the

results.

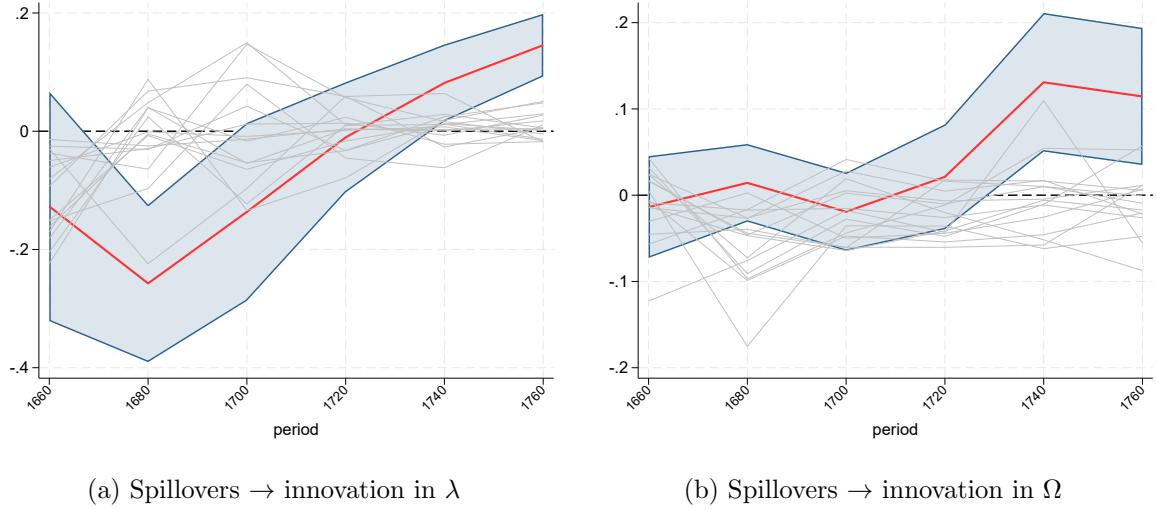


FIGURE 12: Placebo spillovers from plausibly unrelated fields

Notes: The figure reports placebo results from plausibly unrelated spillovers. These are defined as *drama*, *poetry*, *antiquities*, *amusements*, *foreign languages*, *printing*, *prophecies*, *stories*, *superstition*, *architecture*, *art*, *biographies*, *moral tales*, *state affairs*, and *travel descriptions*. Placebo coefficients are reported as gray lines, while the spillovers from propositional and prescriptive knowledge are reported as red lines. Coefficients estimated from equation 7. Panel a) shows the association between spillovers and innovation in prescriptive knowledge (λ) and panel b) shows the association between spillovers and innovation in propositional knowledge (Ω). Standard errors shown for spillovers from propositional and prescriptive knowledge and clustered at the publication year level.

We find that for none of the plausibly unrelated fields does the relationship between spillovers and innovativeness increase over time. Furthermore, none of the plausibly unrelated fields yields positive coefficients of the same size as the positive spillovers from $\Omega \rightarrow \lambda$ and $\lambda \rightarrow \Omega$ for the period 1740–1759 and 1760–1779. We can also use this information for statistical inference by computing Fisher-exact p-values using $p = \frac{1 + \sum_{j=1}^J \mathbf{1}(|\hat{\tau}_j| \geq |\hat{\tau}_{\text{prop/pres}}|)}{1+J}$ where J denotes the number of placebo coefficients. The p-value equals the share of placebo effects that are at least as extreme in absolute value as the estimated coefficient for propositional and prescriptive knowledge. The test is non-parametric and does not rely on standard errors or asymptotic approximations. Its validity rests on the exchangeability assumption—that under the null, propositional and prescriptive knowledge are indistinguishable from the placebo fields in the absence of a true spillover effect. For figure 12, this yields a Fisher exact p-value of 0.067.

Placebos from all fields

Given that the previous choice of plausibly unrelated placebos leaves some room for subjective judgment, we next include all 44 subject classes in the ESTC as placebos to demonstrate the robustness of the results. Note that this includes clearly related subject classes such as

seventeenth and eighteenth century. See also Hornung (2014) and Becker, Rubin and Woessmann (2024) for potential influences of religion or religious minorities on technological innovations. However, religion is included in the next placebo specification in figure 13.

medicine, *biology*, or *military* (see appendix 9 for a list of all subject classes).²⁹ Yet, as shown in figure 13, spillovers from propositional and prescriptive knowledge still dominate spillovers from *all other fields*.

In both graphs, none of the spillovers from all other fields is larger than the spillovers for propositional and prescriptive knowledge for 1760–1779. Fisher exact p-values are reported in table 1. The coefficients for the period 1760–1779 have a p-value of 0.022. For 1740–1759, there are, respectively, 2 and 1 placebo coefficients that are larger in absolute value than the original coefficients. Nevertheless, even in the presence of that, we can report p-values of 0.065 and 0.043 respectively. Furthermore note that the only positively larger coefficient is found in panel b) for Ω . Moreover, this is a spillover from *medicine* that could have been considered part of propositional knowledge, but was excluded as a life science (see section 3.4).³⁰

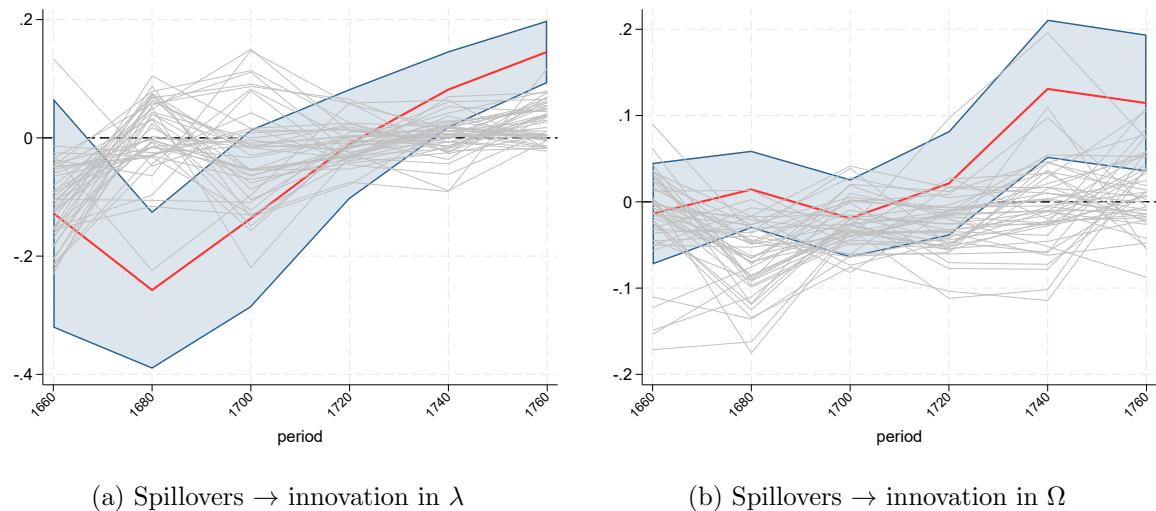


FIGURE 13: Placebo spillovers from all ESTC fields

Notes: The figure reports placebo results from spillovers in all ESTC that are not part of prescriptive or propositional knowledge. Placebo coefficients are reported as gray lines, while the spillovers from propositional and prescriptive knowledge are reported as red lines. Coefficients estimated from equation 7. Panel a) shows the association between spillovers and innovation in prescriptive knowledge (λ) and panel b) shows the association between spillovers and innovation in propositional knowledge (Ω). Standard errors shown for spillovers from propositional and prescriptive knowledge and clustered at the publication year level.

²⁹Other classes that are plausibly related include e.g. philosophy and political philosophy. Many writers such as Thomas Hobbes or David Hume were clearly responding to changes in contemporary science and actively referencing it. Likewise, the topics of *classical education* or *educational theory*, do not actively form a part of the hard sciences, but might plausibly reference them. Other topics such as *music* or *curiosities and wonders* might contain objects studied by science — concepts such as vibration, pitch, and resonance were actively studied during the Scientific Revolution and Isaac Newton treated both sound and colors analogously in his *Opticks*. Furthermore, other fields like *legal* or *administrative* might increase the value of e.g. patents through different channels.

³⁰As eighteenth century medicine was heavily drawing on innovations in other fields of propositional knowledge, such as e.g. discoveries from electricity or the chemical analysis of mineral waters, the existence of spillovers is not surprising.

TABLE 1: Fisher exact p-values from placebo tests

| Period | Original spill. coeff. | Exact p-value | Placebo spill. J | # as ex- treme |
|---|---------------------------|------------------|-----------------------|-------------------|
| <i>Panel A: Spillovers → innov. in λ</i> | | | | |
| 1740–1759 | 0.082 | 0.065 | 45 | 2 |
| 1760–1779 | 0.146 | 0.022 | 45 | 0 |
| <i>Panel B: Spillovers → innov. in Ω</i> | | | | |
| 1740–1759 | 0.131 | 0.043 | 45 | 1 |
| 1760–1779 | 0.114 | 0.022 | 45 | 0 |

Notes: The table reports Fisher-style exact p -values comparing the estimated science-to-technology spillover coefficients with placebo spillover coefficients from all other 44 fields that are neither a subset of propositional nor prescriptive knowledge. The p -value is calculated as $p = \frac{1 + \sum_{j=1}^J \mathbf{1}(|\hat{\tau}_j| \geq |\hat{\tau}_{\text{prop/pres}}|)}{1 + J}$, where J is the number of placebo coefficients. Panel a) reports statistics for spillovers into λ and panel b) reports statistics for spillovers into Ω corresponding to panel a) and b) in figure 13.

Altogether, the placebo results show that the relationship between spillovers from propositional and prescriptive knowledge cannot be explained by trends in language that also would have affected spillovers from other fields. The results from placebo spillovers from all fields in the ESTC further illustrates the importance of the propositional and prescriptive channel in comparison to other knowledge related channels that might impact innovation, e.g. originating from legal innovation or administrative practices.

Robustness tests

Lastly, the paper conducts a series of robustness tests, showing that results hold irrespective of parameter choices in the innovation and spillover function from equation 3 and 4. We further conduct a heterogeneity analysis for spillovers from different fields in Ω and λ and distinguish between micro- and macro-inventions. Lastly, we also account for field size compositional bias.

- Different parameters in innovation measure.** For the calculation of the innovation index, we compare each title to the k most similar forward and backward titles. The innovation index for the baseline results was calculated with $k = 20$. We assume that this corresponds to the size of individual topics/sub-fields that are relevant for comparison. To show that results are robust to other values of k , figure 28 and 29 show the baseline results from figure 10 for $k = 10$, $k = 20$, and $k = 30$. Appendix figure 30 and figure 31 further changes k and ρ from received spillover equation 5 simultaneously to $k, \rho = 10$, $k, \rho = 20$, and $k, \rho = 30$. Results are highly robust to changing the parameters for the innovation and received spillover index.
- Different length of time window for innovation and spillover index.** Additionally,

appendix G.3 reports robustness when changing the length of the time windows, t, τ , for the calculation of the innovation and received spillover index from equation 3 and 5. Note that changing the time window also changes the interpretation of the innovation index. Shorter time windows place more weight on titles with a short-run impact, while longer time windows place more weight on inventions with a long-run impact. Following Mokyr (1992), this corresponds to the distinction between micro- and macro-inventions. Also note that by construction, t determines the length of the sample, $[1600 + t, 1800 - t]$, for which the innovation index is defined. Figure 32 and 33 report results for $t, \tau = 10$, $t, \tau = 20$, and $t, \tau = 30$. We find that results are generally stronger for shorter t, τ , indicating that spillovers between Ω and λ mainly led to innovation in micro-inventions.

3. **Heterogeneity across outcome distribution** To test whether the results are driven by the upper quantiles of innovation or evenly distributed, we report results from a quantile regression in appendix figure 27. Coefficients appear relatively evenly distributed across the outcome distribution, although there is some evidence that coefficients are stronger at the tails of the distribution.
4. **Heterogeneity analysis by spillovers from different fields.** The framework from equation 7 also allows for separately estimating coefficients for spillovers from individual fields in Ω and λ . Appendix figure 35 and 36 report results. We find that positive spillovers at the end of the eighteenth century were strongest for *applied physics, astronomy, and chemistry* from Ω and *navigation, and scientific instruments* from λ .
5. **10-year coefficients.** To show that results are robust to shorter time intervals interacted with the spillover coefficients, appendix figure 34 also reports results for 10-year period coefficients.
6. **Accounting for compositional bias and full 1620–1780 period.** A potential concern is compositional bias from the changing number of titles per field over time. This holds especially for earlier periods with fewer titles per field and higher compositional variation. To account for this, appendix figure 37 reports results with additional subject \times year fixed effects. Trends do not change and the positive effects post-1740 remain significant. Figure 38 additionally reports robustness for including the full time period of 1620–1780. Again, the trends do not change. Note, however, that given the few underlying observations for the period 1620–1660 (see figure 3), results pre-1660 should be interpreted cautiously.
7. **Robustness to patent data.** As argued in section 3.4, access to patent specifications in the eighteenth century was costly, uneven, and sparingly used. Therefore, patents are

excluded from the calculation of the spillover measure for all baseline specifications.³¹ Figure 39 reports robustness to including patents into the spillover measure. Trends remain similar, although the 1760–1779 coefficient becomes insignificant. Likewise, we show robustness to entirely excluding patents from our definition of prescriptive knowledge. Results for $\Omega \rightarrow \lambda$ spillovers are reported in figure 40. Again, trends remain similar.

7.2 Mechanism

7.3 Upper-tail human capital

To examine the mechanism behind the previous results, we begin by asking what made knowledge spillovers possible in the first place. A key point is that authors' specialization alone was not sufficient. Incorporating knowledge spillovers required access across epistemic domains. For spillovers to occur, specialized artisans needed exposure to scholarly ideas, just as scholars depended on insights from artisanal techniques. Historically, these channels of exchange were limited.

Here, Mokyr (2002, 2016, 2021) has highlighted the importance of upper-tail human capital and the networks of the Republic of Letters for bridging the gap between propositional and prescriptive knowledge. Upper-tail human capital would have been needed to access knowledge far from one's expertise, while Enlightenment networks would have connected scholars or practitioners to different spheres of knowledge than their own expertise. The importance of this channel has been shown in a burgeoning literature on upper-tail human capital.³² Based on this argument, this section tests whether proxies for upper-tail human capital and enlightenment networks can explain the strength of knowledge spillovers in 1740–1779, the moment the spillover mechanism turned positive. As proxies for upper-tail human capital and enlightenment networks, we use authors' occupation and education, as well as authors' membership in academic and other enlightenment societies.

We extract information on authors' education and occupation from ESTC book titles (see data section 3.1 and appendix section A.1). The approach builds on the logic that authors with prestigious occupations or memberships, such as the Royal Society, were likely to report these when publishing their books. Concretely, we extract information on authors' membership in

³¹ Also note that Patents are only available post-1700 (see section 3.2). Therefore, including patents in the calculation of the spillover index, would add an undesirable discontinuity to the spillover measure. Therefore, results from appendix figure 39 should be interpreted carefully.

³² See e.g. Squicciarini and Voigtländer (2015), Mokyr (2016), Dittmar and Meisenzahl (2020), Maloney and Valencia Caicedo (2022), Mokyr, Sarid and van der Beek (2022), Kelly, Mokyr and Ó Gráda (2023), Hanlon (2025), and Cinnirella, Hornung and Koschnick (2025). Following Mokyr (2009, 2016, 2018, 2021), we define upper-tail human capital as an endowment of skills that could be used to advance the knowledge frontier:

“‘upper-tail human capital,’ that is, the skills and knowledge of the best scientists, artisans, engineers, mechanics, and physicians (...) the idea that the envelope of useful knowledge is pushed forward by a relatively small number of people” (Mokyr, 2018)

the Royal Society, England's first and most prestigious scientific society, as well as membership in other enlightenment societies, consisting of economic societies like the Society of Arts in London or the Dublin Society (Stapelbroek and Marjanen, 2012; Howes, 2020), the Linnean Society, the Society of Antiquaries, various agricultural societies, and the Philosophical Society in Philadelphia. We further use information on authors who were engineers or had an academic career. Lastly, we use data on university students from Koschnick (2025) to capture authors who enrolled at either the university of Oxford or Cambridge.³³

To test how authors' networks, occupation, and education could have impacted spillovers, we formulate the following model:

$$\text{Received spillover}_{B \rightarrow A}(v_{ijt}) = \sum_{p \in \text{Occ}} \theta_p \mathbf{1}\{\text{Occ}_{ijt} = p\} + \mathbf{X}'_{ijt} \zeta + \gamma_j + \alpha_t + \varepsilon_{ijt} \quad (8)$$

where the dependent variable, Received spillover_{B → A}(v_{ijt}), captures the strength of received spillovers from equation 5. It is transformed using the natural logarithm. The main explanatory variable, Occ_{ijt}, is the set of authors' affiliations, occupations, and education, provided as a set of indicator variables. It includes membership in the *Royal Society*, or other *enlightenment societies*, being an *engineer*, having been *enrolled at university*, or having had an *academic career*. The model further includes word count controls, X'_{ijt}, and subject and year fixed effects, γ_j and α_t. Because we want to understand the determinants of spillovers during the period when spillovers were associated with innovation, we estimate the model for the time period 1740–1779.

Overall, we expect that membership in the Royal Society or other enlightenment societies would have created access to more knowledge, improving the search function over knowledge domains (A₂ from section 2).³⁴ Likewise, a university education would have endowed students with the skills to more easily access and apply formalized and complex knowledge (A₁ from section 2). Moreover, engineers would have been able to access both theory and practice, helping them to combine ideas (A₂), and to apply theories with precise measurement tools and to set the research agenda through puzzling facts or practical needs (B₂, B₃ from section 2).³⁵

³³In the seventeenth and eighteenth century, these were the only English universities.

³⁴See also Mokyr (2005), Zanardello (2024), Cinnirella, Hornung and Koschnick (2025), and de la Croix, Sceba and Zanardello (2025) for the role of enlightened academies and societies in economic history.

³⁵See also de Pleijt, Nuvolari and Weisdorf (2020), Mokyr, Sarid and Van Der Beek (2022), Maloney and Valencia Caicedo (2022), and Hanlon (2025) for the role of the engineer in economic history.

TABLE 2: Determinants of spillovers from propositional (Ω) to prescriptive knowledge (λ)

| | Time frame: 1740–1779 | | | | | |
|-------------------------------|-----------------------|------------------------|---------------------|-----------------------|----------------------|------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Spillover | Spillover | Spillover | Spillover | Spillover | Spillover |
| Fellowship in Royal Society | 0.0912*** (0.0186) | | | | | 0.0447** (0.0190) |
| Other enlightenment societies | | 0.0716*** (0.00584) | | | | 0.0595*** (0.00566) |
| Engineer | | | 0.0417* (0.0212) | | | 0.00379 (0.0212) |
| University enrollment | | | | 0.0507*** (0.0132) | | 0.0252** (0.0117) |
| Academic career | | | | | 0.107*** (0.0164) | 0.0634*** (0.0158) |
| Word count controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Subject class fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 2911 | 2911 | 2911 | 2911 | 2911 | 2911 |
| R-squared | 0.22 | 0.25 | 0.21 | 0.22 | 0.22 | 0.25 |

Notes: The table shows coefficients from estimating equation 8 via OLS. The dependent variable is the received spillover index from equation 5 for spillovers from propositional (Ω) to prescriptive knowledge (λ). It is transformed using the natural logarithm. The main explanatory variables are a set of indicator variables capturing authors' membership, occupations, and education. Column 2-5 then consecutively presents results for each indicator variable. Column 6 then reports a horse race with all explanatory variables. The model further controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Standard errors clustered by publication year. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 2 reports results for spillovers from propositional to prescriptive knowledge ($\Omega \rightarrow \lambda$) and table 3 reports results for spillovers from prescriptive to propositional knowledge ($\lambda \rightarrow \Omega$). For table 2, we find significant effects for all groups. However, effects are largest for membership in the Royal Society and other enlightenment societies as well as for academic careers. Being a fellow of the Royal Society is associated with an increase in authors' chances of receiving (incorporating) spillovers from propositional knowledge by 9.12%. The coefficient for membership in an enlightenment society is of a comparable magnitude at 7.16%. Likewise, being an academic was associated with a 10.7% increase in receiving a spillover from propositional knowledge. Running a horse race regression in column 6 confirms the previous findings, with the largest and significant coefficients found for membership in the Royal Society, other enlightenment societies, and an academic career.

TABLE 3: Determinants of spillovers from prescriptive (λ) to propositional knowledge (Ω)

| | Time frame: 1740–1779 | | | | | |
|-------------------------------|-----------------------|------------------------|---------------------|----------------------|-----------------------|------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Spillover | Spillover | Spillover | Spillover | Spillover | Spillover |
| Fellowship in Royal Society | 0.0608*** (0.0123) | | | | | 0.0268** (0.0125) |
| Other enlightenment societies | | 0.0595*** (0.00610) | | | | 0.0545*** (0.00654) |
| Engineer | | | 0.0539* (0.0288) | | | 0.0267 (0.0291) |
| University enrollment | | | | 0.0235** (0.0100) | | 0.0136 (0.00936) |
| Academic career | | | | | 0.0404*** (0.0104) | 0.00282 (0.0112) |
| Word count controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Subject class fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 2798 | 2798 | 2798 | 2798 | 2798 | 2798 |
| R-squared | 0.39 | 0.41 | 0.39 | 0.39 | 0.39 | 0.41 |

Notes: The table shows coefficients from estimating equation 8 via OLS. The dependent variable is the received spillover index from equation 5 for spillovers from prescriptive (λ) to propositional knowledge (Ω). It is transformed using the natural logarithm. The main explanatory variables are a set of indicator variables capturing authors' membership, occupations, and education. Column 2-5 then consecutively presents results for each indicator variable. Column 6 then reports a horse race with all explanatory variables. The model further controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Standard errors clustered by publication year. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

These findings are mirrored by the results from table 3 that present results for spillovers from prescriptive to propositional knowledge. Here, we also find significant effects of a similar size for all groups. However, in the horse race specification only memberships in the Royal Society and other enlightenment societies remain significant, revealing the importance of learned societies and networks for bridging propositional and prescriptive knowledge.

7.4 The scientific method

After having examined how upper-tail human capital helped facilitate knowledge spillovers, this section investigates whether there were complementary factors that were necessary to turn knowledge spillovers into innovation. Several scholars have argued that the scientific method (Landes, 1969, 1998; Woottton, 2015), the Newtonian Revolution (Jacob, 1997, 2014), or precise measurement (Ó Gráda, 2016; Kelly and Ó Gráda, 2022) were crucial for the creation of new

knowledge during the Industrial Revolution. According to these accounts, such developments transformed the way knowledge was produced. Careful observation and measurement of nature, combined with the systematic testing of theories against empirical evidence contributed to a more reliable knowledge base, the key to applying knowledge to innovation.

To test how these new methods aided the adoption of new knowledge in the innovation process, we interact the spillover measure from model 7 with embedding space similarity to these three revolutions in methods. Similar to the approach in Garg et al. (2018) and Ash, Chen and Naidu (2025), we capture each concept through a set of descriptive terms. For example, the scientific method is characterized through *observation*, *measurement*, *mathematical formalization*, or *experiment*, Newtonian mechanics is characterized through *Newtonian mechanics*, *force*, *momentum*, or *impulse*, and precise measurement is characterized through *precise measurement*, *precise instruments*, *standardized scales* or *calibration*. The full set of terms is documented in appendix table 19. These terms are then projected into the *SteamBERT* embedding space. We then calculate the average cosine similarity between each title and each list of concept terms.

To test whether these methods were complementary to knowledge spillovers in the innovation-generating process, we formulate the following model with $S_{ijt} := \text{Received spillover}_{B \rightarrow A}(v_{ijt})$ and $M_{ijt} := \text{Similarity to Method}_{ijt}$:

$$\text{Innovation}_{ijt}^A = \beta_1 S_{ijt} + \beta_2 M_{ijt} + \beta_3 S_{ijt} M_{ijt} + \beta_4 \mathbf{X}'_{ijt} \boldsymbol{\zeta} + \gamma_j j + \alpha_t + \varepsilon_{ijt}. \quad (9)$$

The dependent variable is the innovation index from equation 3. Then, the main set of explanatory variables are the interaction terms between the spillover measure from equation 5 and the similarity measures to the *scientific method*, *Newtonian mechanics*, and *precise measurement*. As before, the innovation and spillover measures are transformed using the natural logarithm. To ease the interpretation of the coefficients, the similarity measures are standardized to z-scores. As before, the model also contains a vector of control variables, \mathbf{X}'_{ijt} , incl. a title's level and quadratic word count as well as subject and year fixed effects, γ_j and α_t .

TABLE 4: Complementary factors for spillovers from propositional (Ω) to prescriptive knowledge (λ)

| | Time frame: 1740–1779 | | | | |
|---|-----------------------|----------------------|-----------------------|-----------------------|-----------------------|
| | (1) | (2) | (3) | (4) | (5) |
| | Innov. index | Innov. index | Innov. index | Innov. index | Innov. index |
| Spillover $\Omega \rightarrow \lambda$ | 0.0960*** (0.0197) | 0.104*** (0.0240) | 0.0925*** (0.0198) | 0.102*** (0.0212) | 0.0825*** (0.0256) |
| Similarity to Newtonian mechanics terms | | -0.00397 (0.0111) | | | -0.0131 (0.0173) |
| Spillover $\Omega \rightarrow \lambda \times$ Similarity to Newtonian mechanics | | 0.0111 (0.0163) | | | -0.0142 (0.0266) |
| Similarity to precise measurement terms | | | 0.00220 (0.00932) | | 0.0110 (0.0151) |
| Spillover $\Omega \rightarrow \lambda \times$ Similarity to precise measurement | | | 0.0255* (0.0130) | | 0.0411* (0.0210) |
| Similarity to scientific method terms | | | | -0.000319 (0.0109) | -0.00200 (0.0223) |
| Spillover $\Omega \rightarrow \lambda \times$ Similarity to scientific method | | | | 0.0189 (0.0150) | -0.00961 (0.0318) |
| Word count controls | Yes | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes |
| Subject class fixed effects | Yes | Yes | Yes | Yes | Yes |
| Observations | 3047 | 3047 | 3047 | 3047 | 3047 |
| R-squared | 0.19 | 0.20 | 0.20 | 0.20 | 0.20 |

Notes: The table shows coefficients from estimating equation 9 via OLS. The dependent variable is the innovation index from equation 3. The main set of explanatory variables are the interaction terms between the spillover measure from equation 5 from Ω to λ and the similarity measures to the *scientific method*, *Newtonian mechanics*, and *precise measurement*. The innovation and spillover measures are transformed using the natural logarithm, similarity measures are transformed using z-scores. The model further controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Standard errors clustered by publication year. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 4 reports results for spillovers from propositional to prescriptive knowledge and table 5 reports results for spillovers from prescriptive to propositional knowledge. First, in table 4, we find a significant and positive coefficient for the interaction term between spillovers and similarity to *precise measurement*. The coefficient also remains significant in the horse race between all similarity measures in column 5. This indicates that *precise measurement* was a complementary factor to receiving knowledge spillovers - being able to use precision measurement in the implementation of technology made the adoption of propositional knowledge more productive (A_3 from section 2) (see also Ó Gráda, 2016; Kelly and Ó Gráda, 2022). The channels seem to be able to explain a relevant part of the original spillover effect, reducing the spillover coefficient

by 14% in the horse race specification in column 5. In contrast, *Newtonian mechanics* and the *scientific method* seem to have been less important for the implementation of propositional knowledge into new techniques in λ .

TABLE 5: Complementary factors for spillovers from prescriptive (λ) to propositional knowledge (Ω)

| | Time frame: 1740–1779 | | | | |
|---|-----------------------|-----------------------|----------------------|-----------------------|---------------------|
| | (1) Innov. index | (2) Innov. index | (3) Innov. index | (4) Innov. index | (5) Innov. index |
| Spillover $\lambda \rightarrow \Omega$ | 0.130*** (0.0328) | 0.0676* (0.0360) | 0.0836** (0.0350) | 0.0914** (0.0342) | 0.0656* (0.0376) |
| Similarity to Newtonian mechanics terms | | 0.0258* (0.0137) | | | 0.0308 (0.0203) |
| Spillover $\lambda \rightarrow \Omega \times$ Similarity to Newtonian mechanics | | 0.0581*** (0.0176) | | | 0.0471* (0.0259) |
| Similarity to precise measurement terms | | | 0.0109 (0.0189) | | -0.0567 (0.0418) |
| Spillover $\lambda \rightarrow \Omega \times$ Similarity to precise measurement | | | 0.0484** (0.0224) | | -0.0353 (0.0526) |
| Similarity to scientific method terms | | | | 0.0273 (0.0183) | 0.0476 (0.0475) |
| Spillover $\lambda \rightarrow \Omega \times$ Similarity to scientific method | | | | 0.0608*** (0.0222) | 0.0497 (0.0608) |
| Word count controls | Yes | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes |
| Subject class fixed effects | Yes | Yes | Yes | Yes | Yes |
| Observations | 1358 | 1358 | 1358 | 1358 | 1358 |
| R-squared | 0.25 | 0.26 | 0.26 | 0.26 | 0.27 |

Notes: The table shows coefficients from estimating equation 9 via OLS. The dependent variable is the innovation index from equation 3. The main set of explanatory variables are the interaction terms between the spillover measure from equation 5 from λ to Ω and the similarity measures to the *scientific method*, *Newtonian mechanics*, and *precise measurement*. The innovation and spillover measures are transformed using the natural logarithm, similarity measures are transformed using z-scores. The model further controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Standard errors clustered by publication year. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Next, table 5 reports interaction terms between spillovers from prescriptive to propositional knowledge. Here, we find that individually (column 1–5), all interaction terms are positive and significant. This indicates that all three methods, *Newtonian mechanics*, *precise measurement*, and the *scientific method*, were important for integrating insights from λ into the stock of knowledge in Ω . Here it seems that Newtonian theory (see Jacob, 1997, 2014) helped to integrate empirical observations into new productive theory. Likewise, the scientific method with its focus

on experimentation as a means to reject false theories seems to have been important (B_1 from section 2) (see also Landes, 1998; Wootton, 2015). We can think of the practical work on water wheels, discussed in section 2, where engineers like John Smeaton conducted experimental work on water wheels that led to the rejection of predictions from theory, crucially helping to develop the theory of hydrodynamics. Altogether, all three channels reduce the spillover coefficient by 50% in the horse race specification in column 5.

Altogether, we find that for incorporating propositional into prescriptive knowledge, precise measurement seems to have played the largest role. In contrast, the process of incorporating prescriptive into propositional knowledge was strongly complemented by the scientific method and the Newtonian Revolution. This shows the complex operation of the feedback loop process, running through a range of channels and resting on recent innovations in methods that originated from the Scientific Revolution.

8 The *Lexicon technicum* as an exogenous shock to access costs to propositional knowledge

The previous section has explored the changing relationship between knowledge spillovers and innovation over time. Its key finding is that at the beginning of the eighteenth century, a deep structural transformation in the British knowledge economy took place, leading to the emergence of a positive feedback loop process. This section further investigates the dynamics at this point of change and presents causal evidence for the existence of innovation-inducing spillovers from propositional knowledge as early as the 1710s and 1720s.

Concretely, the section exploits the publication of the *Lexicon technicum*, the first British scientific and technical encyclopedia in 1704, as an exogenous shock. Its introduction significantly reduced access costs to propositional knowledge, thereby increasing the likelihood of knowledge spillovers from Ω to λ and from λ to Ω . The *Lexicon technicum* was the first British encyclopedia to move beyond brief definitions of scientific and technical terms and to provide substantive explanations of concepts, principles, and methods.³⁶ The scope of the project was significant and foreshadowed the scientific and technical categories of Diderot and d'Alembert's *Encyclopédie* (Kafker, 1981a; Squicciarini and Voigtländer, 2015). Given Harris' scientific interest in mathematics, the *Lexicon technicum* mainly contained topics in propositional knowledge (with a large focus on mathematics, but also astronomy, applied physics and chemistry). Overall, the *Lexicon technicum* contained 1049 entries on propositional subjects, while only covering

³⁶There existed a limited stock of prior encyclopedias such as Chauvin's *Lexicon Rationale* or *Thesaurus Philosophicus*. Yet all of these often did not venture beyond the definition of words. As put by John Harris himself, these dictionaries included “Simple Terms, so that you are told what a Dog, a Cat, a Horse and a Sheep is ; which (...) may [only] be useful to some Persons who did not know that before (...).”

152 on prescriptive subjects (see appendix table 11).³⁷ By making such a large amount of propositional knowledge easily available, the *Lexicon technicum* significantly reduced the cost of access to propositional knowledge and improved readers' search function over propositional knowledge (channel A_2 in section 2). In fact, the spread of the *Lexicon technicum* was considerable. The first volume already included a list of over 900 subscribers (Russell, 2020) and became the standard encyclopedia for the following decades (Bradshaw, 1981*b,a*).³⁸

In the following, we employ a difference-in-differences model to identify the effect of knowledge spillovers from propositional knowledge (Ω) in the *Lexicon technicum* to prescriptive knowledge (λ) in the ESTC. We can measure the strength of these spillovers empirically using the spillover index from equation 4. To identify topic-specific effects from the *Lexicon technicum*, we move to the sub-field unit level of ESTC title i (see appendix section F and appendix table 14 for the definition of sub-fields). Here we exploit, that not all sub-fields received positive knowledge spillovers from $\Omega \rightarrow \lambda$.

Thus, treatment is either defined (i) as the strength of knowledge spillovers post 1704 or (ii) as a binary indicator if a field received a positive knowledge spillover post 1704. Another assumption is that units treated through spillovers from Ω were not simultaneously treated by *Lexicon technicum* entries in λ . Here, we exploit the relatively brief coverage of technical subjects in the *Lexicon technicum* and exclude all sub-topics in the ESTC that were covered in the *Lexicon technicum*. We then estimate the following difference-in-differences model:

$$\text{Innovation}_{ijt} = \sum_{\tau=1685-1689}^{1735-1749} \beta_\tau (\text{Spillover } (\Omega \rightarrow \lambda)_{ijt} \times \eta_\tau) + \zeta_i + \alpha_t + \varepsilon_{ijt} \quad (10)$$

where the dependent variable is defined as the innovation index from equation 3 in sub-field i at year t . The dependent variable is transformed using the natural logarithm. Treatment is defined as differences in the strength of the spillover index from equation 4 between $\Omega \subset \text{Lexicon}$ and $\lambda \subset \text{ESTC}$ interacted with indicator variables for 5-year periods. We estimate two specifications: a) with continuous treatment and b) binary treatment that is positive for all sub-classes that received any positive knowledge spillovers. The model further includes sub-field i and year t fixed effects.

To rule out that new knowledge from λ confounded the estimated spillover, we exclude all sub-topics in λ that were directly covered by the *Lexicon technicum*. As argued before, we can exploit the fact that the *Lexicon technicum* only covered a small set of prescriptive topics in depth (Kafker, 1981*b*). For our baseline results, we exclude all sub-topics with less than 5 entries

³⁷Harris' core expertise was in mathematics and the natural sciences. He had published extensively on mathematics and was a Fellow of the Royal Society. Yet in contrast, he had few personal qualifications in the technical arts. As argued by Kafker (1981*b*), this lack of technical knowledge is apparent in the technical entries.

³⁸Here, the *Lexicon technicum* even competed with Chamber's *Cyclopaedia* after its publication in 1728 (Bradshaw, 1981*b,a*).

based on the reasoning that 5 entries would hardly suffice to fully cover any subjects in depth. Later we report robustness to changing this criterion to anything between 0 and 10 entries.

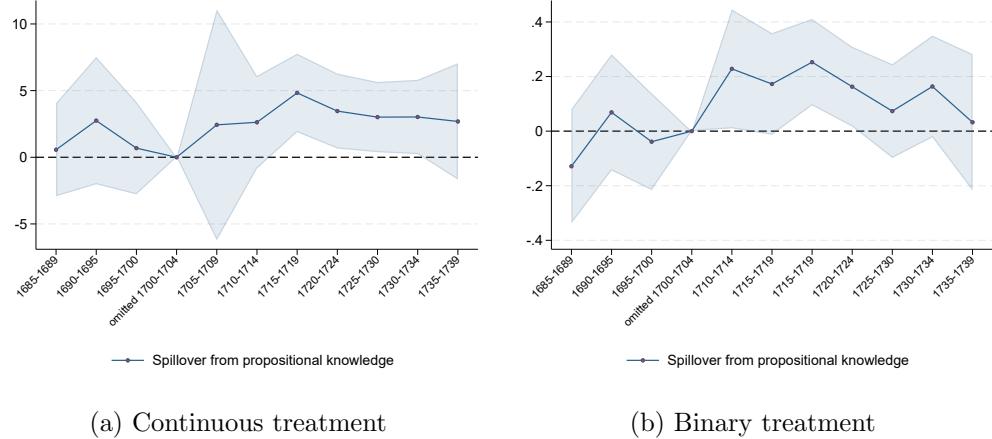


FIGURE 14: Difference-in-differences results for the *Lexicon technicum*: Spillovers from propositional to prescriptive knowledge ($\Omega \rightarrow \lambda$)

Notes: The graph presents the effect of spillovers from propositional knowledge (Ω) from the *Lexicon technicum* on innovation in prescriptive knowledge (λ) in the ESTC as estimated using the difference-in-differences model from equation 10. $N = 468$. Standard errors clustered at the topic level. Confidence intervals shown at the 90% level.

Figure 14 presents results for spillovers from propositional knowledge from the *Lexicon technicum* to prescriptive knowledge (λ) in the ESTC. Treatment starts in the post-1704 period when the *Lexicon technicum* was published.³⁹ Sub-panel a) reports results for continuous treatment, and sub-panel b) reports results for binary treatment. Both specifications indicate a significant increase in the innovativeness of treated sub-classes after the publication of the *Lexicon technicum*. In the continuous specification we find that an increase in knowledge spillovers by 1 standard deviation (corresponding to a 3% increase at the mean) would have led to a 9.8% increase in the average innovativeness of treated sub-classes.⁴⁰ In the binary specification, we find that an increase of knowledge spillovers by 1 standard deviation (corresponding to a 33% increase at the mean) would have led to a 5.3% increase in the average innovativeness of treated sub-classes.⁴¹ This is a highly relevant effect for the impact of a single publication. Yet, given the prominent role of encyclopedias in the history of thought (Darnton, 1973, 1987; Bradshaw, 1981a; Israel, 2001) or innovation and economic growth Squicciarini and Voigtländer (2015), the effect appears plausible.

We can further note that after 15 years, treatment effects decrease over time. Here, we

³⁹We assume here that treatment was not yet effective in 1704 which appears plausible since the *Lexicon technicum* was released in summer 1704 and any publications incorporating its content would have faced a significant lag until the publication was prepared for the printing press.

⁴⁰Assuming an avg. treatment effect over time of $\beta = 3.15$.

⁴¹Assuming an avg. treatment effect over time of $\beta = 0.16$.

should note the confounding effect of the publication of Ephraim Chamber's *Cyclopædia* in 1728. In comparison to the *Lexicon technicum*, the *Cyclopædia* had a significantly broader coverage of subjects. Therefore, it is likely that many of the previously untreated sub-classes received spillover-treatment after the publication of Harris's *Cyclopædia*.

We further conduct the following robustness tests:

1. **Changing sub-topic cluster hyperparameters:** Sub-topics for the ESTC are derived from a HDBSCAN approach similar to Grootendorst (2022) (see data section 3.1). Using embedding similarity, all *Lexicon technicum* entries are then assigned to the closest ESTC sub-topic in the embedding space. This approach has the advantage that clusters are based on the structure of the embedding space and capture deep semantic information. Moreover, in contrast to e.g. k-means, where the researcher chooses the number of clusters, using HDBSCAN, the number of clusters itself is derived from the structure of the embedding space. Nonetheless, the identification of clusters depends on a set of hyperparameters, documented in appendix F. Figure 25 shows that changing these hyperparameters does not lead to relevant changes in the regression results as the allocation of clusters changes (see section F for details). Additionally, figure 25 also reports results for changing the clear control group criterion, that consecutively excludes topics with more than 0, 2, 5, 8, and 10 entries in λ in the *Lexicon technicum*. Overall, figure 25 presents results for 65 different combinations of hyperparameters and clear control group criteria. Altogether, results remain stable and close to the baseline results. Altogether, the majority of the results have a smaller p-value than the baseline, with only a small number of coefficients above 0.1 (see figure 26).
2. **Controlling for subject-specific trends:** Appendix figure 45 reports further robustness for adding subject specific linear trends to account for compositional bias.
3. **Accounting for other shocks to knowledge production:** It is important to carefully consider other shocks to knowledge production that could have occurred simultaneously. Here, the 1714 Longitude Prize might constitute an obvious confounder. To demonstrate robustness, appendix figure 44 presents results for excluding the subject class of *navigation*; estimated coefficients hardly change.

Overall, these results provide causal evidence of the existence of innovation-inducing spillovers in the eighteenth century. They further highlight one specific channel through which positive spillovers could have worked, namely, codification of and decreasing access costs to knowledge. These were important goals of the Enlightenment. The *Lexicon technicum* was only the first of a series of encyclopedias in Britain that steadily grew in scope and accuracy, ultimately culminating in Diderot and d'Alembert's *Encyclopédie* in France.

Nonetheless, this section only highlights one of many channels through which the spheres of propositional and prescriptive knowledge could have become integrated as a part of the “Industrial Enlightenment” (Mokyr, 2002). Other channels include scientific and economic societies, specialized journals, public lectures, and the mobility of scientists and technical experts. These, together with the effect of other encyclopedias, would be fruitful routes for future research.

9 Conclusion

This paper has provided the first quantitative test of the feedback loop hypothesis proposed by Mokyr (2002), who has argued that the emergence of self-sustained modern growth was caused by a structural change in the relationship between propositional and prescriptive knowledge. Using two new innovation and spillover measures constructed from a semantically rich, historically fine-tuned BERT model, the paper has estimated the relationship between knowledge spillovers between propositional and prescriptive knowledge and innovation between 1600 and 1800 in England. The evidence presented shows that knowledge spillovers were negative or neutral before the eighteenth century but began to turn positive in the 1720s and were firmly positive by the 1760s. This timing fits well with the predictions from the Mokyrian account and supports the idea of a structural break in the knowledge economy during the Industrial Enlightenment.

The paper has further shown that spillovers from propositional knowledge were positively associated with patent innovation and patent citations at the end of the eighteenth century. Thus, the paper documents that the positive knowledge spillovers also extended to the real economy. Moreover, it has shown that upper-tail human capital, especially for Royal Society fellows and engineers, played an instrumental role in facilitating these spillovers. Moreover, revolutions in methods, such as *Newtonian mechanics*, *precise measurement* and *the scientific method* seem to have been complementary factors to knowledge spillover. Finally, using the publication of the *Lexicon technicum* (1704) as an exogenous shock to access costs to propositional knowledge, the paper has provided causal evidence on the existence of knowledge spillovers from propositional to prescriptive knowledge. The results for the *Lexicon technicum* also highlight the importance of codified knowledge and encyclopedias for bridging the gap between propositional and prescriptive knowledge.

Altogether, the paper provides empirical support for the Mokyrian thesis. It presents evidence of the emergence of a positive relationship between knowledge spillovers and innovation at the end of the eighteenth century. It also presents evidence that upper-tail human capital played a facilitating role and has highlighted the importance in the revolution in methods produced as part of the Scientific Revolution. Altogether, these results provide crucial quantitative backing to Mokyr (2002)'s account of the transition towards modern self-sustained growth. Thereby, the results found in this paper significantly add to our understanding of the origin of modern

economic growth.

The paper further introduces a framework to estimate spillovers between different corpora of text that can be applied to a variety of other settings. At the same time, the paper has shown how to move beyond simple associations by conducting systematic placebo tests to rule out confounders or by applying difference-in-differences analysis to singular knowledge shocks. Overall, the new NLP measures of innovation and spillovers in combination with causal methods can help us gain a more structural understanding of the forces within either the historical or modern knowledge economy.

References

- Acemoglu, Daron, Ufuk Akcigit, and William R Kerr.** 2016. “Innovation network.” *Proceedings of the National Academy of Sciences* 113 (41): 11483–11488.
- Allen, Robert C.** 2009. *The British industrial revolution in global perspective*. Cambridge: Cambridge University Press.
- Almelhem, Ali, Murat Iyigun, Austin Kennedy, and Jared Rubin.** 2023. “Enlightenment ideals and belief in progress in the run-up to the industrial revolution: A textual analysis.” , (16674).
- Ash, Elliott, Daniel L Chen, and Suresh Naidu.** 2025. “Ideas have consequences: The impact of law and economics on american justice.” *The Quarterly Journal of Economics* qjaf042.
- Baten, Joerg, and Jan Luiten Van Zanden.** 2008. “Book production and the onset of modern economic growth.” *Journal of Economic Growth* 13 (3): 217–235.
- Becker, Sascha O., Jared Rubin, and Ludger Woessmann.** 2024. “Religion and Growth.” *Journal of Economic Literature* 62 (3): 1094–1142.
- Billington, Stephen.** 2019. “Patents, machine learning, and war: invention during the British Industrial Revolution.”
- Billington, Stephen, and Alan Hanna.** 2020. “That’s classified! Inventing a new patent taxonomy.” *Inventing a New Patent Taxonomy (May 1, 2020)*.
- Billington, Stephen D.** 2021. “What explains patenting behaviour during Britain’s Industrial Revolution?” *Explorations in Economic History* 82: 101426.
- Bouscasse, Paul, Emi Nakamura, and Jón Steinsson.** 2025. “When did growth begin? New estimates of productivity growth in England from 1250 to 1870.” *The Quarterly Journal of Economics* 140 (2): 835–888.
- Bradshaw, Lael Ely.** 1981a. “Ephraim Chambers’ Cyclopaedia.” In *Notable encyclopedias of the seventeenth and eighteenth centuries: Nine predecessors of the encyclopédie*, edited by Frank A. Kafker. Taylor Institution.
- Bradshaw, Lael Ely.** 1981b. “John Harris’s Lexicon technicum.” In *Notable encyclopedias of the seventeenth and eighteenth centuries: Nine predecessors of the encyclopédie*, edited by Frank A. Kafker. Taylor Institution.

- Capecchi, Danilo.** 2013. “Over and undershot waterwheels in the 18th century. Science-technology controversy.” *Advances in Historical Studies* 2 (03): 131.
- Chiopris, Caterina.** 2024. “The Diffusion of Ideas.” Manuscript draft.
- Cinnirella, Francesco, Erik Hornung, and Julius Koschnick.** 2025. “Flow of ideas: Economic societies and the rise of useful knowledge.” *The Economic Journal* 135 (669): 1496–1535.
- Clark, Gregory.** 2007. *A Farewell to Alms*. Prince: Princeton University Press.
- Cohen, H Floris.** 1994. *The scientific revolution: a historiographical inquiry*. Chicago: University of Chicago Press.
- Constant, Edward W.** 1983. “Scientific theory and technological testability: Science, dynamometers, and water turbines in the 19th Century.” *Technology and Culture* 24 (2): 183–198.
- Curtis, Matthew JD, and David de la Croix.** 2023. “Seeds of Knowledge: Premodern Scholarship, Academic Fields, and European Growth.” Centre for Economic Policy Research, CEPR discussion paper DP18321.
- Darnton, Robert.** 1973. “The Encyclopédie wars of prerevolutionary France.” *The American Historical Review* 78 (5): 1331–1352.
- Darnton, Robert.** 1987. *The business of Enlightenment: a publishing history of the Encyclopédie, 1775–1800*. Harvard University Press.
- de la Croix, David, Rossana Scebba, and Chiara Zanardello.** 2025. “Flora, Cosmos, Salvatio: Pre-modern Academic Institutions and the Spread of Ideas.” Centre for Economic Policy Research (CEPR) CEPR Discussion Paper 20569.
- de Pleijt, Alexandra, Alessandro Nuvolari, and Jacob Weisdorf.** 2020. “Human Capital Formation During the First Industrial Revolution: Evidence from the use of Steam Engines.” *Journal of the European Economic Association* 18 (2): 829–889.
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova.** 2019. “Bert: Pre-training of deep bidirectional transformers for language understanding.” 4171–4186.
- Dittmar, Jeremiah.** 2019. “The Economic Origins of Modern Science: Technology, Institutions, and Markets.” *Working Paper*.
- Dittmar, Jeremiah, and Ralf Meisenzahl.** 2021. “The research university, invention, and industry: Evidence from german history.” mimeo, London School of Economics.

- Dittmar, Jeremiah E.** 2011. “Information technology and economic change: the impact of the printing press.” *The Quarterly Journal of Economics* 126 (3): 1133–1172.
- Dittmar, Jeremiah E, and Ralf R Meisenzahl.** 2020. “Public goods institutions, human capital, and growth: Evidence from German history.” *The Review of Economic Studies* 87 (2): 959–996.
- Gao, Tianyu, Xingcheng Yao, and Danqi Chen.** 2021. “Simcse: Simple contrastive learning of sentence embeddings.” *arXiv preprint arXiv:2104.08821*.
- Garg, Nikhil, Londa Schiebinger, Dan Jurafsky, and James Zou.** 2018. “Word embeddings quantify 100 years of gender and ethnic stereotypes.” *Proceedings of the National Academy of Sciences* 115 (16): E3635–E3644.
- Gaukroger, Stephen.** 2020. “Encyclopedias and Encyclopedic Knowledge.” *Encyclopedia of Early Modern Philosophy and the Sciences*, edited by Dana Jalobeanu and Charles T. Wolfe, 1–5. Cham: Springer International Publishing.
- Grajzl, Peter, and Peter Murrell.** 2024. “Quiet revolutions in early-modern England.” *Public Choice* 200 (3): 357–381.
- Grajzl, Peter, and Peter Murrell.** 2025. “Innovation Under Suppression: Censorship’s Effect on Cultural Production in Early-Modern England.” Manuscript.
- Graney, Christopher M.** 2010. “The telescope against Copernicus: Star observations by Riccioli supporting a geocentric universe.” *Journal for the History of Astronomy* 41 (4): 453–467.
- Graney, Christopher M, and Timothy P Grayson.** 2011. “On the telescopic disks of stars: a review and analysis of stellar observations from the early seventeenth through the middle nineteenth centuries.” *Annals of science* 68 (3): 351–373.
- Grootendorst, Maarten.** 2022. “BERTopic: Neural topic modeling with a class-based TF-IDF procedure.” *arXiv preprint arXiv:2203.05794*.
- Hall, Alfred Rupert.** 1974. “What did the Industrial Revolution in Britain owe to science?” In *Historical perspectives - Studies in English thought and society in honour of J.H. Plumb*, edited by J. H. Plumb and Neil Kendrick. London: London Europa.
- Hanlon, W Walker.** 2025. “The Rise of the Engineer: Inventing the Professional Inventor During the Industrial Revolution.” *The Economic Journal* 135 (670): 1749–1781.

- Hanlon, W. Walker, Stephan Heblisch, Ferdinando Monte, and Martin B. Schmitz.** 2022. “A Penny for Your Thoughts.” National Bureau of Economic Research NBER Working Paper 30076.
- Hornung, Erik.** 2014. “Immigration and the diffusion of technology: The Huguenot diaspora in Prussia.” *American Economic Review* 104 (1): 84–122.
- Howes, Anton.** 2020. *Arts and Minds: How the Royal Society of Arts Changed a Nation*. Princeton: Princeton University Press.
- Howse, Derek.** 1986. “Navigation and Astronomy the first three thousand years.” *Culture, Theory and Critique* 30 (1): 60–86.
- Israel, Jonathan I.** 2001. *Radical Enlightenment: Philosophy and the Making of Modernity, 1650–1750*. Oxford: Oxford University Press.
- Jacob, Margaret C.** 1997. *Scientific culture and the making of the industrial West*. Oxford: Oxford University Press.
- Jacob, Margaret C.** 2014. *The first knowledge economy: Human capital and the European economy, 1750–1850*. Cambridge: Cambridge University Press.
- Kafker, Frank A.** 1981a. *Notable encyclopedias of the seventeenth and eighteenth centuries: Nine predecessors of the encyclopédie*. Oxford: Taylor Institution.
- Kafker, Frank A.** 1981b. “The Encyclopédie in relation to the nine predecessors.” In *Notable encyclopedias of the seventeenth and eighteenth centuries: Nine predecessors of the encyclopédie*, edited by Frank A. Kafker. Taylor Institution.
- Kauffman, Stuart, José Lobo, and William G Macready.** 2000. “Optimal search on a technology landscape.” *Journal of Economic Behavior & Organization* 43 (2): 141–166.
- Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy.** 2021. “Measuring technological innovation over the long run.” *American Economic Review: Insights* 3 (3): 303–320.
- Kelly, Morgan, and Cormac Ó Gráda.** 2020. *Connecting the Scientific and Industrial Revolutions: The Role of Practical Mathematics*. Working Paper Series, No. WP20/17, University College Dublin.
- Kelly, Morgan, and Cormac Ó Gráda.** 2022. “Connecting the Scientific and Industrial Revolutions: The Role of Practical Mathematics.” *The Journal of Economic History* 82 (3): 841–873.

- Kelly, Morgan, and Cormac Ó Gráda.** 2016. “Adam Smith, Watch Prices, and the Industrial Revolution *.” *The Quarterly Journal of Economics* 131 (4): 1727–1752.
- Kelly, Morgan, Joel Mokyr, and Cormac Ó Gráda.** 2023. “The mechanics of the Industrial Revolution.” *Journal of Political Economy* 131 (1): 59–94.
- Koschnick, Julius.** 2025. “Teacher-directed scientific change: The case of the English Scientific Revolution.” European Historical Economics Society EHES Working Paper 274.
- Landes, David S.** 1969. *The unbound Prometheus technological change and industrial development in Western Europe from 1750 to the present*. Cambridge: Cambridge University Press.
- Landes, David S.** 1998. *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*. New York: W. W. Norton.
- Lincoln, Paula Yuhr.** 1983. “John Donne’s use of numbers.” *University of Louisville, Electronic Theses and Dissertations* 183.
- MacLeod, Christine.** 1989. *Inventing the industrial revolution: The English patent system, 1660-1800*. Cambridge University Press.
- MacLeod, Christine.** 1991. “The Paradoxes of Patenting: Invention and Its Diffusion in 18th- and 19th-Century Britain, France, and North America.” *Technology and culture* 32 (4): 885–910.
- Maloney, William F, and Felipe Valencia Caicedo.** 2022. “Engineering Growth.” *Journal of the European Economic Association* 20 (4): 1554–1594.
- Manjavacas Arévalo, Enrique, and Lauren Fonteyn.** 2021. “MacBERTh: Development and Evaluation of a Historically Pre-trained Language Model for English (1450-1950).” 23–36. Association for Computational Linguistics.
- Mathias, Peter.** 1972. “Who unbound Prometheus?” In *Science and Society, 1600-1900*, edited by Mathias Peter, 129–151. Cambridge: Cambridge University Press.
- Meisenzahl, Ralf R, and Joel Mokyr.** 2011. “The rate and direction of invention in the British industrial revolution: Incentives and institutions.” In *The rate and direction of inventive activity revisited*, 443–479. Chicago: University of Chicago Press.
- Merton, Robert K.** 1938. “Science, technology and society in seventeenth century England.” *Osiris* 4: 360–632.
- Milgrom, Paul, Yingyi Qian, and John Roberts.** 1991. “Complementarities, momentum, and the evolution of modern manufacturing.” *The American Economic Review* 81 (2): 84–88.

- Mokyr, Joel.** 1992. *The lever of riches: Technological creativity and economic progress*. Oxford: Oxford University Press.
- Mokyr, Joel.** 2002. *The gifts of Athena: Historical origins of the knowledge economy*. Princeton: Princeton University Press.
- Mokyr, Joel.** 2005. “The intellectual origins of modern economic growth.” *The Journal of Economic History* 65 (2): 285–351.
- Mokyr, Joel.** 2009. *The Enlightened economy an economic history of Britain 1700-1850*. New Haven: Yale University Press.
- Mokyr, Joel.** 2016. *A culture of growth: the origins of the modern economy*. Princeton: Princeton University Press.
- Mokyr, Joel.** 2018. “Bottom-up or top-down? The origins of the Industrial Revolution.” *Journal of Institutional Economics* 14 (6): 1003–1024.
- Mokyr, Joel.** 2021. “The holy land of industrialism: Rethinking the Industrial Revolution.” *Journal of the British Academy* 9: 223–47.
- Mokyr, Joel.** 2025. “Culture Versus Institutions in the Great Enrichment.” In *Handbook of New Institutional Economics*, 2 edition, edited by Claude Ménard and Mary M. Shirley. Springer.
- Mokyr, Joel, Assaf Sarid, and Karine Van Der Beek.** 2022. “The wheels of change: Technology adoption, millwrights and the persistence in Britain’S industrialisation.” *The Economic Journal* 132 (645): 1894–1926.
- Mokyr, Joel, Assaf Sarid, and Karine van der Beek.** 2022. “The Wheels of Change: Technology Adoption, Millwrights and the Persistence in Britain’S Industrialisation.” *The Economic Journal* 132 (645): 1894–1926.
- Moser, Petra.** 2005. “How do patent laws influence innovation? Evidence from nineteenth-century world’s fairs.” *American economic review* 95 (4): 1214–1236.
- Musson, Albert Edward, and Eric Robinson.** 1969. *Science and technology in the industrial revolution*. Milton Park: Taylor & Francis.
- Nuvolari, Alessandro, and Valentina Tartari.** 2011. “Bennet Woodcroft and the value of English patents, 1617-1841.” *Explorations in Economic History* 48 (1): 97–115.
- Ó Gráda, Cormac.** 2016. “Did science cause the industrial revolution?” *Journal of Economic Literature* 54 (1): 224–239.

- Pomeranz, Kenneth.** 2000. *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton, NJ: Princeton University Press.
- Prize Committee.** 2025. “Scientific Background to the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2025.” The Committee for the Prize in Economic Sciences in Memory of Alfred Nobel.
- Raymond, Joad.** 2003. *Pamphlets and pamphleteering in early modern Britain*. Cambridge: Cambridge University Press.
- Reimers, Nils, and Iryna Gurevych.** 2019. “Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks.” Association for Computational Linguistics.
- Reynolds, Terry S.** 1983. *Stronger than a hundred men - a history of the vertical water wheel*. Baltimore: John Hopkins University Press.
- Romer, Paul M.** 1986. “Increasing returns and long-run growth.” *Journal of political economy* 94 (5): 1002–1037.
- Romer, Paul M.** 1990. “Endogenous technological change.” *Journal of political Economy* 98 (5(2)): 71–102.
- Rosenberger, Lukas, W. Walker Hanlon, and Carl Hallmann.** 2024. “Innovation Networks in the Industrial Revolution.” National Bureau of Economic Research NBER Working Paper 32875.
- Russell, Terence M.** 2020. *The Encyclopaedic Dictionary in the Eighteenth Century: Architecture, Arts and Crafts. Volume 1: John Harris and the Lexicon Technicum*. Abingdon, UK: Taylor & Francis Group. ProQuest Ebook Central.
- Schofield, Robert E.** 1957. “The industrial orientation of science in the Lunar Society of Birmingham.” *Isis* 48 (4): 408–415.
- Schofield, Robert E.** 1963. *The Lunar Society of Birmingham: a social history of provincial science and industry in eighteenth-century England*. Oxford: Clarendon Press.
- Scultetus, Anke H., J.Leonel Villavicencio, and Norman M. Rich.** 2001. “Facts and fiction surrounding the discovery of the venous valves.” *Journal of Vascular Surgery* 33 (2): 435–441.
- Shapin, Steven, and Simon Schaffer.** 1985. *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton, NJ: Princeton University Press.

- Squicciarini, Mara P, and Nico Voigtländer.** 2015. "Human capital and industrialization: Evidence from the age of enlightenment." *The Quarterly Journal of Economics* 130 (4): 1825–1883.
- Stapelbroek, K, and J Marjanen.** 2012. *The Rise of Economic Societies in the Eighteenth Century: Patriotic Reform in Europe and North America*. Palgrave Macmillan UK.
- Stewart, Larry.** 1986. "Public lectures and private patronage in Newtonian England." *Isis* 77 (1): 47–58.
- Stewart, Larry.** 1992. *The rise of public science: rhetoric, technology, and natural philosophy in Newtonian Britain, 1660-1750*. Cambridge: Cambridge University Press.
- Stewart, Larry.** 2007. "Experimental spaces and the knowledge economy." *History of science* 45 (2): 155–177.
- Stewart, Larry, and Paul Weindling.** 1995. "Philosophical threads: natural philosophy and public experiment among the weavers of Spitalfields." *The British journal for the history of science* 28 (1): 37–62.
- Stokes, Donald E.** 1997. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington, D.C.: Brookings Institution.
- Taylor, Eva Germaine Rimington.** 1957. *The haven-finding art - A history of navigation from Odysseus to Captain Cook*. New York: Abelard-Schuman Limited.
- Trajtenberg, Manuel.** 1990. "A penny for your quotes: patent citations and the value of innovations." *The Rand journal of economics* 172–187.
- Valleriani, Matteo.** 2009. "The Transformation of Aristotle's Mechanical Questions: A Bridge Between the Italian Renaissance Architects and Galileo's First New Science." *Annals of Science* 66 (2): 183–208.
- Viollet, Pierre-Louis.** 2017. "From the water wheel to turbines and hydroelectricity. Technological evolution and revolutions." *Comptes Rendus Mécanique* 345 (8): 570–580.
- Woodcroft, Bennet.** 1854a. *Chronological Index of Patents Applied for and Patents Granted: From March 2, 1617 (14 James I.) to October 1, 1852 (16 Victoriae)*. London: George Edward Eyre and William Spottiswoode.
- Woodcroft, Bennet.** 1854b. *Subject-Matter Index of Patents of Inventions, March 2, 1617 to October 1, 1852*. London: George Edward Eyre and William Spottiswoode.

Wootton, David. 2015. *The invention of science: a new history of the scientific revolution.* London: Allen Lane.

Wrigley, Edward Anthony. 2010. “Energy and the English Industrial Revolution.”

Zanardello, Chiara. 2024. “Early Modern Academies, Universities, and Growth.” *WP, IRES/LIDAM, UCLouvain.*

Appendices

A Data

A.1 Variable and source descriptions

Text sources

English Short Title Catalogue. The English Short Title Catalogue (ESTC) kindly shared by the British Library with the author. The cleaned ESTC version is taken from Koschnick (2025) and includes translated titles and has been cleaned for duplicates. Topics assignments are taken from Koschnick (2025). Topics are based on higher-order classes based on subjects assigned by expert librarians. Koschnick (2025) further trained a BERT model on the labeled ESTC corpus to predict missing labels. Further descriptions and model evaluations are provided in Koschnick (2025). The ESTC catalogue includes 285,985 titles over the time period 1600–1800. These cover the universe of all printed works in England as well as all works printed in the English language elsewhere.

Patents. Patent for the time period 1700–1851 are obtained from the *Chronological Index of Patents Applied for and Patents Granted* compiled by Bennett Woodcroft in 1854a. They are merged with patent statistics from Nuvolari and Tartari (2011). Overall, the patent dataset includes 12,722 patent short descriptions.

Lexicon technicum. First published in 1704 by John Harris, the *Lexikon technicum* was the first scientific and technical English encyclopedia. The encyclopedia was published in two volumes in 1704 and 1720. Raw scans of the book are obtained from the Smithsonian Libraries (<https://library.si.edu/digital-library/book/lexicontechnicu1harr> and <https://library.si.edu/digital-library/book/lexicontechnicu2harr>). Optical character recognition was performed using amazon’s *textract*. Additionally, a `llama3:8b-instruct-q4_K_M` acted as a proof reader, checking OCR mistakes with long-s characters. Separation of entries was conducted using regular expressions based on capitalized keywords. Next, subject classes are predicted using the ESTC-trained BERT model from Koschnick (2025). To separate short definitions from substantive entries, all entries with a prediction certainty of less than 70% are excluded in all main specifications.

Variable descriptions

Innovation index. The innovation index measures how much a title introduces new ideas relative to past publications. For title v_{it_0} , it is defined as:

$$\text{Innovation}_{it_0} = \frac{f(v_i, \mathcal{F}_{t_0, \tau})}{f(v_i, \mathcal{B}_{t_0, \tau})},$$

where $f(\cdot)$ is the cosine similarity function between embeddings obtained from the fine-tuned *SteamBERT* model, $\mathcal{F}_{t_0, \tau}$ is the set of future titles, and $\mathcal{B}_{t_0, \tau}$ the set of past titles within a $\tau = 20$ -year window. Titles with $\text{Innovation}_{it_0} > 1$ are more similar to future than to past texts and hence classified as innovative.

Received spillovers. The received spillover index quantifies the extent to which a title in field A draws on prior work from field B :

$$\text{Received spillover}_{A \rightarrow B}(v_{it}) := \frac{f^k(v_{it}, \mathcal{B}_{t, \tau}^A)}{f^\rho(\mathcal{B}_{t, \tau}^B, \mathcal{B}_{t, \tau}^A)}.$$

where $\mathcal{B}_{t, \tau}^B$ is the set of past titles in B , and \mathcal{M}_{it}^A are the $k = 20$ most similar titles to v_{it} in A published in the previous $\tau = 20$ years. Values above one indicate that v_{it} is unusually close to prior work in B relative to its own field.

Received spillovers from propositional knowledge. Received spillovers are calculated using the above formula from equation 5. For each title in v_{it} in prescriptive knowledge, spillovers are calculated to the backward pool, $\mathcal{B}_{t, \tau}^A$, of $\mathcal{P} := \text{applied physics, astronomy, mathematics, chemistry, and encyclopedias}$. Spillovers from propositional knowledge are then averaged as $\frac{1}{|\mathcal{P}|} \sum_{A \in \mathcal{P}} \text{Received spillover}_{A \rightarrow B}(v_{it})$.

Received spillovers from prescriptive knowledge. Received spillovers are calculated using the above formula from equation 5. For each title in v_{it} in prescriptive knowledge, spillovers are calculated to the backward pool, $\mathcal{B}_{t, \tau}^A$, of $\mathcal{P} := \text{technology in trades, technology in agriculture, navigation, and scientific instruments}$. Spillovers from propositional knowledge are then averaged as $\frac{1}{|\mathcal{P}|} \sum_{A \in \mathcal{P}} \text{Received spillover}_{A \rightarrow B}(v_{it})$.

Patent citations. Woodcroft patent citations obtained from Nuvolari and Tartari (2011).

Occupations of ESTC authors. Occupational information derived from ESTC titles. In a first step we identified information on authors using a regex-routine. Overall, we find author information for 47% of all titles within prescriptive and propositional knowledge. Then we extracted post-comma additional information on authors. This covers 32% of all authors. Based on these strings, we identified the following groups: *engineer, medical career, academic career, fellow of the Royal Society and membership in other enlightenment societies*. *Fellow of the Royal Society* includes all mentions of *F.R.S* or *fellow of the Royal Society*. Enlightenment societies consist of economic societies like the Society of Arts in London or the Dublin Society (Stapelbroek and Marjanen, 2012; Howes, 2020; Cinnirella, Hornung and Koschnick, 2025), the Linnean Society, the Society of Antiquaries, various agricultural societies, and the Philosophical Society in Philadelphia. *Engineer* includes all mentions of *engineers, mechanics, surveyors, hydrographers and projectors*. *Academic* includes all mentions of *Teacher, librarian, reader of,*

*lecturer, professor, M.A., Fellow of....*⁴² Finally, going beyond the title information, authors are matched to students enrolled at the universities of Oxford and Cambridge based on Koschnick (2025).

Similarity to new methods. Similarities to the three methodological revolutions of *Newtonian mechanics*, *precise measurement*, and the *scientific method*. Each concept is captured through a list of essential terms that are presented in table 19. Similarities are then calculated at the title level and averages across each term.

A.2 Training corpus

TABLE 6: Summary statistics for training corpus

| Corpus | Number of publications | Average word length | Average character length |
|-------------|------------------------|---------------------|--------------------------|
| Patents | 12722 | 17.4 | 105.1 |
| ESTC titles | 9047 | 55.8 | 346.8 |
| Combined | 21769 | 33.4 | 205.6 |

Notes: The table reports summary statistics for the corpus used for fine-tuning the MacBERT base model.

⁴²An extra rule ensures that fellows at colleges are separated from fellows of the Royal Society. The approach further accounts for different capitalization

TABLE 7: Summary statistics for training corpus by group

| Panel | Group | Num. publications | Avg. word length | Avg. character length |
|--|------------------------------------|-------------------|------------------|-----------------------|
| <i>Patents (1700–1850) (Industry)</i> | | | | |
| | Agriculture | 418 | 16.8 | 99.9 |
| | Carriages | 789 | 17.4 | 104.9 |
| | Chemicals | 1089 | 18.1 | 110.5 |
| | Clothing | 320 | 18.4 | 113.1 |
| | Construction | 633 | 19.5 | 118.8 |
| | Engines | 1577 | 17.0 | 101.8 |
| | Food | 698 | 16.9 | 102.0 |
| | Furniture | 642 | 16.0 | 97.1 |
| | Glass | 120 | 17.3 | 105.0 |
| | Hardware | 817 | 16.4 | 96.7 |
| | Instruments | 582 | 18.6 | 114.4 |
| | Leather | 214 | 17.3 | 102.5 |
| | Manufacturing | 664 | 15.2 | 90.7 |
| | Medicines | 281 | 20.0 | 123.1 |
| | Metallurgy | 662 | 17.1 | 100.7 |
| | Military | 242 | 17.5 | 101.3 |
| | Mining | 76 | 20.2 | 116.4 |
| | Paper | 465 | 14.9 | 90.8 |
| | Pottery | 272 | 20.7 | 125.8 |
| | Ships | 563 | 19.1 | 114.1 |
| | Textiles | 1598 | 17.5 | 107.4 |
| <i>ESTC (1600-1800)(Subject class)</i> | | | | |
| | Applied physics | 822 | 42.1 | 264.0 |
| | Astronomy | 907 | 56.5 | 344.8 |
| | Chemistry | 370 | 43.2 | 269.5 |
| | Encyclopedias and dictionaries | 1352 | 76.4 | 480.3 |
| | Mathematics | 1345 | 56.0 | 355.1 |
| | Navigation | 964 | 53.7 | 329.8 |
| | Scientific instruments | 367 | 46.4 | 288.6 |
| | Technical instructions Agriculture | 1288 | 54.4 | 335.1 |
| | Technical instructions Trades | 1632 | 52.5 | 321.9 |

Notes: The table reports summary statistics for the corpus used for fine-tuning the MacBERT base model.

TABLE 8: Summary statistics variables in ESTC

| | Mean | Std.Dev. | Min | Max | Obs |
|--|---------|----------|---------|----------|------|
| Innovation index | 0.8959 | 0.3075 | -2.1665 | 12.4140 | 5835 |
| Spillover $\Omega \rightarrow \lambda$ | -0.7095 | 0.2233 | -3.3201 | 0.8570 | 4226 |
| Spillover $\lambda \rightarrow \Omega$ | -0.8289 | 0.4153 | -5.6251 | 4.4139 | 4757 |
| Fellowship in Royal Society | 0.0265 | 0.1606 | 0.0000 | 1.0000 | 9024 |
| Medical career | 0.0227 | 0.1490 | 0.0000 | 1.0000 | 9024 |
| Engineer | 0.0107 | 0.1031 | 0.0000 | 1.0000 | 9024 |
| Academic career | 0.0506 | 0.2193 | 0.0000 | 1.0000 | 9024 |
| Word count | 55.8649 | 49.2084 | 1.0000 | 527.0000 | 9024 |
| Observations | 9024 | | | | |

Notes: The table reports summary statistics for the English Short Title Catalogue (ESTC). Note that the innovation index is by definition only measured for the period 1620–1780. Additionally, the spillover index $\Omega \rightarrow \lambda$ is only defined for subset λ and the spillover index $\lambda \rightarrow \Omega$ is only defined for subset Ω .

TABLE 9: List of all subject classes in

| Subject classes |
|---|
| Administrative, Alchemy, Almanacs, Amusements, Antiquities, Applied physics, Architecture, Art, Astrology, Astronomy, Biography, Biology, Chemistry, Church administration, Classical education, Curiosities and wonders, Drama, Economic societies, Economics, Education, Encyclopedias and dictionaries, Exploration, Foreign languages, Geography, History, Legal, Mathematics, Medicine, Mercantile, Military, Military Wars, Moral tales, Music, Navigation, Philosophy, Poetry, Political philosophy, Printing and book trades, Prophecies, Religious, Religious Catholicism, Religious Judaism, Religious Sects, Religious Sermons, Scientific instruments, Societies, State affairs, Stories, Supernatural, Technical instructions Agriculture, Technical instructions Trades, Travel descriptions, University learning, University matters |

Notes: Subject classes predicted with the model from [Koschnick \(2025\)](#) to align with ESTC subject classes.

TABLE 10: Summary statistics variables in patents

| | Mean | Std.Dev. | Min | Max | Obs |
|--|---------|----------|----------|----------|-------|
| Innovation index | 1.0444 | 0.3103 | -20.1120 | 6.1440 | 7704 |
| Patent citations | 1.7805 | 1.1901 | 1.0000 | 21.0000 | 2027 |
| Spillover $\Omega \rightarrow \lambda$ | -0.7968 | 0.3553 | -6.9299 | -0.0524 | 3828 |
| Spillover $\lambda \rightarrow \Omega$ | -0.8179 | 0.3846 | -7.4516 | -0.1117 | 3836 |
| Word count | 18.6138 | 18.0181 | 3.0000 | 472.0000 | 11755 |
| Observations | 11755 | | | | |

Notes: The table reports summary statistics for patents. Note that the innovation index is by definition only measured for the period 1620–1780. Additionally, the spillover index $\Omega \rightarrow \lambda$ is only defined for subset λ and the spillover index $\lambda \rightarrow \Omega$ is only defined for subset Ω . Patent citations obtained from [Nuvolari and Tartari \(2011\)](#).

 TABLE 11: Distribution of subject classes in *Lexicon technicum*, split by propositional and prescriptive knowledge

| Type of knowledge | Subject class | Number entries |
|-------------------------|------------------------------------|----------------|
| Propositional knowledge | Mathematics | 687 |
| | Astronomy | 242 |
| | Applied physics | 60 |
| | Chemistry | 60 |
| Prescriptive knowledge | Technical instructions Trades | 94 |
| | Navigation | 37 |
| | Scientific instruments | 12 |
| | Technical instructions Agriculture | 9 |

Notes: Subject classes predicted with the model from [Koschnick \(2025\)](#) to align with ESTC subject classes.

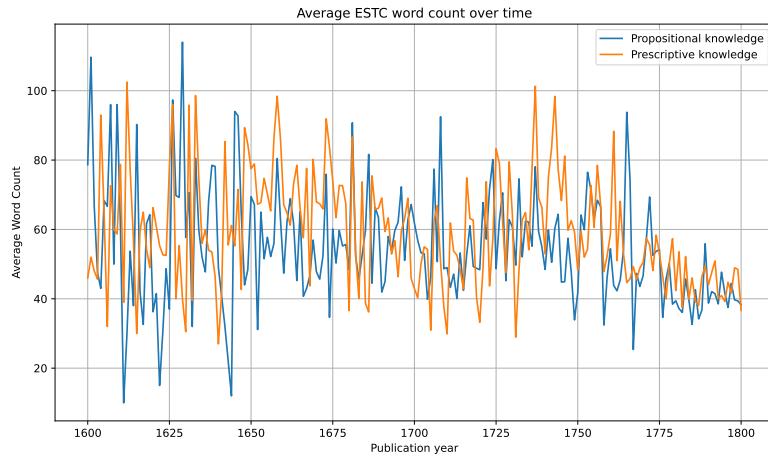


FIGURE 15: Average of word-count of ESTC titles over time

Notes: Titles classified as propositional knowledge are defined as being included in the following subject classes: *Applied physics, chemistry, mathematics, astronomy, encyclopedias and dictionaries*. Prescriptive knowledge is defined over the subject classes of *Technical publications in trades and industry, technical publications in agriculture, Navigation, and Scientific instruments*. Subject classes are taken from Koschnick (2025).

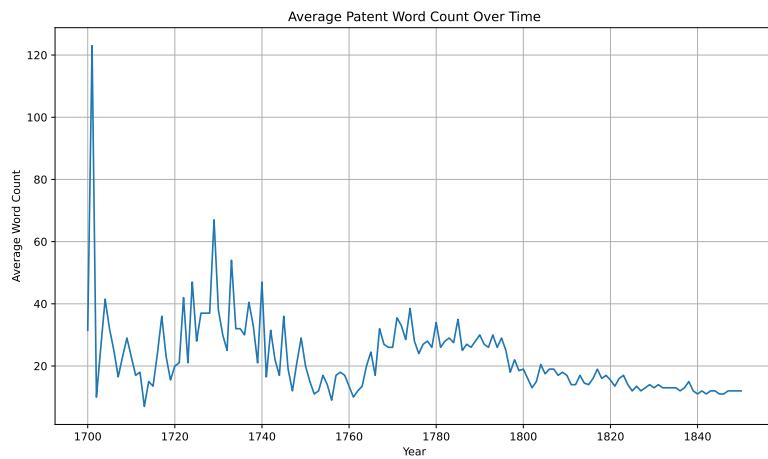


FIGURE 16: Average of word-count of short patent descriptions over time

Notes: Patent descriptions taken from Woodcroft (1854a).

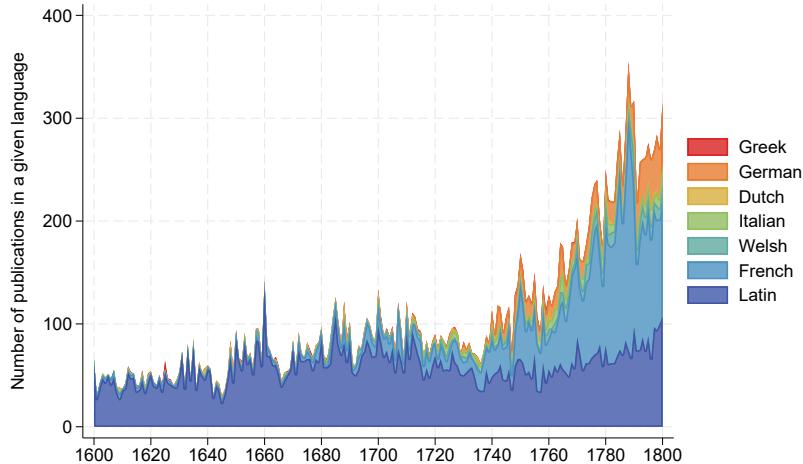


FIGURE 17: ESTC titles published in foreign languages over time

Notes: Number of titles in the English Short Title Catalogue published in non-English languages. Figure reproduced from Koschnick (2025). In Koschnick (2025), languages are identified using fasttext and GoogleTranslate.

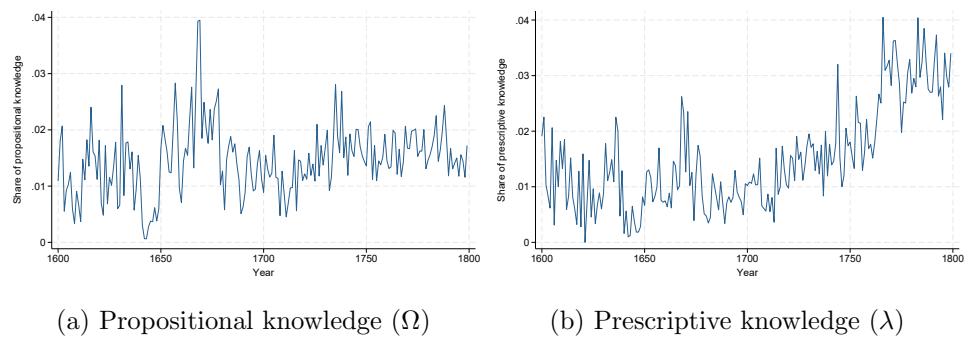


FIGURE 18: Share of propositional (Ω) and prescriptive knowledge (λ) over time

Notes: Share of propositional and prescriptive knowledge as the share of all published ESTC titles and patents. Propositional knowledge is defined as the set of titles in the fields of *applied physics, astronomy, mathematics, chemistry, and encyclopedias*. Prescriptive knowledge is defined as the set of titles in the fields of *technical publications, navigation, scientific instruments, and patents*.

B Fine-tuning procedure

For fine-tuning, the paper uses a SimCSE approach. Intuitively, SimCSE works by treating each sentence as its own training signal: The model sees two copies of the same sentence and learns that they should be embedded close together. At the same time, all other sentences in the batch act as contrasts, so the model learns to push their embeddings farther apart. This makes the model learn a stable embedding space where sentences with similar meaning naturally cluster, even without labels.

For fine-tuning, the paper proceed as follows. First, the training sample is constructed by combining titles in scientific and technical classes from table 7 with all patents. Next, patent descriptions with fewer than three words are dropped. The paper then performs an unsupervised SimCSE on the corpus of text.

The model is trained for two epochs with a multiple negatives ranking loss, a batch size of 32, a learning rate of 2×10^{-5} , and weight decay 0.01. First, all random seeds are fixed, non-deterministic CUDA kernels disabled, and the full software environment (including `torch`, `transformers`, and `sentence-transformers` versions) are logged to ensure exact replicability across runs. The resulting fine-tuned model, *SteamBERTh*, provides domain-adapted embeddings for eighteenth-century scientific and technical language and is used throughout the paper to measure semantic similarity, and to construct spillover measures. The next appendix sections presents fine-tuning evaluative statistics.

C Fine-tuning evaluation of *SteamBERTh*

This section evaluates the performance of the fine-tuned *SteamBERTh* in comparison to the base *MacBERTh* model by relying on descriptive statistics of the embedding space.

First, table 12 reports *mean nearest neighbor cosine similarity*, *principal component variance*, and *rank overlap* of the embedding space of the training corpus for both the base *MacBERTh* and *SteamBERTh*. We can see that the embedding space of the base model shows severe signs of anisotropy, while the fine-tuned model is significantly more isotropic. Moreover, the k-rank comparisons show that the embedding spaces from both models are highly different.

As further illustration of the efficacy of the fine-tuning process, figure 19–20 shows a UMAP embeddings space projection of the embedding space with subject class clusters denoted by different colors. We observe that the fine-tuning procedure significantly changes the positioning of subject classes and industry groups within the embedding space. Additionally, nearest neighbour cosine similarities by industry/subject class from table 21 show that the fine-tuning increased the semantic coherence within groups.

Overall, we find that the fine-tuned *SteamBERTh* models seems to have higher internal

consistency for the ESTC and patent data employed in this analysis.

TABLE 12: Evaluative statistics

| Metric | Base MacBERTh model | Fine-tuned model |
|---------------------------|---------------------|------------------|
| Mean NN-cosine similarity | 0.974 | 0.655 |
| PC-1 variance | 0.198 | 0.033 |
| Rank overlap $k = 5$ | – | 0.119 |
| Rank overlap $k = 10$ | – | 0.071 |

Notes: The table reports a range of evaluative statistics comparing the embedding space for patents and ESTC training classes of the base and fine-tuned data. Mean NN-cosine reports the nearest neighbour cosine similarity and pc1-variance reports the fraction of total variance in the embedding matrix explained by the first principal component. Lower values in NN-cosine similarity and indicate a less collapsed or anisotropic embedding space. Likewise, the lower PC-1 variance indicates a more isotropic embedding space. Rank overlap measures are reported for top-5 and top-10 neighbors. The values indicate systematic changes in the embedding space after fine-tuning.

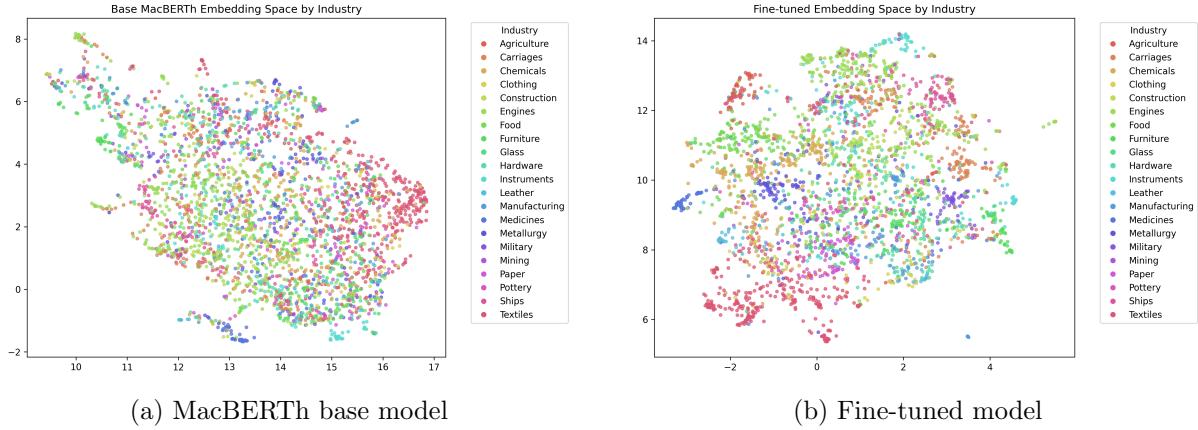


FIGURE 19: Patents UMAP embeddings space projection by industry

Notes: The figure compares a UMAP (Uniform Manifold Approximation and Projection) of the embedding space for patents between the MacBERTh base model and the fine-tuned SteamBERT model. Different colors indicate industry classes from [Nuvolari and Tartari \(2011\)](#).

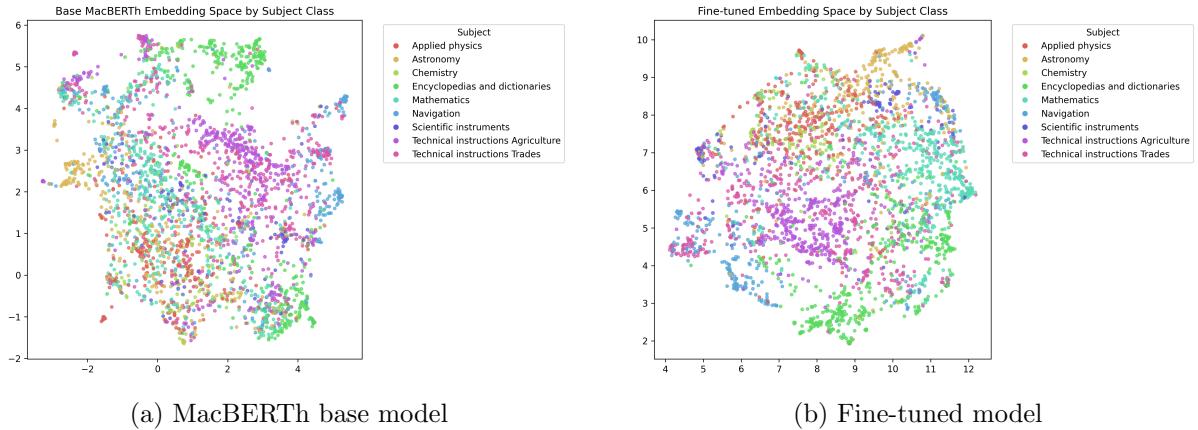


FIGURE 20: ESTC title UMAP embeddings space projection by subject class

Notes: The figure compares a UMAP (Uniform Manifold Approximation and Projection) of the embedding space for English Short Title Catalogue (ESTC) titles between the MacBERTh base model and the fine-tuned SteamBERT model. Different colors indicate subject classes from [Koschnick \(2025\)](#).

FIGURE 21: Nearest neighbor cosine scores for base and fine-tuned models

| Metric | Base MacBERTh model | Fine-tuned model |
|-----------------|---------------------|------------------|
| <i>Patents:</i> | | |
| 5-NN accuracy | 0.468 | 0.599 |
| <i>ESTC:</i> | | |
| 5-NN accuracy | 0.619 | 0.650 |

Notes: The table reports nearest neighbor cosine scores for base and fine-tuned models as measures of the semantic coherence within industry/subject class groups. Higher scores indicate increased semantic coherence within groups.

D Validation of innovation measure

This section addresses the question whether using SteamBERTh, rather than BERT models that are trained on modern data, produces efficiency gains in capturing historical patterns of innovation. To test this, we formulate the following empirical model:

$$\text{Patent citations}_{it} = \beta \text{Innov meas}_{it} + X'_{it}\zeta + \alpha_t + \varepsilon_{it} \quad (11)$$

The dependent variable is the count of patent citations, as an established measure of patent innovativeness in the literature (Trajtenberg, 1990; Nuvolari and Tartari, 2011; Billington, 2021). The main independent variable, Innov meas_{it} , denotes the innovation measure from equation 3 for patent i at year t . For calculating the innovation index from equation 3, we use the embedding space from three BERT text similarity models for comparison: a) *SteamBerth* from section 5 b) *all-MiniLM-L6-v2*, the most downloaded sentence similarity model on HuggingFace, and c) *paraphrase-mpnet-base-v2* (Reimers and Gurevych, 2019), a popular larger $109M$ parameter model. The model further includes the level and quadratic value of patent word counts i , captured by X'_{it} , as well as year fixed effects α_t .

Table 13 reports estimates for the time period of 1720–1800, which corresponds to the sample period of the combined ESTC and patent data.⁴³ Appendix table 14 further presents results for nineteenth century patents, 1800–1841. We find that for the period of 1720–1800, the *SteamBERTh*-based innovation measure has a higher R^2 than the two models trained on contemporary data. Additionally, in a horse race specification (columns 4—5), including the *SteamBERTh* measure drives the coefficients on the other innovation measures close to zero, leaving *SteamBERTh* as the only statistically significant predictor. This suggests that *SteamBERTh* not only reduces historical bias but also outperforms modern models when applied to eighteenth century data.

In contrast, table 14 shows that for the nineteenth century, the two models trained on modern data outperform *SteamBERTh*. Since both the training data for *MacBERTh* as well as the fine-tuning training data for *SteamBERTh* place a heavy weight on the early modern period, this is not surprising. The results suggest that historically fine-tuned models can outperform contemporary LLMs for specific time periods, but perform less well outside their historical training period.

⁴³Given that the patent data starts in 1700 and the innovation index is by definition not defined for the first twenty years of the sample, we get 1720 as a starting year.

TABLE 13: Model performance, innovation index and historical patent citations

| | Log Woodcroft patent citations (1720–1799) | | | | |
|---|--|----------------------|------------------|---------------------|---------------------|
| | (1) | (2) | (3) | (4) | (5) |
| | Pat. cit. | Pat. cit. | Pat. cit. | Pat. cit. | Pat. cit. |
| <i>Historical model:</i> | | | | | |
| Log SteamBERT-based innovation index | | 0.217*** (0.0779) | | 0.211** (0.0912) | 0.229** (0.0938) |
| <i>Comparison — contemporary models:</i> | | | | | |
| Log all-MiniLM-L6-v2-based innovation index | | 0.167* (0.0987) | | 0.0157 (0.116) | |
| Log paraphrase-mpnet-base-v2-based innovation index | | | 0.193 (0.132) | | -0.0394 (0.159) |
| Word count controls | Yes | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes |
| Observations | 1838 | 1838 | 1838 | 1838 | 1838 |
| R-squared | 0.109 | 0.106 | 0.106 | 0.109 | 0.109 |

Notes: The table shows coefficients from estimating equation 11 via OLS for the sample period 1720–1799. The dependent variable is patent citations from Nuvolari and Tartari (2011). The main explanatory variables is the innovation measure from equation 3, with text similarity calculated with three different BERT models, *SteamBERT*, *all-MiniLM-L6-v2* and *paraphrase-mpnet-base-v2* (Reimers and Gurevych, 2019). The innovation measure is calculated for the full sample of patents, 1700–1851. Because the innovation measure mechanically needs a 20 year pre- and post-period for comparison, it is calculated for the period 1720–1841. Results for the sample period 1800–1841 are reported in appendix figure 14. The model contains year fixed effects and controls for the level and quadratic value of patent word counts. Robust standard errors in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

TABLE 14: Model performance, innovation index and historical patent citations, 1800–1841

| | Log Woodcroft patent citations (1800–1841) | | | | |
|---|--|----------------------|----------------------|----------------------|---------------------|
| | (1) | (2) | (3) | (4) | (5) |
| | Pat. cit. | Pat. cit. | Pat. cit. | Pat. cit. | Pat. cit. |
| <i>Historical model:</i> | | | | | |
| Log SteamBERT-based innovation index | 0.171*** (0.0417) | | | 0.0771 (0.0507) | 0.0775 (0.0500) |
| <i>Comparison — contemporary models:</i> | | | | | |
| Log all-MiniLM-L6-v2-based innovation index | | 0.322*** (0.0693) | | 0.246*** (0.0846) | |
| Log paraphrase-mpnet-base-v2-based innovation index | | | 0.462*** (0.0973) | | 0.361*** (0.117) |
| Word count controls | Yes | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes |
| Observations | 6684 | 6684 | 6684 | 6684 | 6684 |
| R-squared | 0.177 | 0.178 | 0.178 | 0.178 | 0.179 |

Notes: The table shows coefficients from estimating equation 11 via OLS for the sample period 1800–1841. The dependent variable is patent citations from Nuvolari and Tartari (2011). The main explanatory variables is the innovation measure from equation 3, with text similarity calculated with three different BERT models, *SteamBERT*, *all-MiniLM-L6-v2* and *paraphrase-mpnet-base-v2* (Reimers and Gurevych, 2019). The innovation measure is calculated for the full sample of patents, 1700–1851. Because the innovation measure mechanically needs a 20 year pre- and post-period for comparison, it is calculated for the period 1720–1841. The model contains year fixed effects and controls for the level and quadratic value of patent word counts. Robust standard errors in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

TABLE 15: Top-20 innovative titles in applied physics

| Author | Publication year | Title |
|---------------------------------------|------------------|---|
| Marshall, William, Engraver. | 1640 | [The four elements. 4 engraved plates.] Earth. (Aire. Fire. Water.) |
| Powell, Thomas, 1608-1660. | 1651 | Elements of optics: a new, easy, and convenient method developed. With several drafts (for fuller elucidation) attached to the footnotes. |
| Burgersdijk, Franco Petri, 1590-1635. | 1650 | Physical college, debates 32 absolute; the whole natural philosophy problem has to be put forward. Author M. Francone Burgersdicio, Professor of Philosophy With the syllabus of the discussions, the nominal names of the respondents. |
| Boyle, Robert, 1627-1691. | 1661 | Some physiological essays at different times and occasions were written by Robert Boyle... translated from English to Latin. |
| Bacon, Francis, 1561-1626. | 1620 | Francis de Verulamius, Chancellor of England, Great Instauration |
| Bacon, Francis, 1561-1626. | 1620 | The great establishment of Francis de Verulam, chancellor of England supreme. |
| — | 1660 | An excerpt out of a book, shewing, that fluids rise not in the pump, in the siphon, and in the barometer, by the pressure of the air, but propter Fugam vacui. At the occasion of a dispute, in a coffee-house, with a doctor of Physick. |
| Boyle, Robert, 1627-1691. | 1661 | Certain physiological essays, written at distant times, and on several occasions: by the Honourable Robert Boyle. |
| Bradshaw, Ellis. | 1649 | A nevv and cleer discovery, of the true, and proper, natural cause, of the ebbing and flowing of the main sea: Convincingly held forth, both from Scripture and reason: so as any rational man, may easily apprehend, the proper cause on its flunct [sic] motion: and that it is not the Moon, as some have imagined, and gone about to prove. Written by Ellis Bradshawe of the parish of Boulton in the county of Lancaster, husbandman. |
| Hobbes, Thomas, 1588-1679. | 1662 | Physical Problems Of gravity. Cap.I. Of the marine tides Chapter II Of the vacuum. Chapter III Of heat and light. Chapter IV Of hard and soft. Cap.V. Of rain, wind, and other varieties of heaven. Chapter IV VI. Of the types of motions Chapter IV VII. Two propositions are also adjoined concerning the duplication of the cube and the dimension of the circle. |
| Kepler, Johannes, 1571-1630. | 1653 | John Kepler ... Dioptrice: or, a demonstration of those things which happen to the sight and sight for the sake of the sights, not so long ago. The prefaced Epistle of Galilee to the gods, which, after the publication of the starry news, through the aid of an eye-witness, were discovered in heaven, new and admirable. Also an examination of the preface of John Pena of Gaul in the optics of Euclid, on the use of optics in philosophy. |
| Boyle, Robert, 1627-1691. | 1660 | New experiments physico-mechanicall, touching the spring of the air, and its effects, (made, for the most part, in a new pneumatical engine) written by way of letter to the Right Honorable Charles Lord Vicount of Dungarvan, eldest son to the Earl of Corke. By the Honorable Robert Boyle Esq; |
| Descartes, René, 1596-1650. | 1649 | A discourse of a method for the wel-guiding of reason, and the discovery of truth in the sciences. Being a translation out of that famous philosopher Renaldus Des Cartes. |
| Line, Francis, 1595-1675. | 1661 | A treatise on the inseparability of bodies; in which experiments concerning the vacuum, both Torricellia and Magdeburgica and Boyliana, are examined, and it is shown that, by discovering their true cause, it is impossible that a vacuum can be naturally Whence also the Aristotelian opinion on rarefaction is demonstrated, both against the theorists of emptiness and of corpuscles. There was added the most difficult solution of that Aristotle's problem concerning two wheels; which, although, are very unequal, yet they describe equal orbits. By Francisco Linus |
| Gassendi, Pierre, 1592-1655. | 1653 | The astronomical institution of Peter Gassendi, according to the hypotheses of both ancients and moderns. To him came Galileo, Galileo's starry proclamation, and John Kepler's Dioptric. |
| Hobbes, Thomas, 1588-1679. | 1661 | A natural dialogue, or a conjecture about the nature of air, taken from experiments recently held in London at Gresham College and also about the duplication of a cube. Author of Tho: Hobbes Malmesb |
| Hooke, Robert, 1635-1703. | 1661 | An attempt for the explication of the phænomena observable in an experiment published by the Honourable Robert Boyle, Esq; in the XXXV. experiment of his epistolical discourse touching the aire. In confirmation of a former conjecture made by R.H. |
| Ross, Alexander, 1591-1654. | 1645 | The philosophicall touch-stone: or Observations upon Sir Kenelm Digbie's Discourses of the nature of bodies, and of the reasonable soule. In which his erroneous paradoxes are refuted, the truth, and Aristotelian philosophy vindicated, the immortality of mans soule briefly, but sufficiently proved. And the weak fortifications of a late Amsterdam ingeneer, patronizing The soules mortality, briefly slighted. By Alexander Ross. |
| Charleton, Walter, 1619-1707. | 1654 | Physiologia Epicuro-Gassendo-Charltoniana: or A fabrick of science natural, upon the hypothesis of atoms, founded by Epicurus, repaired [by] Petrus Gassendus, augmented [by] Walter Charleton, Dr. in Medicine, and physician to the late Charles, monarch of Great-Britain. The first part. |
| — | 1663 | A short compendium of the new and much enlarged sea-book, or, Pilots sea-mirror, containing the distances and thwart courses of the Eastern, Northern, and VWestern navigation; also the courses and distances of the straits, or Mediterranean seas. With the tide-tables, and the full and change of the moon, for eight years. Newly enlarged and amended, by several experienced navigators. And now for the benefit and encouragement of our sea-men, translated into English; and calculated according to 20 leagues for a degree. By L. Childe, Esq; |

Notes:

E Comparing similarity functions

This section validates the following two similarity functions for the calculation of the innovation index from section 4.2.

Average similarity index:

$$f(v_{it}, Y) = \frac{1}{n} \sum_{y' \in Y} sim(v_{it}, y)$$

K-top similarity index:

$$f(v_{it}, Y; k) = \frac{1}{k} \sum_{y \in \text{Top}_k \{sim(v_{it}, y') : y' \in Y\}} sim(v_{it}, y), \quad k := \min\{k, |Y|\}$$

Using both measures, we create two innovation indices and plot their distribution in figure 22. We see that the average similarity index is significantly more likely to produce extreme values, while the k-top index is well-centered around 1 (see figure 23).

The problem of extreme values can be mitigated by using the natural logarithm, see figure 24. However, note that for the average similarity index values equal to or below zero had to be excluded.

Next, we test the performance of the innovation indices as predictors of patent citations using the model from equation 11 from section 5.2. We see that the innovation index based on the k-top similarity function outperforms the innovation index based on the average similarity function. The innovation index based on the k-top similarity function has a larger R-squared value. Additionally, its coefficient is a significant predictor of patent citations, while the innovation index based on the average similarity function is insignificant within this setting.

Overall, the results of this section indicate that the k-top similarity function yields an innovation index with a tighter distribution and more predictive power for the test case of patent citations. Therefore, for the other validation exercises as well as for the main empirical framework in section 6, the paper adopts the k-top innovation index for the calculation of innovation and knowledge spillovers.

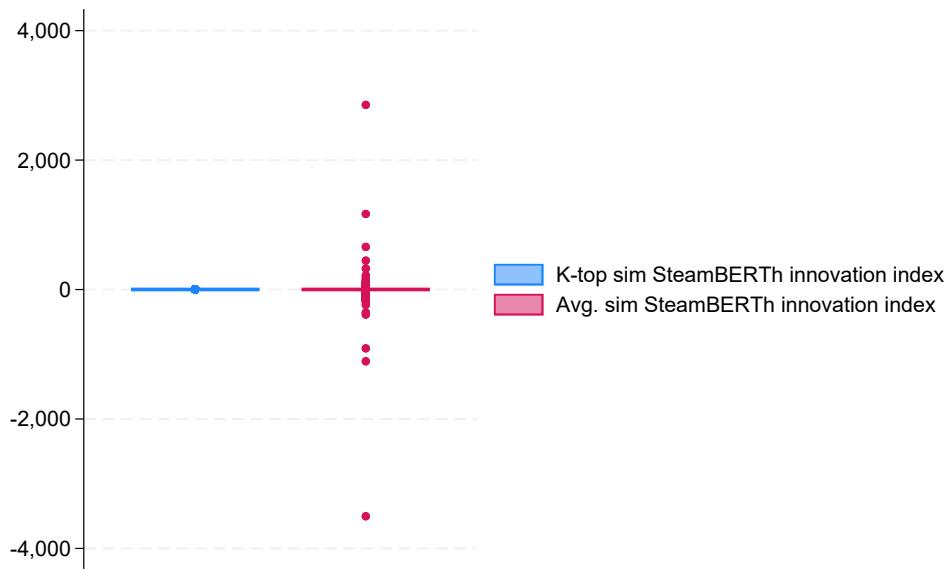


FIGURE 22: Boxplot of innovation index using (1) the k-top and (2) the average similarity function

Notes: The k-top similarity function is defined in equation 6 in section 4.2. Throughout the paper, we use the k-top similarity function for the calculation of the innovation index.

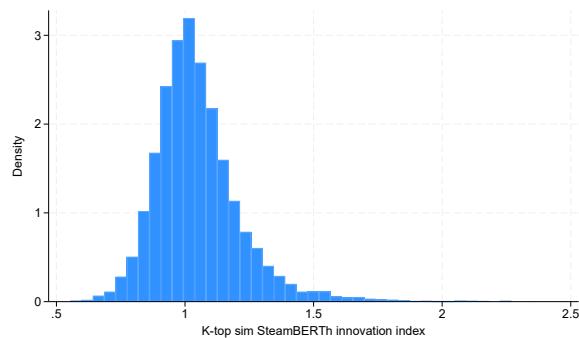
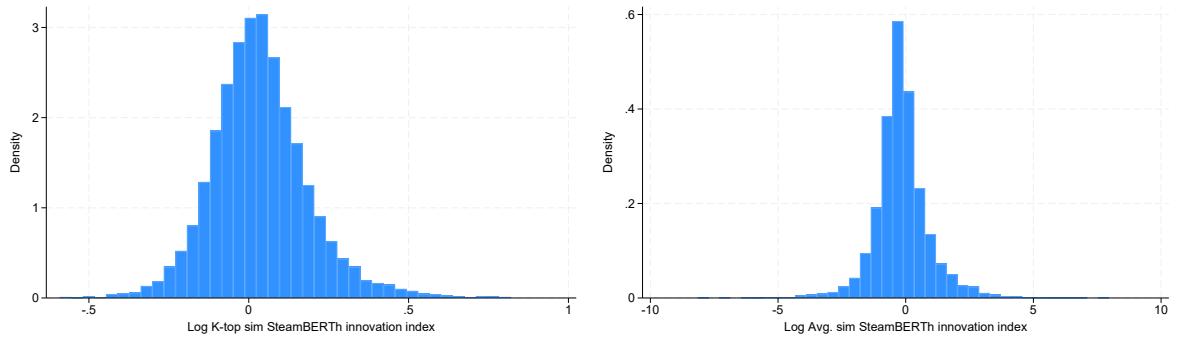


FIGURE 23: Histogram of level of innovation index using the k-top similarity function

Notes: The k-top similarity function is defined in equation 6 in section 4.2.



(a) Log innovation index using the k-top similarity function (b) Log innovation index using the average similarity function

FIGURE 24: Histogram of natural logarithm of the innovation index using (1) the k-top and (2) the average similarity function.

Notes: The k-top similarity function is defined in equation 6 in section 4.2. Throughout the paper, we use the k-top similarity function for the calculation of the innovation index.

TABLE 16: Predicting patent citations: Comparison of innovation index using (1) the k-top and (2) the average similarity function

| | Log Woodcroft patent citations (1720–1799) | |
|---|--|--------------------|
| | (1) | (2) |
| | Pat. cit. | Pat. cit. |
| Log K-top sim SteamBERTH innovation index | 0.176** (0.0780) | |
| Log Avg. sim SteamBERTH innovation index | | 0.0304 (0.0212) |
| Word count controls | Yes | Yes |
| Year fixed effects | Yes | Yes |
| Observations | 1787 | 1787 |
| R-squared | 0.111 | 0.110 |

Notes: The table shows coefficients for estimating equation 11 via OLS. The dependent variable is patent citations from Nuvolari and Tartari (2011). The main explanatory variables is the innovation measure from equation 3 using (1) the k-top and (2) the average similarity function. The model contains year fixed effects and controls for the level and quadratic value of patent word counts. The model is estimated for the sample period of 1720–1799. Applying the natural logarithm to values below zero for the innovation index based on the average similarity function leads to excluded observations. To make the models comparable, we apply the same sample restriction to the model in column 1 and 2. Robust standard errors in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

F Unsupervised clustering of sub-topics

In order to find unsupervised clusters for ESTC titles and *Lexicon technicum* entries, we formulate the following three aims

1. Clusters should be grouped by semantic similarity, not by similarity in style
2. Clusters should identify similar groups on the embedding space, even when the number of clusters varies slightly
3. Cluster size should be as detailed as possible, yet not exceed ca. 15 sub-clusters per subject to leave sufficient within cluster variation for estimation
 - (a) Likewise the number of sub-clusters should vary dynamically with the size of the subject class

In order to achieve these aims, the paper implements the following HDBSCAN-based approach, similar to Grootendorst (2022):

First, titles in the ESTC are clustered by subject class using a deterministic HDBSCAN procedure. Embeddings are derived from using the fine-tuned *SteamBERTh* model. In order to remove noise, the paper first reduces dimensionality using a principal component analysis (PCA) with full SVD and 50 components.⁴⁴ The resulting embeddings are L2-normalized so that Euclidean distance equals cosine similarity. HDBSCAN is then applied with Euclidean distance, with an adaptive `min_cluster_size` equal to $\max(6, \gamma \times N_{\text{titles}})$ (with $\gamma=0.015$) to maximize coverage while avoiding over-fragmentation. Noise points are deterministically re-assigned using `approximate_predict` if their posterior probability exceeds 0.15, and residual outliers are absorbed into the nearest centroid when cosine similarity exceeds 0.4. For each resulting topic, keywords are extracted using a CountVectorizer and TF-IDF weighting on aggregated term counts, yielding interpretable topic labels (see table 17).

Next, the paper assigns entries from the *Lexicon technicum* to the same ESTC sub-classes. For each subject class, entries are first encoded using *SteamBERTh* and reduced via the subject-specific PCA fitted on the ESTC titles. The normalized embeddings are then compared to the normalized centroids of the ESTC topics. Each entry is assigned to the nearest centroid based on cosine similarity, producing one-to-one mappings between ESTC sub-topics and *Lexicon technicum* entries. This procedure allows measuring the overlap of eighteenth-century printed topics with early scientific encyclopedic knowledge, ensuring consistent topic spaces across both corpora.

To ensure reproducibility, all sources of non-determinism were disabled across Python, BLAS, and CUDA backends, with fixed seeds, single-thread execution, and a stable input ordering. Texts were normalized, encoded using the *SteamBERTh* model from section 5, and converted into contiguous, rounded 64-bit vectors to eliminate floating-point drift.

While this method offers a state-of-the-art approach of assigning subject topics based on the topology of the embedding space (Grootendorst, 2022), it remains important to test the robustness of the results when changing hyperparameters. Figure 25 reports results when changing PCA dimensions (effectively influencing the granularity of clusters), changing, and minimum cluster size parameter γ , as well as HDBSCAN post labeling (post) and minimum cosine similarity for centroid adoption (absorb). Additionally, figure 25 reports coefficients when changing the maximum number of entries of topics in λ in the *Lexicon technicum* for the clean sample criterion (see section 8). Panel a) reports results for continuous treatment and panel b) reports results for binary treatment. Overall, the vast majority of coefficients remain stable between specifications and significant. Only rare single combinations as maximum number of Lexicon

⁴⁴Starting with a PCA-dimensionality reduction, helps to ensure stability across difference cluster specification by removing noise from the original embeddings.

entries > 10 and $\gamma = 0.1$, lead to a large reduction in the coefficient. Figure 26 shows that the majority of the results have a smaller p-value than the baseline, with only a small number of coefficients above 0.1. Overall, these results indicate relative stability to the specification of hyperparameters.

TABLE 17: Sub-clusters per subject class

TABLE 18: HDBSCAN sub-topics in ESTC by subject, document counts, and TF-IDF labels

| Subject class | Sub-class | # ESTC docs | Labels (TF-IDF) |
|-----------------|-----------|-------------|---|
| Mathematics | 0 | 16 | analysis, amp, mercury, game, theodosius, fluxes, various, art, trade, algebra |
| Mathematics | 1 | 10 | collection, collection mathematical, thomas simpson, simpson, mathematical, containing, essays, choice, variety, mathematica |
| Mathematics | 2 | 62 | use, arithmetic, tables, plain, logarithms, questions, rules, containing, numbers, trigonometry |
| Mathematics | 3 | 18 | professor, professor mathematics, university, mathematics, edinburgh, arithmetic, book keeping, keeping, stirling, aberdeen |
| Mathematics | 4 | 21 | analyst, defence, letter, philalethes, philalethes cantabrigiensis, author analyst, cantabrigiensis, author, mathematicians, fluxions |
| Mathematics | 5 | 28 | professor, dr, elements, isaac, professor mathematics, archimedes, added, mathematics, royal, theorems |
| Mathematics | 6 | 15 | table, table artificial, 10, artificial sines, secants, sines tangents, sines, unit, logarithms numbers, artificial |
| Mathematics | 7 | 23 | trigonometry, spherical, plain, elements, spherical trigonometry, sphere, plain spherical, sailing, plano, sphere plano |
| Mathematics | 8 | 45 | arithmetic, decimal, vulgar, vulgar decimal, fractions, use, arithmetic, writing, method, new |
| Mathematics | 9 | 40 | arithmetic, cocker, decimal, john, corrected, edition, new, plain, french, fractions |
| Astronomy | 0 | 6 | neighboring, birth lord, year gracious, lord heyland, heyland jesus, 40 degrees, american calender, neighboring countries, german, high german |
| Astronomy | 1 | 13 | orrery, mr, william, curious, machine, william hudson, hudson, shop, description, half |
| Astronomy | 2 | 13 | geography, navigation, astronomy, useful, globes, uses, introduction, use, astronomy geography, easie |
| Astronomy | 3 | 10 | newton, 39, isaac newton, isaac, book isaac, public schools, newton translated, translated english, cambridge public, newton 39 |
| Astronomy | 4 | 9 | noon, equation, clock, table, rules, day year, days, time, day, adjusting |
| Astronomy | 5 | 15 | dissertation, theory earth, theory, sykes, earth, defence, mr, dr burnet, mentioned phlegon, dissertation eclipse |
| Astronomy | 6 | 27 | astronomy, royal, professor, university, fellow royal, royal society, fellow, society, professor astronomy, course |
| Astronomy | 7 | 10 | natural philosophy, mathematical demonstrations, philosophy notes, containing mathematical, compendious natural, compendious, notes containing, notes, philosophy, demonstrations |
| Astronomy | 8 | 24 | hour, dials, dialling, declination, dyals, sun, lines, latitude, sorts, calculated |
| Astronomy | 9 | 13 | sun, moon, new, tables, places, earth, place, observations, motions, assimil |
| Astronomy | 10 | 14 | motion, nature, essay, sea, comets, properties, divine, causes, earth, cause |
| Astronomy | 11 | 7 | deluge, origin, world, theory, method, theory earth, old, order, earth, way |
| Astronomy | 12 | 16 | year, leap, leap year, year lord, prognostication, prognostication year, edinburgh, new prognostication, lord, city |
| Astronomy | 13 | 22 | eclipse, great, happen, years, eclipses, sun, visible, year, london, remarkable |
| Astronomy | 14 | 17 | eclipse, passage, eclipse sun, total, sun, great eclipse, shadow, passage shadow, decription passage, great |
| Applied physics | 0 | 6 | dr, nobility gentry, moral, mineral, political moral, unveil, drink, drink mineral, epson, waters |
| Applied physics | 1 | 7 | list, council, society, royal society, royal, list royal, navigators, list present, request, council continued |
| Applied physics | 2 | 7 | illustrious society, acta, society, memoirs, copper, plates, copper plates, illustrious, illustrated copper, germany |
| Applied physics | 3 | 7 | papers, year, transactions, philosophical transactions, general heads, heads, abridged, philosophical, year 1732, end year |
| Applied physics | 4 | 11 | dr, remarks, reply, distinct, indistinct vision, indistinct, essay distinct, distinct indistinct, morgan, mr |
| Applied physics | 5 | 9 | experiments, perform, course, experiments perform, explain, hauksbee, hydrostatal pneumatical, lectures, pneumatical, hydrostatal |

| Subject class | Sub-class | # ESTC docs | Labels (TF-IDF) |
|-------------------------------|-----------|-------------|---|
| Applied physics | 6 | 10 | light, colours, reflections, inflections, newton, isaac newton, isaac, books author, algarotti, colors light |
| Applied physics | 7 | 14 | weather, wind, years, alterations, rain, causes, sic, rational, wet, cause |
| Applied physics | 8 | 11 | rohaulti, jacob, jacob rohaulti, french, translated, des, latine, reviewed, translated reviewed, physics jacob |
| Applied physics | 9 | 9 | water, experiments, works, nature, manner, best, fountains, water works, seats, springs |
| Applied physics | 10 | 20 | philosophy, astronomy, professor, university, lectures, professor mathematicks, natural philosophy, mathematicks, cambridge, fellow royal |
| Applied physics | 11 | 15 | coll, physics, keill, reg, amp, oxford, amp reg, oxon, lessons, true physics |
| Applied physics | 12 | 9 | power, electricity, nature, experiments, properties electricity, properties, shewn, electrical power, william watson, resistance |
| Applied physics | 13 | 11 | experiments, electricity, machines, curious, pump, description, air pump, surprizing, performing, subjects |
| Chemistry | 0 | 7 | course, chemistry, compleat course, course chymistry, compleat, chymistry, gresham college, gresham, performed, course chemistry |
| Chemistry | 1 | 12 | acid, water, nature, dr, physicians, quicksilver, letter, alkali, remarks, properties |
| Chemistry | 2 | 12 | translated, chemistry, original, latin, shaw, boerhaave, notes, leyden, peter shaw, peter |
| Scientific instruments | 0 | 6 | quadrant, description, short description, ring dial, astrolabe inverted, ring, particular astrolabe, quadrant particular, astrolabe, table artificial |
| Scientific instruments | 1 | 6 | reasons, clocks, watches, clockmakers, hutchinson, invention, movement invented, stones, hutchinson property, act |
| Scientific instruments | 2 | 13 | clock, clock work, work, watch clock, watch, treatise, artificial clock, treatise watch, pendulum, clock maker |
| Scientific instruments | 3 | 7 | sea, quadrant, horizon, instrument, latitude, description use, use, description, taking, invented |
| Scientific instruments | 4 | 8 | sun, weather, time, account, day, sun place, rising setting, declination, rising, use |
| Scientific instruments | 5 | 28 | microscope, objects, london, description, reflecting, street, magnifying, cuff, new, pocket |
| Navigation | 0 | 11 | navigation, war, parliament, member, fishery, view, letter, act, seamen, trade |
| Navigation | 1 | 9 | st, church, act, parish church, city, said, saint mary, saint, church st, mayor |
| Navigation | 2 | 10 | instructions, tide table, sailing fighting, fighting, sailing, table, tide, ye, correct tide, bridge |
| Navigation | 3 | 19 | confirming, act confirming, act, esq, county, napier, agreement, common, john, lands |
| Navigation | 4 | 10 | act, thousand, reign, pounds, seven, tenth, king william, granted, reign king, thousand pounds |
| Navigation | 5 | 36 | coasts, sea, pilot, describing, harbours, sea coasts, ports, english pilot, english, coasts capes |
| Navigation | 6 | 19 | longitude, longitude sea, sea, discovery, finding longitude, discovery longitude, humbly, honourable, finding, method |
| Navigation | 7 | 14 | navigation, art, art navigation, useful, atkinson, new, tables, enlarged useful, mariner compass, compass |
| Navigation | 8 | 26 | tables, navigation, plain, use, sun, latitude, table, new, mercator, containing |
| Technical instructions Trades | 0 | 6 | mechanick, great things, mechanick engines, shewing great, doctrine, doctrine handy, handy, exercises doctrine, mechanick exercises, engines |
| Technical instructions Trades | 1 | 9 | reasons humbly, reasons, humbly offered, humbly, brandy, offered, prohibiting distilling, prohibiting, passing, corn |
| Technical instructions Trades | 2 | 13 | patent, granted, blank, king, ink, majesty, invention, making, reign, patent great |
| Technical instructions Trades | 3 | 11 | water fresh, fresh, making, water, sea, sea water, salt, making sea, walcot, art making |
| Technical instructions Trades | 4 | 9 | logarithms, tables natural, tables, differences, 000, sines, minute, minute quadrant, quadrant, natural |
| Technical instructions Trades | 5 | 11 | sheathing, lead, lead sheathing, wood, england, wood sheathing, shipwrights, advertisement, shipwrights england, plainly |
| Technical instructions Trades | 6 | 16 | shipwrights, ships, ships vessels, building, better, england, vessels, breeding, better breeding, trade |
| Technical instructions Trades | 7 | 17 | gold, silver, curiously, varnishing, glass, art, painting, colours, plates, copper plates |
| Technical instructions Trades | 8 | 10 | history, ancients, moderns, useful, curious, pleasant, metals, added, english, observations |
| Technical instructions Trades | 9 | 9 | dorigny, pleased, signior, near, pleased grant, graciously pleased, lincoln, grant, graciously, patent |
| Technical instructions Trades | 10 | 15 | wines, liquors, vintners, sorts, art, wine, art mystery, wine coopers, mystery vintners, vintners wine |
| Technical instructions Trades | 11 | 15 | candyng, art, sorts, preserving, flowers, making, newest, cookery, directions, confec- tioner |
| Technical instructions Trades | 12 | 12 | draining, fens, levels, river, level, navigation, taken, survey, level fens, john grundy |
| Technical instructions Trades | 13 | 28 | practical, measuring, gauging, edition, rule, timber, art, easy, containing, corrected |
| Technical instructions Trades | 14 | 26 | measuring, timber, land, practical, inches, surveying, new, feet, tables, chain |
| Technical instructions Trades | 15 | 21 | measuring, gauging, solids, rule, superficies, artificers, inches, carpenters, plain, art |

| Subject class | Sub-class | # ESTC docs | Labels (TF-IDF) |
|------------------------------------|-----------|-------------|--|
| Technical instructions Agriculture | 0 | 8 | logos oikonomikos, libri, logos, xenophōntos logos, influence, xenophōntos, spreading, oikonomikos, rustica libri, rustica |
| Technical instructions Agriculture | 1 | 6 | time, seeds, fruit, figures, best, ripe, flower, month, curious, seasons |
| Technical instructions Agriculture | 2 | 8 | sold, catalogue, grass, seeds, seeds sold, london, gray, sun, sold william, strand london |
| Technical instructions Agriculture | 3 | 19 | bees, bee, monarchy, written, translated, art, new, colonies, admirable, common hives |
| Technical instructions Agriculture | 4 | 8 | fruit, trees, use, pruning, kitchen, fruit trees, planting, art pruning, manner fruit, sorts |
| Technical instructions Agriculture | 5 | 18 | husbandry, society, considerations, promoting, improvement husbandry, agriculture, improvement, husbandry trade, promoting agriculture, considerations promoting |
| Technical instructions Agriculture | 6 | 15 | flax, seed, raising flax, flax seed, raising, hemp, flax hemp, directions, ireland, directions raising |
| Technical instructions Agriculture | 7 | 13 | ellis, william ellis, timber, william, hertfordshire, improved, farmer, tree improved, tree, practical |
| Technical instructions Agriculture | 8 | 27 | wheat, hertfordshire, sowing, new, invented, matters, great, author, husbandry, month |
| Technical instructions Agriculture | 9 | 12 | horse, farrier, horses, farriery, management horses, cure, journey, gentleman, diseases, pocket farrier |
| Technical instructions Agriculture | 10 | 22 | horses, farrier, diseases, distempers, breeding, cure, directions, ordering, incident, jockey |

Notes: The table lists HDBSCAN sub-topics identified within each ESTC subject class (1685–1750).

Each sub-cluster groups titles by semantic proximity in the embedding space. Labels correspond to tf-idf keywords per cluster.

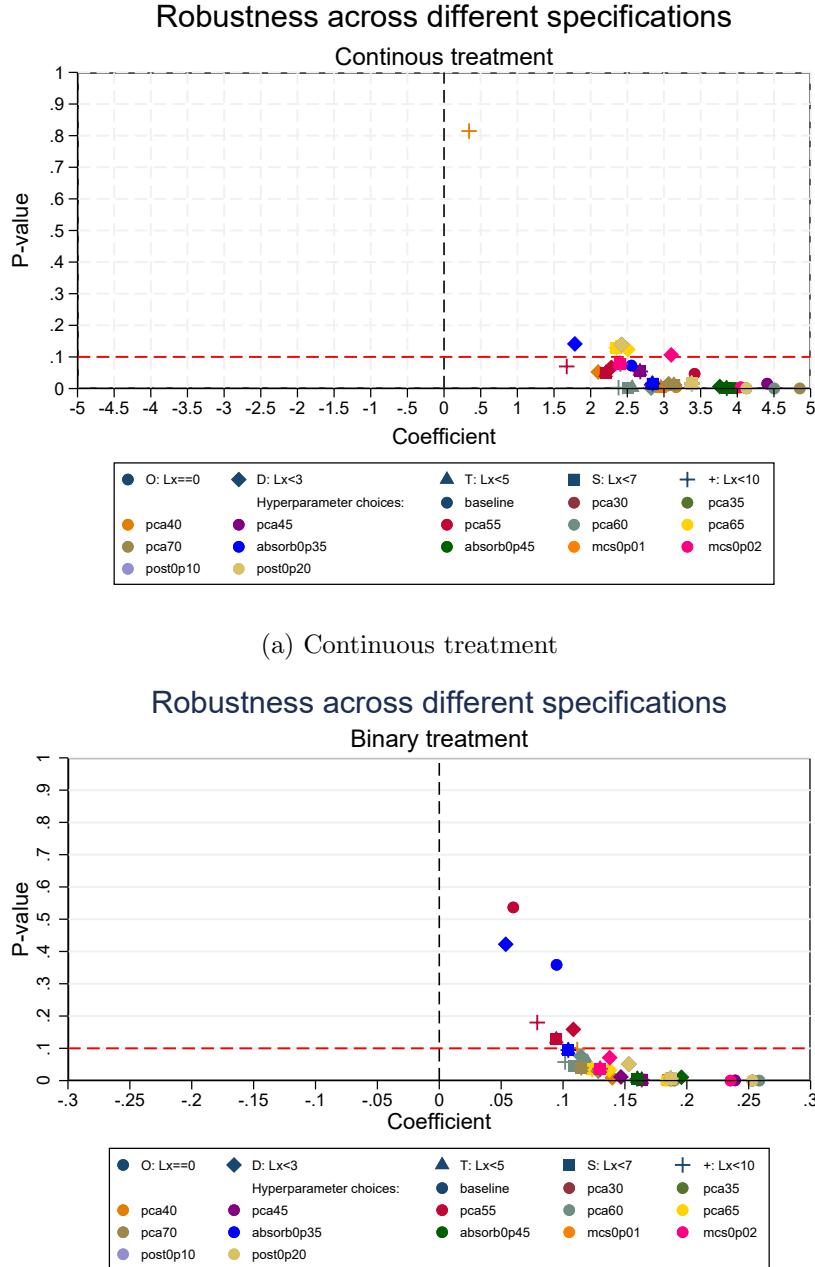


FIGURE 25: Robustness for different clustering-hyperparameters and regression specifications

Notes: The figure reports robustness to a) specifying different hyperparameters for identifying sub-clusters and b) different strictness when excluding sub-topics in λ that were also covered in the *Lexicon technicum*. First, for a) different colours denote different choices of hyperparameters for calculating spillovers. Here, we show robustness to different components in the PCA dimensionality reduction (pca), minimum cluster size parameter γ (mcs), HDBSCAN post labeling (post), and minimum cosine similarity for centroid adoption (absorb). Finally, for b), different symbols either denote a maximum of 0, 3, 5, 7, or 10 entries in the *Lexicon technicum* in λ that are tolerated for a sub-topic to be unaffected. To condense information, coefficients are reported for a 2×2 DiD model, similar to equation 10, where treatment is specified for the time frame of 1705–1725 (the period before the publication of Chambers' *Cyclopaedia*).

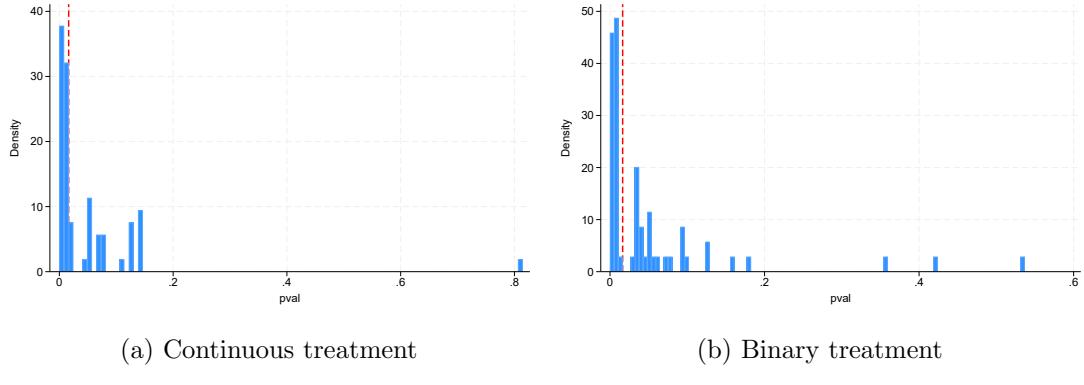


FIGURE 26: Distribution of p-values for different clustering-hyperparameters and regression specifications

Notes: The figure reports the distribution of p-values for the robustness results using the hyperparameter and regression specifications from figure 25.

G Robustness: Main results

G.1 Quantile regression for 1760–1779 coefficients

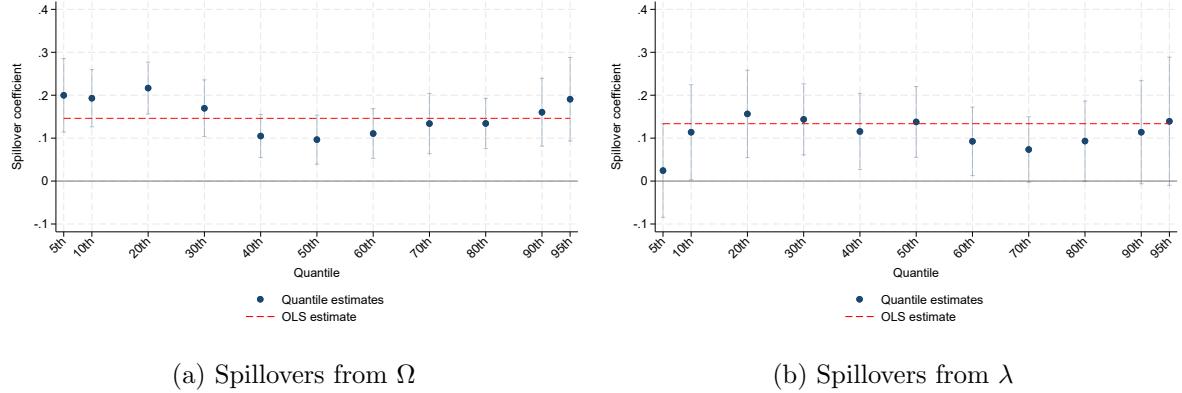


FIGURE 27: Quantile regression for 1760–1779 coefficients

Notes: The figure reports estimated coefficients for the period in the model from equation 7 using a quantile regression approach. First, the figure reports OLS coefficients corresponding to figure 10. Then it reports quantile regression coefficients for 5th, 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, 90th, and 95th quantiles. Bootstrapped standard error reported at the 90% level.

G.2 Different k and p in innovation index formula

G.2.1 $\Omega \rightarrow$ innovation in λ

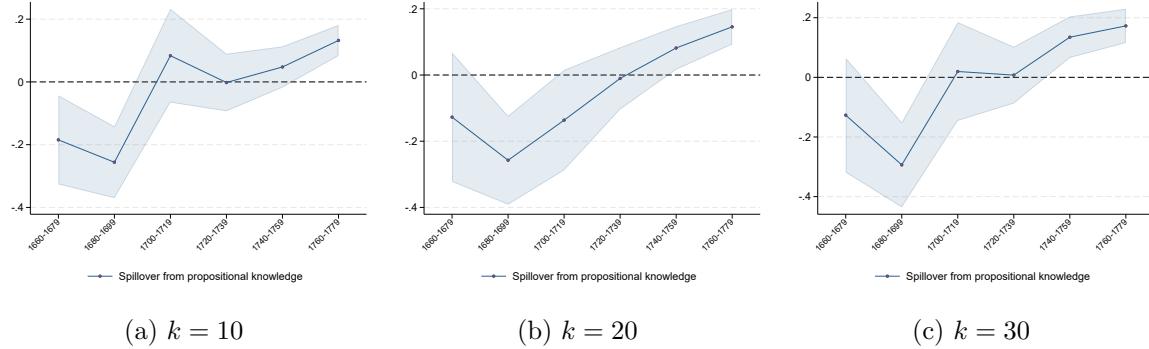


FIGURE 28: Robustness for different values for k in the calculation of the innovation index, Spillovers in $\Omega \rightarrow$ innovation in λ

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The plotted independent variable is an interaction term between the received spillover index from equation 5 and twenty year time periods. The model further controls for the word count of titles and includes year fixed effects. Panel a), b), and c) report results when changing the k parameter controlling the top- k comparisons for calculating the innovation index from equation 7. Standard errors clustered at the publication year level.

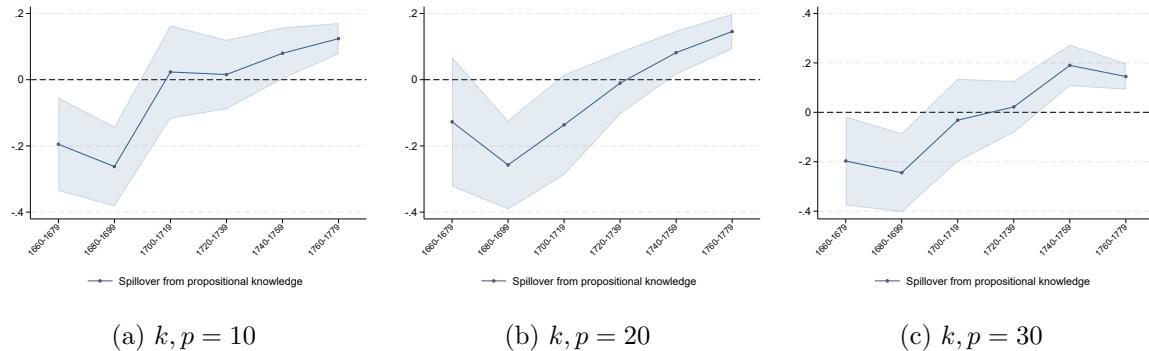


FIGURE 29: Robustness for different values for k, p in the calculation of the innovation index, Spillovers in $\Omega \rightarrow$ innovation in λ

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The plotted independent variable is an interaction term between the received spillover index from equation 5 and twenty year time periods. The model further controls for the word count of titles and includes year fixed effects. Panel a), b), and c) report results when changing the k parameter controlling the top- k comparisons for calculating the innovation index from equation 7. Paralelly, we also change the p parameter controlling the size of the within-subject backwards counterfactual group p from equation 5. Standard errors clustered at the publication year level.

G.2.2 $\lambda \rightarrow$ innovation in Ω

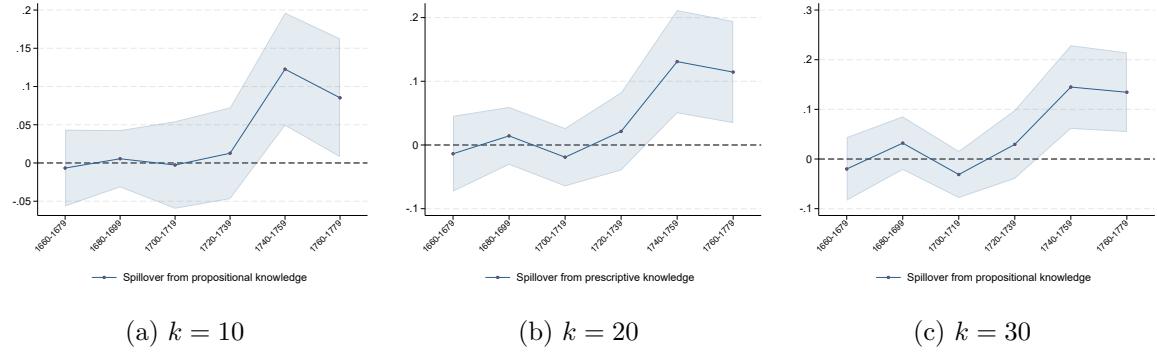


FIGURE 30: Robustness for different values for k in the calculation of the innovation index, Spillovers in $\lambda \rightarrow$ innovation in Ω

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The plotted independent variable is an interaction term between the received spillover index from equation 5 and twenty year time periods. The model further controls for the word count of titles and includes year fixed effects. Panel a), b), and c) report results when changing the k parameter controlling the top- k comparisons for calculating the innovation index from equation 7. Standard errors clustered at the publication year level.

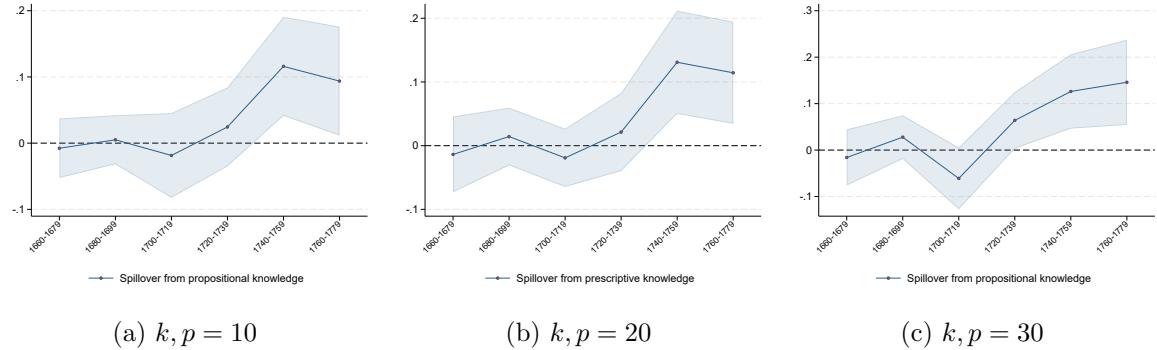


FIGURE 31: Robustness for different values for k, p in the calculation of the innovation index, Spillovers in $\lambda \rightarrow$ innovation in Ω

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The plotted independent variable is an interaction term between the received spillover index from equation 5 and twenty year time periods. The model further controls for the word count of titles and includes year fixed effects. Panel a), b), and c) report results when changing the k parameter controlling the top- k comparisons for calculating the innovation index from equation 7. Paralelly, we also change the p parameter controlling the size of the within-subject backwards counterfactual group p from equation 5. Standard errors clustered at the publication year level.

G.3 Longer t, τ time windows for calculation of innovation and spillover index

G.3.1 $\Omega \rightarrow$ innovation in λ

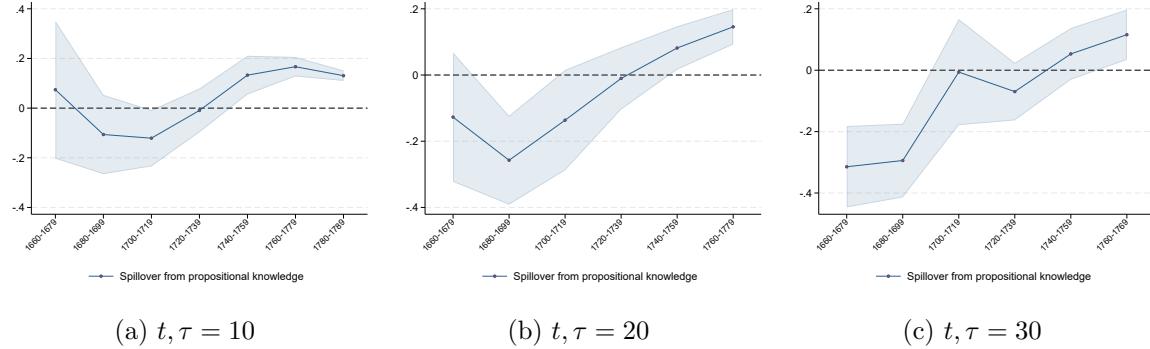


FIGURE 32: Robustness for different values of the time windows t, τ for the calculation of the innovation and spillover index, Spillovers in $\Omega \rightarrow$ innovation in λ

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The plotted independent variable is an interaction term between the received spillover index from equation 5 and twenty year time periods. The model further controls for the word count of titles and includes year fixed effects. Panel a)–c) report results when changing the t and τ parameter for the backward and forward comparison windows in equation 3 and 5. Panel b) reports the baseline results with $t, \tau = 20$. Standard errors clustered at the publication year level.

G.3.2 $\lambda \rightarrow$ innovation in Ω

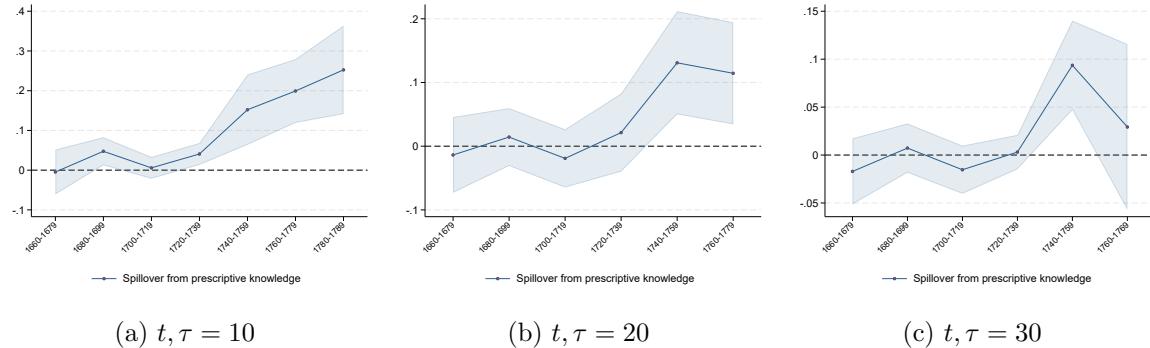


FIGURE 33: Robustness for different values of the time windows t, τ for the calculation of the innovation and spillover index, Spillovers in $\lambda \rightarrow$ innovation in Ω

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The plotted independent variable is an interaction term between the received spillover index from equation 5 and twenty year time periods. The model further controls for the word count of titles and includes year fixed effects. Panel a)–c) report results when changing the t and τ parameter for the backward and forward comparison windows in equation 3 and 5. Panel b) reports the baseline results with $t, \tau = 20$. Standard errors clustered at the publication year level.

G.4 10 year intervals for coefficients

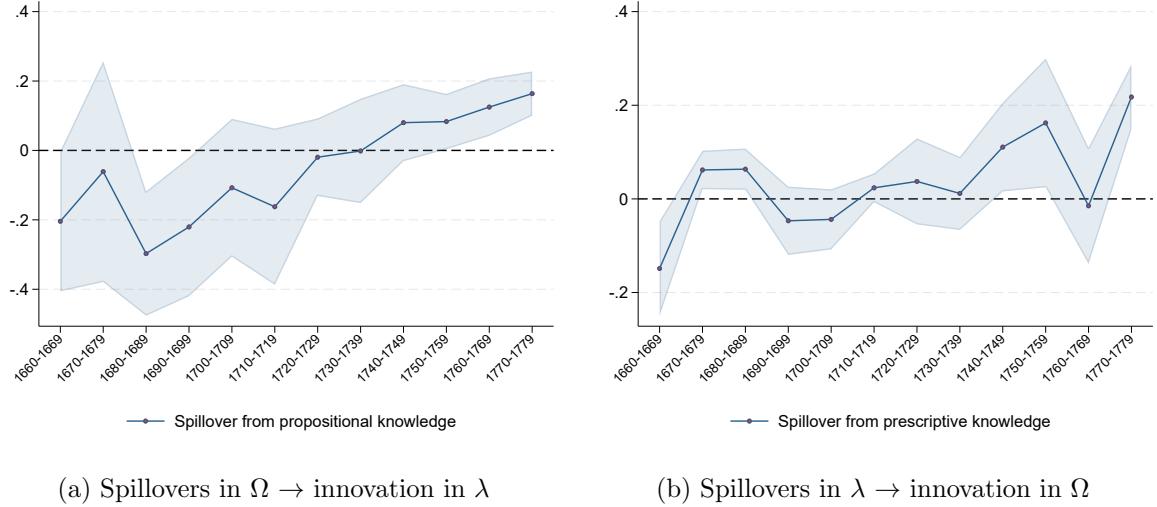


FIGURE 34: 10 year intervals — Spillovers from prescriptive (λ) and propositional knowledge (Ω) \rightarrow innovation

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The graph plots the coefficients from interacting the received spillover index from equation 4 with ten-year time periods. The model further controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Standard errors clustered at the publication year level.

G.5 Spillovers from individual fields

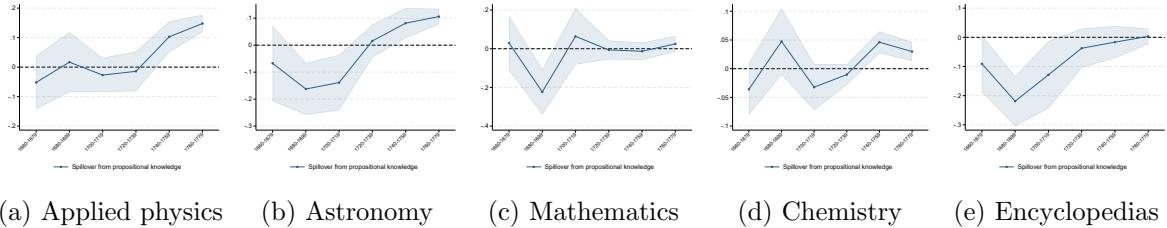


FIGURE 35: Results for individual spillovers from Ω

Notes: The figure reports the estimated coefficients from equation 7. Instead of estimating the association between spillovers from *all* fields in Ω and innovation in λ , this figure reports individual results for spillovers from each field in Ω . Panel a) reports spillovers from applied physics, panel b) from astronomy, panel c) from mathematics, panel d) from chemistry, and panel e) from encyclopedias. Standard errors clustered at the publication year level.

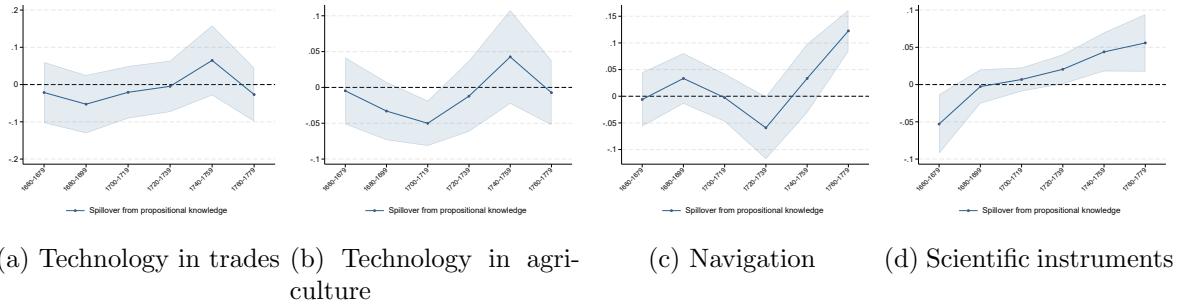


FIGURE 36: Results for individual spillovers from λ

Notes: The figure reports the estimated coefficients from equation 7. Instead of estimating the association between spillovers from *all* fields in λ and innovation in Ω , this figure reports individual results for spillovers from each field in Ω . Panel a) reports spillovers from technology in trades, panel b) from technology in agriculture, panel c) from navigation, and panel d) from scientific instruments. Standard errors clustered at the publication year level.

G.6 Accounting for compositional bias

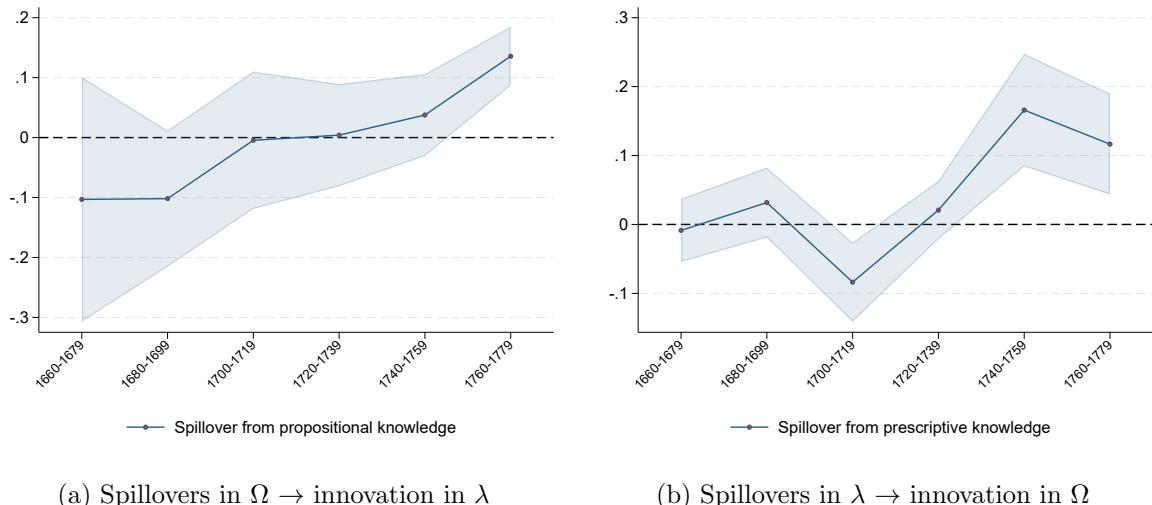


FIGURE 37: Subject \times year fixed effects — Spillovers from prescriptive (λ) and propositional knowledge (Ω) \rightarrow innovation

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The graph plots the coefficients from interacting the received spillover index from equation 4 with twenty-year time periods. The model further controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Additionally, to account for compositional effects, the model further includes subject class \times year fixed effects. Standard errors clustered at the publication year level.

G.7 Long-run results

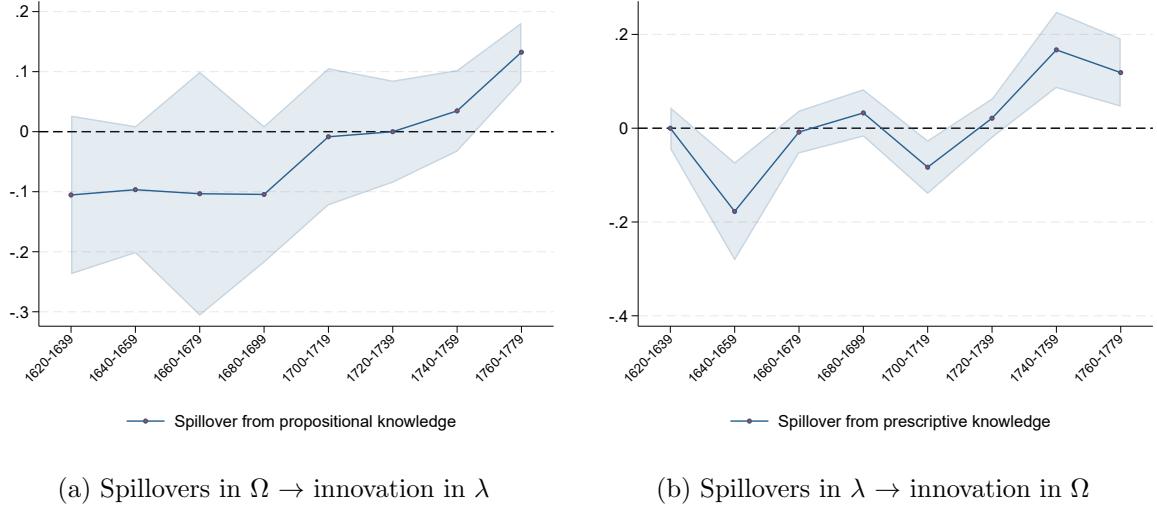
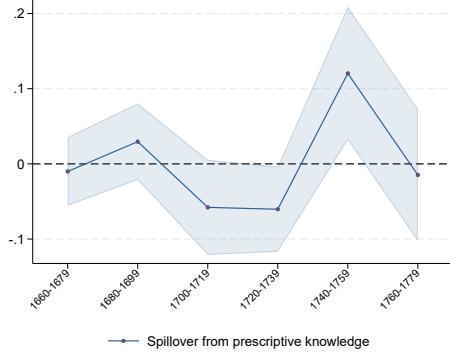


FIGURE 38: 10 year intervals — Spillovers from prescriptive (λ) and propositional knowledge (Ω) \rightarrow innovation

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The graph plots the coefficients from interacting the received spillover index from equation 4 with ten-year time periods. The model further controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Additionally, to account for compositional effects, the model further includes subject class \times year fixed effects. Standard errors clustered at the publication year level.

G.8 Spillover measure incl. patent discontinuity

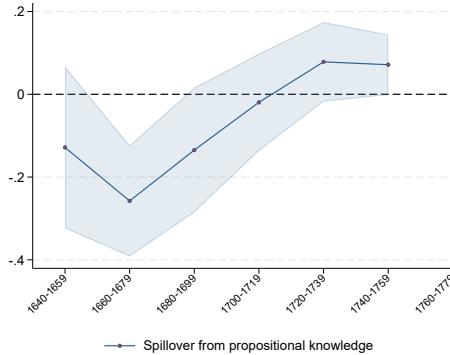


(a) Spillovers in $\lambda \rightarrow$ innovation in Ω

FIGURE 39: Spillovers from propositional (λ) and propositional knowledge (Ω) \rightarrow innovation — incl. post 1700 patents into the calculation of spillovers

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The graph plots the coefficients from interacting the received spillover index from equation 4 with twenty-year time periods. In contrast to the baseline, that excludes patents from the calculation of the spillover index, this figure reports robustness to including patents to the calculation of the spillover index. The model further controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Because patent data is not available before 1700 (see data section 3.2), adding patent data to the calculation of the spillover index creates an undesirable discontinuity. To econometrically mitigate this discontinuity, we further add subject class \times year fixed effects to the specification. Standard errors clustered at the publication year level.

G.9 Excluding patents from set of prescriptive knowledge



(a) Spillovers in $\Omega \rightarrow$ innovation in λ

FIGURE 40: Spillovers from propositional (Ω) to prescriptive knowledge (λ) — excluding patents from set of prescriptive knowledge

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The graph plots the coefficients from interacting the received spillover index from equation 4 with twenty-year time periods. To show robustness, we exclude patents from the set of prescriptive knowledge (λ). The model further controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Standard errors clustered at the publication year level.

G.10 Patents: 10-year periods

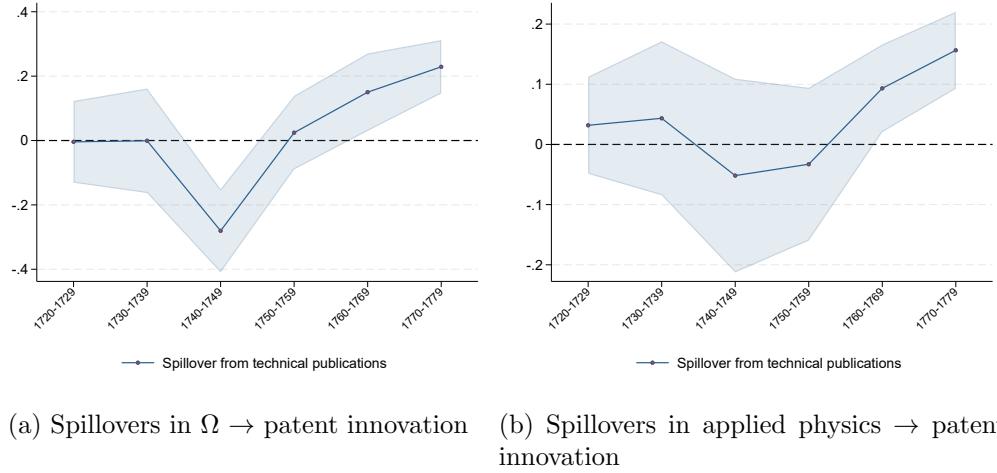


FIGURE 41: Spillovers from propositional knowledge (Ω) and applied physics \rightarrow patent innovation — 10 year periods

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The graph plots the coefficients from interacting the received spillover index from equation 4 with twenty-year time periods. Panel a) shows results for spillovers from the full set of Ω . Panel b) shows results for spillovers from applied physics. Dependent and independent variables are transformed using the natural logarithm. The model controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Given that the patent data only starts in 1700 and the necessary comparison period of $[-\tau, \tau]$, $\tau = 20$ for the innovation index from equation 3 the model is only estimated on the post 1720 sample. Standard errors clustered at the publication year level.

G.11 Patents: Patent citations as alternative outcome

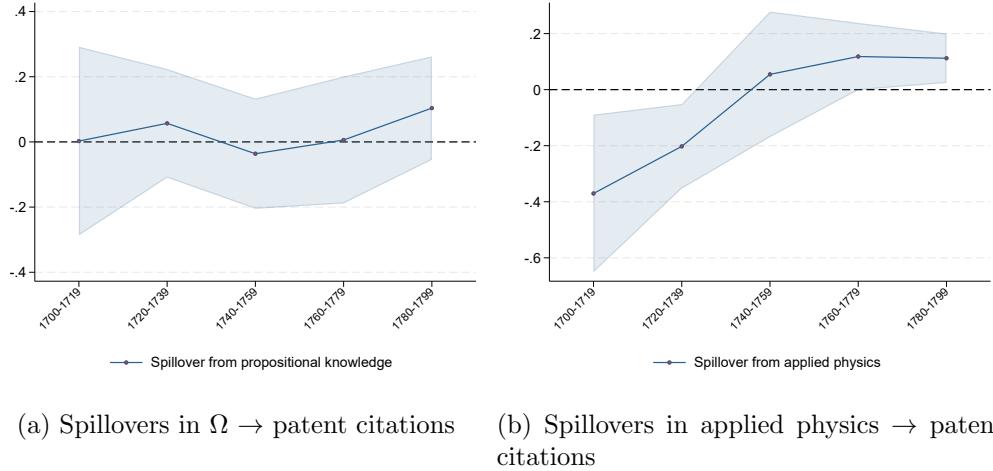


FIGURE 42: Spillovers from propositional knowledge (Ω) and applied physics \rightarrow patent citations

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is patent citations from Nuvolari and Tartari (2011). The graph plots the coefficients from interacting the received spillover index from equation 4 with twenty-year time periods. Panel a) shows results for spillovers from the full set of Ω . Panel b) show results for spillovers from applied physics. Dependent and independent variables are transformed using the natural logarithm. Note that the Woodcroft citation index from Nuvolari and Tartari (2011) defines lacking patent citations beyond Woodcroft as “1”, thereby implicitly applying a $\log(x + 1)$ transformation. The model controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Using patent citations as the dependent variable allows us to use the full 1700–1800 sample period.

G.12 Patents: By-industry results

This section presents by-industry results for the association between patents and received spillovers from Ω and applied physics. We estimate a model similar to equation 7, with the adjustment of interacting received spillovers with indicator variables for each industry from Nuvolari and Tartari (2011):

$$\text{Innovation}_{ijt}^A = \sum_{p \in \text{Industry}} (\beta_p \cdot \text{Received spillover}_{B \rightarrow A}(v_{ijt}) \times I_p) + \mathbf{X}'_{ijt} \zeta + \delta_j + \alpha_t + \varepsilon_{ijt} \quad (12)$$

the model specification is identical to equation 7, with the difference of dropping time interaction terms and adding industry interaction term, I_p . The explanatory variable captures received spillovers from either propositional knowledge or applied physics. Results are shown in figure 43.

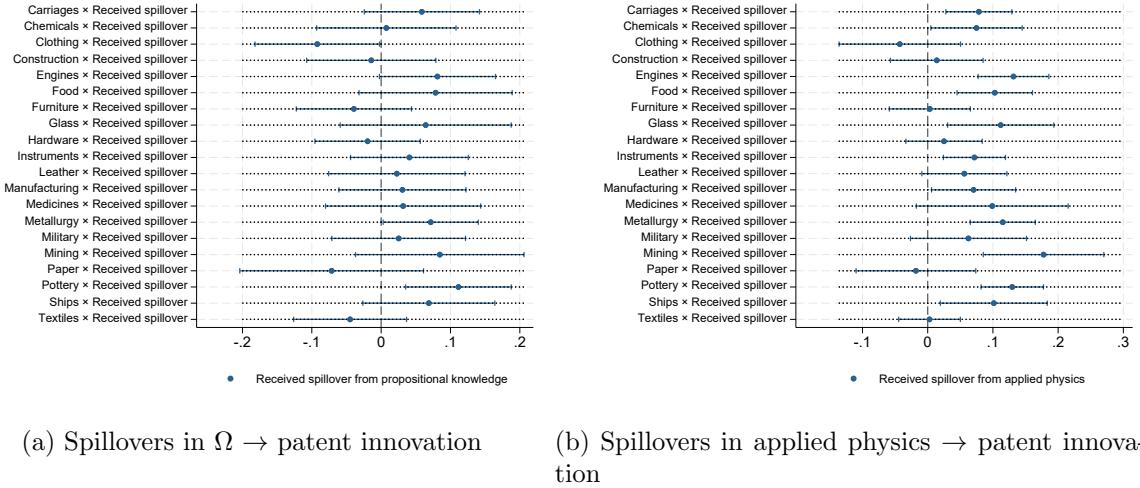


FIGURE 43: By-industry results: The association between innovation and spillovers form Ω and applied physics

Notes: The figure reports the estimated coefficients from equation 7. The dependent variable is the innovation index from equation 3. The graph plots the coefficients from interacting the received spillover index from equation 4 with twenty-year time periods. The model further controls for the level and quadratic count of words in titles and patents and includes year fixed effects. Standard errors clustered at the publication year level.

G.13 List of terms for cosine similarity to methods

TABLE 19: Conceptual term lists for revolutions in methods

| Newtonian mechanics | Precise measurement & toolset | Scientific method |
|---|--|--|
| <ul style="list-style-type: none"> • Newtonian mechanics • laws of motion • force, acceleration, momentum, mass • gravity, center of gravity • levers, pulleys • projectile motion, uniform motion • resistance of media, friction • impact, pressure • hydrostatics, statics, dynamics • centripetal force, centrifugal force • impulse • fluid resistance, flow of fluids | <ul style="list-style-type: none"> • precise measurement • precise instruments • precision instruments • exact measurement • standardized scales • calibration | <ul style="list-style-type: none"> • observation • measurement • mathematical formalization • experiment |

Notes: The table reports the list of terms used for calculating title-level similarity to the three methods of *Newtonian mechanics*, *precise measurement*, and the *scientific method*. The terms are projected into an embedding space using the SteamBERTH model from section 5.1, and average cosine similarities between titles and the terms in each method-set are then calculated.

G.14 Lexicon technicum

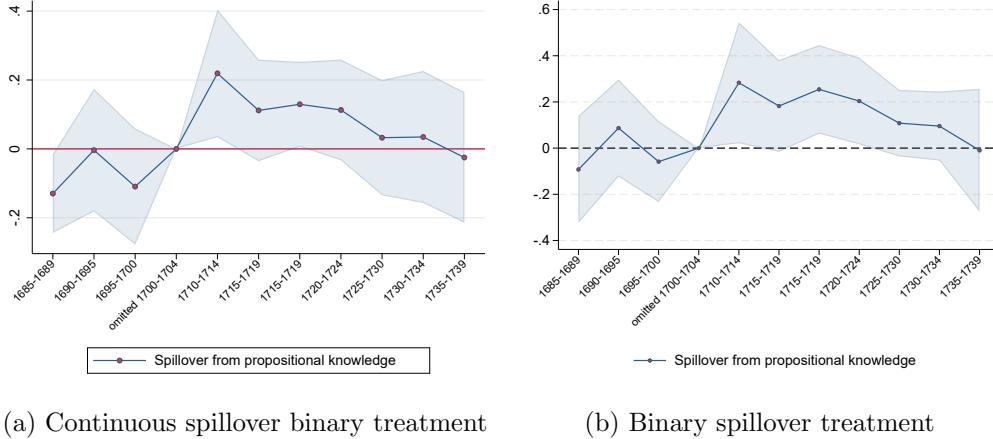


FIGURE 44: Difference in differences results for the *Lexicon technicum*: Spillovers from propositional to prescriptive knowledge ($\Omega \rightarrow \lambda$) — exluding *navigation*

Notes: The graph presents the effect of spillovers from propositional knowledge (Ω) from the *Lexicon technicum* on innovation in prescriptive knowledge (λ) in the ESTC as estimated using the difference-in-differences model from equation 10. To test robustness to the Longitude prize of 1714, this specification excludes the subject class of *navigation*. $N = 468$. Standard errors clustered at the topic level. Confidence intervals shown at the 90% level.

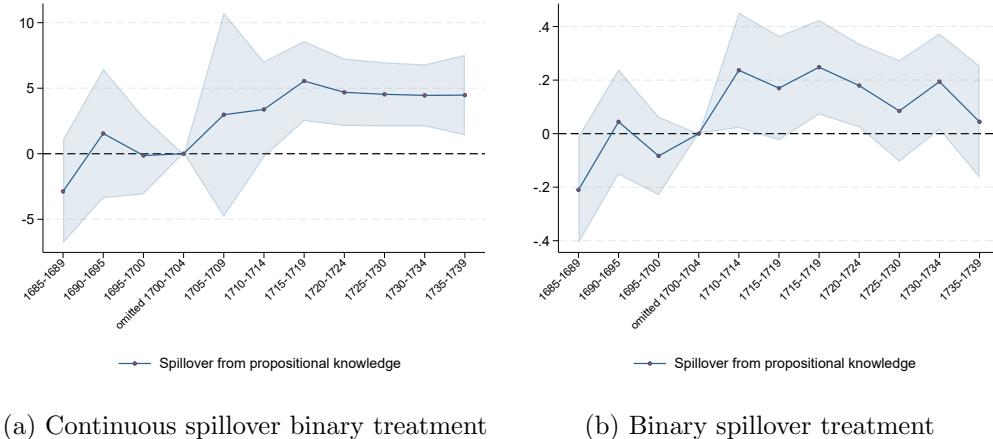


FIGURE 45: Difference in differences results for the *Lexicon technicum*: Spillovers from propositional to prescriptive knowledge ($\Omega \rightarrow \lambda$) — incl. subject specific linear trends

Notes: The graph presents the effect of spillovers from propositional knowledge (Ω) from the *Lexicon technicum* on innovation in prescriptive knowledge (λ) in the ESTC as estimated using the difference-in-differences model from equation 10. The current specification further includes subject specific linear trends. $N = 468$. Standard errors clustered at the topic level. Confidence intervals shown at the 90% level.