Spatial Networks and The Diffusion of Ideas

Caterina Chiopris*

Harvard University

Click here for the latest version and the appendix.

Abstract. The rise of knowledge and the flow of innovation have long been recognized as an important factor for economic development. How do increased spatial connections affect the generation and diffusion of ideas? Intuitively, a denser network should increase the number of novel ideas and augment their diffusion. I study knowledge production in Germany in the 19th century, relying on the universe of bibliographic records and novel railway statistics, among other original data, as well as cutting-edge machine learning and topology to measure ideas. I show that the railroad network increased the creation of new ideas, contributing to 11% of the increase in knowledge production. Scholars' mobility led to the formation of specialized clusters, and thus to cities' specialization. New ideas are formed by combining ideas coming from cities connected by the railroad. However, with a denser network, new ideas diffused less on average. This was a by-product of specialization: with the railroad, groups of scholars could focus on narrower topics and co-locate with similar professionals; they learnt more from similar groups, but became disconnected from dissimilar ones. The findings shed light on the causes for specialization in knowledge production, on the organization of modern science, and on the diffusion of information in dense networks.

Keywords: innovation, infrastructure, history of science
JEL Classification: N73, N93, O31, O18, P00, I23, R4

*Ph.D. Candidate, Political Economy and Government, Harvard University, chiopris@g.harvard.edu
This project would not exist without Alberto Alesina's very early encouragement. Many appreciated suggestions, critiques and encouragement were provided by Melissa Dell, Jeffry Frieden, Edward Glaeser, Torben Iversen, and Marco Tabellini, and many seminar and conference participants. Dominik Flügel provided outstanding research assistance.

1 Introduction

Man does not create material things, he only creates ideas.[1]

Ideas drive innovation and growth, fuel rebellions and social movements, define identities and social cleavages, shape actors' perception of their interests, and affect policy.[2] Ideas matter especially as they diffuse: new technologies need to be adopted widely to drive growth, social movements need extensive support to be successful, theological ideas become a religion or culture when they are pervasive, and so forth.[3] The possibility for ideas and knowledge to diffuse has hence arguably been fundamental for economic growth as well as democratization.[4]

How are new ideas created? And how do they diffuse? This paper investigates the creation and diffusion of ideas in the context of knowledge production, where 'knowledge' follows the broadest possible definition, referring to the content of any printed text. Specifically, I study cities of the German-speaking world in the 19th century, which historians have identified as a defining moment for the development of the modern production of knowledge. To study the effect of a major shock to idea diffusion, I exploit the introduction of the railroad network, and thereby trace the effect of spatial connections on the specialization, production and transmission of new ideas. The patterns emerging will come to characterize the contemporary way of producing knowledge: the separation into clear fields, the co-location of professionals, the knowledge spillovers occurring in cities. I argue that territorial integration, enabled by the railroad network, was a leading cause of these defining trends of modernity.

Previous literature on knowledge, ideas and their diffusion has either studied single successful ideas, or the diffusion within a given area of knowledge, or kept the diffusion mechanism fixed.[5] We know little about what happens in aggregate with increased connections,

[1] Francis Bacon, cited in Webster, Palangkaraya and Spurling (2021).

[2] The following is a brief overview of the literature on how ideas can affect real outcomes. For technological innovation, scientific progress, and economic growth see Romer (1992); Weitzman (1998); Jones (2021). For rebellions, social movements, and revolutions see Tarrow (1996); McAdam, Tarrow and Tilly (2003); Benford and Snow (2000); Weyland (2009). For collective and political identities see Ash, Mukand and Rodrik (2021); Abdelal et al. (2006), for social cleavages Hunter (1992); Inglehart (2015). For economic policy: Hall (1993), foreign policy: Goldstein and Keohane (1993); Simmons and Elkins (2004), judicial decisions: Ash, Chen and Naidu (2022), politics: Hall (2020), and the diffusion of policies: Shipan and Volden (2012). For interests, or how ideas shape how actors understand their interests, and determine the scope of objective functions, see Rodrik and Hall (2018) for an informal discussion.

[3] For technology diffusion, see: Bloom et al. (2021); Hall (2004); for social movements: Lohmann (1994); for religion: Becker et al. (2020). Kuran (1998)'s explanation of changes in ethnic norms depends crucially on a diffusion function. Underlying Kitschelt (1986), broad movements require the dissemination of information and ideas.

[4] See Mokyr (2002, 2016) for how the diffusion of knowledge contributed to Europe's growth, and an example of democratization waves is Acemoglu et al. (2019).

[5] For instance, recent literature studies one idea at a time, such as democracy (Tabellini and Magistretti,

for all ideas in all fields, and without selecting on ideas' success *ex post*. The conventional wisdom, backed by standard work in network theory, predicts that with more connections, ideas diffuse more. Once connected, cities would converge to the same set of salient issues.

In a simple model, I show instead that incorporating the generation of ideas and knowledge spillovers in a network drastically changes these predictions. Increased connections imply a larger market for the production of ideas: once knowledge from elsewhere is more easily accessible, creators have incentives to specialize, and focus on a narrower field of expertise. At the same time, spatial integration facilitates the co-location and communication of idea creators active in similar fields. With easier access to more knowledge, idea producers are better able to select content most similar to theirs. With increased mobility, they are better able to sort along substantive lines, by moving to clusters of similar individuals. Specialized clusters form, and cities diverge, as scholars from connected cities will move to their respective cluster.

Note that it is not clear *ex ante* that we should expect an increase in specialization, which in turn would drive innovation and change patterns of diffusion. The parallel with an expansion of a 'knowledge market' makes the rationale for specialization familiar. However, the conventional wisdom explaining why knowledge (science especially) becomes more specialized points to the burden of knowledge (Jones, 2009). With more accumulated knowledge, innovators need to learn more to make a contribution, and need to narrow their expertise to counterbalance the educational burden. Specialization, according to this view, is a function of existing past accepted knowledge, and not of access to knowledge being produced in the present, as this paper's model suggests.

The empirical expectations can be summarized as follows. With denser spatial connections and the expansion of cities' market access, knowledge production should increase in volume as well as innovativeness. This paper will focus on two specific mechanisms that could link connections to the increase in knowledge. First, scholars' mobility: similar experts would be better able to co-locate, and cities would specialize. A long-standing tradition stressing the importance of knowledge spillovers (Arrow, 1962; Romer, 1986), especially in cities (Marshall, 1890; Glaeser et al., 1992; Iversen and Soskice, 2020) suggests that the clustering of specialists would contribute to the rise of innovation. The second mechanism is ideas' mobility: if ideas are better able to travel, new ideas can emerge by combining existing ideas from different cities. Note, however, that increased specialization and increased connections imply that new ideas would diffuse more within their fields, but not necessarily

2022; Simmons, Dobbin and Garrett, 2008) or liberalism (Simmons, Dobbin and Garrett, 2006), Darwinism (Arold, 2022; Giorcelli, Lacetera and Marinoni, 2022), Protestantism (Cantoni, 2012), the recent MeToo movement (Levy and Mattsson, 2022), or diffusion within science and technology (Giorcelli and Li, 2021), especially measured through patents (Bloom et al., 2021).

overall, as scholars in each city can focus on adopting ideas from their own field, and ignore dissimilar disciplines.

To test the hypotheses outlined above, Germany in the 19th century is especially adequate, not only for the unique proliferation of ideas and for emergence of the modern system of knowledge production. It also saw the first *serious* attempt by contemporaries at recording every publication in a country (Domay, 1987), in a time when the communication of ideas occurred chiefly in printing. This is especially important to capture the most complete possible picture of a setting's intellectual landscape, including ideas in all fields, even those that did not end up diffusing or persisting. Furthermore, it witnessed a shock in connections significant enough to affect cities' networks and diffusion patterns in all fields, i.e., the introduction of the railroad.

The empirical analyses rely on the collection and digitization of several original data sources: the universe of bibliographies (records of all publications) compiled by contemporaries; the universe of the current bibliographic records of all German, Austrian and Swiss libraries, research institutes, universities etc., containing 22 million items; the histories of all the main publishing houses and distributors, including the technology used and their specialization; the approval and opening of all lines in the first 100 years of the railroad; various variables on trade, passenger volumes, mail, telegraphs for each line; and the specific reason for delays in construction, used to build counterfactual lines. I further collected city-level data such as population, printing houses, newspapers; as well as the presence of universities, significant scientific discoveries, and 230,000 standardized biographies of notable individuals.

To operationalize the concept of an idea into a measurable unit, an idea is defined as the specific topic of any publication with an informative title, academic or not – in this sense, I use a broad definition of 'knowledge'. For measurement, note that a new idea can be reflected in language in different ways: a new word is created, or previously disjoint words are combined. The German language is especially adequate to measure the first: because of the use of compounds, new words are very frequent when a new concept arises.[6] The first measure hence relies on vocabulary expansion. Second, for the combination of existing words, I measure how long it takes for two concepts to co-occur since their respective appearance.[7] To measure this efficiently, I use topology methods applied to the lifetime of knowledge gaps in a concept network, described in more details below.[8]

[6] As an example, with the invention of the cylinder printing press, no new words are created in English, but in German the word *Zylinderdruckmaschine* appears.

[7] E.g. it famously took 54 years (1828-1882) for the words *Bakterium* and *Tuberkulose* to appear together.

[8] Note that another possibility for innovation to be reflected in text are semantic changes – a swift alteration of a word's meaning. For instance, the word *Zelle* was mostly used to describe a monastic cell until the cell in biology was discovered. To measure this phenomenon, I am currently in the process of clustering contextual embeddings (based on Sentence Transformers) to identify significant changes in meaning. Note

To identify the effect of networks (in this case, the railroad network) on knowledge production, the analysis follows the most recent work on estimating the causal effect specifically for infrastructure (Borusyak and Hull, 2023). New railroad lines were clearly not built at random; more 'central' locations were more likely to be connected. However, construction delays were frequent, and often occurred for reasons unrelated to the line itself. The identification strategy exploits this exogenous variation in timing of construction to recenter the treatment effect. By permuting completion status, the centrality effect of lines that were selected into treatment, but were not built for exogenous reasons, is subtracted from the observed treatment effect, attenuating selection concerns.

First, I test whether increased market access induced by the introduction of the railroad and the reduction of travel time led to an increase in knowledge production, where I define market access following Donaldson and Hornbeck (2016). I find that the railroad can explain roughly 11% of the increase in the number of publications. Not only more titles were published, but also in more disciplines, and on more topics. The increase in quantity and diversity is accompanied by an increase in novelty as well: the share of new words over the total number of words increases. Furthermore, new knowledge is produced faster, as the average time to close a knowledge gap (to make a connection between concepts that will eventually be made) decreases significantly.

Next, to check that the mechanisms proposed are indeed are work, I establish that the railroad did affect the movement of notable individuals and of ideas. Using event study designs, I find that with the railroad ideas travel faster: conditional on being adopted in other cities, the average time to reach new destinations decreases significantly. Notable individuals, on the other hand, tend to move more: they are associated with a higher number of towns throughout their lifetime if the railroad exists in their towns of birth. An interesting side result concerns the background of who becomes a notable individual: after one generation since the introduction of the railroad, the family of origins become more diverse – descriptively, less likely to be from the upper classes. In other words, mobility had a democratizing effect on the intellectual elites: a larger fraction of the population has access to, and can produce, new knowledge.

Having established that the railroad did, broadly speaking, affect the movement of scholars and ideas, I test whether specialization increases. At the individual level, over time, individuals are more likely to specialize – they are less likely to be active in multiple fields, or be polymaths. While they move more often, they are not more likely to move to larger cities necessarily, but to cities that have a larger presence of individuals in a similar profession – they cluster along substantive lines. Cities' publications are more concentrated on

however that drastic semantic changes occur relatively infrequently.

certain fields, and on certain subfields within those fields. Among all city-pairs, connected cities become more *dissimilar* in the topics their publications are on.
I check whether different measures of market access are correlated with and affect specialization at the city level. The standard measure for market access (a function of cities' population) is correlated with specialization in idea production, but it is not robust to subtracting the centrality effect described above. On the other hand, intellectual market access, a function of the other cities' intellectual production, has a significant effect even after re-centering. Access to the intellectual stock (which would support the 'burden of knowledge' hypothesis) is not significant. In other words, access to current new ideas being produced elsewhere – an increase in the market size – leads to specialization in knowledge creation.

The second mechanism expects ideas' mobility to affect knowledge production. I find that once two cities become connected, scholars in those cities are more likely to produce new ideas combining concepts coming from the other city, explaining roughly 16% of new combinations. Furthermore, within the same discipline, new ideas are more likely to diffuse; and the ideas that are able to diffuse initially are more likely to persist.
However, specialization means that there is an increased sorting of ideas as well: new ideas do diffuse more within their field, but diffuse to a smaller number of cities, and diffuse less to other disciplines, consistent with the formation and crystallization of disciplines. As the physical constraint of proximity for information acquisition is less binding, information is selected along substantive lines. Knowledge is more strictly separated into fields, and those fields are less prone to cross-pollination. Clusters communicate more frequently with other clusters in the same field, but less with other fields.

Last, I provide suggestive evidence that the documented fragmentation of knowledge is related to further real outcomes. Cities' specialization within the sciences is correlated with tangible, relevant discoveries in the same field in the same city. As mentioned above, a stricter separation into disciplines and subfields did not only occur within the sciences, but also in the humanities and the social sciences. Interestingly, the change in the organization of the social and legal sciences affected the training of bureaucrats, lawyers and legislators. A stronger specialization on the 'state sciences' is correlated with a higher probability of training high-ranking public officials. Legal scholars also narrowed their expertise: this was reflected in the legislation they were asked to draft, and in turn affected implementation and enforcement. In other words, the changes in knowledge production affected not only science, but also the state and its institutions.

This study contributes to several strands of the literature. First, it provides novel insights on the mechanisms of diffusion of information. In recent years, with the rise of the Internet and social media, work on information diffusion has focused on online polarization, bias in information consumption, the spread of misinformation, and so forth.[9] The common finding is that participation in such networks biases and narrows individuals' perceptions. The processes described in this paper may remind the reader of similar dynamics: the possibility to acquire a significantly larger amount of information leads to a narrowing of one's focus and perspective. Recent literature in this area disagrees on what exactly leads to this dynamic; but whether the source is biased algorithms, homophily, or individuals' previous values and beliefs, the exact mechanisms are tied to the specific functioning and interactions within social media or the internet. What distinguishes the findings of this paper, beside the breadth of the ideas considered, is that the results are not tied to a similarly specific intermediary: the narrowing of the focus is a direct function of the increased amount of information accessible.

Second, this study contributes to a growing literature in historical political economy. Previous work set in the same time period (the 19th century) stresses the role of human capital as a source for industrialization and growth. For instance, Becker, Hornung and Woessmann (2011) and Squicciarini and Voigtländer (2015) study the development of human capital and their link to growth and industrialization. Hanlon (2022) and Maloney and Valencia Caicedo (2022) the rise of the engineer; Dittmar and Meisenzahl (2022) the emergence of research universities, while Serafinelli and Tabellini (2022) the role of markets and institutions for creativity. This paper takes human capital as a pre-condition, and focuses instead on the *content* of knowledge. It adds a new factor – individuals' mobility – and shows the important consequences it had for the further democratization of knowledge production and the expansion of the intellectual elites. Compared to work on historical scientific advancement, which usually relies on the effect of foreign ideas or immigrants,[10] this study does not involve external sources of knowledge creation, but explains endogenous, domestic ideas generation.

Taken together, the results challenge three commonly held notions: that networks and integration increase diffusion and convergence; that ideas diffuse linearly across space (or purely as a function of their medium of diffusion); and that specialization in knowledge is a product of the burden of knowledge alone.

The rest of the paper is organized as follows. In the next section, I present the existing theoretical expectations on the diffusion of information in a network, and I show that integrating a knowledge production process in a model of diffusion reverses such expectations.

[9] For a recent review, see Zhuravskaya, Petrova and Enikolopov (2020). For some examples on Facebook: Bakshy, Messing and Adamic (2015), Levy (2021), for Twitter: Halberstam and Knight (2016).

[10] For example, Moser and San (2022) investigates the role of immigration; both Moser, Voena and Waldinger (2014) and Becker et al. (2021) study German academics and innovators as they emigrate during the Nazi regime; Abramitzky and Sin (2014) explore the import of foreign knowledge after the end of Communism, Borjas and Doran (2012) the influx of Soviet mathematicians to the US specifically, while Giorcelli and Li (2021) the role of external alliances; Iaria, Schwarz and Waldinger (2018) international knowledge flows.

In section 3, I provide a brief overview of the historical background. I describe my original datasets in section 4 and the methods for the text analysis and identification strategy in section 5. I present my results and robustness checks in section 6. Section 7 discusses the consequences of specialization. Section 8 concludes with implications and open questions.

2 A Model of Idea Generation and Diffusion

The goal of this section is to explore what one would expect, theoretically, for the creation and diffusion of ideas once cities' connections become more dense. First, I set up a simple model of diffusion to adapt standard results from network theory to my setting. Next, I complement the model by including a mechanism of knowledge production which includes local and global knowledge spillovers.

The main mechanisms of the model are simple. In an extreme scenario with no connections between cities, every knowledge producer is a generalist: they can only rely on their own knowledge, thus need to cover every possible topic, but can only do so in a superficial way. With connections, knowledge can flow from other cities, and idea creators can focus on a narrower field in a deeper way, delegating the rest of knowledge to producers in other cities. Intuitively, with connections, we hence expect increased specialization, which is the first hypothesis that will be tested.

Knowledge spillovers play an important role. Following a long tradition in urban economics, if two knowledge producers co-locate in the same city, additional knowledge will be created through positive externalities. This has an important consequence: with more spatial connections, the costs of moving will decrease for scholars, and they will move to cities with creators of similar expertise. Cities will thus specialize, as they attract scholars from a specific field, and additional knowledge will be created thanks to the spillover effects. The second hypothesis is therefore that knowledge production increases with connections.
What are the implications for diffusion? I show that with increased connections, diffusion decreases, which is the third hypothesis that will be tested. With scholars and cities specializing, the chances of knowledge producers adopting an idea from a specific field is lower, as their focus is narrower. Niches form, that are well connected across cities, but that ignore the expertise created in other fields. While the simple model of diffusion predicts networks to increase the overall transmission of information, a model that incorporates field-specific selection of information predicts the opposite.

The following provides the intuition and results of the model, which can be found in the Theoretical Appendix. The empirical implications are in section 2.3.

2.1 Atomistic Diffusion with Naïve Learning

Assume that in a number of cities there are innovators who are engaged in creating new knowledge (ideas). Suppose that all of knowledge can be divided into two areas, referred to as fields or disciplines in what follows. For simplicity, each city starts with one idea creator in each field.

Ideas can move between cities as long as there is a physical connection (a positive link) between them, and someone active in the same discipline to adopt them. At time 0, every innovator creates a new idea. Each idea travels to all other cities that have a series of positive links starting from the city of its creator. At time 1, ideas will thus be adopted in every city that can be reached through positive links from the origin. Intuitively,

Proposition A1. *Ideas diffuse to a (weakly) higher number of cities when the number of positive links increases.*

Proof: in the Appendix. For what follow, define M the adjacency matrix for the cities' connections, and $\eta(M)$ the number of cities an idea would reach on average.

With more connections, ideas have a higher chance of reaching any other city in the network. More positive links imply more possible paths connecting any two cities, and thus a higher chance of ideas traveling between them. As long as two cities have an uninterrupted sequence of connections between them, they will have the same ideas adopted at time 1.

Proposition A2. *With an additional positive link, the probability that an idea is present in any two cities (weakly) increases.*

Proof: in the Appendix.

A denser network means that the distribution of ideas present at time 1 in two different cities is more likely to be the same. More links lead to uniformity. Consistent with what has been shown elsewhere, a network in which there exists a path between all pairs of nodes (cities in this case) will reach convergence (Golub and Jackson, 2010): all cities adopt all ideas.

2.2 Idea Generation with Knowledge Spillovers

The approach just outlined abstracts from how new ideas come to life. More realistically, innovators in a given field choose their area of expertise, and produce new ideas within it. They face a trade-off between the coverage ($k > 0$) and the depth ($d \geq 0$) of their knowledge: they can either be experts in a narrower field, or generalists with a more superficial expertise. Each producer p is hence time constrained such that $k_p + d_p \leq T$.

The demand for knowledge is exogenous, coming from societal actors. Such demand is for all topics in both fields, where fields are denoted by $D \in \{1, 2\}$. Innovators provide answers based on the knowledge they have available. Society, however, rewards a deeper answer compared to a shallow one. The value V each producer receives from responding to society's questions q is hence:

$$V(q) = \begin{cases} 1 + d_c^{\alpha} \mathbb{1}[D_q = D_p] & \text{if } q \in A_p \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where A_p is the knowledge available to producer p, $\alpha < 1$ depicts the decreasing returns from investing in depth, and $D_q = D_p$ holds if society's question lies in the producer's field of expertise.

How do innovators respond to society's need for knowledge? Knowledge spillovers are central to this process. A knowledge producer can answer society's questions either with their own expertise (k_p), or through the knowledge coming from global or local spillovers. Global spillovers refer to knowledge produced in other cities (A_{-c}) that reaches a specific creator. They depend on the connections a city has with other towns; define the share of outside knowledge $\tau(M)$. The more the links, the higher the global spillovers.

Local spillovers (π) refer to the knowledge created by other scholars in the same town (N_c). To capture the notion that human capital accumulates faster in cities due to frequent interactions allowing rapid learning (Glaeser, 1999; Iversen and Soskice, 2020), proximity to other scholars in the same field comes with the transmission of additional depth in knowledge. Summing up, the knowledge available to idea producer p in city c is given by:

$$A_{pc} = k_p + \pi(N_c) \cdot \sum_{N_c - 1} k_s + \tau(M) \cdot A_{-c} \tag{2}$$

To understand the effect of connections, consider first the hypothetical, extreme, scenario where there are no positive links between cities ($\tau(M) = 0$). What is known is the knowledge created in a city, and nothing else; and because there are no connections, scholars cannot move. But societal actors have a demand for all possible knowledge, and the two scholars in each city will have to answer all of society's questions. It follows that:

Lemma B1. *With high transmission costs, all scholars are generalists: if $\tau = 0$, then $k_p = 1$ and $d_p = 0$.*

Proof: in the Appendix.

In other words, both scholars in each city will expand the breadth of their knowledge in their respective field as much as possible, sacrificing depth.

What happens if links were created between cities ($\tau(M) > 0$)? Part of the necessary knowledge would be coming from the outside, without innovators needing to be actively engaged in all topics. They become more focused, and depth of knowledge can hence increase.

Lemma B2. *If $\tau > 0$, $k < 1$, $d > 0$. As connections increase, scholars specialize.*

Proof: in the Appendix.

Understanding the causes of specialization is crucial: a centuries-long scholarly tradition has stressed how specialization increases productivity and therefore growth. This model's driving force for specialization is similar to Adam Smith's familiar explanation: specialization is a function of market size. In this case, what defines 'market size' are connections allowing global spillovers.

Note, however, that the causes for specialization in knowledge are far from obvious. The conventional wisdom is that a narrow expertise (especially in science) is a function of the burden of knowledge: specialization is necessary as scholars are encumbered with more and more to learn. Specialization is hence a function of the volume of previously produced knowledge, not current market size or active connections. Both hypotheses will be tested in the empirical section.

Adding another layer to the model, innovators also face a choice on whether to move to another city. Assuming there are costs from moving, increased connections across cities would arguably decrease such costs. The benefits from moving are linked to the local knowledge spillovers mentioned above: being close to other creators in the same field comes with positive externalities. With little connections, the costs of moving would outweigh the externalities, and innovators do no move. With more links, past a certain point of connectivity, (at least some) creators would solve this trade-off differently, as the knowledge spillovers become more valuable than the moving costs.

Note the implications: creators would move to cities with similar scholars. They can access knowledge from other fields thanks to global spillovers, which will be sufficient if connections are high enough – the same conditions that leads them to move. Different cities will hence attract innovators in different fields. In each city, scholars will produce knowledge in one field only, and not the other.[11]

Proposition B1. *As connections increase, scholars are more likely to move, and cities specialize in one discipline or the other.*

Proof: In the Appendix.

The patterns of idea diffusion would also change. Now, cities are active in one field only. If two cities become newly connected, a new idea may not diffuse to the other city, if there is none active in the same field to adopt. Past the threshold of connections in which cities

[11] Note that congestion costs prevent everyone from moving to the same city.

specialize (τ^*, see Appendix), as scholars pay attention only to their own discipline, the diffusion of ideas would decrease.

Proposition B2. *Above a certain threshold of connections, ideas diffuse to a smaller number of cities:* $\frac{\eta(M)}{N_c}$ *instead of* $\eta(M)$.

Proof: In the Appendix.

2.3 Summary of Empirical Expectations

Following a standard model of diffusion, it can be expected that:

(A1) *New connections increase the diffusion of ideas:* with more links, information has more routes to reach any agent or city.

(A2) *With more connections, cities converge on the topics or ideas they adopt*: more links imply that the chance of any two cities to receive and adopt the same information increases.

Including a knowledge generation process with local and global spillovers, one can expect instead:

(B1) *As connections increase, scholars specialize*: if part of the necessary knowledge is produced elsewhere and becomes available in their city, writers can focus on less topics, and increase the depth of the knowledge they produce.

(B2) *With sufficient connections, scholars move to other cities*: with lower transportation costs, they can exploit local knowledge spillovers by moving to the same city as others in the same field.

(B3) *With more connections, cities specialize*: as scholars in similar fields cluster in the same place, the knowledge produced in a given city is more specialized.

(B4) *Increased connections let cities diverge*: since there are gains from being in the same city as similar scholars, each city will attract knowledge producers from different fields.

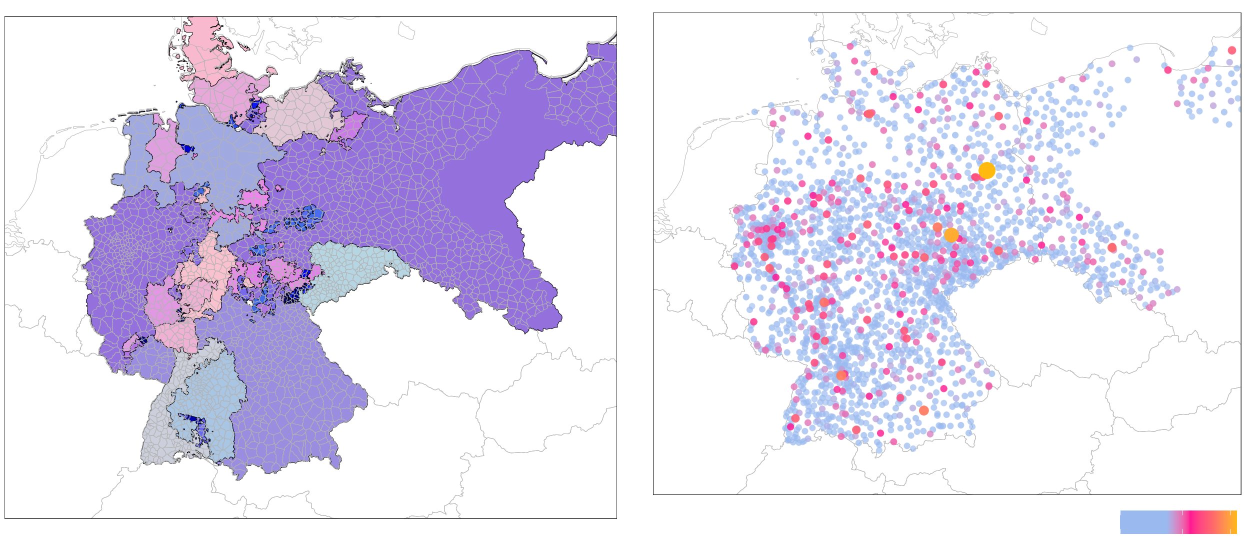
(B5) *Beyond a point of connectivity, diffusion decreases*: for any given topic, there are less cities directly involved in knowledge production in the same field to adopt its ideas.

Each of the hypotheses above will be tested separately; however, the predictions from the expanded model can be grouped into three categories: (1) With more connections, specialization increases. (2) With more connections, knowledge production increases. (3) With more connections, overall diffusion decreases, but increases within field.

3 Historical Background

The empirical analysis relies on spatial differences in publications to measure local specialization, innovation, and diffusion of new knowledge. It is hence fundamental to understand the functioning of the publication industry of the time, as well as the broader intellectual landscape that defined German cities in the 19th century.

A distinguishing feature of German knowledge production – historically and today – is the distribution over space of major centers of learning and publishing. As a comparison, in the early 19th century, France and England had their academics, intellectuals and publishing houses highly concentrated in one and three cities, respectively. The German case looked vastly different: renowned universities were scattered across its territory (see Figure A4 in the Appendix), all main towns had their publishing houses, both major and minor cities a circle of intellectuals (Pfeiffer, 1928). This was a natural consequence of the historical fragmentation into a multitude of counties, duchies, free cities, principalities, bishoprics, etc: there was no push toward intellectual centralization. Even after the collapse of the Holy Roman Empire, with the German Confederation (Figure 1a) and the eventual unification under the German Empire, the federalist structure sustained a decentralized and diffused system of creating ideas.



(a) German States in 1820, with historical cities borders.

(b) Books were published substantially even in medium-sized and small towns.

Figure 1: Political fragmentation and presence of publishing activities.

The publishing industry had adapted to decentralization. Throughout the 18th century, a sophisticated system had been put in place to support local publishing. Commercial enterprises replaced universities and monasteries as the main loci of publishing (Kapp et al., 1886). Retailers emerged to take over distribution from publishers, splitting the risk of sales and increasing the territorial reach of publishing houses (Fischer, 1903; Widmann, 1952, p. 118). Small traders took over the transportation of books. Commissioners started to represent multiple companies at the annual book fairs, and to operate between fairs as well (Pfeiffer, 1928, p. 129). This led to an increase in the flow of traffic between publishers and retailers (Widmann, 1952, p. 118). The retail system enabled authors and publishers to sell widely. Writers, on the other hand, facing issues of supervision with their publishers, had strong incentives to publish as close as possible to their residence (Pfeiffer, 1928). While Berlin and Leipzig were major publishing centers, publishing occurred substantially even in medium-sized and smaller towns, as portrayed in Figure 1b.

The technological advances of the early 19th century allowed even smaller towns to print locally. As steam-operated machines replaced hand-operated ones, printing became ten times faster – and cheaper. The inventor of the first machine-operated press was a German typographer and printer, Friedrich König (1774 - 1833). His *Zylinderdruckmaschine*, perfected between 1812-1818, exploited steam power to replace human labor, and the rotatory motion of cylinders to increase efficiency: they were soon sold in all major cities (Wilkes, 2004). By the 1830s, multiple producers of printing machines were selling their steam-operated apparatuses to printing houses across Germany (Wilkes, 2004).

Another important necessary condition for the creation of knowledge is the education system. In the late 18th century, Prussia became one of the first states in Europe to have free and compulsory primary schooling. Literacy rates rose substantially as a consequence. Educational reforms were made a priority after the defeat in the Napoleonic wars.
Intellectual historians and sociologists have long identified Germany in the late 19th century as the origin of the crystallization of disciplines with which we are so familiar today (Rádl, 1909; Ben-David, 1971). This required the previous development of the modern research university (Dittmar and Meisenzahl, 2022). In the early 19th century, Wilhelm von Humboldt transformed education with his idea of *Bildung* – academic education should form individuals as well-rounded, independent thinkers, rather than providing vocational skills. After he founded the University of Berlin, combining teaching and research, new universities and institutes of higher education were created across the territories of the confederation (Rüegg, 2004). The research university was founded on the idea of *Wissenschaft* (scholarship, research, science), the active search for new knowledge. Universities began to mold into their contemporary form, with the creation of seminars, the incorporation of research as part of the training, and the conferral of standardized doctorate degrees (Watson, 2010).

While a necessary component, the presence of the research universities *alone* cannot explain the strict separation of fields and subfields (Habinek, 2010), and the subsequent city-specific specialization.

The reader should also note the relevance of the intellectual contribution of this time period. Before the Nazi regime eliminated 60,000 of its scholars, and emigration further weakened its universities, German scientists discovered non-Euclidean geometries and real analysis, the tuberculosis and cholera bacteria, started biogeography, produced aspirin, founded modern genetics, and developed electrical engineering. The list goes on, while the arts, humanities and social sciences reached similar achievements (Watson, 2010).

The main efforts to build the railroad hence took place when the primary and secondary system of education were established, universities had been built across the country, and most towns had local, fast printing.

It is worth underlying that before the railroad, internal migration barely existed: until the 1820, 2% of the population ever moved, 80% of which would move within 50 km of their origin (Tolloi, 2010). The introduction of the railways changed these patterns dramatically (McIntosh, 1997). The very first railroad line to be opened dates back to 1835, connecting Nürnberg to Fürth. However, construction initially proceeded slowly. As a reference, in the first decade 680 kilometers of lines were opened, while 13,700 kilometers were constructed between 1865-1875 alone (Heinze and Kill, 1988). Opening a railway was far from straightforward, and often interrupted by financial issues and difficulty in capital accumulation, uncertainty in route determination and, in the first stage especially, in obtaining public concessions.

The rationale for construction varied. The railroad could serve local economic interests, such as connecting mines to cities and ports, or it would connect major cities across states. In other cases, it was located so to influence international trade routes, or for military-strategic purposes, such as the movement of the army (Heinze and Kill, 1988; Fremdling, 1975). In the later stages, after unification, the further expansion of the network was intended to improve accessibility in all regions, and connect the rural periphery (Fremdling, 1975). The gradual expansion of the railroad is captured in Figure 2. By the beginning of the 20th century, a total of 63,794 km of railways had been constructed.

4 Data

The key outcome of interest are ideas, and for measurability, written (printed) ideas. I rely on a novel dataset of bibliographies and library catalogs for measures of intellectual production. I build another original dataset describing the railroad network for a shock in

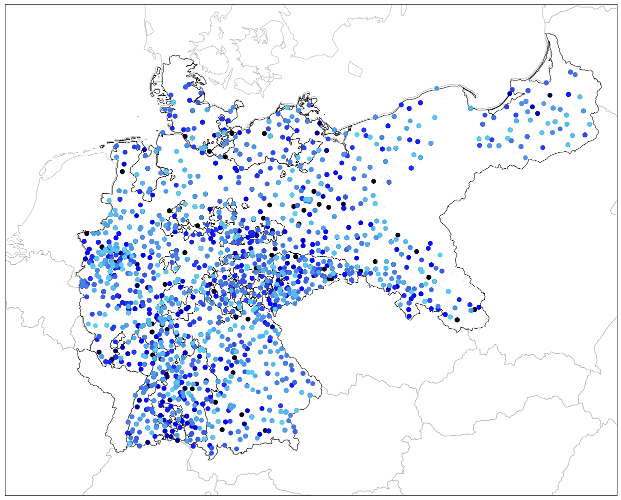


Figure 2: The expansion of the railroad between 1835-1914. Darker circles indicate cities being connected to the network earlier; light blue circles are cities connected later (1880-1914).

cities' spatial connections. I outline these two sources of data next, together with additional data I collected.

Bibliographic Records: Gemeinsamer Verbünde Index. A key source of data comes from the German Collective Library Consortium (the union of all German, Austrian, and German-speaking Swiss libraries and institutes), which operates a comprehensive, unified catalog renowned for its completeness (Conradt and Lohrum, 2017). The catalog contains the bibliographic records of all public libraries, private institutes and research institutions, and university collections. It is composed of bibliographic entries for books, journals, periodicals, newspapers,[12] maps, music, sound recordings, pamphlets, films, letters, and other forms of media.[13] For the period of interest (1750-1914), there are about 22 million entries. Each record includes details such as the author's name, the language, the title of the work, publication data, editor and publishing house, subject classifications, description through keywords, and the physical description as well as location within the participating libraries. The exact information extracted from each entry (MARC tags) is described in the Appendix.

In the Appendix, I also report a series of validation tests to ensure that the descriptive

[12]While the book format was still the main form of publication, especially in the early period, specialized journals become more prominent toward the end of the period. The catalog also includes the German library of digitized journals and newspapers, citing volumes and the prominent articles and papers within those.

[13]Among these, I have excluded anything that is not strictly text (for instance: pictures, music, and maps) and private letters. Most of the analyses focuses on publications written in German or Latin (then translated into German), as the focus is on domestic production. Most analysis also exclude novels, as they do not have a precise subject classification, and the titles are often non-informative.

patterns within the data are consistent with intellectual historians' expectations. Figure 3 below confirms the expected transition in knowledge production in the main publishing cities of the time. With time, the mass of publications shifts critically from religion, the arts and humanities toward the natural and social sciences.

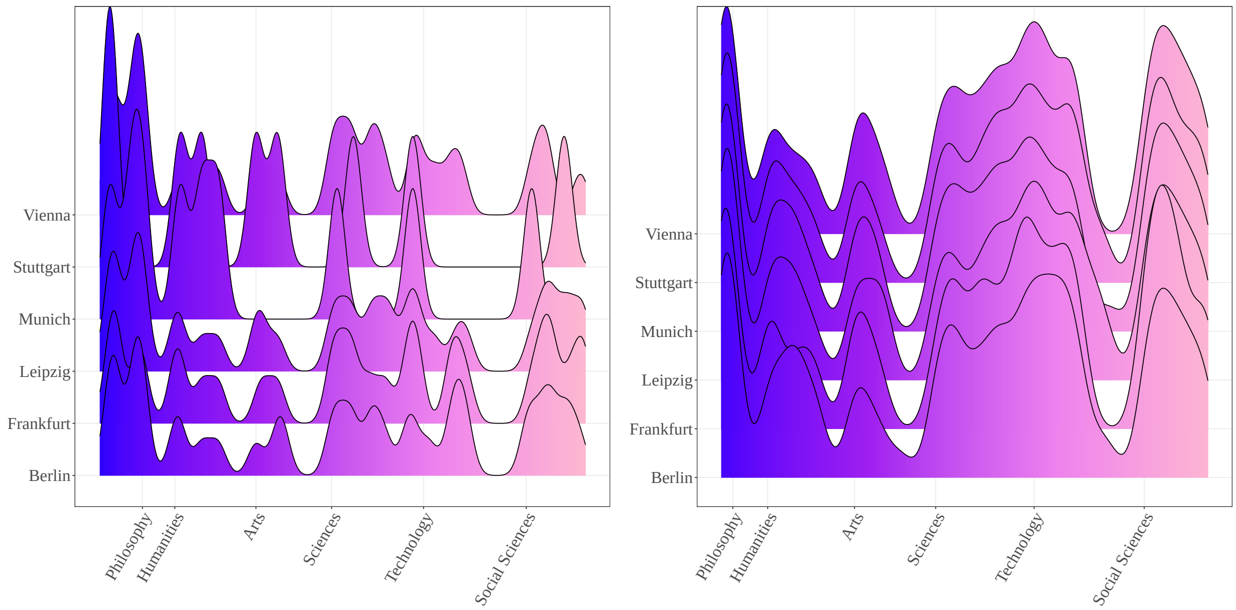


Figure 3: Relative changes in topic distribution between 1750 (left) and 1900 (right) for the main publishing cities. By the end of the time period, a larger share of publications is in science and and technology (light purple) and social sciences (pink), compared to the previous dominance of the humanities (blue).

City names. Within the catalog, the same city can be reported with several different names. To match each name to its unique city, I collect all versions of cities' names found in the German Biographies (described below), which are geolocated, and I merge them with each city's historical territory from Bogucka, Cantoni and Weigand (2019). I refer to Graesse (1861) *Orbis Latinus* for the Latin names and synonyms of cities.

Vollständiges Bücher Lexicon (Index Locupletissimus Librorum) by Christian Gottlob Kayser (and successors). To understand whether library catalogs are reflective of the contemporaries' perception of relevant publications, I collected and digitized the first complete series of biographies, as compiled by a team of contemporaries starting in 1750. This series is recognized as the first serious attempt to record all yearly publications in a country (Domay, 1987). A sample page is in Figure can be found in the Appendix.

Histories of printers, editors and book traders. Given the importance of the publishing industry for the analysis, and how it could affect the results, I furthermore collected and digitized a novel data source (*Deutsche Buchdrucker, Deutsche Buchhändler*), the original collection compiled in 1903 with the sponsoring of the German Association of Booksellers

(*Börsenverein der Deutschen Buchhändler*). Across 6 volumes of 1200 pages, detailed histories are provided of all the major active printers, editors, and retailers in German cities of the previous centuries. A sample page can be found in the Appendix. From such descriptions, I recorded: the name, location of the firm, year of opening and transferal of ownership, the technology used, and the specialization (if any).

Railroad Openings. The existing literature considering the effect of the German railroad (Hornung, 2015) uses information reconstructed from maps covering only a few years, which is not sufficient to use the current methods in causal inference for infrastructure construction. As described below, it is necessary to rely on appropriate counterfactuals. I therefore collected and digitized the original source behind the maps, the *Handbuch der deutschen Eisenbahnen*. This describes, for every year, the opening of new lines, including every single stop, the railroad company, the law through which it was approved, and the exact date in which it was opened. I gathered this information for roughly 10,000 stops (or lines expansions), i.e., the connections opened in the first 100 years of history of the railroad. For every stop, I adjusted the name of the station to match the city names mentioned above.

Railroad Statistics. Additionally, I merged this data with the yearly railroad statistics (*Deutsche Eisenbahn-Statistik*), collected by the directorate of the Association of German Railway Administrations (*Verein Deutscher Eisenbahn-Verwaltungen*) since 1850, and based on precise surveys of each existing railroad companies. The information entailed in the 18,000 pages I collected includes trade volumes, passengers by class, military, mail volumes, telegraphs (length of lines, number of apparatuses by types), and crucially, a description of construction and construction delays, and the reasons for the delays. A sample page of such descriptions is in the Appendix.

Handbook of the Natural Sciences At a time when patents either did not exist, or were not widely used or reliable, a more adequate source for innovation used in the literature are detailed collections of technical and scientific discoveries recognized by the scientific community. I hence digitized the *Handbuch der Naturwissenschaften*, specifically German inventions starting in 1800, reporting the place of the invention, and if not recorded, the main place of operation of the inventor, the year, the inventor, and the description of the innovation, based on which the field of the invention or discovery was assigned.[14]

City-level data. A widely used source in the literature is the German City Book (*Deutsches Städtebuch*), containing information of 2,390 towns, which define my sample. From this source, I digitized panel data for each town's population, as well as the number of printing houses and newspapers, and the years they were founded.

I integrated this information with the founding of each university and institute of higher

[14] I hired an engineering and a physics student for the assignment into scientific fields.

education, with the specialization of the university, or departments founded, as reported in Rüegg (2004).

Biographies. I furthermore scraped roughly 230,000 biographies of notable individuals born between 1700 and 1900 from the *Deutsche Biographien*.[15] The latter consists of an harmonization of all the major collections of biographies in the German language. Based on the scraped information, I created a database with the name, description of the profession, life dates, religion, description of family background, connections to other individuals, and relevant cities for each individual. When useful, I complemented this information with the Wikipedia page of the retrievable authors in the bibliographies.

Protest data and other controls. Data on protest and insurrections was collected by Tilly (1980), and controls for Prussia such as literacy, industrialization, agricultural output are from the Prussian Economic History Database (Becker et al., 2014).

5 Methods: Natural Language Processing, Topological Data Analysis, and Identification Strategy

5.1 Measuring Ideas

A key challenge for the purposes of the project lies in operationalizing and measuring "ideas" as accurately as possible. To do so, I measure ideas in three different ways. First, to measure the extent of specialization, I rely on labels assigned to each title by thousands of trained experts, i.e., the subject classifications provided in the bibliographic record. As a baseline, I use the topics following the *Basisklassifikation* system, which is the standard one used in the catalog of the Library Consortium. The labels follow a hierarchical system, in which closer numbers indicate topics closer in substance. For instance, within medicine (discipline 44), 44.44 refers to parasitology, and 44.45 to immunology; or within law (86), 86.44 covers constitutional law, 86.47 administrative law. As a robustness check, I use a different classification system, the second most used one (*Regensburger Verbundklassifikation*).

A second measure takes advantage of linguistic properties of the German language: the frequent use of concatenated compounds. This means that, compared with English, a new concept can be represented by a new word, and new words are much more frequent. For reference, the vocabulary of the bibliographic records is composed by 600,000 unique words, while text data from the same time period in England (Eggers and Spirling, 2014) has about 100,000 unique words. For a concrete example, consider the invention of the new

[15] As a reference, similar studies covering multiple centuries have information on about 40,000 notable individuals (Tabellini and Magistretti, 2022; Chaney, 2022).

printing press described above. Once ‘cylinder’, ‘printing’, and ‘press’ are known, no new word is created in English. In German it was, as ‘*Zylinderdruckmaschine*’. When using this approach, I identify and remove common words by looking at words’ frequency and exploiting the first 50 years of publications I collected (1750-1800). New words capturing new ideas are then almost exclusively nouns, which are less prone to change in their meaning over time compared to other lexical categories (Hamilton, Leskovec and Jurafsky, 2016). As a robustness check, I cluster the static word embeddings of new words using $k-$means clustering, where the 300-dimensional embeddings are trained using word2vec, on the entire corpus of German titles, plus 800,000 titles in Latin translated into German.

5.2 Measuring Knowledge Gaps

One of the analyses below aims at measuring the lifetime of conceptual gaps, where the closure of such gaps can be considered a new idea – i.e., a previously unexplored connection between words is made. To do so, I exploit recent methods in topological data analysis, specifically the use of persistent homology to measure the creation and filling of gaps in a conceptual network, based on previous work in topology and computational linguistics (Carlsson, 2009; Christianson, Sizemore Blevins and Bassett, 2020). The process is as follows. First, I build a concept network (or “filtration”): based on the text data, I measure the first time a word is introduced (that is, a new node is created in the concept network) and the first time it co-occurs with any other node (a new edge appears in the network). With the times of introduction and co-occurrence, the relative distance between any two points is measured – thus creating a sparse matrix with about 65,000,000 distances. The goal is to understand how long it takes between the formation of a new gap (a disconnected component is added to the network, or a network gap appears through new nodes) and the time it is filled (the previously disconnected component is linked to an existing component). I refer to this as the lifetime of a knowledge gap. An intuitive representation is in Figure 4. As a concrete example, it took 54 years for ‘tuberculosis’ and ‘bacterium’ to co-occur in a text.

To measure this lifetime, distance in time is translated as distance in a simplicial complex. The purpose of persistent homology is to identify gaps between a point cloud, in a way is robust to perturbations of input data, and is independent of dimensions and coordinates (Otter et al., 2017). The lifetime of gaps between the points, or the words in the concept network, is measured by filling convex hulls of increasing size between the points.[16] The time of death of a knowledge gap is then given by the size of the convex hull required for the gap

[16] In technical terms, representing the concept network as a sequence of Vietoris-Rips simplicial complexes; in simple words, creating circles or spheres of increasing sizes around the points.

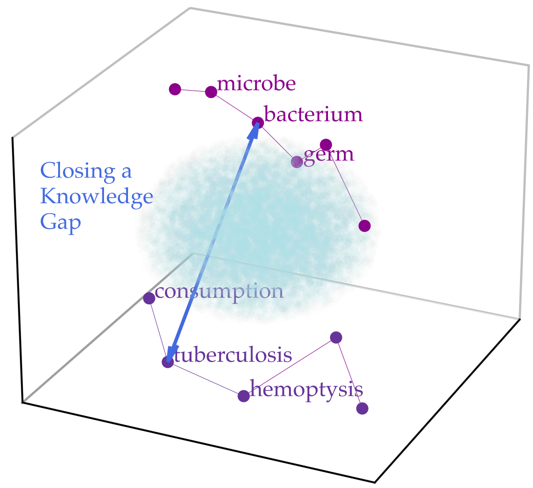


Figure 4: Visualization of the closing of a knowledge gap: the distance between the two previously disconnected components indicates the time between a knowledge gap appears and when it is filled through a connection.

to be filled.[17]

5.3 Identification Strategy

To study the effect of the railroad, one can think of two different approaches. First, understanding its introduction as a one-time event, and consider whether a city is connected to network or not, with the immediate associated lowering of transportation costs. Second, as in previous literature on the effect of the railroad (Donaldson and Hornbeck, 2016), one can think of the railroad network as a continuum: being connected to the railroad has varying intensities, depending on how fast one can get to all other cities, weighted by their size; and as the network expands and it gets faster to reach more cities, the treatment intensity increases.

To measure the first, I use a standard event study design, with city-specific time trends:

$$Y_{c,t} = \sum_{T} \alpha_T D^T_{c,t} + \beta_c + \lambda_t + \gamma_c \cdot t + \nu_{c,t} \tag{3}$$

With $y_{c,t}$ the outcome of interest for city c at time t, T the relative time since the arrival

[17] Note that some gaps are never filled: some connections are either not made within the time period studied, or are not reasonable connections. Such gaps have infinite lifetime and are excluded from the analysis; it is useful to note that their number remains quite stable throughout the century.

of the railroad, β_c city fixed effects, and λ_t time fixed effects, and $\gamma_c \cdot t$ city-specific time trends. In the appendix, I also report estimates adjusted following De Chaisemartin and d'Haultfoeuille (2020) and Sun and Abraham (2021) to account for contamination from heterogeneous treatment effects.

In some specifications, the unit of analysis is not a single city, but rather any pair to cities, and the treatment is whether they are directly connected. The specification is hence modified as:

$$Y_{ck,t} = \sum_T \alpha_T D^T_{ck,t} + \beta_{ck} + \lambda_t + \nu_{ck,t} \tag{4}$$

With $Y_{ck,t}$ the similarity between city c and city k, and β_{ck} the city-pair fixed effects.

To estimate differential effects of railroad lines with different characteristics, I use triple-differences of the form:

$$Y_{c,t} = \sum_T \alpha_{1T} D^T_{c,t} + \sum_T \alpha_{2T} D^T_{c,t} \cdot X_c + \beta_c + \lambda_t + \nu_{c,t} \tag{5}$$

with X_c the category of interest (such as whether city c is on a line with high volumes of trade relative to passengers).

For the second, continuous, approach, the standard measure of market access (Donaldson and Hornbeck, 2016) is to consider the sum of the connections to all other cities, weighted positively based on their population size, and inversely proportional to their relative distance. I estimate the effect of standard market access, and I also adapt this measure to my purposes, i.e., I measure intellectual market access for city c at time t:

$$MA_{ct} = \sum_{k=1}^{N} \frac{Y_k}{\tau(Network_t, Loc_c, Loc_k)} \tag{6}$$

where N is the total number of other cities, Y_k is either population size, intellectual production (number of books in the previous decade), or the intellectual stock (number of books published since 1750) in city k. $\tau(Network_t, Loc_c, Loc_k)$ is a function of the travel time between c and k based on the state of the railroad network at time t. The specific functional form follows Borusyak and Hull (2023).[18] The travel time is calculated for the shortest path between two cities using Dijkstra's algorithm.

The identification strategy relies on the most recent approach in the literature on non-random exposure to random shocks (Borusyak and Hull, 2023). The motivation is straight-

[18] The functional form accounts for the elasticity of movement with respect to distance: $\tau_{ck}^{-\theta}$, where τ is the distance between two cities, and θ the elasticity.

forward: cities that have higher exposure to market access, i.e., that are more central to the network, may be significantly different from cities are more peripheral.

The counterfactual for railroad lines that were actually build can be constructed based on random draws among the lines that were planned to be built, but were not for exogenous reasons. As mentioned in the previous section, I collected detailed information on the approval date of each line, the time it was actually built, and if there were delays, the reason for the delays or missed completion. Once a line was already approved and planned for construction, the reasons for delays are usually linked to causes that are indeed not related to the line itself. For instance, the line between Cologne and Minden was interrupted multiple times since 1856 because of floods from the Rhein river; the line between Krefeld and Nimwegen (Nijmegen) was delayed in 1858 because of interruptions in the continuation of the line in the Netherlands; the line Aussig-Teplitz was delayed in 1864 because of legal issues related to the ownership of metal-producing factory; and so forth. Roughly 15% of the lines planned between 1850 and 1870 experienced significant delays, with 10% of the lines taking more than 10 years to be completed, consistent with the secondary literature stressing complications in construction (Heinze and Kill, 1988). Figure 5 shows standardized mean differences in pre-treatment number of publications, population size, market access, and distance to Berlin between the cities that witnessed delays in construction versus cities that do not. Such differences are small and not statistically significant, alleviating the concern of selection of cities experiencing delays.

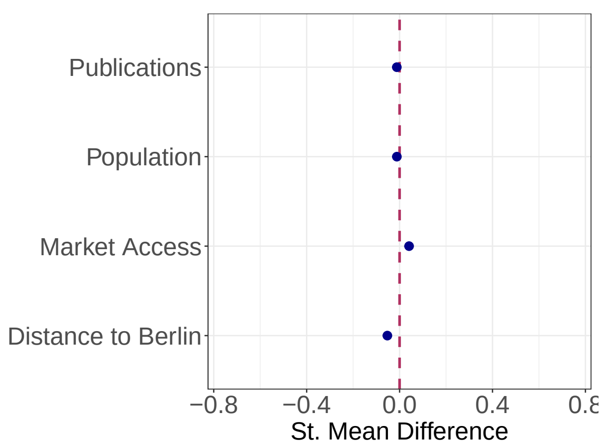


Figure 5: Standardized mean difference between cities that saw their railway lines completed on time versus those that witnessed delays, for pre-treatment number of publications, population, market access, and distance to Berlin.

Given this set of counterfactuals, one can "recenter" the treatment effect: from the observed effect in the units that did receive the treatment, the effect measured from random draws of the counterfactual lines (the expected treatment) is subtracted, with the goal of removing the omitted variable bias of being a city that is central within a network. In practice, the 'recentering' is executed by repeatedly permuting treatment status between completed and delayed lines, and using the difference with the average of the market access measures over all permutations as an instrument.

6 Results

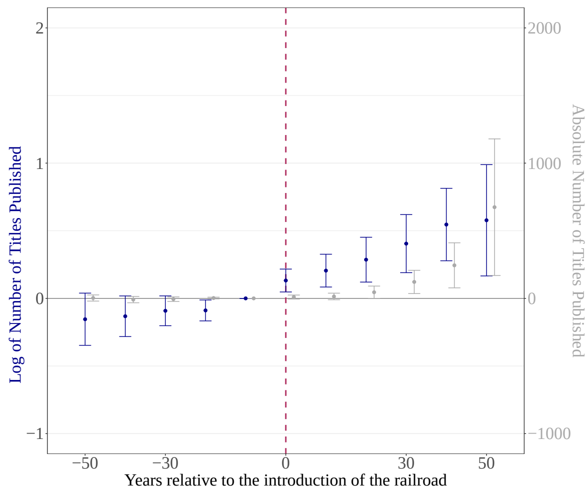
In the following section, empirical evidence is presented showing how the introduction of the railroad increased knowledge production and innovation. Next, two mechanisms are explored, which investigate the consequences of creators' and ideas' mobility. Last, alternative explanations are considered, and descriptive evidence for the consequences of specialization for real outcomes is presented.

6.1 Increase in Knowledge

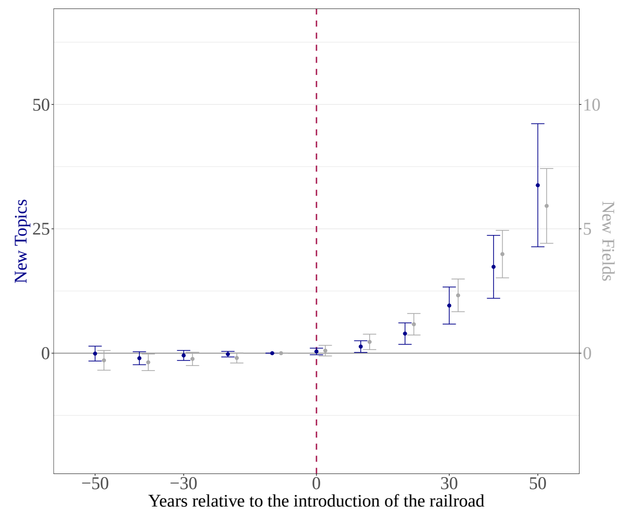
Did the railroad increase knowledge production? The first analysis tests whether cities' increase in market access stemming from the introduction of the railroad increased the number of publications at the city level. Table 1 reports the results from regressing the log change in number of publications in each city between 1850 and 1870 (the main period of the expansion of railroad, and right before unification) on the log change in market access in the same period, controlling for region fixed effects, pre-treatment number of publications, population size, centrality, and distance to Berlin. While the first column reports the results from an OLS regression, the second column uses recentered market access as an instrument for the increase in market access, suggesting an elasticity of 8-12%. With an average market access increase of 1.02 log points, and an overall increase in number of publications of 0.69, the results indicate that the increase in market access thanks to the railroad explain roughly 11% of the increase in knowledge production.

To follow the dynamic effect of the introduction of the railroad over time, figure 6a follows specification (3), where the outcome is the number of unique titles published in a given city, measured before and after a city is connected to the railroad for the first time, including for city and decade fixed effects, and city-specific time trends. In this specification as well, the quantity of publications does increase over time: on average, a connected city has about 50 more publications after three decades of being connected to the network, where the gray lines in the figure indicating the raw number of titles.

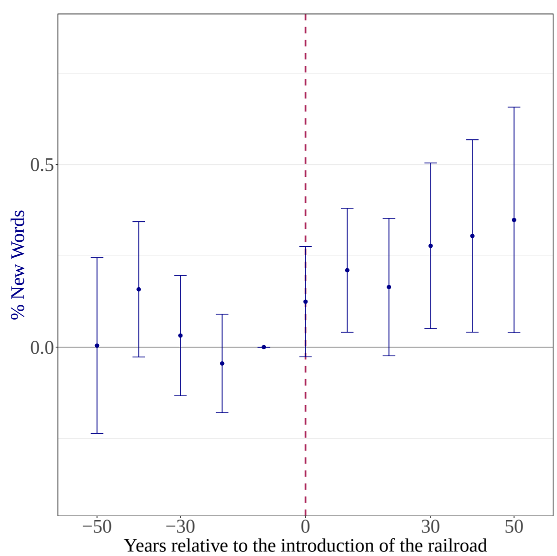
The increase in intellectual production is not only related to the quantity, but also to



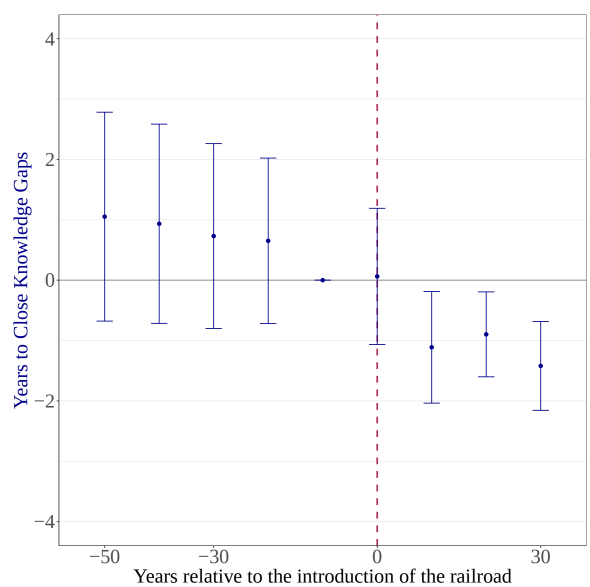
(a) Number of titles published before and after the introduction of the railroad, raw and per-capita.



(b) Number of fields and topics published before and after the introduction of the railroad.



(c) Share of new words created before and after the introduction of the railroad.



(d) Number of years to close knowledge gaps.

Figure 6: Increase in knowledge and new ideas with increased connections.
Differential effect relative to the introduction of the railroad on: the number of titles published in a given city (top-left), the number of fields (gray) and topics (blue) published in a given city (top-right), the number of new words as a fraction of total words, multiplied by 100 (bottom-left), and the number of years to connected previously disconnected words. The definition of fields and topics follows the *Basisklassifikation*. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level. Regression tables are in the Appendix. Further robustness checks are also in the Appendix.

	Dependent Variable: Growth in Number of Titles	
	OLS	*Recentered*
MA Growth	0.1213***	0.084***
	(0.0358)	(0.0135)
Region FE	✓	✓
Controls	✓	✓
Number of cities	2114	2114
First-stage F-Stat		90.9

Table 1: Log change in cities' market access between 1850 and 1870, and log change in the number of publications at the city level. The specifications include region fixed effects and pre-treatment controls. The first column reports the results from an OLS regression, while the second column reports results from recentered market access as an instrument for change in market access, following (Borusyak and Hull, 2023).

the diversity and innovativeness of the content produced. Figure 6b reports the number of topics on which publications are written each decade before and after the introduction of the railroad: after three decades, eight new topics appear in connected cities on average, fifteen after four, and so forth. The number of fields increases similarly. I next measure the share of new words used in titles in each city (rather than the absolute number – as there are more titles published, the chance of any word to be new is mechanically higher). This number significantly increases as well, as reported in Figure 6c.

Next, I study how new ideas form conceptual gaps, and how long it takes for those gaps to be filled. Gaps are closed once previously disconnected components of the conceptual network are connected, and it is reasonable to expect that substantively unrelated components should not (and never) be connected. These are treated as having infinite distance, and are excluded from the analysis, as are the knowledge gaps that are not closed within the period of study.

I measure the lifetime (creation, and time until closure) of knowledge gaps at the city-decade level, meaning based on whether a new idea (a node in the conceptual network) first appears in a given city in a given decade and creates a knowledge gap, and when that knowledge gap is also closed within the city.[19] The results are reported in Figure 6d, where the outcome is the average lifetime length of a gap. It decreases over time, suggesting conceptual gaps being filled more quickly. These results are again supportive of an increase in knowledge production, specifically an increase in the speed in creating innovation, conceptualized as the combination of existing ideas in ways previously unexplored.

[19]The gaps created earlier in the century have more time to be filled within the time period of study (120 years at first versus 20 at the end), which could bias the results so that the knowledge gaps created toward the end are filled less frequently, but in a shorter time. I hence only consider the gaps that are closed within 30 years, and stop the analysis with the 1880s, to ensure consistency across time.

The picture is quite clear: access to the network and a larger market increase idea production, its diversity, and increases innovation. Note that this confirms quantitatively previous research in intellectual history pointing out to the railroad as one of the enablers of knowledge production in this time period (Pfeiffer, 1928).

6.2 Mechanisms

The main mechanisms considered are a consequence of scholars' and ideas' mobility. As a preliminary step, it is worth assessesing whether scholars and ideas indeed traveled differently after the connection to the railroad.

Figure 7a reports the results of an event study following specification 3, where the outcome is the number of years it takes for a new idea to reach other cities, conditional on reaching them. This measure captures the speed of idea diffusion: after the railroad, ideas spread significantly faster, reaching other cities a decade earlier than they would have otherwise.

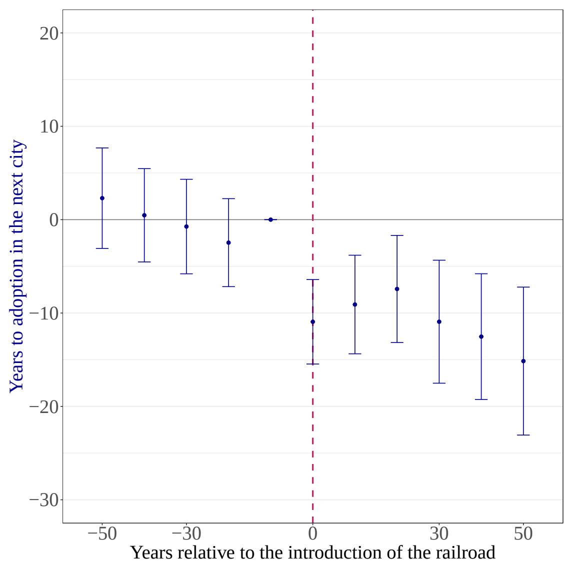
Figure 7b reports instead the number of destinations that notable individuals move to, aggregated at the city of birth, after the railroad is introduced in their towns of origins. Note that, as notable individuals tended to be quite mobile compared to the rest of the population even before the railroad, it is not clear that its introduction would significantly affect them. Figure 7b suggests that their mobility did indeed increase. Even taking into account city-specific time trends, on average they move to one additional city in their lifetimes, with the effect increasing over time.

6.2.1 Mechanism 1: Scholars' Mobility and Specialization

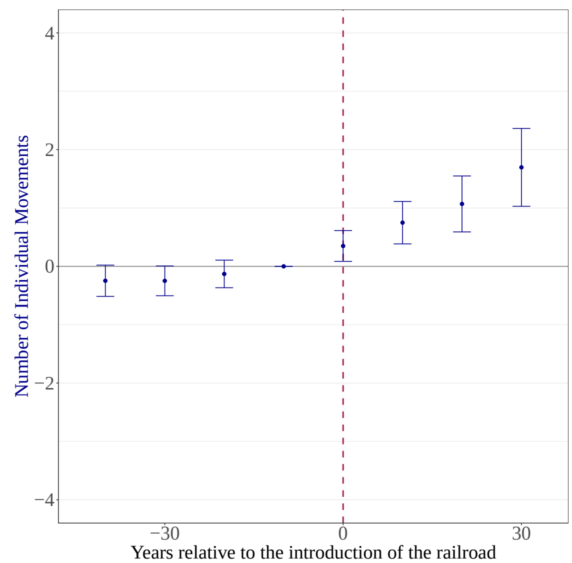
Having established that the railroad did have a considerable effect on knowledge and individuals' mobility, the following analyses test whether scholars' mobility facilitated the co-location of specialists, allowing cities' specialization in knowledge production and local knowledge spillovers.

If scholars were more likely to move, where did they tend to move to? A reasonable expectation would be that individuals would move to larger cities. To check whether this is the case, I rank each city based on the number of scholars related to that city in the three decades surrounding an author's decade of birth, and I test whether individuals become more or less likely to move to a larger intellectual hub – which is correlated to moving to a larger city. This does not seem to be the case, as shown in Figure 8a.

Another possibility, suggested by the theory outlined above, is that scholars would be more likely to move to a cluster of other intellectuals active in a similar field. To do so, I use the biographies' standardized description of their profession. To group them into fields, I create



(a) Number of years it takes for a new idea to reach other cities.



(b) Number of destinations notable individuals move to.

Figure 7: Increase in ideas' speed of diffusion and scholars' mobility. The left figure shows the differential number of years it takes for a new idea, measured through new words, to be used in a different city compared to the city it first appears in, relative to the time when the city of origin is connected to the railroad network. The right figure shows the differential number of movements to other cities in an individual's lifetime, relative to the time the railroad is introduced in their town of birth. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level. Regression tables are in the Appendix.

professional categories that reproduce the disciplines used in the bibliographic information, and I assign individuals to a field based on keywords related to a profession found in their description. For each field, I rank cities based on the number of scholars active in that field in the three decades around an author's birth, as above.

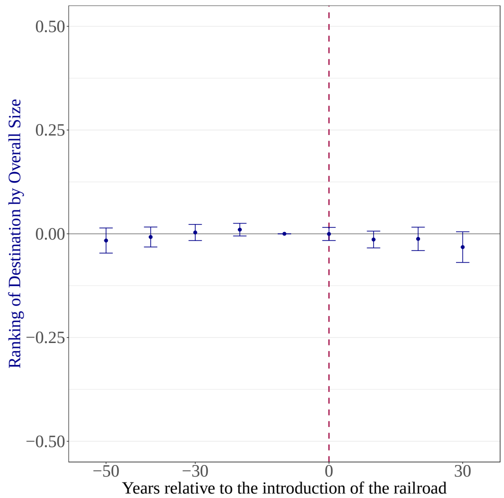
The outcome is the difference in field-specific rank between the city of destination and the city of origin. Figure 8b shows that with the advent of the railroad, scholars are more likely to move to a city that has a larger cluster of scholars in the same discipline. Taken together, the results suggest that with the railroad scholars are more likely to 'spread out', and move not necessarily to larger cities, but to cities where there is a higher concentration of similarly minded individuals. With mobility, clusters of specialized individuals are able to form and crystallize.

As a concrete example, consider the life trajectory of physicist Erasmus Kittler, the son of a tailor from Schwabach, outside of Nürnberg. He first received his training in physics in Nürnberg and Münich; but the town he permanently moved to was indeed smaller and less-known (Darmstadt). There, he joined a newly born university for engineering, and founded the first electrical engineering department in the world, training some of the most successful engineers and inventors of the next generation.

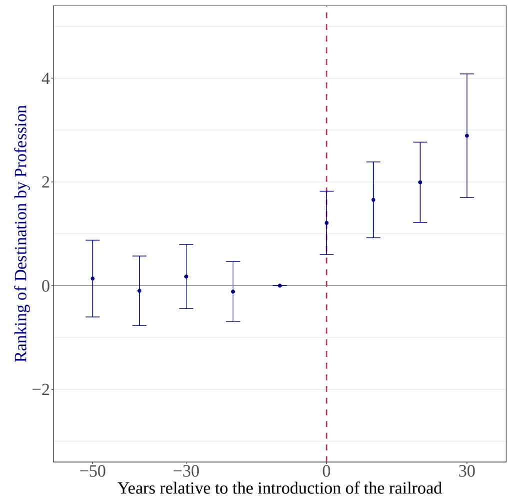
This example suggests that mobility may also have an effect on the composition of idea producers. Half a century before, it would have been hard to imagine the son of a tailor becoming one of the founding fathers of electrical engineering. A side analysis hence seeks to understand changes in individuals' familial background. To do so, I extract from the early paragraphs of the biographies the sentences describing their family, their mother, or their father. Sentence-level embeddings from pre-trained multilingual S-BERT (Reimers and Gurevych, 2019) are then run on these descriptions. The same analysis is done with two separate sources of data: the biographies from the *Deutsche Biographien*, as well as the Wikipedia pages of all writers for whom I can match name, last name, and birth and death date to the information in the bibliographies. To measure whether individuals come from a more or less traditional background, I calculate the cosine similarity between the sentence embedding of any scholar of the 19th century, with those of scholars from the 18th century (which I use as a reference to define a traditional background). The results from the *Biographien* are in Figure 8c, from Wikipedia in the Appendix. In both sources of data, the results are consistent with diversity increasing one generation after the introduction of the railroad. In other words, after 20-30 years since the advent of the railway in a given city, scholars born in that city tend to have family's descriptions that are less similar to those of scholars born in the previous century, who tended to come from the nobility and the upper classes. Mobility had a democratizing effect on the intellectual elites, expanding the range of individuals who could access and produce knowledge.

In sum, scholars tend to move and cluster with individuals in the same profession. With the exposure to a larger market, do individuals specialize more? Do cities tend to specialize as well?

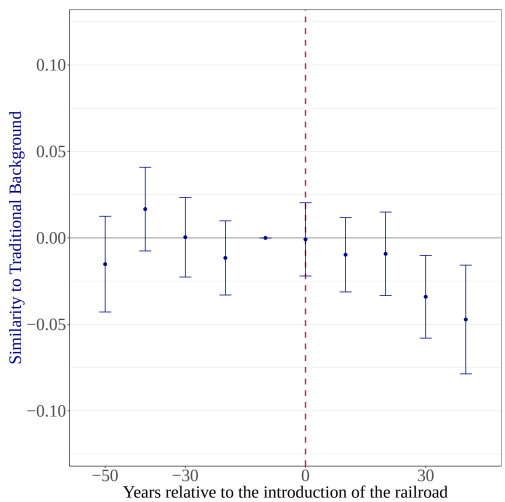
Figure 9a reports the results of an event study where the outcome is the probability of any individual to be active in multiple fields. Figure 9b reports instead city-level specialization. As cities are connected to a larger market, we would expect them to start specializing in certain fields of knowledge. To measure concentration, I estimate the Hirschman-Herfindahl Index of disciplines within a city, and of topics within those disciplines. In other words, I calculate to what extent publications are mostly centered around a few fields (or subfields), versus evenly spread across many. The results suggest a clear increase in specialization in both cases after access to the railroad network. In line with suggestive evidence, the measurement confirms, for instance, that Göttingen has a high concentration in mathematics and theoretical physics; Weimar literary; Stuttgart, Munich, Darmstadt and Aachen in



(a) Destination city's rank based on total number of individuals.



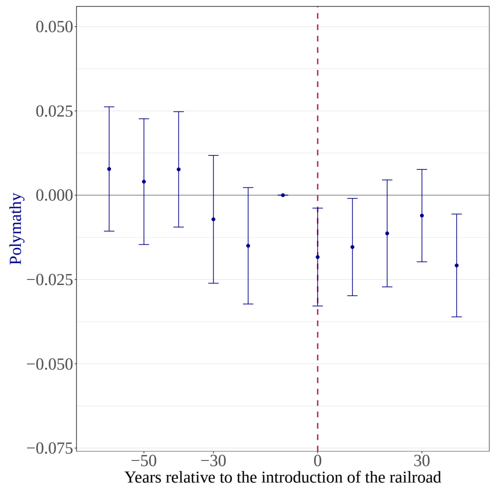
(b) Destination city's rank based on individuals in the same profession.



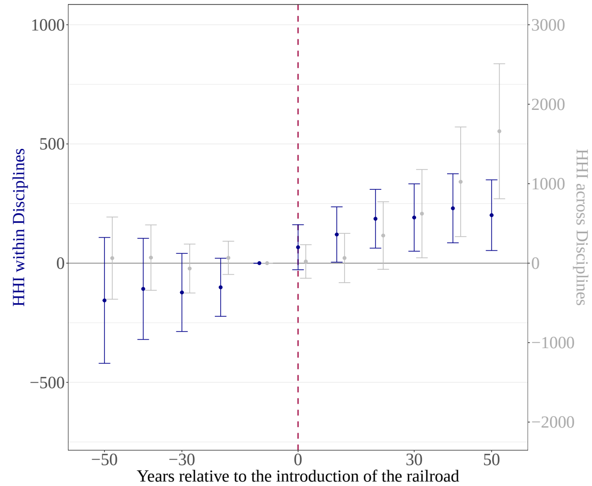
(c) Similarity of scholars' family background to the traditional background in the 18th century.

Figure 8: Scholars' movements and family background.
The top-left figure shows the differential effect on the rank of cities individuals move to, where cities are ranked based on their size, relative to the introduction of the railroad. The top-right figure shows instead the difference in ranking of each individual's professional cluster, where the ranking is calculated based on the number of individuals active in the same profession in the three decades around each person's birth. The bottom figure shows the differential effect in the similarity to a traditional background, calculated extracting the description of individuals' family of origin, and using S-BERT sentence embeddings to measure the cosine similarity with the typical family background of the late 18th century. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level. Regression tables are in the Appendix.

technology and engineering, Berlin in political science, physics and medicine.



(a) Individuals' activity in multiple fields.



(b) Concentration of publications on disciplines, and topics within disciplines.

Figure 9: Specialization of individuals and cities with increased connections.
The figures show the differential effect relative to the timing of the railroad on: whether an individual is active in multiple groups of professions (left), and on the right, the city-level specialization on fields (gray) and topics (blue). Specialization is calculated based on HHI. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level. Regression tables are in the Appendix.

To improve identification, I follow the strategy outlined in 6. For every city, I estimate the effect of traditional market access (based on city size), or intellectual market access (based on number of publications) and intellectual stock (based on the historical number of publications). The results are in Table 2. While market access is correlated with specialization, the results become not significant when removing the selection effect of being centrally located. Access to the intellectual stock is not significant – the 'burden' of previous knowledge does not seem to affect current specialization. Access to other cities' contemporary intellectual production, on the other hand, is correlated with specialization and is significant even after recentering the treatment. The results show that exposure to a larger market through increased connections leads to significant increase in cities' specialization in knowledge production. As scholars became more productive when active within a cluster, specialization and agglomeration effects played an important role in the increase in knowledge, with the railroad contributing to an increase of 10% in cities' specialization compared to the pre-treatment average.

	Dependent Variable: Knowledge Production Specialization					
	OLS			*Recentered*		
	(1)	(2)	(3)	(4)	(5)	(6)
Market Access	258.8			213.6		
	(140.3)			(181.1)		
Intellectual Access		283.4			190.4	
		(107.1)			(108.2)	
Access to Intellectual Stock			115.4			100.3
			(195.4)			(101.2)
Region FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Number of cities	2085	1290	1458	2085	1290	1458
First Stage F-Stat				28.62	37.99	35.73

Table 2: Market access, intellectual access, intellectual stock and knowledge production specialization. Specialization is measured through a HHI index of concentration of publications. The market access measure weighs every other city in the network based on city size, and negatively based on distance. Intellectual market access is a function of the number of publications. Intellectual stock weighs the historical number of publications. Columns (1)-(3) report the results from OLS regressions. Columns (4)-(6) report a recentered treatment effect following Borusyak and Hull (2023).

6.2.2 Mechanism 2: Ideas' Mobility and Diffusion

Another channel through which the railroad could have contributed to an increase in innovation is if ideas diffuse to other cities, and new combinations of existing ideas are formed thanks to the additional diffusion. To test whether this is the case, specification 4 is used; where the unit of observation are city-pairs. The outcome is the number of new ideas formed by combining ideas stemming from the two different cities, before and after the railroad connects the two cities directly. The significant increase, reported in Figure 10, indicates that the railroad explains 16% of cross-cities new combinations of concepts.

Next, I check whether the railroad increases the diffusion of ideas. If a new idea appears in a given city before or after the railroad, in how many other cities does that idea, or that new language or topic, appear in the next decade?

Figure 11a shows the number of cities new ideas are adopted in. Counter to the intuitive expectation that new ideas would diffuse to a larger number of cities, the results point to a different pattern: on average, a new word is on average 5% less likely to be adopted in other cities.

Similarly, one would expect intuitively that as any two cities become directly connected through the railroad, they would become more similar: as knowledge and ideas can travel faster between cities, they would start writing on the same type of topics, as they would consider the same issues to be salient. I check whether this is the case by comparing any

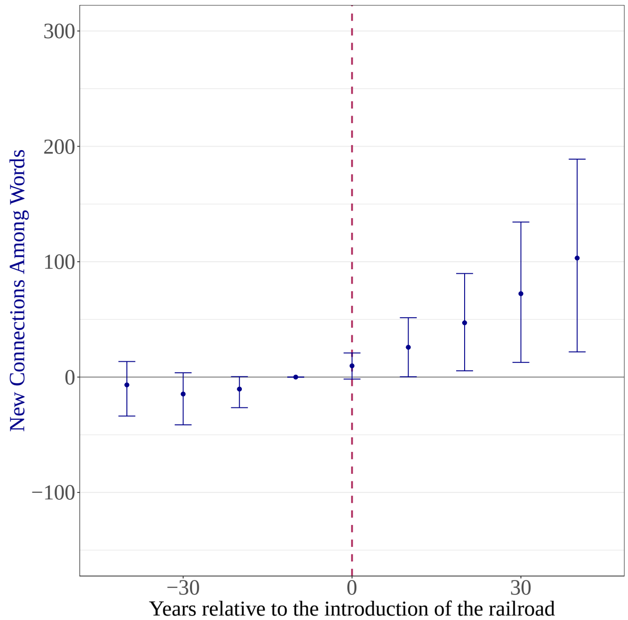
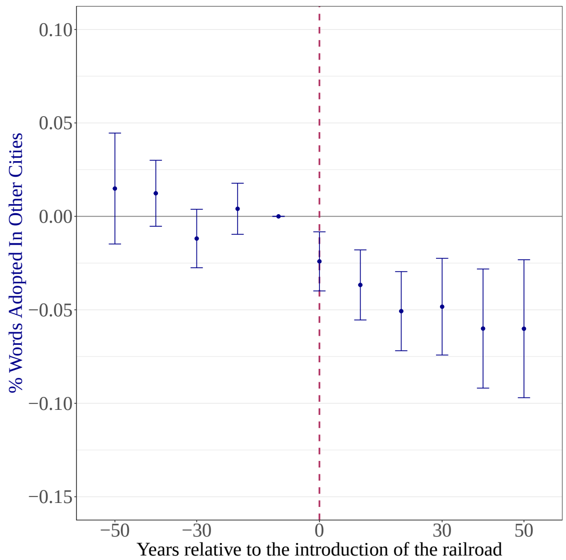


Figure 10: New ideas as combinations of words coming from two different cities, before and after those two cities were directly connected through the railroad.

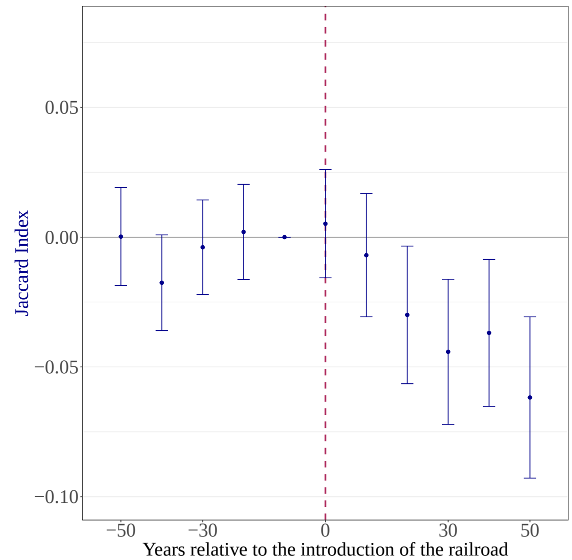
pair of cities following specification 4, where the outcome is the Jaccard Index of the fields in which the two cities have publications, i.e., the intersection of joint disciplines, over the union of all disciplines among the two. In this case as well, the pattern is clear: cities become less similar after they are directly connected, as shown in Figure 11b; the share of joint disciplines decreases by 5%. As an example, consider the town of Jena and Weimar. Before they were connected by the train, both cities mostly wrote on generally in the humanities. Afterwards, scholars concentrated in Jena for law and philosophy; in Weimar for literature and architecture.

Next, one can think of what sustains specialization: as the costs of reaching other scholars in similar fields decreases, one can circumscribe their area of expertise and still have a sustainably large audience. At the same time, if exposed to a much larger amount of information, scholars would tend to adopt only the information that is closest to their own ideas. Increased specialization should hence come with increased learning and communication across cities within the same discipline. I check whether this is the case by studying whether new words and new ideas are adopted more often within the same discipline after the connection to the railroad. The results are in Figure 11c, and they are consistent with increased learning and more diffusion of ideas within the same field.

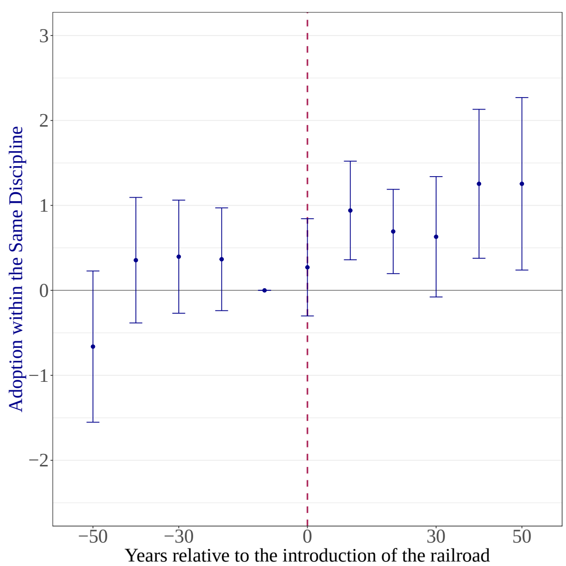
What happens to learning across fields within the same city? In Figure 11d, the outcome is the number of times a new word created in a certain discipline is adopted in any other discipline within the same city. The trend is not increasing: while there is initially no



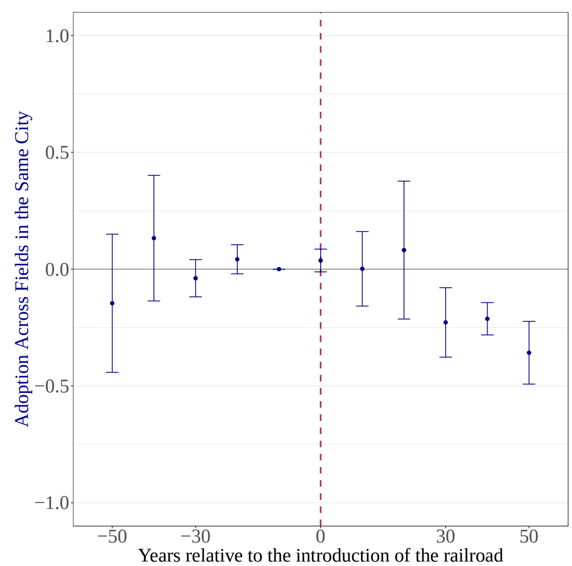
(a) Number of cities where a new idea is adopted.



(b) Similarity between two connected cities before and after a direct connection.



(c) Adoption of new ideas within the same field, in other cities.



(d) Adoption of new ideas within the same city, in other disciplines.

Figure 11: Aggregate patterns of diffusion and convergence. Differential effect relative to the time of the railroad on adoption of new ideas (top-left), where adoption is calculated based on whether a new word appears in other cities compared to the city where it first appeared. In the top-right, the Jaccard Index between any two cities calculates the intersection of common fields present in the two cities over the union of all fields present in the two cities. The bottom-left changes the adoption measure, where adoption is considered only within the same field. The bottom-right shows the differential effect of adoption of new words within the same city, but across disciplines. All estimates control for city (or city-pairs in the top-right) and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level. Regression tables are in the Appendix.

effect, within-city, across-disciplines diffusion decreases after 30 years. This is consistent with the notion that the railroad makes the constraint of physical proximity to be less binding for cross-disciplinary learning. Following the logic above, with easier communications with similar scholars (both within the same city, and with the ideas coming from other cities), it is less likely that ideas will be adopted from other fields, no matter how physically close.

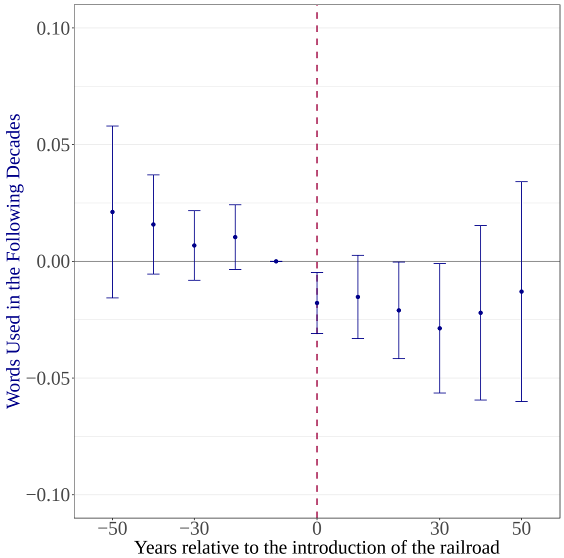
Taken together, the results suggest that with the expansion of the market for ideas, specialization in knowledge production increases. New ideas are formed thanks to field-specific diffusion of information. While the evidence points to an increase in innovation and productivity, it also identifies a decrease in diffusion: different fields create their own language and their own niche, and stop learning from dissimilar fields.

6.3 Selection of Ideas that Diffuse

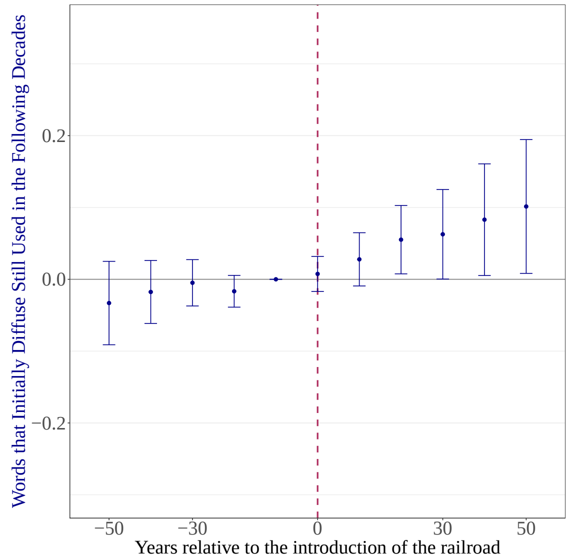
The results presented so far are suggestive of diffusion of ideas occurring less overall, more within their respectively fields, but with each new idea impacting a smaller share of other intellectuals. Yet, the intuition coming from well-known anecdotal evidence would suggest otherwise. Darwin's theory of evolution soon became very well known, and affected science more broadly. Marxist ideas and ideology spread across the country, despite censors' efforts.

Note that the results presented above were the average across all new ideas. Restricting the analysis to the 'successful' ideas, meaning those that did spread sufficiently within the first decade of their first appearance, the patterns are quite different. If new ideas spread at first, they are more likely to diffuse even more over time, as plotted in Figure 12b. Figure 12a plots the persistence of new ideas on average and if they diffuse initially (12b). On average, new ideas are less likely to persist; but if they do spread initially, they are more likely to do so. This suggests a starker difference in diffusion between 'general purpose' and field-specific knowledge coming from network density. A clear example of this difference within the political realm are the cases of nationalism and feminism. While feminist ideas appeared in 1848, they failed to diffuse extensively until the very end of the century. At the same time, nationalistic ideas consistently increased their diffusion (despite changing forms) throughout the century.

Summing up, while increased connections lead to an atomization of intellectual production, and the chances of success (diffusion and persistence) for a given idea decrease, at the same time, the returns from initial success are larger. Especially ideas that are able to adapt and be adopted across different disciplines are less likely to be replaced by other novel ideas. It is interesting to note how most of previous research has focused (very reasonably and quite naturally) on successful ideas, those that spread quickly and persist; while the dynamics are



(a) Persistence of ideas, on average.



(b) Persistence of successful ideas.

Figure 12: Persistence of initially successful ideas.
Differential effects relative to the time of the railroad on whether new words are still used in the following four decades since their first appearance. On the left, all new words are included. On the right, only words that are adopted in other cities within their first decade are included. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level. Regression tables are in the Appendix.

very different for the ideas produced in the process of normal science in the sense of Kuhn (1962).

6.4 Robustness and Ruling Out Alternative Hypotheses

I next conduct a series of robustness checks to address possible measurement errors and to test whether other mechanisms are driving the results. The analyses are reported in the Appendix.

First, it could be that the increase in knowledge production and the resulting growth in publications is due to urbanization and economic growth alone. To test this, I run the analyses in per-capita terms, where the population data is from the *Deutsches Städtebuch*. I find that while urbanization certainly was an important factor contributing to knowledge production, it does not entirely explain the increase in publications.

Second, it could be that the changes observed are driven by changes in the publishing industry. One explanation would be an increased concentration of publishing houses in fewer cities – this is not the case, as more and more publishing houses open even in smaller centers.

Alternatively, it could be that publishing houses tend to specialize, meaning they become more likely to circumscribe the topics they accept for publication, for reasons different from the specialization of local scholars. To check this, I use the detailed description of companies involved in the book industry, where I recorded whether a publishing house specialized on certain disciplines, and if so, which ones. The vast majority of publishing houses was somewhat specialized before the advent of the railroad as it was after. Descriptively, what seems to happen is instead a change in the fields of specialization, specifically a move away from religious texts toward scientific ones – in lines with previous evidence and qualitative descriptions (Pfeiffer, 1928).

Next, it could be that the increase in production is driven by the improvement in printing technology. As mentioned in the historical background, the main change from manual to steam-operated machines occurred at the beginning of the century, before the railroad was introduced. Nonetheless, to check that this is not the main mechanism, I study the technology used by each publisher as reported in their histories, and find no significant differences in timing of adoption following the railroad.

Another possible confounder is related to the use of the specific subject classification used mainly in the German Library Consortium Index. It could be that this specific hierarchical structure labels publications in a systematically different way toward the beginning and the end of the period. To check whether this is the case, I run again the analysis using a different classification system that was developed in another time period. I furthermore exploit tools created for inter-libraries exchanges that enable me to translate labels in one system into another.

Last, noise may come from writers publishing in a different city as that of their residence. One should note that there were strong incentives not to do so, especially since printing and publishing houses had spread to more or less all cities in the sample. Even if one wanted to reach a wider audience, communication with one's editor was frequent and costly if distant; while the system of book transfer in place (as described in the historical background) was such that even if one published locally could have books delivered more easily across cities. Nonetheless, it was sometimes necessary to publish in another city, especially in earlier periods in order to avoid censorship. To check to what extent this confounds the results, I compare the authors' biographies and their place of residence with the place of their publications, and estimate the number of publications printed in a city that were done by someone who did not live there. Such cities tended to be the larger cities (Berlin, Leipzig, and Hamburg because of its more liberal approach to censorship); intuitively, in smaller towns authorship was mostly local. I check the difference between residence and place of publication by matching authors' names and birth dates from the bibliographies with the biographies, and can identify roughly 20% of all authors. Note that these would be the most likely individuals to publish in a city different from that of their residence; yet, roughly 82%

of publications occur in residence. Furthermore, the results from an event study suggest that authors were not more likely to publish in a different city compared to that of their residence after the introduction of the railroad. Last, excluding the cities in which non-resident authors were most likely to publish from the main analysis does not affect the results.

An open question is what determines which field a city would specialize in. Giving a precise answer is not trivial, and still in progress. However, descriptively, correlating cities' specialization before and after the railroad points to interesting patterns: while cities initially focusing on the humanities tended to switch away from them, specialization in science and technology is more persistent, as shown in the Appendix.

7 Consequences of Specialization

The following section stresses the importance of specialization of knowledge production for real outcomes (scientific discoveries) and state institutions (the bureaucracy and legislation). First, Figure 13 confirms what the literature in urban economics and innovation would predict: with the co-location of specialists and knowledge spillovers in cities, technical innovation should increase (Krugman, 1991; Glaeser et al., 1992; Glaeser, 1999; Iversen and Soskice, 2020). There is indeed a positive correlation between a city's specialization within the natural sciences and technology and the number of major scientific discoveries occurring in the same city.

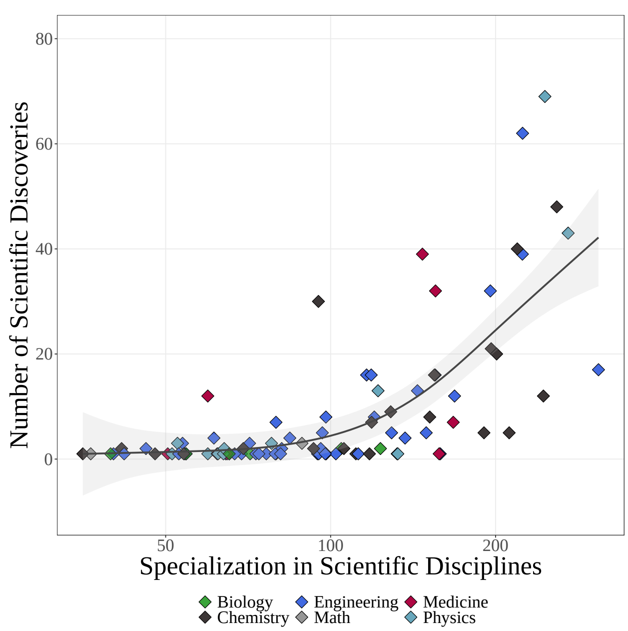


Figure 13: Specialization (squared share of publications) and major scientific discoveries by city; excluding major cities.

7.1 State Sciences, the Bureaucracy, and Legislation

As mentioned above, the trend toward specialization resulting from larger market exposure affected not just the natural sciences, but overall intellectual production. Of particular interest is the effect of specialization for the development of what became the 'legal sciences', and the 'state sciences'. Previous work in history, intellectual history and comparative law has pointed out how the specialization within these fields affected the organization, training and expertise of the bureaucracy; the reach of the state through the development of statistics; the drafting of legislation and the precision of legal codification. I address each below.

The Sciences of the State. Much has been written on the Prussian bureaucracy, whose model came to dominate the German Empire after the unification in 1871: the traditional stress has been on its efficiency and system of meritocracy, based on training and competitive examinations (Dorn, 1931; Bleek, 1972), its hierarchical structure (Toye, 2006) and internal chains of agency and command (Kiser and Schneider, 1994), on the status and privileges granted to the *Beamten*, its civil servants (Kocka, 1981; Esping-Andersen, 1990, p. 59), and its later backing of the Nazi regime (Caplan, 2014).

The bureaucracy was so relevant for the state and the military that Prussia came to be described as a *Beamtenstaat*, a 'state of public officials' (Beck, 1992, p. 263), and the Prussian administration as a 'paradigm for the forms and modifications of the life of the modern state as such' (Hintze, Oestreich and Hartung, 1970, p. 16).

It is interesting to follow the development of public officials' training. Since the 17th century, 'cameral studies' existed for the education of the bureaucracy; they then developed into the better formalized 'state sciences' (*Staatswissenschaften*) (Lindenfeld, 1997, p. 11-46). The expansion of the discipline in the 19th century was reflected in the increasing number of seminars dedicated to producing new knowledge specifically in the state sciences, such as Engel's seminar in Berlin, and Hildebrand's in Jena (Lindenfeld, 1997, p. 215-216), and the creation of a doctorate in state sciences, which would consist of economics, finance, statistics, politics, and state law (Lindenfeld, 1997, p. 234).

By the 1870s, the education served to train professionals within the public service was described as a highly articulated, specialized systems of knowledge (Lindenfeld, 1997, p. 262). Within this system, different cities and their universities, specialized institutes and technical schools, had different focuses or approaches to methods. For examples, Strassburg's school was centered on statistics and stressed empirical observations, in contrast to Berlin's focus on theory, especially on labor; and Vienna's influence by the emerging Austrian school of economics; Koenigsberg's emphasis was on legal positivism within constitutional law, versus Heidelberg's emphasis on merging the legal and political sciences (Lindenfeld, 1997, p.205-260).

The last three decades of the century witnessed further specialization within the state sciences. Administrative law transformed from a subcategory of state law into its own field, with an increasing relevance for bureaucrats' training. Specialized institutions were created for the continuing education of public officials. The studies of the national economy (*Nationalökonomie, or Volkswirtschaft*) versus the private economy (*Betriebswirtschaft*) were separated, with a later further separation of the world economy (*Weltwirtschaft*).
Volkswirtschaftspolitik (economic policy) was further subdivided, notably with a distinction between social, trade, agricultural, transportation, commercial, and colonial policy, among others (Lindenfeld, 1997, p. 286-288). Social policy developed specifically to design labor contracts, factory legislation, and relations with trade unions; agricultural policy trained students differently from agricultural studies, being focused also on farmer indebtedness, land distribution, and the flight of agricultural laborers (Lindenfeld, 1997, p. 288-292). Insurance science was established; and while lectures were offered in mathematics, economics, statistics, and law, their purpose was specifically to study the design of social insurance programs. Those trained within the discipline later joined the apparatus implementing Bismarck's social reforms and insurance policies (Lindenfeld, 1997, p. 293).
At the same time, the expansion of statistics as a discipline led to the increase in the interest and capacity of scholars and public officials to monitor the population and the economy. Research from statisticians (especially from the Strassburg school mentioned above) was coupled with censuses and surveys of increasingly greater volumes and details (Oberschall, 1997).

The specialization was not only a matter of training: it was then reflected into the specialization of officials' competencies within the bureaucratic apparatus (Lindenfeld, 1997, p.264-315). In fact, the type of training officials received was closely matched to the type of career they were entitled to and they would pursue (Kocka, 1981). As the German Empire became the first modern welfare state, by providing sickness insurance, accident insurance and old-age and disability pensions (1883-1889), it had trained professionals in social policy to draw from for administration. As industrialization required the development of regulatory bureaucracies, the expertise existed to support these processes; and the organizational capacities of the bureaucracy had spillovers into the organization of firms (Kocka, 1981).
While in the 18th century the German description of governing was summed up with the word *Klugheit* (cleverness), *Fachkompetenz* (field expertise) became the defining word of the bureaucracy by 1900 (Beck, 1992; Lindenfeld, 1997).[20] And not surprisingly, one of the

[20]Note, however, that increased specialization and meritocracy do not necessarily lead to bureaucratic quality, an issue debated in the literature on the German bureaucracy of the Empire (see Pollard, 1990;

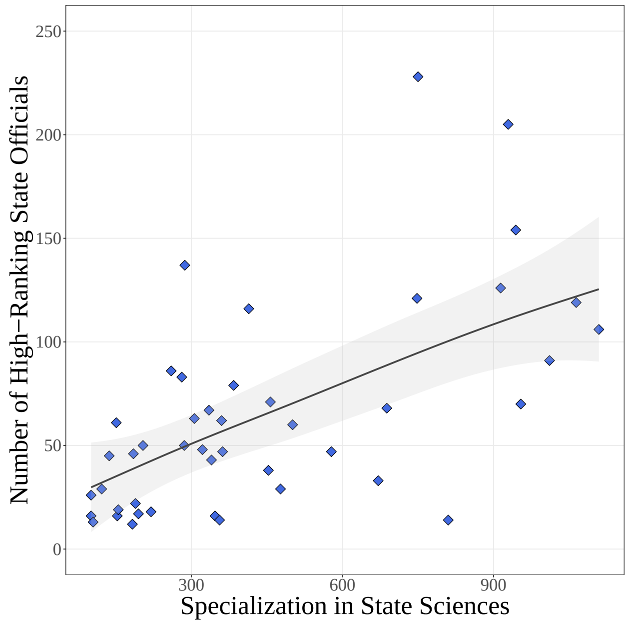


Figure 14: Specialization in the state sciences and training of individuals who will become high-ranking officials; excluding major cities.

features of Weberian bureaucracy is the 'functional division of labor, and well-defined areas of jurisdiction' (Weber, 1968, p. 223).

Civil Law and Property Rights. The development of the German Civil Code (Bürgerliches Gesetzbuch), together with the Commercial Code, also provides insights into the consequences of specialization, and is relevant for two reasons. First, it consolidated and elaborated the principles that define the German legal tradition of property rights, which have been shown to affect further development (Berkowitz, Pistor and Richard, 2003; Porta, Lopez-de Silanes and Shleifer, 2008). Second, it is widely recognized by comparative legal scholars as the most important codification effort since the Napoleonic code, has been subsequently used as the paradigm for civil law codification for several European and Asian countries. It has been described by contemporaries as 'the most carefully considered statement of a nation's laws that the world has ever seen' (Higgins, 1905).

The code, in development since 1881 and effective since 1900, is still in force (with amendments) today, and covers property law and property rights, family and inheritance law, torts and obligations, and the general principles applied to merchants, further specified in the Commercial Code, which was developed and became effective in tandem.

A comparison with the Napoleonic code makes the implications of the German legal specialization clear (Golding, 1971). Only four jurists were charged to draft the Napoleonic

Andersen, 2018; Gillis, 1968, for details).

code: two members of the Cour de Cassation, and two of the Council of State. More than 30 committee members were instead in charge of the first and second draft of the German Civil Code (Freund, 1899). Each member of the committee was chosen based on their expertise; for instance, Alexander Achilles, trained in Halle an der Saale, was a specialist in property and mortgage; Bernhard Windscheid, trained in Berlin and Bonn, in roman civil law; Gottlieb Planck, trained in Göttingen, in family law; Hermann Struckmann, trained in Heidelberg, in estate administration; Robert Bosse in occupational safety and labor protection, and so forth (Haferkamp, 2009).

The specialization, a product of the legal training of the time, is reflected in the 'scientific draftsmanship' of the text (Golding, 1971, p. 313). While the Napoleonic code 'cannot give more than an approximate indication of the law' (Golding, 1971, p. 314), the German code was compiled with (almost excessive) thoroughness and 'purports to deal comprehensively with the minutiae of juridical problems' (Golding, 1971, p.315).[21] At the same time, the doctrine provided a much clearer distinction between the private sphere and the role of the state than in the past (Haferkamp, 2008), also reflecting a clearer separation of fields within the law. Extending the comparison to the economic consequences of legal traditions, countries following the German tradition score overall better in terms of securing property rights, especially in terms of contract enforcement (Porta, Lopez-de Silanes and Shleifer, 2008; Berkowitz, Pistor and Richard, 2003) compared to those following the French tradition.

8 Conclusion

As spatial connections increase territorial integration, one could expect convergence – as cities become connected, ideas diffuse more, and cities become more similar in what they consider relevant for writing.

This study provides evidence that, if one considers new ideas in all fields of knowledge, the patterns are more nuanced: 'successful' ideas are indeed more likely to become widespread, and reach even more remote cities – in this historical setting, nationalism comes to mind as an obvious example. However, the changes in the total distribution of topics and ideas suggest instead that with increased connections, ideas start diffusing only within specific fields, which become increasingly narrow. Scholars with similar interests have lower costs in clustering, and communicating with similar clusters, thus ideas within fields travel faster. But those ideas also tend to stay in those clusters: they are less likely to reach other fields.

These results contribute to our understanding of the origins of the modern knowledge

[21] The expertise came with a downside: the German code is so technical to be considered incomprehensible outside the legal profession, much unlike the French one (Golding, 1971, p. 315)

economy. The patterns observed mark the incipit of today's specialization and co-location of specialists in cities, and the ensuing increase in productivity and innovation. Interestingly, increased specialization seem to have affected not only the natural sciences, which see an increase in discoveries and relevant contributions. The legal and state sciences specialize as well, and qualitative evidence points to an effect on legislation as well as the competence of the bureaucracy.

However, these dynamics open up more questions. Consider the German bureaucracy: while specialized, it suffered from a lack of responsiveness (Andersen, 2018). Is there a trade-off between the two? As for the sciences and knowledge production in general, specialization is nowadays so structural that it is taken for granted – none would question the separation of knowledge and research into disciplines and subfields. Yet, while specialization increases productivity, it comes possibly at a cost, especially as it further progresses – missing out on riskier contributions, but with higher possible quality (Fleming, 2007), which could explain why recently research effort has not been matched by research output (Bloom et al., 2020). The conventional wisdom asserts that specialization is due to the burden of knowledge – as knowledge increases, one needs to train longer or narrower to be an expert in a field (Jones, 2009). This has an important implication: specialization is inevitable with the growth of knowledge. The results above suggest that it may not be the case.

Last, a clear path for future research lies in connecting these findings with the development of political ideas specifically. Descriptively, social and political movements throughout the 19th century followed a trend similar to the one found in this paper: fragmentation, the creation of movement-specific language and narratives, and the development of networks across cities. To what extent were spatial connections responsible for the development of modern ideologies? How was this reflected in the establishment of political parties? Future work will address such issues.

References

Abdelal, Rawi, Yoshiko M Herrera, Alastair Iain Johnston and Rose McDermott. 2006. "Identity as a Variable." *Perspectives on politics* 4(4):695–711.

Abramitzky, Ran and Isabelle Sin. 2014. "Book translations as idea flows: The effects of the collapse of Communism on the diffusion of knowledge." *Journal of the European Economic Association* 12(6):1453–1520.

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo and James A Robinson. 2019. "Democracy does cause growth." *Journal of political economy* 127(1):47–100.

Andersen, David Delfs Erbo. 2018. "Does meritocracy lead to bureaucratic quality? Revisiting the experience of Prussia and Imperial and Weimar Germany." *Social Science History* 42(2):245–268.

Arold, Benjamin W. 2022. "Evolution vs. Creationism in the Classroom: The Lasting Effects of Science Education." *ifo Working Papers* .

Arrow, Kenneth J. 1962. "The economic implications of learning by doing." *The review of economic studies* 29(3):155–173.

Ash, Elliott, Daniel L Chen and Suresh Naidu. 2022. "Ideas have consequences: The impact of law and economics on american justice.".

Ash, Elliott, Sharun Mukand and Dani Rodrik. 2021. "Economic Interests, Worldviews, and Identities: Theory and Evidence on Ideational Politics.".

Bakshy, Eytan, Solomon Messing and Lada A Adamic. 2015. "Exposure to ideologically diverse news and opinion on Facebook." *Science* 348(6239):1130–1132.

Beck, Hermann. 1992. "The social policies of Prussian officials: the bureaucracy in a new light." *The Journal of Modern History* 64(2):263–298.

Becker, Sascha O, Erik Hornung and Ludger Woessmann. 2011. "Education and catch-up in the industrial revolution." *American Economic Journal: Macroeconomics* 3(3):92–126.

Becker, Sascha O, Francesco Cinnirella, Erik Hornung and Ludger Woessmann. 2014. "iPEHD—the ifo Prussian economic history database." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 47(2):57–66.

Becker, Sascha O, Volker Lindenthal, Sharun Mukand and Fabian Waldinger. 2021. "Persecution and escape: professional networks and high-skilled emigration from Nazi Germany.".

Becker, Sascha O, Yuan Hsiao, Steven Pfaff and Jared Rubin. 2020. "Multiplex network ties and the spatial diffusion of radical innovations: Martin Luther's leadership in the early reformation." *American Sociological Review* 85(5):857–894.

Ben-David, Joseph. 1971. "The Socientist's Role in Society".(Ed.) Alex Inkeles.".

Benford, Robert D and David A Snow. 2000. "Framing processes and social movements: An overview and assessment." *Annual review of sociology* 26(1):611–639.

Berkowitz, Daniel, Katharina Pistor and Jean-Francois Richard. 2003. "Economic development, legality, and the transplant effect." *European economic review* 47(1):165–195.

Bleek, Wilhelm. 1972. "Von der Kameralausbildung zum Juristenprivileg: Studium, Prüfung und Ausbildung der höheren Beamten des allgemeinen Verwaltungsdienstes in Deutschland im 18. und 19. Jahrhundert." *(No Title)* .

Bloom, Nicholas, Charles I Jones, John Van Reenen and Michael Webb. 2020. "Are ideas getting harder to find?" *American Economic Review* 110(4):1104–1144.

Bloom, Nicholas, Tarek Alexander Hassan, Aakash Kalyani, Josh Lerner and Ahmed Tahoun. 2021. "The diffusion of disruptive technologies.".

Bogucka, Edyta P., Davide Cantoni and Matthias Weigand. 2019. "Princes and Townspeople: A Collection of Historical Statistics on German Territories and Cities. 1: City Locations and Border Maps." *Harvard Dataverse, V3, UNF:6:WJzeigKxzLFC07zCufxlYQ==* .
URL: *https://doi.org/10.7910/DVN/ZGSJED*

Borjas, George J and Kirk B Doran. 2012. "The collapse of the Soviet Union and the productivity of American mathematicians." *The Quarterly Journal of Economics* 127(3):1143–1203.

Borusyak, Kirill and Peter Hull. 2023. "Non-random exposure to exogenous shocks: Theory and applications." *Econometrica* .

Callaway, Brantly and Pedro HC Sant'Anna. 2021. "Difference-in-differences with multiple time periods." *Journal of econometrics* 225(2):200–230.

Cantoni, Davide. 2012. "Adopting a new religion: the case of Protestantism in 16th century Germany." *The Economic Journal* 122(560):502–531.

Caplan, Jane. 2014. Bureaucracy, politics and the National Socialist state. In *The Shaping of the Nazi State*. Routledge pp. 236–258.

Carlsson, Gunnar. 2009. "Topology and data." *Bulletin of the American Mathematical Society* 46(2):255–308.

Chaney, Eric. 2022. "Religion and the rise and fall of Islamic science." *Working Paper* .

Christianson, Nicolas H, Ann Sizemore Blevins and Danielle S Bassett. 2020. "Architecture and evolution of semantic networks in mathematics texts." *Proceedings of the Royal Society A* 476(2239):20190741.

Conradt, Volker and Stefan Lohrum. 2017. "Gemeinsamer Verbünde-Index, gemeinsamer Fernleihindex: ein Projekt der AG-Verbundsysteme.".

De Chaisemartin, Clément and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review* 110(9):2964–2996.

Dittmar, Jeremiah and Ralf Meisenzahl. 2022. "The research university, invention, and industry: Evidence from german history.".

Domay, Friedrich. 1987. "Bibliographie der nationalen Bibliographien= Bibliographie mondiale des bibiographies nationales= A world bibliography of national bibliographies.".

Donaldson, Dave and Richard Hornbeck. 2016. "Railroads and American economic growth: A "market access" approach." *The Quarterly Journal of Economics* 131(2):799–858.

Dorn, Walter L. 1931. "The Prussian bureaucracy in the eighteenth century." *Political Science Quarterly* 46(3):403–423.

Eggers, Andrew C and Arthur Spirling. 2014. "Electoral security as a determinant of legislator activity, 1832–1918: New data and methods for analyzing british political development." *Legislative Studies Quarterly* 39(4):593–620.

Esping-Andersen, Gosta. 1990. *The three worlds of welfare capitalism.* Princeton University Press.

Fischer, Gustav. 1903. *Grundzüge der Organisation des deutschen Buchhandels.* Vol. 41 G. Fischer.

Fleming, Lee. 2007. "Breakthroughs and the 'long tail'of innovation." *MIT Sloan Management Review* .

Fremdling, Rainer. 1975. *Eisenbahnen und deutsches Wirtschaftswachstum, 1840-1879: ein Beitrag zur Entwicklungstheorie und zur Theorie der Infrastruktur.* Vol. 2 Gesellschaft für westfälische Wirtschaftsgeschichte.

Freund, Ernst. 1899. "New German Civil Code." *Harv. L. Rev.* 13:627.

Gillis, John R. 1968. "Aristocracy and bureaucracy in nineteenth-century Prussia." *Past & Present* (41):105–129.

Giorcelli, Michela and Bo Li. 2021. "Technology Transfer and Early Industrial Development: Evidence from the Sino-Soviet Alliance.".

Giorcelli, Michela, Nicola Lacetera and Astrid Marinoni. 2022. "How does scientific progress affect cultural changes? A digital text analysis." *Journal of Economic Growth* pp. 1–38.

Glaeser, Edward L. 1999. "Learning in cities." *Journal of urban Economics* 46(2):254–277.

Glaeser, Edward L, Hedi D Kallal, Jose A Scheinkman and Andrei Shleifer. 1992. "Growth in cities." *Journal of political economy* 100(6):1126–1152.

Golding, Graham. 1971. "A critical comparison of the schemes of the french and german civil codes." *Irish Jurist (1966-)* 6(2):305–322.

Goldstein, Judith and Robert Owen Keohane. 1993. *Ideas and foreign policy: beliefs, institutions, and political change*. Cornell University Press.

Golub, Benjamin and Matthew O Jackson. 2010. "Naive learning in social networks and the wisdom of crowds." *American Economic Journal: Microeconomics* 2(1):112–49.

Graesse, Johann Georg Theodor. 1861. *Orbis latinus; oder, Verzeichniss der lateinischen benennungen der bekanntesten städte etc: meere, seen, berge und flüsse in allen theilen der erde, nebst einem deutsch-lateinischen register derselben. Ein supplement zu jedem lateinischen und geographischen wörterbuche*. Schönfeld.

Habinek, Jacob. 2010. "State-building and the Origins of Disciplinary Specialization in Nineteenth Century Germany.".

Haferkamp, Hans-Peter. 2008. "The science of private law and the state in nineteenth century Germany." *The American Journal of Comparative Law* 56(3):667–690.

Haferkamp, Hans-Peter. 2009. Bürgerliches Gesetzbuch. In *Handwörterbuch des Europäischen Privatrechts*, ed. Jürgen Basedow, Klaus J. Hopt und Reinhard Zimmermann (eds.). Mohr Siebeck Verlag.

Halberstam, Yosh and Brian Knight. 2016. "Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter." *Journal of public economics* 143:73–88.

Hall, Bronwyn H. 2004. "Innovation and diffusion.".

Hall, Peter A. 1993. "Policy paradigms, social learning, and the state: the case of economic policymaking in Britain." *Comparative politics* pp. 275–296.

Hall, Peter A. 2020. *The political power of economic ideas: Keynesianism across nations*. Princeton University Press.

Hamilton, William L, Jure Leskovec and Dan Jurafsky. 2016. Cultural shift or linguistic drift? comparing two computational measures of semantic change. In *Proceedings of the conference on empirical methods in natural language processing. Conference on empirical methods in natural language processing*. Vol. 2016 NIH Public Access p. 2116.

Hanlon, W Walker. 2022. "The rise of the engineer: Inventing the professional inventor during the Industrial Revolution.".

Heinze, G Wolfgang and Heinrich Kill. 1988. "The Development of the German Railroad System." *The Development of Large Technical Systems* .

Higgins, A Pearce. 1905. "The Making of the German Civil Code." *Journal of the Society of Comparative Legislation* 6(1):95–105.

Hintze, Otto, Gerhard Oestreich and Fritz Hartung. 1970. *Staat und Verfassung: gesammelte Abhandlungen zur allgemeinen Verfassungsgeschichte*. Goettingen.

Hornung, Erik. 2015. "Railroads and growth in Prussia." *Journal of the European Economic Association* 13(4):699–736.

Hunter, James Davison. 1992. *Culture wars: The struggle to control the family, art, education, law, and politics in America*. Avalon Publishing.

Iaria, Alessandro, Carlo Schwarz and Fabian Waldinger. 2018. "Frontier knowledge and scientific production: evidence from the collapse of international science." *The Quarterly Journal of Economics* 133(2):927–991.

Inglehart, Ronald. 2015. *The silent revolution: Changing values and political styles among Western publics*. Princeton University Press.

Iversen, Torben and David Soskice. 2020. *Democracy and prosperity: Reinventing capitalism through a turbulent century*. Princeton University Press.

Jones, Benjamin F. 2009. "The burden of knowledge and the "death of the renaissance man": Is innovation getting harder?" *The Review of Economic Studies* 76(1):283–317.

Jones, Charles I. 2021. "Recipes and Economic Growth: A Combinatorial March Down an Exponential Tail.".

Kapp, Friedrich, Johann Adolf Goldfriedrich, Albrecht Kirchhoff and Alfred Friedrich von der Leyen. 1886. *Geschichte des deutschen Buchhandels*. Vol. 1 Börsenverein der deutschen Buchhändler.

Kiser, Edgar and Joachim Schneider. 1994. "Bureaucracy and efficiency: an analysis of taxation in early modern Prussia." *American Sociological Review* pp. 187–204.

Kitschelt, Herbert P. 1986. "Political opportunity structures and political protest: Anti-nuclear movements in four democracies." *British journal of political science* 16(1):57–85.

Kocka, Jurgen. 1981. "Capitalism and bureaucracy in German industrialization before 1914." *The Economic History Review* 34(3):453–468.

Krugman, Paul. 1991. "Increasing returns and economic geography." *Journal of political economy* 99(3):483–499.

Kuhn, Thomas. 1962. "The structure of scientific revolutions.".

Kuran, Timur. 1998. "Ethnic norms and their transformation through reputational cascades." *The Journal of Legal Studies* 27(S2):623–659.

Levy, Ro'ee. 2021. "Social Media, News Consumption, and Polarization: Evidence from a Field Experiment." *American Economic Review* 111(3):831–70.

Levy, Roee and Martin Mattsson. 2022. "The effects of social movements: Evidence from# MeToo." *Available at SSRN 3496903* .

Lindenfeld, David F. 1997. *The practical imagination: The German sciences of state in the nineteenth century*. University of Chicago Press.

Lohmann, Susanne. 1994. "The dynamics of informational cascades: The Monday demonstrations in Leipzig, East Germany, 1989–91." *World politics* 47(1):42–101.

Maloney, William F and Felipe Valencia Caicedo. 2022. "Engineering Growth." *Journal of the European Economic Association* 20(4):1554–1594.

Marshall, Alfred. 1890. *Principles of economics*. London, MacMillan.

McAdam, Doug, Sidney Tarrow and Charles Tilly. 2003. "Dynamics of contention." *Social Movement Studies* 2(1):99–102.

McIntosh, Terence. 1997. *Urban decline in early modern Germany: Schwäbisch Hall and its region, 1650-1750*. Vol. 62 Univ of North Carolina Press.

Mokyr, Joel. 2002. *The gifts of Athena*. Princeton University Press.

Mokyr, Joel. 2016. *A culture of growth: The origins of the modern economy*. Princeton University Press.

Moser, Petra, Alessandra Voena and Fabian Waldinger. 2014. "German Jewish émigrés and US invention." *American Economic Review* 104(10):3222–55.

Moser, Petra, Parsa Sahar and Shmuel San. 2022. "Immigration, science, and invention: Evidence from the 1920s quota acts." *Unpublished manuscript* .

Oberschall, Anthony. 1997. *Empirische Sozialforschung in Deutschland, 1848-1914*. Alber Reihe Kommunikation.

Otter, Nina, Mason A Porter, Ulrike Tillmann, Peter Grindrod and Heather A Harrington. 2017. "A roadmap for the computation of persistent homology." *EPJ Data Science* 6:1–38.

Pfeiffer, Hermann. 1928. *Der deutsche Buchhandel, seine Organisation und Einrichtungen*. C. Dünnhaupt.

Pollard, Sidney. 1990. *Typology of industrialization processes in the nineteenth century*. Taylor & Francis.

Porta, Rafael La, Florencio Lopez-de Silanes and Andrei Shleifer. 2008. "The economic consequences of legal origins." *Journal of economic literature* 46(2):285–332.

Rádl, Emanuel. 1909. *Geschichte der biologischen Theorien: seit dem Ende des siebzehnten Jahrhunderts*. Vol. 2 W. Engelmann.

Reimers, Nils and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.
URL: *https://arxiv.org/abs/1908.10084*

Rodrik, Dani and Peter Hall. 2018. "Distinguishing ideas and interests.". Last Accessed on July 5, 2023.
URL: *https://rodrik.typepad.com/dani$_r$odriks$_w$eblog/*2018/01/*more − on − distinguishing − ideas − and − interests − an − exchange − with − peter − hall.html*

Romer, Paul M. 1986. "Increasing returns and long-run growth." *Journal of political economy* 94(5):1002–1037.

Romer, Paul M. 1992. "Two strategies for economic development: using ideas and producing ideas." *The World Bank Economic Review* 6(suppl_1):63–91.

Rüegg, Walter. 2004. *A history of the university in Europe: Volume 3, universities in the nineteenth and early twentieth centuries (1800–1945)*. Vol. 3 Cambridge University Press.

Serafinelli, Michel and Guido Tabellini. 2022. "Creativity over time and space: A historical analysis of European cities." *Journal of Economic Growth* 27(1):1–43.

Shipan, Charles R and Craig Volden. 2012. "Policy diffusion: Seven lessons for scholars and practitioners." *Public Administration Review* 72(6):788–796.

Simmons, Beth A, Frank Dobbin and Geoffrey Garrett. 2006. "Introduction: The international diffusion of liberalism." *International organization* 60(4):781–810.

Simmons, Beth A, Frank Dobbin and Geoffrey Garrett. 2008. "The global diffusion of markets and democracy.".

Simmons, Beth A and Zachary Elkins. 2004. "The globalization of liberalization: Policy diffusion in the international political economy." *American political science review* 98(1):171–189.

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.

Sun, Liyang and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics* 225(2):175–199.

Tabellini, Marco and Giacomo Magistretti. 2022. "Economic Integration and the Transmission of Democracy.".

Tarrow, Sidney. 1996. "Social movements in contentious politics: A review article." *American Political Science Review* 90(4):874–883.

Tilly, Richard. 1980. *Kapital, Staat und sozialer Protest in der deutschen Industrialisierung: gesammelte Aufsätze*. Vol. 41 Vandenhoeck & Ruprecht.

Tolloi, Philipp. 2010. *Das Bürger-und Inwohnerbuch der Stadt Brixen von 1500-1593: Edition und Kommentar*.

Toye, John. 2006. *Modern bureaucracy*. Number 2006/52 WIDER Research Paper.

Watson, Peter. 2010. *The German genius: Europe's third renaissance, the second scientific revolution and the twentieth century*. Simon and Schuster.

Weber, Max. 1968. "Economy and Society: An Outline of Interpretive Sociology.".

Webster, Elizabeth, Alfons Palangkaraya and Tom Spurling. 2021. "What drives firm innovation? A review of the economics literature." *A review of the economics literature (October 29, 2021)* .

Weitzman, Martin L. 1998. "Recombinant growth." *The Quarterly Journal of Economics* 113(2):331–360.

Weyland, Kurt. 2009. "The diffusion of revolution:'1848'in Europe and Latin America." *International Organization* 63(3):391–423.

Widmann, Hans. 1952. *Geschichte des Buchhandels*. Harrassowitz.

Wilkes, Walter. 2004. *Buchdruck-Schnellpressen und Endlos-Rotationsmaschinen des 19. Jahrhunderts: eine Dokumentation*. Technische Universität Darmstadt.

Zhuravskaya, Ekaterina, Maria Petrova and Ruben Enikolopov. 2020. "Political effects of the internet and social media." *Annual review of economics* 12:415–438.

A Theoretical Appendix

The following reports detailed assumptions and the proofs of the model outlined in the empirical section.

A.1 Atomistic Diffusion with Naïve Learning

Assume P idea producers (or scholars) live in C cities, with N_c producers per city. Connections exist between cities such that a positive edge between city c and k ($e_{ck} = 1$) indicates that they are directly connected, and $e_{ck} = 0$ otherwise. Connections are independently assigned, meaning $Pr(e_{ck} > 0) = Pr(e_{ck} > 0|e_{jl})$ for any $c, k, j, l \in \{1, ...C\}$.

Scholars operate in one of two disciplines, $D \in \{1, 2\}$. For simplicity, assume that $N_c = 2$ for all cities, and that each city c has one knowledge producer p in $D = 1$, and one producer s in $D = 2$. At $t = 0$, each scholar produces independently an idea i^D_{pc} within their discipline. Suppose that ideas are transmitted along all positive edges within the same discipline. Then, at $t = 1$, each idea will have reached all cities with a path to the city of origins. Those are all cities k for which there exists a sequence of cities j such that $e_{j,j+1} > 0$ for each $j \in \{1, \ldots, J-1\}$.

Proposition A1. *Ideas diffuse to a (weakly) higher number of cities when the number of positive edges increases.*

Proof. Let M be the adjacency matrix of the cities' connections. Define $M^C = \sum_{n=1}^{\infty} M^n$. Then there exists a path between city c and k iff $M^C_{[c,k]} > 0$. If $M' > M$, then $M^{C'} \geq M^C$.

Proposition A2. *With an additional positive edge, the probability that an idea is present in any two cities (weakly) increases.*

Proof. Define I_c and I_k to be the set of ideas present in city c and k respectively at $t = 1$. The probability of any idea from city k, i_{pk}, to reach any other city c is given by $\eta = Pr(i_{pk} \in I_c) = \frac{\sum_c \mathbb{1}(M^C_{[k,c]})}{C}$. As above, this probability weakly increases with an increase in positive edges. Similarly, the probability of the same idea i_{pk} to be present in any two cities is given by

$$Pr(i_{pk} \in I_c \cap I_j) = \left[\frac{\sum_c \mathbb{1}(M^C_{[k,c]})}{C}\right]^2$$

which weakly increases as well with an additional positive edge.

A.2 Idea Generation with Knowledge Spillovers

The approach just outlined ignores the underlying process of knowledge production. In what follows, I incorporate the presence of local and global knowledge spillovers.

As above, assume each city c starts with $N_c = 2$ scholars, one in each discipline. Now consider scholars' relative area of expertise in knowledge production within their discipline D, an interval k^D_p. Scholars face a time constraint T to produce knowledge, and hence a trade-off between the breadth ($k > 0$) and the depth ($d \geq 0$) of their expertise, given the time constraint: $k_p + d_p \leq T$. Normalize so that $T = 1$.

Suppose each scholar receives a question $q \in \Omega^D$, where Ω^D is the set of possible topics within discipline D. Normalize Ω^D so that it lies in the interval $[0, 1]$.[22] Assume a uniform distribution for questions in each discipline, $x \sim U_2(0, 1)$. Each scholar receives payoff $V(q)$ from their answer to the question:

$$V(q) = \begin{cases} 1 + d_c^{\alpha} \mathbb{1}[D_q = D_p] & \text{if } q \in A_p \\ 0 & \text{otherwise} \end{cases} \tag{A1}$$

where $\alpha < 1$, as there are decreasing marginal returns from investing in depth d. Realistically, a scholar p can answer with depth only if the question is within their same discipline, i.e., if $D_q = D_p$.

A_p is the set of knowledge accessible to scholar p. This set is composed of: (1) knowledge p produces themselves, (2) the knowledge the other producer s in the same city transmits to p, (3) knowledge reaching c from other cities:

$$A_{pc} = k_p + \pi(N_c) \cdot \sum_{N_c - 1} k_s + \tau(M) \cdot A_{-c} \tag{A2}$$

$\tau(M)A_{-c}$ are global spillovers. A_{-c} is knowledge produced in all other cities but c, while τ are transportation and coordination costs, a continuous, weakly increasing function of the adjacency matrix M. $\tau'_M > 0$: the share of existing knowledge that reaches c increases with an additional positive edge.

$\pi(N_c) \cdot \sum_{N_c - 1} k_s$ are local knowledge spillovers. $\pi(N_c)$ is the probability that s transmits knowledge to p, and it decreases in the number of scholars present in city c, $\pi = \frac{1}{N_c - 1}$.

Scholars can choose to move to another city at a cost $\kappa(\tau, N_c)$, or stay in the same city at no cost. κ is a continuous, strictly decreasing function of τ and N_c. When connections increase, it is less costly to move, hence $\kappa'_\tau < 0$. As $\tau \to 0$, $\kappa \to \infty$: with no connections, it is impossible to move.

Congestion costs increase as there are more scholars in the same city, hence $\kappa'_{N_c} > 0$.

The sequence of play is as follows:

1. Scholars observe τ.
2. They choose sequentially whether to move to another city.
3. They choose (k_p, d_p).
4. Questions are asked, and payoffs collected.

Lemma B1. *When $\tau = 0$, $k_p = 1$ and $d_p = 0$. With high transmission costs, all scholars are generalists.*

Proof. Suppose $\tau = 0$, and hence $k(\tau) \to \infty$: there are no global spillovers, and no scholar will move. All questions can only be answered with local knowledge. The expected

[22] The normalization of both T and Ω^D to 1 corresponds to assuming that in a given city, there are enough scholars to cover superficially every topic in each discipline.

payoffs for p are:

$$\begin{aligned}\mathbb{E}[V(q)] =& \int_{k_p} q^1[1+d_p^\alpha]dF(q^1) + \int_{k_s} q^2 dF(q^2) \\ =& k_p[1+(1-k_p)^\alpha] + k_s\end{aligned}$$

and similarly for s. It is optimal to set $k_p = 1$ and $d = 0$ (and similarly $k_s = 1$).[23] In other words, both scholars in each city will expand the breadth of their knowledge as much as possible, sacrificing depth.

Lemma B2. *If $\tau > 0$, $k < 1$, $d > 0$. As connections increase, scholars specialize.*

Proof. Suppose $\tau > 0$. For now, consider the case in which τ is not large enough to justify movement across cities; considered below. Define $\tau A_{-c} = \omega \Omega^D$, that is, define the knowledge produced globally as a fraction of each discipline's knowledge set; hence $\omega'_\tau \geq 0$.

It is optimal for both p and s to set $k^* = (1-\omega)$ and $d^* = \omega$. To see why, note that payoffs under k^* and d^* are:

$$\mathbb{E}[V(q)|k_p^*, d_p^*, k_s, \omega] = (1-\omega)[1+\omega^\alpha] + k_s + \sum_D \omega^D \tag{A3}$$

Suppose p increased k_p by a positive amount ϵ, where given the time constraint $k + d = 1$, $0 < \epsilon < \omega$. Then returns would be

$$\mathbb{E}[V(q)|k_p^* - \epsilon, d_p^* - \epsilon, k_s, \omega] = (1-\omega+\epsilon)[1+(\omega-\epsilon)^\alpha] + k_s + \sum_D \omega^D \tag{A4}$$

and A4 - A3 < 0. In fact, subtracting and rearranging: $\epsilon[(\omega-\epsilon)^\alpha] < (1-\omega)[\omega^\alpha + (\omega-\epsilon)^\alpha]$. Rearranging again, $\frac{\epsilon}{1-\omega} < \frac{\omega^\alpha}{(\omega-\epsilon)^\alpha} + 1$, which is always true because $\frac{\omega}{\omega-\epsilon} > 1$ and $\epsilon < 2(1-\omega)$. If instead p decreased k_p by a positive amount ϵ, the expected payoffs would be analogous as in A4, with $-\epsilon$ instead. They are again are smaller than the payoffs in A3.[24]

Proposition B1. *If $\tau \geq \tau^*$, p_c and s_j will move to j and c respectively. As connections increase, scholars are more likely to move, and cities specialize in one discipline or the other.*

Proof. Suppose p_c^1 (producer in discipline 1 in city c) needs to choose whether to move to j or stay in c.

From A3, if they stay, expected payoffs are $(1-\omega)[1+\omega^\alpha] + \omega + 1$. If they decide to move, they know s_j^2 will face the same trade-off and make the analogous decision to move to c. Considering that s_j^2 will move, accessible knowledge in discipline 2 in city j would be reduced to ω^2. And since s_j^2 will move, then it will still be that $N_j = 2$: which means that $\pi = 1$ and $\kappa(N_c) = 0$ (meaning there are no additional congestion costs).
Given that p_j^1 will face the same choice for (k_p, d_p) as p_c^1, expected payoffs from moving are:

[23] Where the FOC for d are: $-d^\alpha + \alpha(1-d)d^{\alpha-1} = 0$, with negative second derivative, hence $d = 0$, and $k = T - d$.

[24] In this case, subtracting and rearranging: $(1-\omega)[(\omega+\epsilon)^\alpha - \omega^\alpha] < \epsilon[1+(\omega+\epsilon)^\alpha]$ and $\frac{\epsilon}{1-\omega} > \frac{(\omega+\epsilon)^\alpha - \omega^\alpha}{1+(\omega+\epsilon)^\alpha}$ which is always true because: $\epsilon > [(\omega+\epsilon)^\alpha - \omega^\alpha]$ while $1-\omega < 1+(\omega+\epsilon)^\alpha$

$$\mathbb{E}[V(q)] = 2 \cdot k_p(1 + d_p^\alpha) + \sum_D \omega^D - \kappa(\tau) \tag{A5}$$

If p_c moves, they can't do worse than choosing d^* as above ($d^* = \omega$).
Assuming $2(1-\omega) \geq 1$,[25] i.e., their joint knowledge covers all Ω^1, p_c moves as long as payoffs from A5 are greater than payoffs from A3, which simplifies to:

$$\omega[2 + \omega^\alpha] + \delta \geq 1 + \kappa(\tau) \tag{A6}$$

where $\delta \geq 0$ is the additional payoff from adjusting d_p optimally to the local spillovers. Note that the LHS is a continuous, weakly increasing function of τ (from above: $\omega'_\tau \geq 0$). The RHS is continuous, strictly decreasing function of τ. It follows that there exists a unique value τ^* that solves:

$$\omega(\tau^*)[2 + \omega(\tau^*)^\alpha] + \eta(\tau^*) - 1 - \kappa(\tau^*) = 0 \tag{A7}$$

As long as $\tau \geq \tau^*$, p originally from city c will move to city j. s, producing knowledge in discipline 2 originally in city j, will move to city c. Cities hence specialize in one discipline or the other.
This would be true for any two pair of cities. One may wonder whether a third producer from city l has incentives to move to city j. Note that they would face a different trade-off: the increase in depth from knowledge spillovers is mitigated by $\pi < 1$, and since $N_c > 2$, there is an additional congestion cost $\kappa'(N_c)$. Their relative weight determines whether each city will end up with $N_c^* \geq 2$.

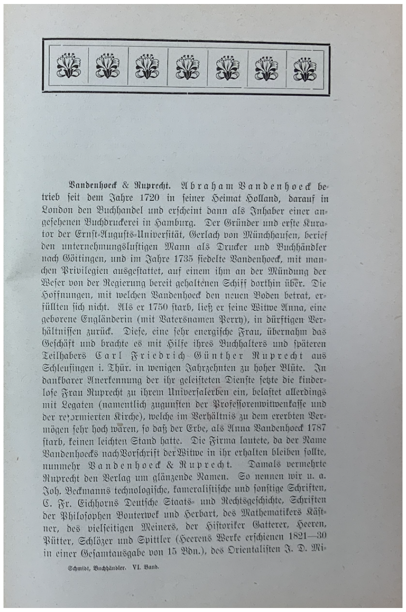
Proposition B2. *If positive links are such that $\tau(M) \geq \tau^*(M^*)$, the probability of an idea to reach any other city drops from η to $\frac{\eta}{N_c^*}$. Above a certain threshold of connections, ideas diffuse to a smaller number of cities.*

Proof. If $\tau < \tau^*$, the probability that an idea i_{cp}^D is adopted in any city j is η, defined under Proposition A1. If $\tau > \tau^*$, the number of cities actively producing knowledge in the same discipline (and hence adopting i^D) is reduced to $\frac{1}{N_c^*}$. The probability of a given idea to be adopted in any another city hence drops from η to $\frac{\eta}{N_c^*}$.

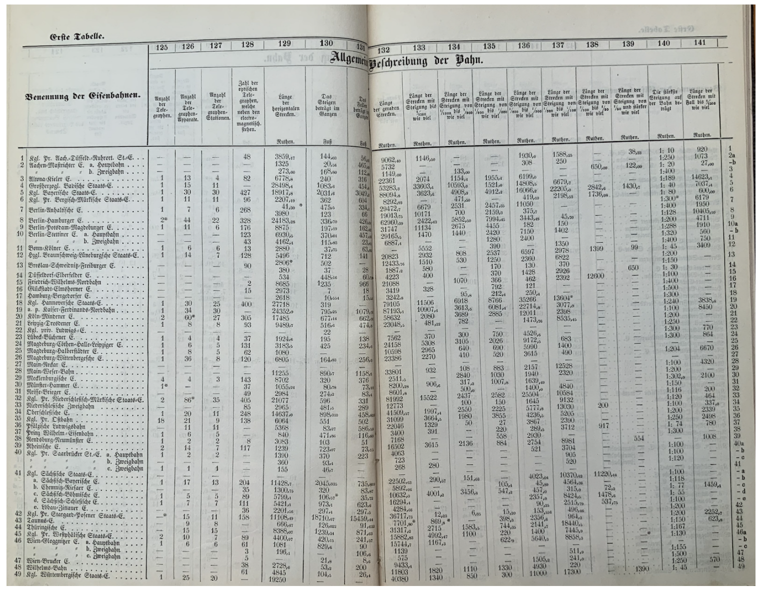
[25] If that's not the case, equation A6 is instead: $\eta + 2\omega + \omega^\alpha(1-\omega) \geq 1 + \kappa(\tau)$ and the point still holds.

B Data Appendix

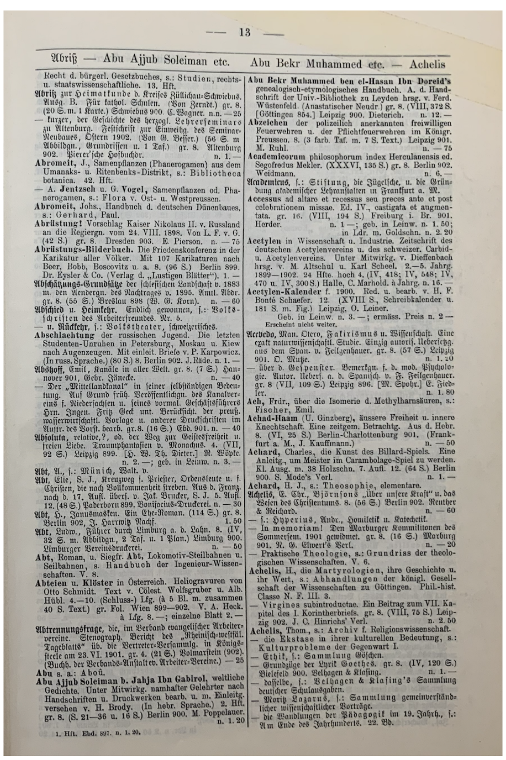
B.1 Sample Pages of Original Historical Data



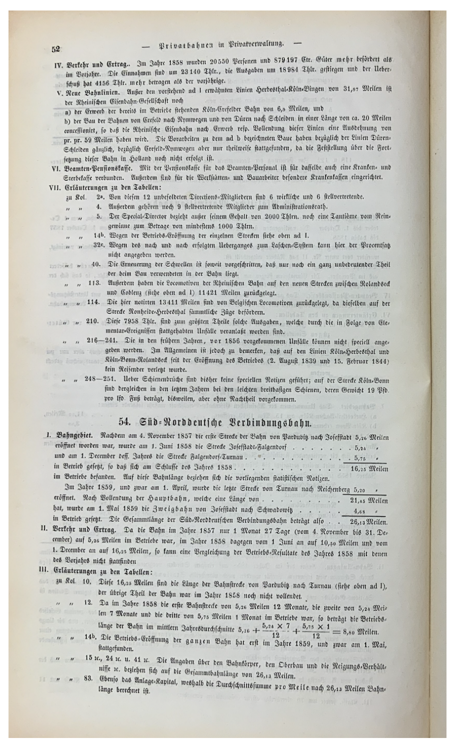
(a) Publishers' histories.



(b) Railroads statistics.



(c) Bibliographies.



(d) Railroads description.

B.2 Increase in Knowledge: Descriptive Trends

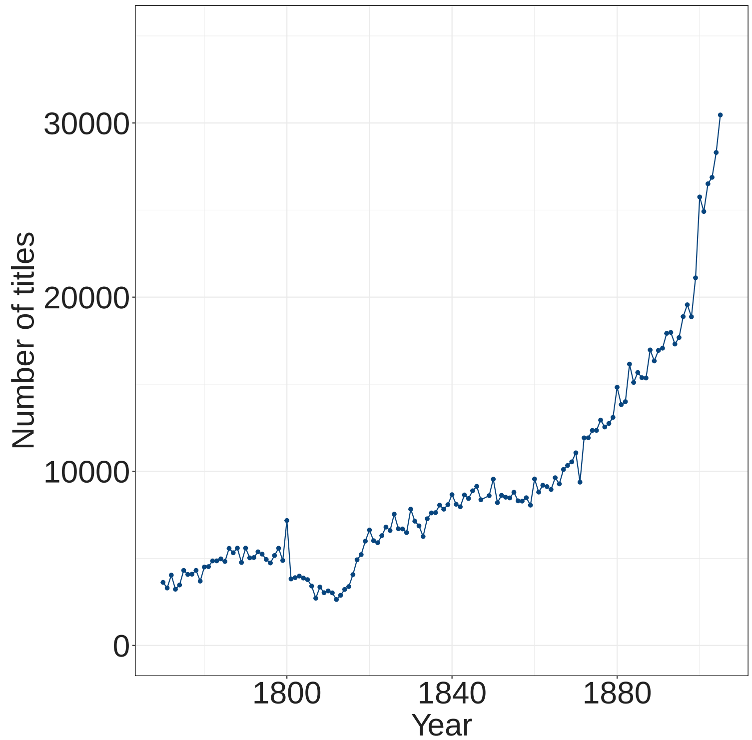


Figure A2: Yearly number of publications.

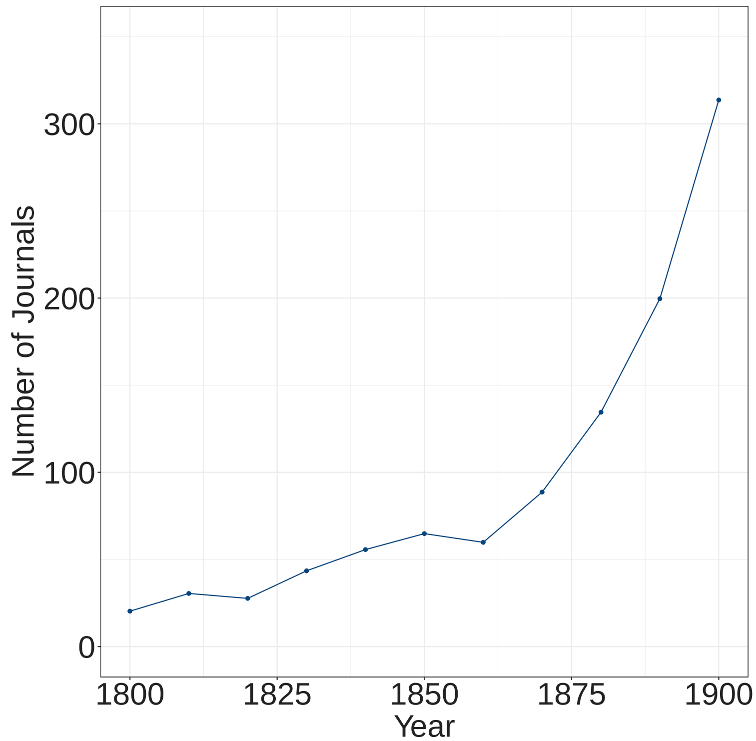


Figure A3: Yearly number of new scientific journals.

B.2.1 Geographical Distribution of Universities

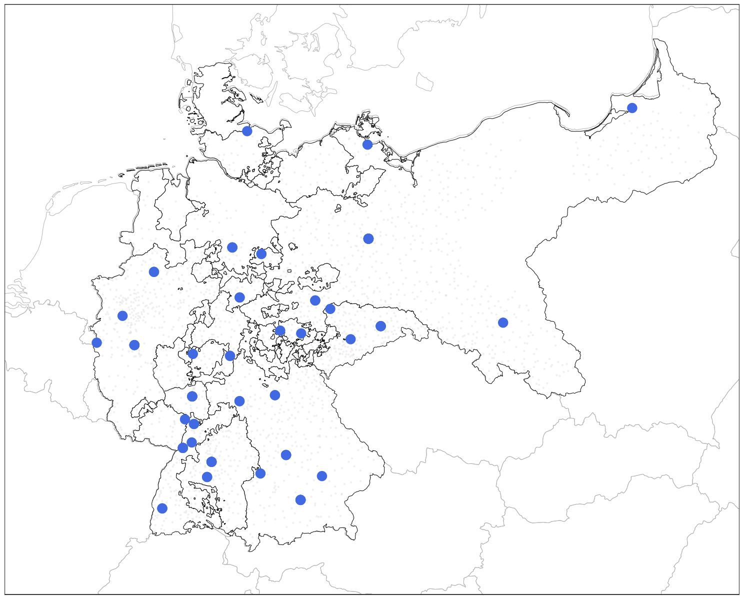


Figure A4: Location of universities in the first half of the 19th century.

B.3 Data Validation

In what follows, I ensure that testable trends that the literature in intellectual history identifies for this period are reflected in the bibliographical data. Specifically, I check whether: (1) current events (such as the 1848 revolutions, wars with France, and other external trends) are reflected in the number of publications. (2) Publications became more generally accessible, noticeable in a drift in the language of writing from Latin to German, especially encouraged after unification. (3) Educational reforms, especially with the spread of compulsory education, should be reflected in the number of school textbooks published. (4) Publications on nationalistic ideas spread over time to most cities. (5) As an example of a specific movement, feminist publications should increase together with the strength of the feminist movement.

B.3.1 Current events are reflected in publications

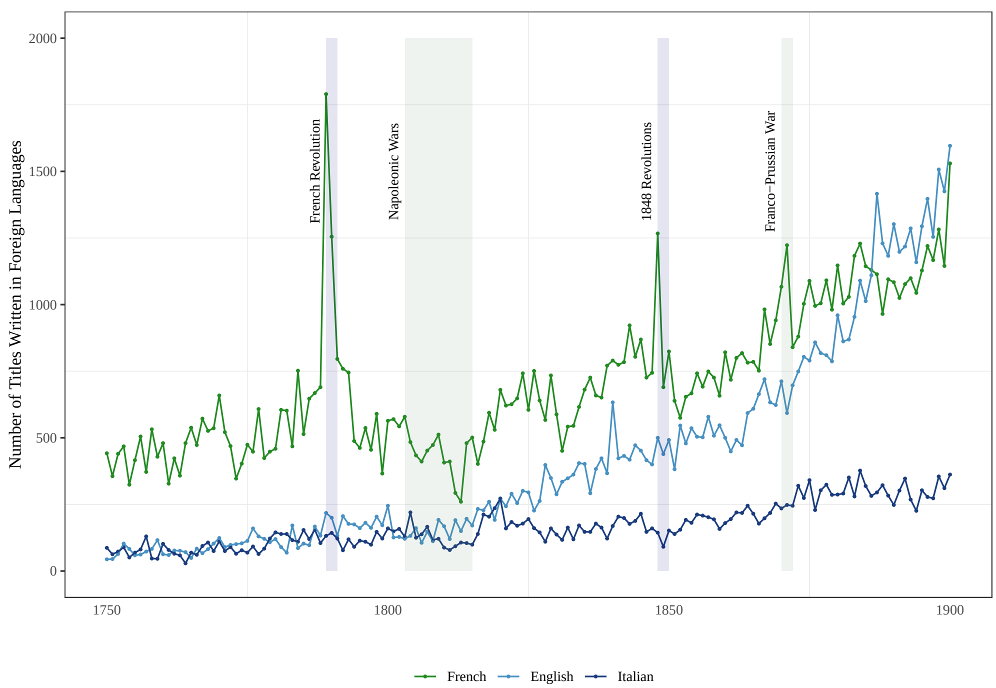


Figure A5: Yearly number of titles written in French, English and Italian. The increasing intellectual importance of the United States is reflected in the increase in number of publications in English (light blue) that penetrated the German market. Given the volatile relationship with France over the 19th century, we would expect the number of publications coming from France (green) to be volatile as well. During periods in which revolutionary ideas came from France (1789 and 1848 especially) we see a peak in the number of French publications. The salience of the (relatively brief) Franco-Prussian war of 1870 is also noticeable, while the prolonged Napoleonic wars led to a temporary drop in intellectual production.

B.3.2 The Democratization of Readership

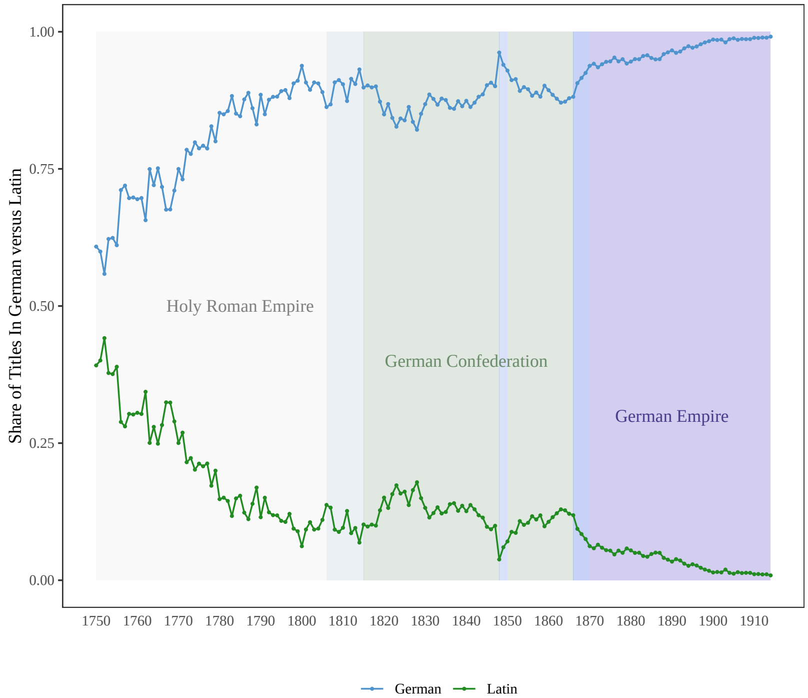


Figure A6: Share of publications within German written in Latin (blue) or German (green). The trends confirm the known phenomenon of the disappearance of Latin in favor of German, which contributed to the expansion of readership, as literacy rates increased over the century.

B.3.3 Educational Reform

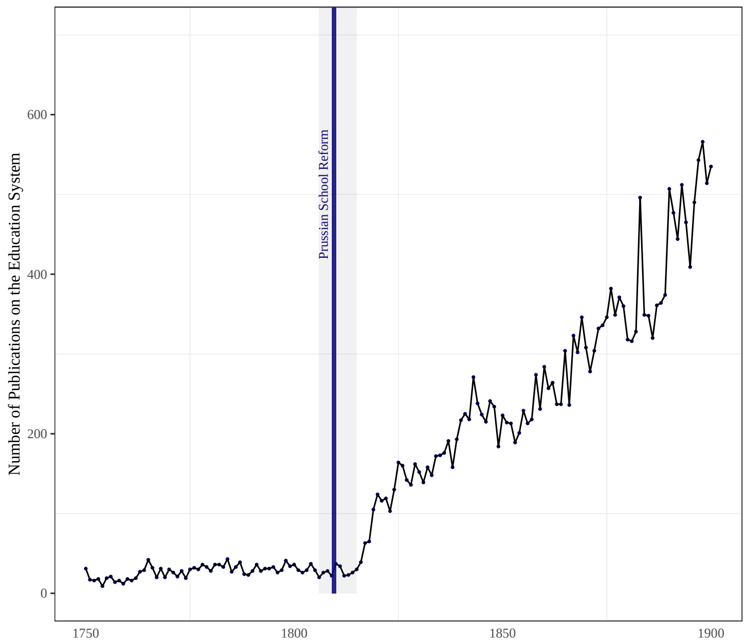


Figure A7: Number of yearly titles on education and the education system. A clear break is visible after Prussia enacted a series of school reforms.

B.3.4 The Spread of Nationalism

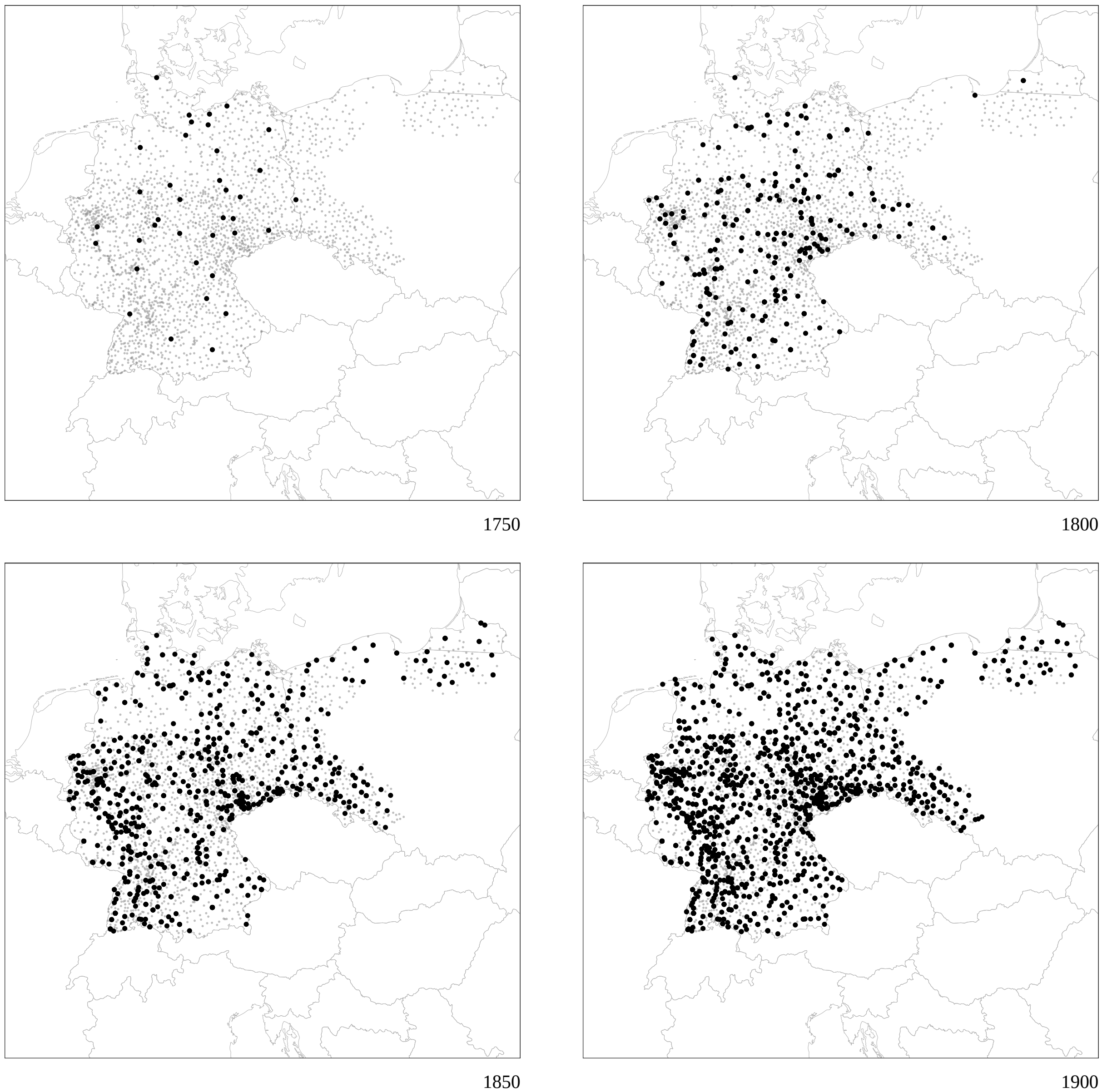


Figure A8: Presence of titles referring to a common German past or identity, which use ‘German history’ (different from state-specific history), ‘German people’ and ‘German nation’.

B.3.5 Building Feminism

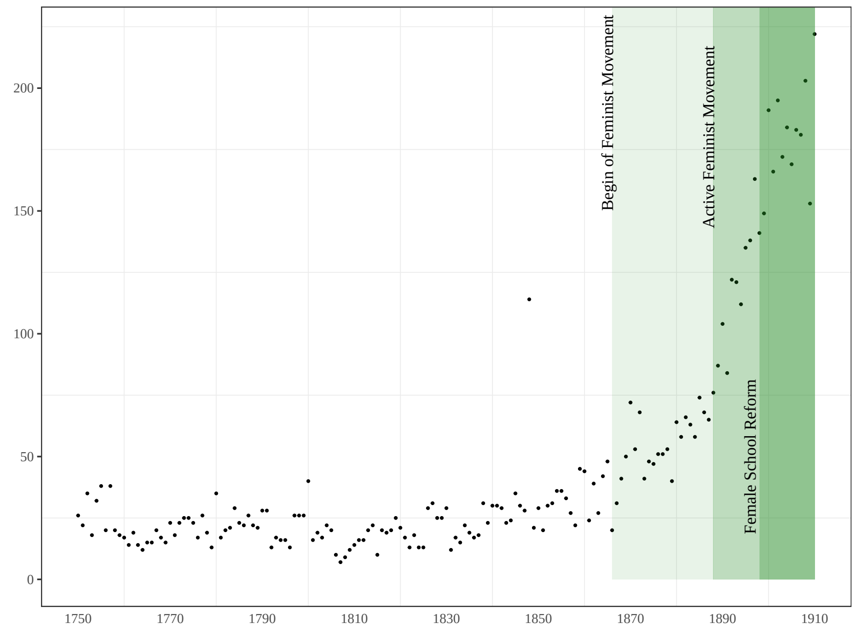


Figure A9: Yearly number of titles related to women (any topic), i.e., that contains the word 'Frau'

B.3.6 MARC Records

MARC Record	Description
041a	Language
035	System control number
100a	Author
100d	Life dates of the author
830a	Uniform title
245a	Title
245b	Title continued
245c	Title continued
246a-c	Varying form of title
250a	Edition
260a	Place of publication
264a	Place of publication
264b	Name of publisher
264c	Year of publication
300a	Extent (pages)
300c	Dimensions (cm)
490	Series statement
500	"Fraktur" (Gothic alphabet)
534a	Original version
534c	Original version - location
700a	Added entry name
700d	Added entry dates
850a	Holding institution
600+	Subject classifications
924	Holding library

Table A1: MARC Record Tags and Descriptions. The labels above have been extracted from the German Library Consortium bibliographic records to build the dataset for the analysis.

C Empirical Appendix

C.1 Movement of Ideas and Individuals

Table A2: Speed of Idea Diffusion & Frequency in Individuals' Movement

	Years for Ideas to Diffuse	Number of Destinations
50 years to railroad	2.30	-0.21
	(2.75)	(0.18)
40 years to railroad	0.47	-0.10
	(2.55)	(0.13)
30 years to railroad	-0.74	-0.07
	(2.58)	(0.09)
20 years to railroad	-2.46	-0.08
	(2.40)	(0.06)
Decade of the railroad	-10.94***	0.35**
	(2.31)	(0.13)
10 years after railroad	-9.09***	0.75***
	(2.69)	(0.19)
20 years after railroad	-7.42***	1.07***
	(2.93)	(0.24)
30 years after railroad	-10.93***	1.70***
	(3.36)	(0.34)
40 years after railroad	-12.53***	4.50**
	(3.44)	(1.82)
City Fixed Effects	✓	✓
Decade Fixed Effects	✓	✓
City-Specific Time Trends	✓	✓

Differential number of years it takes for a new idea, measured through new words, to be used in a different city compared to the city it first appears in, relative to the time when the city of origin is connected to the railroad network, in the left column. Differential number of movements to other cities in an individual's lifetime, relative to the time the railroad is introduced in their town of birth, in the right column. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level.

C.2 Co-location

Table A3: Scholars' Destinations.

	Number of Notable Individuals	Moving to Larger City	Moving to Cluster
-50 years to railroad	-0.017	-0.016	0.137
	(0.191)	(0.015)	(0.377)
-40 years to railroad	0.205	-0.008	-0.099
	(0.167)	(0.012)	(0.342)
-30 years to railroad	0.133	0.003	0.175
	(0.147)	(0.010)	(0.315)
-20 years to railroad	-0.009	0.010	-0.114
	(0.101)	(0.008)	(0.296)
Decade of railroad	0.200*	-0.000	1.210***
	(0.104)	(0.008)	(0.311)
10 years after railroad	0.403***	-0.014	1.655***
	(0.114)	(0.010)	(0.373)
20 years after railroad	0.261	-0.012	1.993***
	(0.155)	(0.014)	(0.395)
30 years after railroad	-0.248	-0.032	2.890***
	(0.404)	(0.019)	(0.608)
City Fixed Effects	✓	✓	✓
Decade Fixed Effects	✓	✓	✓
City-Specific Time Trends	✓	✓	✓

Differential number of scholars who become notable (first column) relative to the time of the railroad, ranking of destination city based on overall size (column 2), and ranking of destination city based on individual's profession. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level.

C.3 Democratizing the Intellectual Elites

Table A4: Diversity of Intellectuals.

	Traditional Background (Biographies)	Traditional Background (Wikipedia)	Distance to City Center
-50 years to railroad	-0.015	0.015	-25.403
	(0.014)	(0.009)	(88.632)
-40 years to railroad	0.017	-0.002	27.288
	(0.012)	(0.008)	(84.024)
-30 years to railroad	0.000	-0.002	158.569
	(0.012)	(0.008)	(98.211)
-20 years to railroad	-0.012	-0.004	-83.815
	(0.011)	(0.008)	(82.392)
Decade of railroad	-0.001	-0.000	-192.443*
	(0.011)	(0.008)	(79.838)
10 years after railroad	-0.010	-0.000	-193.024*
	(0.011)	(0.008)	(84.912)
20 years after railroad	-0.009	-0.007	-179.435*
	(0.012)	(0.009)	(82.003)
30 years after railroad	-0.034***	-0.023**	-202.160*
	(0.012)	(0.011)	(87.745)
40 years after railroad	-0.047***	-0.080***	
	(0.016)	(0.016)	
City Fixed Effects	✓	✓	✓
Decade Fixed Effects	✓	✓	✓
City-Specific Time Trends	✓	✓	✓

Differential effect on coming from a background similar to the traditional family background from the 18th century, relative to the time of the railroad. The first column uses the S-BERT sentence embeddings on the family description from the *Biographien*, the second column uses family-related sentences extracted from the Wikipedia pages of the authors from the bibliographic records. The third column measures the distance between one's location of birth and the centroid of the city's historical territory. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level.

C.4 Specialization

Table A5: Specialization

	HHI Across Disciplines	HHI Within Disciplines
-60 years to railroad	174.43	-243.62
	(368.27)	(188.29)
-50 years to railroad	63.30	-156.27
	(314.17)	(134.43)
-40 years to railroad	69.72	-107.99
	(249.84)	(108.00)
-30 years to railroad	-67.59	-122.90
	(186.99)	(83.40)
-20 years to railroad	66.71	-101.25
	(127.08)	(62.05)
Decade of railroad	21.23	66.79
	(128.29)	(48.10)
10 years after railroad	64.60	120.18**
	(188.14)	(59.07)
20 years after railroad	347.56	186.23***
	(258.34)	(62.72)
30 years after railroad	622.65*	191.37***
	(337.42)	(71.95)
40 years after railroad	1023.82**	230.12***
	(418.86)	(73.77)
50 years after railroad	1659.84***	141.22*
	(516.09)	(75.49)
City Fixed Effects	✓	✓
Decade Fixed Effects	✓	✓
City-Specific Time Trends	✓	✓

Differential effect on city-level specialization on certain disciplines (left) and topics (right) relative to the time of the introduction of the railroad. The definition of disciplines and topics follows the *Basisklassifikation* as a library subject classification system. Specialization is measured through an HHI index. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level.

C.5 Increase in Production

Table A6: Quantity of New Knowledge

	Number of titles	Titles (log)	Number of topics
-60 years to railroad	12.289	-0.199	0.980
	(14.342)	(0.122)	(0.834)
-50 years to railroad	2.738	-0.154	-0.082
	(13.890)	(0.098)	(0.762)
-40 years to railroad	-9.261	-0.132	-1.009
	(14.160)	(0.077)	(0.666)
-30 years to railroad	-5.829	-0.092	-0.447
	(10.237)	(0.056)	(0.513)
-20 years to railroad	2.443	-0.089*	-0.191
	(3.757)	(0.039)	(0.286)
Decade of railroad	9.782	0.132**	0.348
	(8.946)	(0.043)	(0.336)
10 years after railroad	14.085	0.205***	1.332*
	(14.756)	(0.062)	(0.597)
20 years after railroad	45.255	0.286***	3.940***
	(27.826)	(0.085)	(1.100)
30 years after railroad	121.594*	0.405***	9.575***
	(52.194)	(0.110)	(1.900)
40 years after railroad	244.379*	0.546***	17.359***
	(101.166)	(0.137)	(3.225)
50 years after railroad	674.296*	0.578**	33.754***
	(306.942)	(0.210)	(6.306)
Callaway & Sant'Anna	72.273***	0.902***	9.682***
	(17.293)	(0.311)	(1.423)
Sun & Abraham	64.534***	0.768***	8.236***
	(16.934)	(0.0552)	(0.930)
City Fixed Effects	✓	✓	✓
Decade Fixed Effects	✓	✓	✓
City-Specific Time Trends	✓	✓	✓

Differential effect on number of titles produced in a given city in a given decade, log of the tiles, and number of topics following the *Basisklassifikation* relative to the introduction of the railroad. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level. In the bottom of the table, the same analyses follow Callaway and Sant'Anna (2021) and Sun and Abraham (2021) respectively, aggregating over the decades after the introduction of the railroad.

C.6 Innovation

Table A7: Innovation

	% of New Words	Years to Fill Knowledge Gap
50 years before railroad	0.004	1.052
	(0.093)	(1.055)
40 years before railroad	0.158*	0.934
	(0.072)	(1.006)
30 years before railroad	0.032	0.730
	(0.064)	(0.934)
20 years before railroad	-0.045	0.652
	(0.052)	(0.836)
Decade of railroad	0.125*	0.062
	(0.059)	(0.688)
10 years after railroad	0.211**	-1.112**
	(0.066)	(0.564)
20 years after railroad	0.165*	-0.898**
	(0.073)	(0.429)
30 years after railroad	0.277**	-1.420***
	(0.088)	(0.449)
40 years after railroad	0.305**	
	(0.102)	
50 years after railroad	0.348**	
	(0.120)	
City Fixed Effects	✓	✓
Decade Fixed Effects	✓	✓
City-Specific Time Trends	✓	✓

Differential effect on the share or new words (left) multiplied by 100, and average time to make new connections between previously disconnected words (right), relative to the introduction of the railroad. The second regression stops at three decades after the railroad to keep the possible time to make connections consistent across decades. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level.

C.7 Diffusion

Table A8: Diffusion

	Diffusion to other cities	Similarity between connected cities	Diffusion within discipline	Diffusion within city across fields
50 years before railroad	0.0149	0.0002	-0.662	-0.146
	(0.0151)	(0.0096)	(0.454)	(0.151)
40 years before railroad	0.0123	-0.0176	0.355	0.133
	(0.0090)	(0.0094)	(0.377)	(0.137)
30 years before railroad	-0.0119	-0.0039	0.397	-0.0390
	(0.0080)	(0.0093)	(0.340)	(0.041)
20 years before railroad	0.0041	0.0020	0.366	0.0421
	(0.0070)	(0.0094)	(0.309)	(0.032)
Decade of railroad	-0.0241***	0.0052	0.271	0.0370
	(0.0081)	(0.0106)	(0.292)	(0.025)
10 years after railroad	-0.0367***	-0.0070	0.940***	0.00155
	(0.0096)	(0.0121)	(0.296)	(0.081)
20 years after railroad	-0.0507***	-0.0300**	0.693***	0.0815
	(0.0108)	(0.0135)	(0.253)	(0.151)
30 years after railroad	-0.0483***	-0.0442***	0.631*	-0.228***
	(0.0132)	(0.0143)	(0.362)	(0.076)
40 years after railroad	-0.0600***	-0.0369***	1.254**	-0.212***
	(0.0162)	(0.0144)	(0.447)	(0.035)
50 years after railroad	-0.0601**	-0.0618***	1.254***	-0.358***
	(0.0188)	(0.0158)	(0.518)	(0.069)
City Fixed Effects	✓	✓	✓	✓
Decade Fixed Effects	✓	✓	✓	✓
City-Specific Time Trends	✓	✓	✓	✓

Differential effect on the diffusion of new words to other cities, relative to the introduction of the railroad. The outcome of the first column is whether new words appear in other cities. The second is the Jaccard Index of similarity in publication fields between two newly connected cities. The third is whether the same word appears in other cities in the same discipline. The last column has as an outcome the adoption of new words in the same city, but in different fields. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level.

C.8 Persistence

Table A9: Persistence of Ideas

	All Ideas	Initially Successful
50 years before railroad	0.0211	-0.0331
	(0.0188)	(0.0353)
40 years before railroad	0.0158	-0.0176
	(0.0108)	(0.0266)
30 years before railroad	0.0068	-0.0049
	(0.0076)	(0.0196)
20 years before railroad	0.0104	-0.0167
	(0.0071)	(0.0134)
Decade of railroad	-0.0179**	0.0075
	(0.0067)	(0.0148)
10 years after railroad	-0.0153*	0.0278
	(0.0091)	(0.0225)
20 years after railroad	-0.0210*	0.0552*
	(0.0106)	(0.0289)
30 years after railroad	-0.0287*	0.0627*
	(0.0141)	(0.0378)
40 years after railroad	-0.0221	0.0830*
	(0.0190)	(0.0472)
50 years after railroad	-0.0130	0.1014*
	(0.0240)	(0.0566)
City Fixed Effects	✓	✓
Decade Fixed Effects	✓	✓
City-Specific Time Trends	✓	✓

Differential effect on the presence of the same word four decades after their first appearance. The first column considers all new words. The second column considers only the words that were adopted in other cities within one decade since their introduction, and controls for the number of cities they were initially adopted in. All estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level.

C.9 Robustness Checks

C.9.1 More on Accumulated Knowledge & Specialization

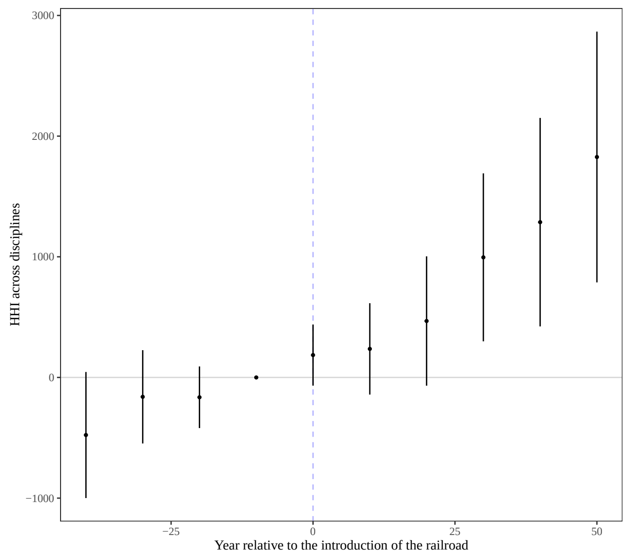


Figure A10: Specialization trends removing science and technology.

If accumulated knowledge was driving specialization, this may be an effect relevant only for science and technology. The same trend holds when removing titles published in science and technology according to the *Basisklassifikation*. The figure reports the results with the same specification as above, with city and decade fixed effects, and clustering standard errors at the city level.

C.9.2 Specialization of Publishers

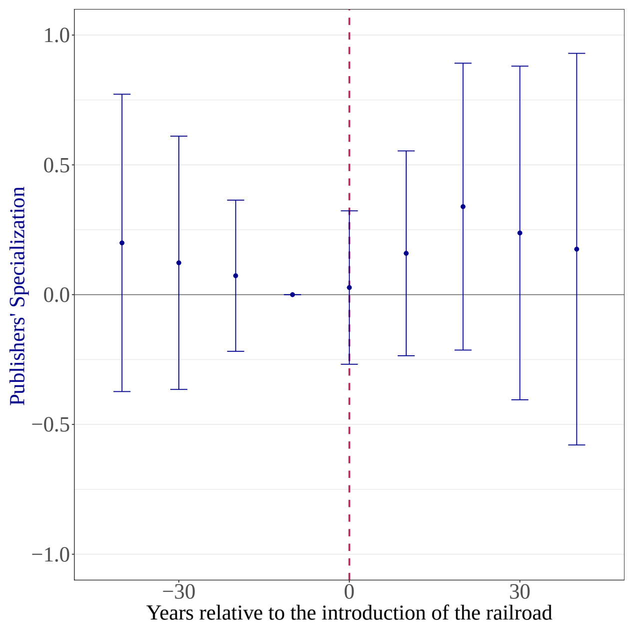


Figure A11: Publishers' specialization.

It could be the case that specialization is driven not by individuals specializing and clustering, but publishing houses themselves specializing. The outcome in the figure is whether a publishing house specializes in any topic as reported by the publishers' histories, specifically the differential effect relative to the introduction of the railroad in the publishers' city.

C.9.3 Location of Publishers

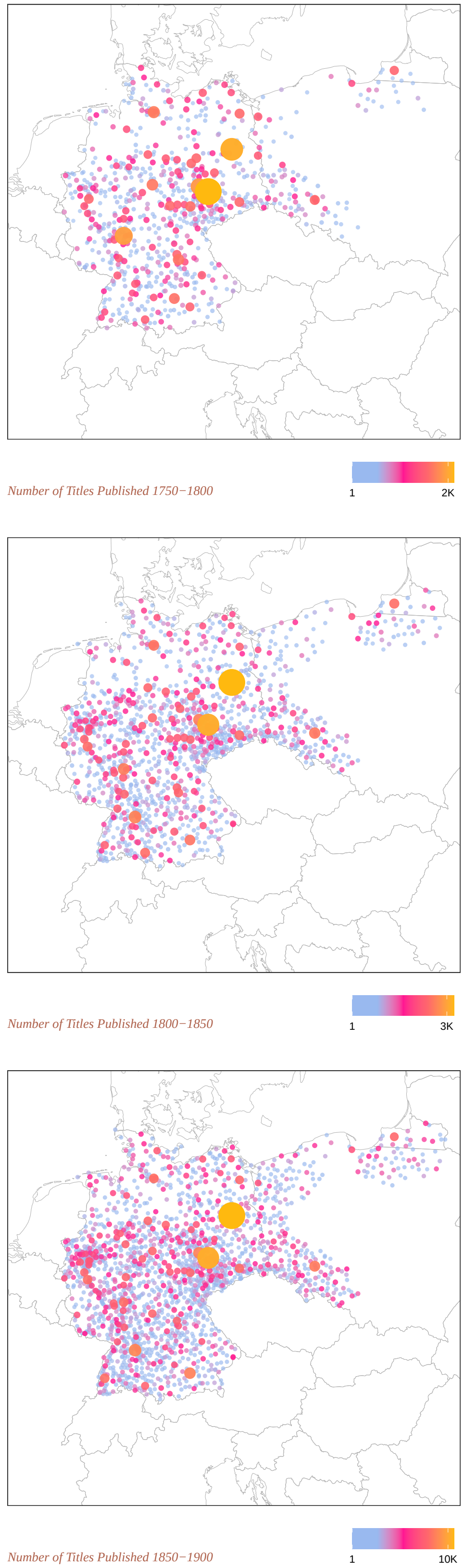


Figure A12: Number of Publishing Houses Present in Each City Over Time.

C.9.4 Advances in Printing Technology

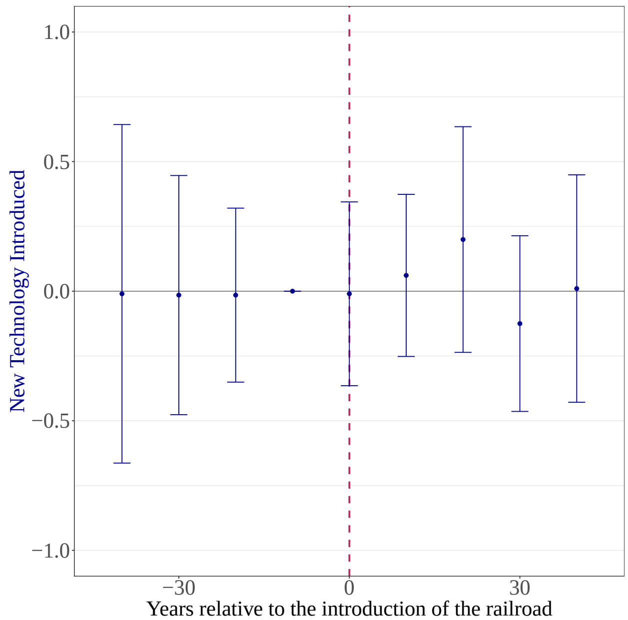


Figure A13: Introduction of New Printing Technology

The figure shows the differential effect relative to the introduction of the railroad on the adoption of new technology (fast printing) by publishing houses. The specification controls for city and decade fixed effects, and standard errors are clustered at the city level.

C.9.5 Non-Resident Authors

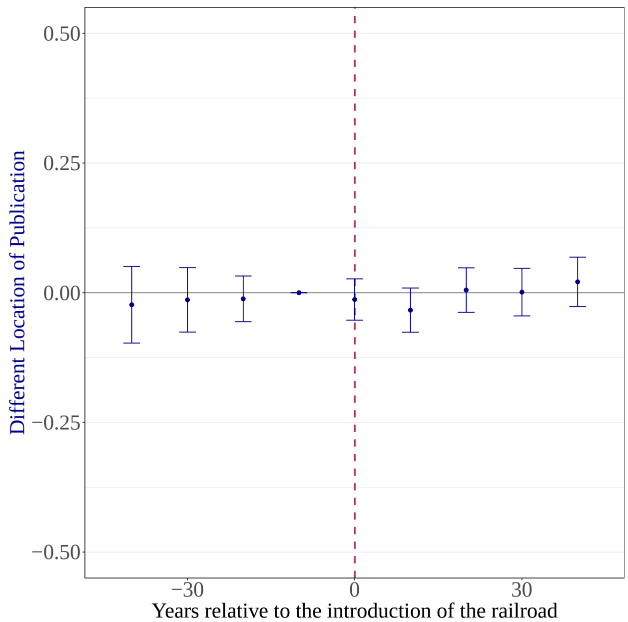


Figure A14: Publishing not in residence.

Noise may come from authors differentially publishing in cities other than their cities of residence before and after the railroad. The figure shows the differential effect on publishing away from one's residence relative to the introduction of the railroad, controlling for city and decade fixed effects, and clustering the standard errors at the city level.

C.9.6 Different Labeling of Topics

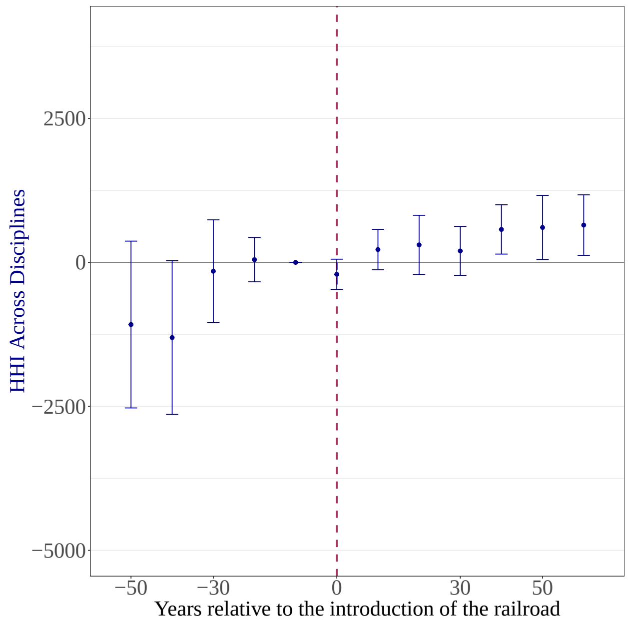


Figure A15: Specialization using different categorization system.

Differential effect on city-level specialization on certain topics (right) relative to the time of the introduction of the railroad. The definition of disciplines and topics follows the *Regensburger Verbundklassifikation* as a library subject classification system, instead of the *Basisklassifikation* as above. Specialization is measured through an HHI index. The estimates control for city and time period (decade) fixed effects, as well as for city-specific time trends. Standard errors are clustered at the city level.

C.9.7 Persistence of Fields

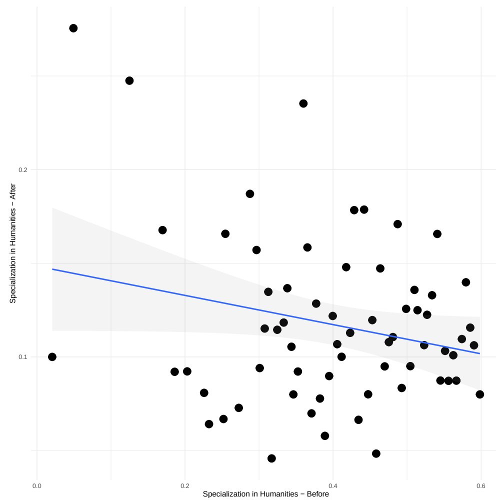


Figure A16: Cities' specialization in the humanities before and after the railroad.

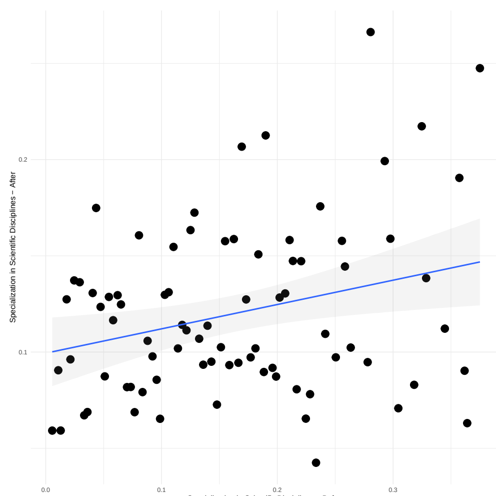


Figure A17: Cities' specialization in the science and technology before and after the railroad.