# Market structure and creative cluster formation: The origins of urban clusters in German literature, 1700-1932

Lukas Kuld,

University of Limerick

Sara Mitchell*,

University of Southern Denmark

4 April 2022

## Abstract

Using yearly data on 153 prominent German authors (1700-1932), we show how changes in the political and economic environment facilitated the formation of literary clusters. Early authors follow general population patterns, leading to geographic dispersion in a patronage system characterized by spatial competition. At the end of the 19th century, authors concentrate in large economic and political capitals. These changes in location patterns mirror trends in political and territorial consolidation and the professionalization of authorship. The last cohort shows large-scale migration into literary centers around the age of 20. Therefore, these literary clusters are not due to changing birth locations.

JEL-codes: N93, R12, Z11

Keywords: urban history, creative clusters, agglomeration, cultural economics, German literature

*Corresponding Author: Sara Mitchell, Department of Economics, University of Southern Denmark, 5230 Odense, Denmark, sabe@sdu.dk.

# 1 Introduction

Many papers observe two key patterns in the location of creative workers: the attraction of large cities and proximity to other creative workers. However, creative clusters are usually taken as a given, with few studies looking at the origins of creative clusters. Our empirical setting and detailed biographical data on German literary authors allow us to provide new insights into the formation of creative clusters, that is their location and dependence on market structure.

Within Western literature, German literature is a rare case of simultaneously having a long timeline and an initially high degree of geographic dispersion. This allows us to observe a long-term concentration process in the location decision of authors. By comparison, Paris and London long dominated as a location for British and French authors (see e.g. Mitchell [2019]), while the US has a relatively short literary history. Furthermore, the creative process studied here, literary writing, is a primarily solitary activity without infrastructure requirements like concert halls or studios. Therefore, authors can react to changing economic conditions more easily. In addition, during the time frame observed the labor market for German writers changed markedly. At the end of this period, writing could provide a sufficient income, while earlier writers often relied on some form of patronage.

To study location decisions in this empirical setting, we have lifetime biographic, publication, and location data on 153 of the prominent writers associated with German literature from 1700 to 1932 (in total 8146 observations with a known location in Germany).[1] We combine these with various datasets on the development and location of book publishing and trade, as well as on urban population, capitals, independent cities, and university cities. These detailed data and the long time horizon allows us to study changes in the attraction of large cities, proximity to other writers, and ultimately the formation of literary clusters. Importantly, the yearly data allows to distinguish moves at different career stages, which is indicative of the changing opportunities provided, for instance by a university town or capital city.

We postulate three stylized stages in the economic framework of literary writing. During the 18th century, most authors rely on patronage for their outcome, which often comes as reward for literary success. After the political consolidation during and after the Napoleonic wars, and during the early industrialization and urbanization in the first half of the 19th century,

employment opportunities were more often found in cities than small university towns and with the new middle classes rather than the rural nobility. However, writing in itself did not provide a sufficient outcome for almost all writers. In contrast, writers who were born after the unification of Germany in 1871 could find employment with newspapers and publishing houses and, if successful, live off the direct income from literary publications. This last phase was delayed by anti-liberal sentiment after the failed 1848 revolution [Estermann and Füssel, 2013]. As a result, we observe a gap of around 20 years with few new writers and a clearly defined generation of 'modern market authors' after this gap.

Our empirical findings are consistent with this stylized framework. We find that, over the 19th and early 20th century, authors become more likely to live in large political and economic capitals. A second concentration process within cities led to half of all authors living in Berlin alone in 1932 (or 2/3 of those inside Germany). The number of authors in Berlin, and to a lesser extent Munich, at the beginning of the 20th century is much higher than in earlier short-lived clusters that relied on some form of patronage, such as Weimar.

Our main empirical finding, however, relates to the location choice of authors over the life-cycle. Not only do we observe fewer moves into literary clusters for earlier cohorts, 18th century authors moved into clusters and prestigious locations like political capitals late in their career. Therefore, employment in such a location can be seen as a reward for a successful literary career. In contrast, the last cohort moves to Berlin and Munich and other large cities during their twenties, before or around the time of their first literary publication.

Therefore, the late clusters in Berlin and Munich resemble the image of a modern creative cluster more closely. That is a thriving cultural scene that attracts young aspiring artists to live, work, and learn together. In contrast, the geographic distribution of early authors more closely resembles that of modern-day university scientist who, by design, are spatially dispersed at different universities. This similarity is consistent with the spatial competition in the labor market of early authors relying on patrons (including universities) and the labor market of university based researchers.

The findings add to the historical evidence on migration, in particular high-skilled migration. We can directly compare our results to evidence of the locations and overall migration levels of other groups of creative workers during the time given by Kelly and O'Hagan [2007] and O'Hagan

and Hellmanzik [2008] for visual artists, Borowiecki [2013] for composers, O'Hagan and Walsh [2017] for philosophers, Mitchell [2019] and Kuld et al. [2021] literary writers (indirectly), and Borowiecki and Dahl [2021] for all creative activity. These studies focus on the clustering of creative workers and show remarkable concentration in key locations. However, by focusing on location choice, we can provide a more complete picture than these studies on aspects such as the link between age and migration and its development over time.

This paper complements large scale historical research on the locations of creative workers (with $n \gg 1000$, see for example, Schich et al., 2014, and Serafinelli and Tabellini, 2022) by providing a more complete picture of individual migration. We show, for instance, that the common empirical restriction to only observe birth and death locations can hide substantial geographic concentration due to life-cycle migration.

This research adds to the literature on how the political and economic environment influences urban development (e.g., Stasavage, 2014; Ades and Glaeser, 1995; Davis and Henderson, 2003; Cox and Figueroa, 2021; Wahl, 2019; Michaels and Rauch, 2018) and the geography of creative and innovative activity (Audretsch and Feldman, 1996; Feldman, 1994; Audretsch and Feldman, 2004; Carlino et al., 2007, among others). To date, the role of political and economic environment in shaping the geography of creative and innovative activity, particularly in historical contexts, has been relatively unexplored.[2,3] To our knowledge, this paper is the first historical study to focus on how these factors influence cluster formation.

The remainder of the paper is organized as follows. Section 2 provides an overview of the political and economic environment in Germany from 1700 to 1932. Section 3 discusses theoretical patterns in location choice and cluster formation. Section 4 describes the dataset and provides descriptive statistics. Section 5 presents empirical findings, and Section 6 concludes.

## 2 Empirical setting

Until the Napoleonic Wars in the early 19th century, the territory that would later constitute the German Empire had a very decentralized political system composed of around 300 effectively independent imperial cities, abbeys, small princely states, and other small territorial states [Fullerton, 2015]. Large imperial cities often restricted longer-distance, interregional migration

and migration from rural areas, while small and medium-sized towns tended to offer more job opportunities and easier migration processes. As a result, immigrants tended to originate from areas within 5 to 10 kilometers of their destination [Oltmer, 2015].[4]

Almost all of the imperial cities, abbeys, and small princely states were annexed into larger states between 1802 and 1815 as a result of the Napoleonic Wars (the number of German states was reduced from almost 300 to 41). This consolidation extended to universities as many of the smaller universities closed, including 13 Catholic universities. The closed universities were partly replaced by newly founded universities in capital cities such as in Berlin in 1810. Earlier, universities were mostly located in small, politically unimportant towns.

Napoleon's final defeat in 1815 was followed by another political and territorial restructuring at the Congress of Vienna. The earlier consolidation and secularization was confirmed in principle though in new shapes reflecting the changed power structure. This led to 37 principalities and 4 free states in 1815 [Fullerton, 2015]. This establishment of the German Confederation transformed migration patterns by allowing citizens to move freely between German states.

During this time, the book market grew relatively slowly, as illustrated in Figure A.1. The German book trade and authorship operated under a patronage system well into the 19th century, lagging other European countries, like the United Kingdom, in the development to a profitable book market [Tatlock, 2010].[5] Under this system, authors competed for a spatially limited number of positions under a patron, including positions in the public administration or at a university, or tutoring the children of nobility and merchants.[6]

By the early 19th century, it was possible to earn a modest income through writing under this system,[7] but few writers before the 1830s were able to make a career solely out of writing. Johann Wolfgang von Goethe (1749-1832) was one of the first authors who was able to make a living from writing, but he was the exception to the rule, as Tatlock [2010] describes:

> "Those who tried in the new age of the market to live by the pen struggled; many, like Theodor Storm, for example, wrote and published on the side...Goethe received from Cotta 65,000 thalers for this 'Ausgabe letzter Hand' (final authoritative edition). By contrast Gottfried Keller, they point out, received for his now canonical novel *Grüner Heinrich* (1853-55; Green Henry) a total of 742 thalers for the five-year span,

1850-55, in which he worked on the novel." (p. 8).[8]

Therefore, it was not uncommon for writers to have secondary employment. Even the literary 'superstar' Goethe worked as an adviser in the political administration of the Duke of Saxe-Weimar, a reward for his literary success. In this way, writing could lead indirectly to an income as tutor, teacher, or counsellor, but writing was of minor importance for direct income, such as royalties, until at least the 1830s.

This low profitability of authorship was due, in part, to lack of copyright protections until the mid-19th century. Because of the decentralized and fragmented political systems, it was difficult to enforce even informal copyright agreements across states. Pirated copies were common, particularly for commercially successful original editions [Moldovanu and Tietzel, 1998].[9] This lack of copyrights led to legal ambiguity regarding rights to the manuscript that led to conflicts between authors and publishers.[10] Publishers usually did not reveal to authors the true number of copies printing in a given edition or the number of editions printed,[11] which made it difficult for authors to negotiate payment based on prints, reprints, sales or to know the extent to which their works were pirated.[12]

However, growing demand from middle-class readership during the first half of the 19th century caused the book market to expand quickly, as illustrated in Figure A.1 [Fullerton, 2015].[13] This gave authors more bargaining power, which enabled them to negotiate better payments and protections. The first copyright was recognized when Prussia introduced copyright legislation in 1837 but offered limited protections to authors due to the culture of piracy and lack of protection in other German territories [Tatlock, 2010].

This rapid market expansion came to an abrupt halt with (ultimately unsuccessful) political revolutions in 1848. Economic instability, political suppression of ideas, and censorship affected all aspects of the publishing industry. Many authors left Germany, and many publishers went bankrupt. After years of rapid growth, the number of new book titles and daily newspaper titles saw a sudden decline, and the number of book traders stagnated. (See Figure A.1.) It took more than a decade for the publishing industry to fully recover [Tatlock, 2010].

The 1870s marked an important turning point. The year 1867 saw the end of perpetual publishing rights and the establishment of a general 30 year copyright [Berman, 1983]. A few years

later, in 1871, copyright was finally enforced throughout the German territories. The same year also saw the unification of Germany and the establishment of the German Empire. This establishment of a German nation-state not only shifted and centralized political power but also resulted in a harmonization of the legal systems, which reduced the costs of migration.[14]

Around the same time, rapid urbanization was accompanied by a growth in incomes and a shorter working week [Fullerton, 2015]. The result was an urban population that had more disposable income, more leisure time, and a taste for reading.[15] Mass book production was made more efficient due to several key advancements in printing technology that reduced printing costs and printing time,[16] the first international copyright laws were introduced just a few years later with the Berne Convention of 1886 [Tatlock, 2010]. The result was a boom in the publishing industry and book trade starting in the early 1870s (see Figure A.1).[17] The number book dealers and retail shops both more than doubled from the mid-1860s to 1890, and the number of book titles published annually increased by 62% over roughly the same period [Fullerton, 2015, p. 138]. This was accompanied by a rapid growth in authorship.

These economic and political changes cemented the shift from a patronage system to a market-based system. Authors were finally operating in a modern, competitive market with a growing demand for books. Author location choice was no longer determined by a high degree of spatial competition, and with the lowering of barrier to migration under German unification, authors had more freedom in terms of location choice. The German territory and population was reduced by around a tenth following World War I, but the German Empire remained a unified nation-state until the rise of the Nazi Party in 1933. Before 1933, the political events did not have a major impact on the book market or author mobility. We end our study in 1932, in the following years the hostile environment of Nazi Germany resulted in emigration of the vast majority of authors in our sample.

## 3 Theoretical patterns in author location and cluster formation

A number of studies have observed geographic clusters of artists across a variety artistic domains (Borowiecki and Dahl, 2021; O'Hagan and Hellmanzik, 2008; Hellmanzik, 2010; Borowiecki and O'Hagan, 2012; Borowiecki, 2013; Borowiecki, 2015; Mitchell, 2019; Kuld et al., 2021, among

others). A common finding is high degree of clustering of creative workers both across artistic periods and artistic domains in relatively few cities. Using insights from these studies and urban economic theory, we make several hypotheses about literary clusters and author location under idealized conditions (i.e., a competitive market with relatively low barriers to internal migration). We then contrast these idealized conditions to the political and economic reality of Germany from the early 18th century until the Third Reich and explain how these factors likely shaped the formation of literary clusters.

Under idealized conditions, authors will tend to co-locate in order to benefit from within-occupation agglomeration economies. In contrast to firms, the geographic concentration of a single artistic industry is likely to be more influenced by intellectual spillover than physical spillovers via the creation, transmission, and diffusion of knowledge between authors (Duranton and Puga, 2004; Ellison and Glaeser, 1997). They also likely benefit from matching with other authors to form collaborative circles that facilitated intellectual exchange.[18]

In addition, authors will tend to co-agglomerate with book publishers, sellers, and other related industry. Industry tends to concentrate geographically due to a variety of centripetal forces, such as backward and forward linkages, thick labor markets, information spillovers [Krugman, 1998]. Industrial agglomeration also reduced a variety of transportation costs: the cost of moving goods, the costs of moving people, and the costs of moving ideas [Ellison et al., 2010]. Co-agglomerated firms may benefit from the sharing of indivisible facilities, a labor pool, and gains from individual specialization; matching between employers and employees, buyers and suppliers, business partners, etc.; and learning via the creation, transmission, and diffusion of knowledge and skill acquisition [Duranton and Puga, 2004].

In the context of authorship, authors are likely to chose to live close to the industries related to the production and sale of books due to the high transportation costs associated with transporting manuscripts or personal travel. Proximity to both suppliers of inputs (such as paper, ink, and quills/pens) and consumers of their works (advertising firms, newspapers, magazines, the reading public) reduces transportation costs, facilitates the exchange of ideas, and reduces the costs of realizing ideas. A dense network of workers engaged in industries related to book production likely increases the probability of a good match between authors and industry gatekeepers, such as publishers or critics.

Furthermore, authors in periphery will tend to migrate to a literary cluster at early stage of career and remain there for the duration of their career. The standard economic way to see migration is a function of expected gains and costs of moving. Typically, labor migration peaks during early adulthood; higher age is associated with more social and economic attachments and less cumulative gains from migration [Becker, 1964]. Furthermore, authors show a high level of skill and specialization. These factors are typically linked with higher mobility as, arguably, specialized jobs are less widely distributed and high levels of education can reduce adaption costs. A dense network of publishers, newspapers, magazines, advertising firms, and other industries that require writing skills would provide a large pool of alternative or supplementary labor market opportunities. Therefore, authors are likely migrate to a cluster when they are young because they face lower relocation costs and will receive a higher pay-off, and they are likely to remain in the cluster for the duration of their career.

Finally, the location of literary clusters will tend to persist over time. New Economic Geography predicts that economic activity will be spatially concentrated and that the patterns of location will persist over time [Krugman, 1998]. These locations can continue to be economically important even after an initial competitive advantage is no longer relevant due to self-reinforcing agglomeration mechanisms [Fujita and Mori, 1996].[19] There is also evidence that this persistence is resistant to major shocks. For example, Schumann [2014] found that the geographic distribution of the population of Germany persisted even after as the resettlement of millions of German expellees after World War II.

While author mobility is not limited by physical infrastructure requirements, firms in related industries (such as book printers, publishing, sellers, editors) do have large physical infrastructure requirements and thus high cost of relocation. Therefore, their patterns of location are likely to persist over time. Given that authors will tend to co-agglomerate with these industries, the location of literary clusters will also tend to persist over time. Furthermore, Mitchell [2019] and Kuld et al. [2021] find these related industry tend to be located in large urban areas that are economic and political centers. As such, literary clusters are likely to develop and persist in large urban areas.

However, these idealized conditions that facilitate these outcomes were not met for most of the period due to the political and economic factors described in the previous section. These

factors likely impacted author mobility in several key ways. First, authors not only competed for literary patronage but also for a limited number of teaching or advisory positions, this system would reduce the attraction of living near other writers. Therefore, we expect authors to be distributed relatively evenly across territories; although, we would still expect to see a greater concentration of authors in the territories famous for their literary patronage or with wealthy patrons or a university that could support multiple authors.

Second, under the patronage system, authors' employment was linked to a patron rather than to a publishing house. Therefore, authors were not able to fully realize the agglomeration gains associated with co-agglomeration with industries related to book production and book sales. Thus, we do not expect to see a high degree of co-agglomeration until after the decline of the patronage system.

Third, a high degree of competition for a limited number of positions in a given location would increase the rate of migration throughout the career as authors would need to move to another location pursue new employment opportunities. However, the high political barriers to migration severely limited migration over long distances. Therefore, during the periods characterized by a patronage system, a high degree of political fragmentation, and high barriers to migration, we expect to see a a high rate of migration over short distances between smaller and medium-sized cities with few or no other authors.

Fourth, due to the spatial competition among authors and a dramatically shifting political environment, we do not expect to observe a stable literary cluster or group literary clusters. Instead, we expect the changes in the location of literary clusters to initially mirror the changes in the political and economic centers, and we expect the location of literary clusters to stabilize with the professionalization of authorship.

Since political and territorial consolidation decreased the barriers to migration, we would expect migration distance to increase with increased consolidation. However, we do not expect to see a major shift in author mobility or migration patterns until after the book market shifted away from a patronage system and towards a competitive book market in the mid-1800s. Specifically, we do not expect the periods of territorial consolidation in 1806 and 1815 to have an immediate impact on author mobility, location choice, or clustering intensity because market structure would not support the formation of literary clusters. Instead, we expect these patterns to shift

relatively slowly during the period of territorial consolidation and reorganization from 1818 to 1870.

Since the German book market was transitioning to a competitive market during the latter half of the 19th century, we expect to see more rapid changes in author mobility, location choice, and clustering intensity after German Unification in 1871. After this period, we expect that authors will move less frequently and over long distances. Because the market structure enabled authors to benefit from proximity to other authors, we expect to see an increased likelihood of locating in a large city and, as a result, a high degree of clustering intensity in a few, large cities.

# 4 Data

Historical data for Germany is often limited. This is partly due to the shifting German borders and subdivisions. For instance, there might be population data on Prussian counties in some years, but this would not be the case for Bavaria or some other smaller states.[20] To address this limitation, we will combine data from a variety of sources in order to outline political and industry developments. However, the coverage remains spotty in several areas.

## 4.1 Author data

We use a unique panel dataset on 153 prominent authors associated with German literature and born in the 18th and 19th centuries. We followed the Mitchell [2019] methodology for data collection. The dataset includes biographic and publication data, as well as the place of residence for every year of an author's life. We use a map of the 1910 German Empire to classify German locations. We identify the longitude and latitude of each location using OSM data by OpenStreetMap contributors [2021]. We calculate all geographic distances based on locality centers. The construction of the dataset and all definitions / classifications used are described in more detail in Supplementary Appendix B.3.1.

To track relative changes over the lifespan, we show statistics for three cohorts, 53 writers born between 1700 and 1785, 43 writers born between 1786 and 1830, and 57 writers born between 1831 and 1899. The years 1785 and 1830 were chosen as cut-off points to reflect the

political development and birth peaks. A writer born in 1786 would turn twenty after the end of Napoleon's influence in Germany and the political reorganization of hundreds of formerly independent territories into large territorial states. A writer born in 1830 would have spent the formative and typically most productive years of his career before the unification in 1871.

The last group, then consists mostly of authors born toward the end of the 19th century, as the failed 1848 revolution might have led to a lost generation of authors (only four authors were born during the 16 years from 1832 to 1848, compared to eg 23 during the last 16 years of the 19th century). However, these groups should not be seen as inherently distinct, homogeneous entities. The cut-off points for these birth cohorts remain random to some degree. Our empirical findings are robust to changing these cut-off point by several years.

We provide summary statistics of the three birth cohorts in Table A.1. Since we truncate the analysis at 1932, we provide summary statistics for the 1831-1899 birth cohort for their entire lives from birth until death, as well as for their lives up to 1932. On average, authors in the first two birth cohorts live to their early 60s, while the average author in the last birth cohort lives around 7 years longer. Authors across all birth cohorts produce, on average, 11-13 works over the span of their working lives (age 18-65). The average author in the 1700-1785 birth cohort relocated 7 times within their lives, while the 1786-1830 birth cohort relocated the most with around 9 relocations.

We summarize the top 10 author locations (in terms of author-year observations) for the periods 1700-1805, 1806-1870, and 1871-1932 in Table A.2. It is important to note that these cities represent the cities in which authors lived for many years but do not necessarily represent cities where authors tend to co-locate.[21] Berlin is the most popular location of residence in every period, but a high degree of clustering does not emerge until the later periods.

This can be compared with Figure A.4, which illustrates the share of authors and the share of the urban population in major cities per year. In all periods, authors are more highly concentrated than urban population. Berlin is consistently the most popular destination for authors from the mid-1800s. By the end of the sample period, Berlin emerges as the only major cluster, with more than 50% of the author sample living in Berlin at its peak. These patterns observed at the end of the sample period are consistent with the findings of Kuld et al. [2021] and Mitchell [2019] regarding the patterns of co-location and clustering intensity of authors in the US and

UK, respectively.

### 4.2 Location type data

To understand how the political and economic environment may impact author location choice, we define five broad categories of locations: capital cities, independent cities, large cities, university cities, and centers of book trade. We choose these categories because they represent political and economic centers, centers of intellect (and patronage), clusters of related industry, and population centers. For capital cities and independent cities, we only include cities in which at least one author in the panel dataset resided for at last one year. Therefore, these location type categories do not include an exhaustive list of cities in Germany. We provide a complete description of these data sources, how our location type data was constructed, and fully lists of the cities and the years in which they were in the respective category in the data appendix in Supplementary Appendix B.3.

We define an author as living in an independent/capital city if she lives in a city that is an independent/capital city in a given year. See Supplementary Appendix B.3.9 for more information about the construction of this list of capital and independent cities. The result is a list of 80 cities, most of which permanently lose their status as a capital or independent city by the end of the Napoleonic era in 1815. See Supplementary Table B.8 for a full list of the cities.

We define an author as living in a large city if she lives in one of the 10 most populous cities in a given year based on the Reba et al. [2018] and Bairoch et al. [1988] historic urban population datasets. See Supplementary Appendix B.3.7 for more information on the construction of the dataset of the most populous cities. The result is a list of 20 cities. The full list of cities is shown in Supplementary Table B.6.

We define an author as living in a university city if she lives in a city with an active university in a given year. However, we exclude cities that were included in the list of large cities and capital cities in order to separately estimate the effect of being a large, capital city from being in a smaller city/town with a university. The result is a list of 46 cities. See Supplementary Appendix B.3.8 for more information about the construction of this list of university cities. See Supplementary Table B.7 for the complete list.

We define an author as living in a center of book trade if she lives in a city listed in the Fullerton [2015] or Rarisch [1976] tables described in Supplementary Appendix B.3.5. As this data is only available for a limited number of years, we define a city as a center of book trade if city was ever included in these lists. We believe this is a reasonable restriction for most of the sample period, as physical infrastructure involved would make it difficult and costly to change location and high upfront investment costs would make it difficult to establish a new book shop in an area without existing economies of scale. The list of cities is provided in Supplementary Table B.5.

For some analyses, we also separately examine Berlin, Munich, and other large cities (excluding Berlin and Munich). Finally, we create the category "important city", which we define as any city included in any of location type category lists defined above, that is, any city that was a capital, independent city, large city, center of book trade, or university city in a given year.

# 5 Empirical Findings

In our empirical results, we begin by focusing on the location choice and concentration process of authors over time. We then show how the location choice over the life-cycle shifted between cohorts of authors. Finally, we provide evidence that the observed location patterns are not due to systematic differences in the author population with respect to birth location, migration intensity, or publication patterns.

We first estimate the probability that an author lives in a given location type or near other writers, and the distance to other writers for each year from 1700 to 1932. Second, we estimate these outcomes for three different birth cohorts over their life-cycle, i.e. at each from 0 to 80. These estimates are based on polynomial logit or Poisson regressions and provided as graphs. We provide alternative specifications by, first, relaxing the functional form assumptions in the polynomial regressions using splines in generalized additive models. Second, we provide all graphs as simple year or age means. All main findings are robust to changing the estimation method. However, the means do not account, for instance, for a change in the age structure or the overall number of authors. Therefore, we focus our discussion on outcomes estimated using regression analysis. For details see the Technical Appendix B.2. We complement the figures by

providing regression tables with dummies for different age groups and birth cohorts.

### 5.1 Location type and major cities over time

To illustrate the attraction of large urban areas and political and economic centers over time, we predict the probability of a 30 year old author being located in a given location type each year from 1700 to 1932. We account for age in our analysis to address changing age structures in our sample over time. However, we do not interact age with other factors. Therefore, a different age would not change the ordering of outcomes or subsequently interpretation of results (See Supplementary Appendix B.2 for a detailed description of the estimation method.). We present the yearly point estimates in graphs to illustrate the predicted outcomes over time in Figure A.5. We truncate the graphs due to a limited number of observations before 1750.

In the 18th century, the author population was geographically dispersed, as reflected in the relatively low probability of being located in most location types compared to later years. Authors were mostly located in towns or cities, though no single place dominated. For instance, no city has more than 9% of observations within Germany before 1805 as shown in Table A.2. However, our list of towns and cities accounts for 71% of observations within Germany in the 18th century, almost the same share as between 1800 and 1932 (74%). This is a considerably higher share than in the general population, where as late as 1871 almost two thirds of the population lived in rural locations [Estermann and Füssel, 2013]. With 35% and 25% respectively, capital and university cities are the most popular location categories, accounting together for 57% of observations before 1800. It is important to note that these two categories have very little overlap in the 18th century (2% of observations), before the establishment of universities in capitals in the early 19th century (e.g. Berlin 1810, Munich 1826, Stuttgart 1818/1829).

The geographic distribution of authors mirrors that of the German printers, publishers, and book sellers, as seen in Figure A.2. The publisher Friedrich Arnold Brockhaus (1772-1823) noted this unusually high geographic dispersion of the book market as he complained about the growing number of authors: "Germany's scribblers write too much and too much of what they write is printed. These excesses are fueled by the fact that there are so many places here which are publishers, in contrast to England and France, where they are confined to London, Edinburgh, and Paris" (as cited in Fullerton, 2015, p. 8).

The geographic distribution of authors is consistent with a model of spatial competition in a patronage system, under which only few positions are available in each location. The relative high share of university and capital cities equally fits this pattern, as courts, nobility and universities were major employers for literary authors. Weimar, the most famous early cluster in German literature, provides an illustrative example for employment and location patterns for successful authors in the 18th century. The exception is the large number of positions given to different authors.

Among the well known authors living in Weimar were Wieland, Goethe, Herder, and Schiller. The first was employed by Duchess Anna Amalia as a tutor for her two sons in 1772, the second given various administrative and ministerial roles by Anna Amalia's son and by then Duke Carl August from 1775, the third given an administrative role due to Goethe's lobbying in 1776 (The Editors of Encyclopaedia Britannica, 2022; Boyle, 2022; Irmscher, 2021). Schiller first moved to Weimar in 1787 before taking up a professorship at the university in nearby Jena [Witte, 2021]. However, Weimar never became a center of publishing or book trade. And though it kept its reputation for historical cultural significance (e.g. the rationale behind the Weimar Republic), later generations of writers would not settle in Weimar. Therefore, the cultural cluster in Weimar was not by itself sustainable in the way Paris or London remained cultural centers for centuries.

Authors are less likely to be located in smaller capitals and independent cities during and after the Napoleonic wars (1803-1815). As mentioned in Section 2, there were several episodes of political and territorial restructuring during the Napoleonic wars, at which time many cities permanently lost their university, or status as a capital or independent city. Therefore, this finding is due, in part, to cities losing their respective status and not necessarily a large-scale migration of authors. After the defeat of Napoleon and the formation of the German Confederation, authors become increasingly likely to be located in large urban areas, capitals, and centers of book trade.

We observe an increase in the frequency of large cities and centers of book trade (as well as a further increase for capitals) throughout the 19th century and the start of the 20th century. The 1871 formation of the German Empire further consolidated the German territories, which facilitated migration to the now fewer and larger political and economic centers. Importantly, local rulers, for instance the kings of Prussia and, later, Bavaria made an effort to attract writers

and artists. This could take different forms such as a direct stipend for Ludwig Tieck in Berlin or by giving professorships at the newly founded universities (e.g., Schelling and Heyse in Munich) [Fromm et al., 2019].

However, this period was also characterized by large-scale urbanization. Therefore, we investigate whether these patterns are associated with large cities or specifically associated with Berlin and, to a lesser degree, Munich. In fact, while we observe a small increase in other large cities at the beginning of the 19th century, the ten largest cities besides Berlin and Munich only account for 12% of observations between 1800 and 1932. The Ruhr, Germany's largest agglomeration in the 20th century and its center of coal mining and steel, had no author living there after 1886. But also other large cities, such as Cologne, Frankfurt or Hamburg were of limited importance as author locations, with overall 0.1% of all German observations in Cologne and 2% in the other two cities. Neither of these three cities were the seat of a secular ruler during the the time considered. In contrast, Berlin and Munich were the capitals of the two largest German states, Prussia and Bavaria, as we exclude Austria in our analysis.

While other large cities (excluding Munich) decline in popularity from the mid-1800s, the attractiveness of Berlin accelerates. In contrast, Munich was an unimportant location for authors until after the establishment of the German Confederation. Munich's popularity accelerated from the mid-1800s, suggesting that the politic of attracting authors was successful. The persistence of Munich as an attractive location for authors was due, in part, to serendipitous timing — the king's efforts coincided with the professionalization of authorship and transition away from the patronage system.

At the end of the 19th century, the literary market has been transformed. As described earlier, the technical and market innovations, author rights, and a larger demand made is possible for more writers to live directly from publishing. This removes constraints in the location choice for authors, though adds new incentives to settle close to publishing houses, critics, and other writers. The strong concentration of writers at the turn of the century in Berlin and Munich, and later only Berlin, fits this pattern.

Berlin was, besides Leipzig, the center of the publishing industry and, at the same time, an artistic center and the center of the newspaper industry. The observed concentration in Berlin therefore mirrors what has been observed much earlier in Paris and London [Mitchell, 2019],

and around the same time in New York [Kuld et al., 2021].

These findings suggests that the concentration process of authors over time is not simply a by-product of urbanization but also due, in large part, to an attractive political, economic, and cultural environment. In fact, most large cities or centers of book trade gained no importance as a location for literary writers.

If we compare the findings to the results from using generalized additive models (Figure B.1) and simple averages (Figure B.7) in Supplementary Appendix B.1, we see that the overall trends are robust to different model specifications. However, we see some interesting divergences. For instance, we see a large drop of authors in large cities following the failed 1848 revolution looking at the simple averages. While we know that this analysis neglects age imbalances (very few authors are born in the years before in the years leading to 1848), it is still interesting to note that large cities, including Berlin, lost a large share of their author population during a time in which new literature was suspect to mainstream culture ([Estermann and Füssel, 2013] note that 'culture' was rather shown by reciting the classics during this time period).

### 5.2 The formation of author clusters

We next analyze patterns in proximity to other authors and illustrate the formation of author clusters over time. To this end, we predict the probability of a 30-year old author living in an area with at least one other author within a 30 km radius, the probability of there being two or more other authors living within a 10 km radius, the distance to the nearest author, and the total number of authors within a 10 km radius. See Supplementary Appendix B.2 for a detailed description of the estimation method. We present the yearly point estimates in graphs to illustrate the predicted outcomes over time. As before, we truncate the graphs at 1750. The results are illustrated in Figure A.6.

Before the Napoleonic wars (1806-1815), authors are likely to be located in an area with a low density of authors in their immediate vicinity (within 10km) but with at least one other author within 30km. This suggests that authors during this period are relatively evenly distributed across space with little co-location of authors. We see small and short-lived clusters of authors emerge, for instance in Weimar, and Göttingen, facilitated by the university in Göttingen and

the Duke in Weimar. However, neither place offered employment opportunities to larger groups of authors.

The probability of an author being located in a cluster with at least two other authors nearby stayed low (less than 50%) until the second half of the 19th century. So although the territorial and political consolidation at the beginning of the 19th century generated some centripetal agglomeration forces, the centrifugal forces associated with the spatial competition of the patronage system continued to dominate. One might characterize this phase as a more concentrated patronage system, in which, in particular, the kings in Berlin and Munich were fostering a local art scene.

This began to change around the time of German Unification in 1871. We observe a sharp increase in the probability of being located in a cluster with two or more authors, and the distance to the nearest author begins to decline. However, it is not until the start of the 20th century that we observe an acceleration of these agglomeration forces. There was a sharp increase in the number of co-located authors, decrease in the distance to the nearest author, and and a dramatic probability of having at least one other author within 30km. We argue that this dramatic shift in proximity to other authors and the formation for author clusters was only made possible though the combination of sufficiently consolidated political and economic forces in relatively few cities and, crucially, the professionalization of authorship.

Again, the results from the alternative estimation techniques largely confirm our findings. The large peak in distance to nearest authors around 1860-1890 in the simple averages is understandable, if we consider the low number of authors born during the 1830s and 1840s. A low number of authors increases the distances all else equal.

### 5.3 Life-cycle patterns in location choice

Our life-cycle analysis has two main objectives. We show that the location patterns are due to migration and not purely reflective of changing general birth and population patterns. Second, we show that the latest cohorts is able to move to an attractive location or cluster much earlier, that is before becoming a successful author. In contrast, earlier authors move to capitals and (small) clusters much later in their career.

With the territorial and political consolidation, there was a convergence of location types by the end of the sample. That is, there was a reduction in the number of capitals and independent cities, and both the general population and book trade became increasingly concentrated in these areas. We have argued that this convergence of location types (specifically in Berlin) was critical for literary cluster formation. However, the patterns we have presented thus far could be an incidental phenomena associated with the rapidly increasing urbanization that occurred over the 19th century, and in the latter half of the century in particular. Therefore, we examine author location patterns over the life-cycle to show that the change in location patterns and proximity to other authors was due to a change in the choice of location over life-cycle and not only reflective of an increased probability of being born in a large, urban capital.

We again begin the life-cycle cycle analysis with plotted probabilities. We estimate the probability of an author being located in a location type or near another author for every year of their life from birth to age 80. See Supplementary Appendix B.2 for more information about the estimation method. We complement this evidence with a secondary regression analysis based on age dummies and shown in Tables A.3, A.4, and A.5.

### 5.3.1 Life-cycle patterns in location type and author concentration

The authors in the earliest birth cohort live most (if not all) of their lives in the highly fragmented Holy Roman Empire and spent their careers under the patronage system. The middle birth cohort lived predominantly in the more consolidated German Confederation, and their careers peaked during the early period of professionalization of authorship. They experienced the ends of the patronage system and the early beginnings of a competitive market. The final birth cohort, born after 1830, experienced the transition from the German Confederation to the unified German Empire. Authors were able to make a living directly from their writing at all stages of their careers, and their careers peaked during the era of mass book production.

Across all birth cohorts, authors were highly mobile over the life-cycle (7-9 re-locations on average, see Table A.1). While not all location types were popular destinations, the concentration of authors in specific location types is predominantly due to inward or outward migration of working-age adults. The attraction of, for example, large cities is not due to a high probability of being born there but instead due to authors moving there at some time during their adulthood.

The core difference — and a key finding — is how these location patterns over the life-cycle changed across birth cohorts.

We see in Figure A.7, that the earliest birth cohort (born 1700-1785) was most likely to locate in rural or small towns and cities during the first half of their career. This birth cohort only moved to capitals, large cities, or centers of book trade towards the end of their career. University cities tended to attract authors during early adulthood (likely for study) and late adulthood (likely for employment). This suggests that capital cities, centers of book trade, and large cities were "rewards" for being an established author and did not provide sufficient economic gains to authors in an early stage in their career.

This is mirrored in the estimates in Table A.3. While the first cohort was slightly more likely to be born in a capital (which were also more numerous), these author's odds of living in a capital between 18 and 40 are only a third ($e^{-1.10}$) of the last cohort's odds. We see similar estimates for large cities and the centers of book trade. In contrast, early authors are more likely to spend the years between 18 and 40 in independent cities (mostly small towns and of secondary political importance, most dissolved in 1803-1806) and in small university cities. For both location types, authors in the first cohort are much more likely to live there as adults. Therefore, these location patterns are not due to birth but reflect migration into these location types.

The second birth cohort (born 1786-1830) is broadly similar to the first cohort in the location pattern over the life-cycle. These authors were still often dependent on secondary work for their living; though, these might now be more likely to be in cities than small university towns or minor capitals. This suggests that, while political and territorial consolidation in the early 19th century facilitated migration, the conditions of the literary market ultimately determined author location choice over the life-cycle.

We observe the most substantial shifts in life-cycle location patterns for the last birth cohort (born 1831-1900). This cohort is much more likely to be in a large city, capital city, or center of book trade compared to previous cohorts. This cohort is also less likely to be located in those types of cities during their childhood and in the later stage of their career compared to the early stages of their career (see Table A.3). In addition, this cohort was less likely to be located in a university city across the life-cycle, signaling an end to the reliance on university positions as a form of patronage.

Authors in all three birth cohorts had similarly low probabilities of being born in Berlin or spending their childhood in Berlin, despite substantial urbanization over the 18th and 19th centuries. This suggests that the high concentration of authors in Berlin in the late 1800s and early 1900s is due to migration into Berlin and not general birth and population patterns. Furthermore, we do not observe similar labor market migration patterns for other large cities. These findings suggest that authors are drawn to large political and economic capitals with a rich cultural milieu rather than large, industrial cities. And, in turn, Berlin was able to accommodate and employ a large number of aspiring and early stage authors, unlike earlier clusters such as Weimar.

Looking at simple means and gam results, we see that the polynomial specifications smooths the steep increase in authors from the last birth cohort moving into large and capital cities, mostly Berlin, around the age of twenty (see Supplementary Appendix Figures B.3 and B.9).

### 5.3.2 Life-cycle patterns in proximity to other authors

We conclude the life-cycle analysis by looking into the impact of the migration on clusters and generally the geographic distance to other writers. To this end, we estimate the probability of there being at least one other author living within a 30 km radius, the probability of there being two or more other authors living within a 10 km radius, the distance to the nearest author, and the total number of authors within a 10 km radius. Following the same approach as in the previous section, we provide plotted estimates over the life-cycles of the three birth cohorts. The results are shown in Figure A.8 and Table A.5.

We find that authors in the early and middle birth cohorts are likely to be located within 30km authors but not within 10km of two or more other authors. This suggests that authors whose careers peaked before the 1871 unification are more evenly distributed across space with few authors co-locating. Importantly, authors in these cohorts only move closer to other authors or to an area with a higher density of authors at a relatively late stage of their career. We argue that this is evidence of spatial competition due patronage system discussed in Section 2. We also argue that there are no substantial differences between the trends of the first and middle birth cohorts because the centripetal agglomeration forces generated by post-Napoleon territorial and political consolidation were not sufficient to outweigh the centrifugal forces posed by the market

structure for authors.

In contrast, the last birth cohort is likely to be around fewer authors and be farther away from authors in their childhood and late adulthood. They move closer to other authors and to an area with a higher density of authors at an early stage of their career and only begin to move away to a lower-density area around age 40. We argue that this substantial shift in proximity to other authors over the life-cycle and the formation for author clusters was only made possible though the combination of sufficiently consolidated political and economic forces in relatively few cities and, crucially, the professionalization of authorship.

These last clusters, therefore, resemble the image of a modern creative cluster most closely. That is a vibrant cultural industry that attracts young aspiring artists to live, work, and learn together.

### 5.4 Migration and publication patterns

In attributing the observed patterns to the shifting political and economic environment, we must address concerns that lives of authors changed systematically over time with respect to their migration patterns or career development. That is, authors may systematically differ over time with respect to the frequency, distance, and probability of migration, the age at which they tend to migrate, or the age at which they begin their careers. The location patterns observed in the previous sections could be a reflection of these systematic differences in authors over time rather than a likely result of the changes in the political and economic environment.

We present predicted outcomes over time in Figure A.9 and predicted outcomes over the life-cycle by birth cohort in Figure A.10 and alternative regressions in Table A.6. We confirm our findings using gams in Figures B.5 and B.6 and simple means in Figures B.11 and B.12 in the Supplementary Appendix.

With respect to our concerns, the main finding is that the age of first publication is relatively similar between the three cohorts (see Figure A.10). In fact, the last cohort publishes slightly later than previous cohorts. Therefore, the higher age at moves to capitals and clusters of earlier cohorts cannot be explained by a later career start.

Despite dramatic improvements in printing technology and a growing demand for books over the 18th and 19th centuries, the probability of publishing is only slightly increasing over time. The strong increase in overall publications is not mainly a reflection of increased individual prolificacy, but is indicative of a market that is able to sustain a larger number of authors.

Relative migration patterns are relatively consistent across cohorts. Authors in all three birth cohorts had the highest probability of migrating at an early stage of their career, indicating that authors across birth cohorts did not exhibit different labor market behavior. However, authors in the last birth cohort were less likely to migrate, a result of arriving early and staying in Berlin, but more likely to move long distances and to locations that are far from their birth location.

# 6 Conclusion

In this paper, we evaluate how the development of literary clusters in Germany via location choice and clustering intensity was influenced by the transition from hundreds of independent states to a unified empire and from a patronage system to a competitive book market. We find that the geographic concentration of authors increased over time, as authors moved greater distances to be closer to other writers and into large capital cities and centers of book trade. This led, ultimately, to the formation of a literary cluster in Berlin. These trends are not due to systematic differences between author cohorts with respect to migration of publication patterns over the life-cycle. The core difference between cohorts is the choice of location: the type of city, its proximity to other authors, and the distance required to move there.

A growing number of economic studies have shown the importance of institutions, political freedoms, and local autonomy in shaping urban environment and clusters of talent. Our findings imply that these factors are not sufficient conditions for agglomeration of creative activity to occur. Market structure and economic incentives also play an important role in the shaping of creative clusters. We have seen evidence for policies successfully creating clusters in Munich and Berlin given the right market conditions. On the other hand, policies to widen the geographic distribution, e.g. in university based research, might mask advantages from clustering by prohibiting industry agglomeration.

Furthermore, we observe a high degree of mobility over the life-cycle for all cohorts. There is

a growing interest in using large-scale historical datasets of notable people to study the economic geography of creativity and innovation (e.g., Serafinelli and Tabellini, 2022). A common approach is to use automatically extract structured data from online databases, which tend to provide only birth and death location. If these patterns of life-cycle mobility that we observe hold for other high skilled workers, studies of the historical clustering of creative and innovative activity could be dramatically underestimating the degree of spatial concentration, particularly for the areas in which people do their most important work.

However, detailed data on the lives and locations of historical scientists, inventors, and other creative and innovative workers is often not available. In this respect, the economic history of the arts may offer important contributions in the future, as the academic interest and study of artists across all artistic domains has generated a breadth of information about their lives. Historical research using individual data on various types of artists could provide valuable insights into creativity over the life-cycle and into innovative activity that lies outside the patent system.

## References


Ades, A. F. and Glaeser, E. L. (1995). Trade and circuses: Explaining urban giants. *The Quarterly Journal of Economics*, 110(1):195–227.

Audretsch, D. B. and Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *American Economic Review*, 86(3):630–640.

Audretsch, D. B. and Feldman, M. P. (2004). Knowledge spillovers and the geography of innovation. In *Cities and Geography*, volume 4 of *Handbook of Regional and Urban Economics*, pages 2713–2739. Elsevier.

Bairoch, P., Batou, J., and Pierre, C. (1988). *Population des villes européennes de 800 à 1850: banque de données et analyse sommaire des résultats (la)*. Librairie Droz.

Becker, G. S. (1964). Human capital: A theoretical and empirical analysis, with special reference to education. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.

Berman, R. A. (1983). Writing for the book industry: The writer under organized capitalism. *New German Critique*, (29):39–56.

Borowiecki, K. J. (2012). Are composers different? Historical evidence on conflict-induced migration (1816–1997). *European Review of Economic History*, 16(3):270–291.

Borowiecki, K. J. (2013). Geographic clustering and productivity: An instrumental variable approach for classical composers. *Journal of Urban Economics*, 73(1):94–110.

Borowiecki, K. J. (2015). Agglomeration economies in classical music. *Papers in Regional Science*, 94(3):443–468.

Borowiecki, K. J. and Dahl, C. M. (2021). What makes an artist? the evolution and clustering of creative activity in the us since 1850. *Regional Science and Urban Economics*, 86:103614.

Borowiecki, K. J. and O'Hagan, J. W. (2012). Historical patterns cased on automatically extracted data: The case of classical composers. *Historical Social Research/Historische Sozialforschung*, pages 298–314.

Boyle, N. (2022). Johann Wolfgang von Goethe. *Encyclopedia Britannica*, 18 March 2022.

Carlino, G. A., Chatterjee, S., and Hunt, R. M. (2007). Urban density and the rate of invention. *Journal of Urban Economics*, 61(3):389–419.

Corporation, M. and Weston, S. (2022). *doParallel: Foreach Parallel Adaptor for the 'parallel' Package*. R package version 1.0.17.

Cox, G. W. and Figueroa, V. (2021). Political fragmentation, rural-to-urban migration and urban growth patterns in western eurasia, 800–1800. *European Review of Economic History*, 25(2):203–222.

Davis, J. C. and Henderson, J. V. (2003). Evidence on the political economy of the urbanization process. *Journal of Urban Economics*, 53(1):98–125.

Deutsche Biographie (2016). Deutsche Biographie. `https://www.deutsche-biographie.de`.

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Cities and Geography*, volume 4 of *Handbook of Regional and Urban Economics*, pages 2063–2117. Elsevier.

Ellison, G. and Glaeser, E. L. (1997). Geographic concentration in us manufacturing industries: a dartboard approach. *Journal of Political Economy*, 105(5):889–927.

Ellison, G., Glaeser, E. L., and Kerr, W. R. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review*, 100:1195–1213.

Encyclopaedia Britannica, Inc. (2016). Encyclopedia Britannica. `www.britannica.com`.

Estermann, M. and Füssel, S. (2013). *5.14 Belletristische Verlage*, pages 164–299. De Gruyter.

Farrell, M. P. (2003). *Collaborative circles: Friendship dynamics and creative work*. University of Chicago Press.

Feldman, M. (1994). *The Geography of Innovation*. Kluwer Academic Publishers, Boston.

Fromm, W., Knedlik, M., and Schellong, M. (2019). *Literaturgeschichte Münchens*. Verlag Friedrich Pustet, Regensburg.

Fujita, M. and Mori, T. (1996). The role of ports in the making of major cities: self-agglomeration and hub-effect. *Journal of development Economics*, 49(1):93–120.

Fullerton, R. A. (2015). *The Foundations of Marketing Practice: A history of book marketing in Germany*. Routledge.

Hanlon, W. W. and Heblich, S. (2021). History and urban economics. *Regional Science and Urban Economics*, page 103751.

Hellmanzik, C. (2010). Location matters: Estimating cluster premiums for prominent modern artists. *European Economic Review*, 54(2):199–218.

Hijmans, R. J. (2021). *geosphere: Spherical Trigonometry*. R package version 1.5-14.

Irmscher, H. D. (2021). Johann Gottfried von Herder. *Encyclopedia Britannica*, 14 December 2021.

Kelly, E. and O'Hagan, J. (2007). Geographic clustering of economic activity: The case of prominent Western visual artists. *Journal of Cultural Economics*, 31(2):109–128.

Krugman, P. (1998). What's new about the new economic geography? *Oxford Review of Economic Policy*, 14(2):7–17.

Kuld, L., Mitchell, S., and Hellmanzik, C. (2021). Manhattan Transfer: Productivity effects of agglomeration in American authorship. *Trinity Economic Papers tep0821, Trinity College Dublin, Department of Economics*.

Kuld, L. and O'Hagan, J. (2019). Location, migration and age: Literary clusters in Germany from mid-18th to early-20th century. *TRiSS Working Paper Series TRiSS-03-2019*.

Literature Online (2016). Literature Online (LION). `https://www.proquest.com/lion`.

Michaels, G. and Rauch, F. (2018). Resetting the urban network: 117–2012. *The Economic Journal*, 128(608):378–412.

Mitchell, S. (2019). London calling? Agglomeration economies in literature since 1700. *Journal of Urban Economics*, 112:16–32.

Moldovanu, B. and Tietzel, M. (1998). Goethe's second-price auction. *Journal of Political Economy*, 106(4):854–859.

O'Hagan, J. and Hellmanzik, C. (2008). Clustering and migration of important visual artists: Broad historical evidence. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 41(3):121–136.

Oltmer, J. (2015). *Handbuch Staat und Migration in Deutschland seit dem 17. Jahrhundert*. Walter de Gruyter GmbH & Co KG.

OpenStreetMap contributors (2021). OSM data retrieved from https://nominatim.openstreetmap.org/. `https://www.openstreetmap.org`.

O'Hagan, J. (2021). Age, mobility and creative output: prominent authors in 18th and 19th century Germany. *International Journal of Cultural Policy*, pages 1–12.

O'Hagan, J. and Walsh, A. (2017). Historical migration and geographic clustering of prominent western philosophers. *Homo Oeconomicus*, 34(1):11–32.

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1):439–446.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rahlf, T. (2015). German time series dataset, 1834-2012. GESIS Data Archive, Cologne. ZA8603 Data file Version 1.0.0, https://doi.org/10.4232/1.12202.

Rarisch, I. (1976). *Industrialisierung und Literatur - Buchproduktion, Verlagswesen und Buchhandel in Deutschland im 19. Jahrhundert in ihrem statistischen Zusammenhang*. Berlin: Colloquium Verlag.

Reba, M., Reitsma, F., and Seto, K. (2018). *Historical Urban Population: 3700 BC - AD 2000*. NASA Socioeconomic Data and Applications Center (SEDAC).

Reba, M., Reitsma, F., and Seto, K. C. (2016). Spatializing 6,000 years of global urbanization from 3700 bc to ad 2000. *Scientific data*, 3(1):1–16.

Schich, M., Song, C., Ahn, Y.-Y., Mirsky, A., Martino, M., Barabási, A.-L., and Helbing, D. (2014). A network framework of cultural history. *Science*, 345(6196):558–562.

Schumann, A. (2014). Persistence of population shocks: Evidence from the occupation of West Germany after World War II. *American Economic Journal: Applied Economics*, 6(3):189–205.

Serafinelli, M. and Tabellini, G. (2022). Creativity over time and space. *Journal of Economic Growth*, pages 1–43.

Singh, J. and Marx, M. (2013). Geographic constraints on knowledge spillovers: Political borders vs. spatial proximity. *Management Science*, 59(9):2056–2078.

St Clair, W. (2004). *The Reading Nation in the Romantic Period*. Cambridge University Press.

Stasavage, D. (2014). Was Weber right? The role of urban autonomy in Europe's rise. *American Political Science Review*, 108(2):337–354.

Tatlock, L. (2010). Introduction: The Book Trade and "Reading Nation" in the Long Nineteenth Century. In Tatlock, L., editor, *Publishing Culture and the "Reading Nation": German Book History in the Long Nineteenth Century*, pages 1–21. Camden House, Rochester.

The Editors of Encyclopaedia Britannica (2022). Christoph Martin Wieland. *Encyclopedia Britannica*, 16 January 2022.

Wahl, F. (2019). Political participation and economic development. evidence from the rise of participative political institutions in the late medieval german lands. *European Review of Economic History*, 23(2):193–213.

Wegge, S. A. (1998). Chain migration and information networks: Evidence from nineteenth-century hesse-cassel. *The Journal of Economic History*, 58(4):957–986.

Wegge, S. A. (2002). Occupational self-selection of european emigrants: Evidence from nineteenth-century hesse-cassel. *European Review of Economic History*, 6(3):365–394.

Wegge, S. A. (2021). Inheritance institutions and landholding inequality in nineteenth-century germany: Evidence from hesse-cassel villages and towns. *The Journal of Economic History*, 81(3):909–942.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wikipedia (2021). Liste der Hochschulen in Deutschland — Wikipedia, die freie Enzyklopädie. [Online].

Witte, W. (2021). Friedrich Schiller. *Encyclopedia Britannica*, 6 November 2021.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, 73(1):3–36.

Zeileis, A. (2006). Object-oriented computation of sandwich estimators. *Journal of Statistical Software*, 16(9):1–16.

# Notes

1. This work builds from the Kuld and O'Hagan [2019] working paper and O'Hagan [2021].

2. A limiting factor for historical studies has been a lack of data on both creative workers and cities. Hanlon and Heblich [2021] provides an excellent discussion of these challenges and an overview of recent innovations in this area.

3. There are several notable exceptions, such as Singh and Marx [2013], Serafinelli and Tabellini [2022], and Borowiecki [2012].

4. During this time, both migration and urban development due to rural to urban migration was relatively limited compared to that of countries such as France, Italy or England [Oltmer, 2015].

5. See, for example, St Clair [2004] and Mitchell [2019] for a discussion of the book market in the UK and Ireland.

6. For instance, Goethe worked in different roles at the ducal court in Weimar, from the age of 26 following the success of his first novel; Schiller became a history professor at the University of Jena in 1789.

7. Authors were typically paid a fixed fee based on the length of the text and received no additional royalties based on the number of sales [Moldovanu and Tietzel, 1998].

8. Similarly, Goethe received around 68 talers per sheet (16 printed pages) for his 1782 epic poem *Hermann and Dorothea*, while the average author received was around 20 talers per sheet [Moldovanu and Tietzel, 1998].

9. For example, Luise Mühlbach (pen name of Clara Mundt, 1814-1873) produced a number of commercially successful novels that had even international successful, but she was not able to profit from the pirated copies in the German territories or from unauthorized translations in the United States [Tatlock, 2010].

10. Authors usually argued that the fixed fee for a manuscript gave the publisher the right to print a single edition, after which the author could enter into new contractual agreements with other publishers regarding later editions. On the other hand, publishers tended to argue that the payment of the fee implied that they purchased the manuscript itself and thus they had the perpetual right to publish it [Moldovanu and Tietzel, 1998].

11. Even the literary superstar Goethe suffered from this problem: "...Vieweg's profit from the first edition of 6,000 books is 2,600 talers. *Hermann and Dorothea* was a best-seller, and till 1830 Vieweg printing (without telling Goethe!) at least another 20,000 copies" (Moldovanu and Tietzel, 1998, p. 857).

12. As Tatlock [2010] notes: "Early in the eighteenth century it was in fact generally the publisher – not the author – who sought privileges from the ruling authorities to prevent pirated copies" (p. 7-8).

13. The printing industry responded in kind after several major technological advancements in printing technology during this period, starting with the invention of the more durable all-metal press in 1800 [Tatlock, 2010].

14. While citizens of the Empire continued to be formally identified by their individual state (e.g., Prussian, Hamburg, or Saxon), their federal citizenship took precedence in migration law. However, it is important to note that the German Empire did not introduce the freedom of movement for the entire population and barriers to movement still existed [Oltmer, 2015].

15. Fullerton [2015] notes that "Already high in 1871, the rate of literacy increased after that. Statistics on recruits into the Imperial Army show that in 1875, 2.5 percent were illiterate, in 1880, 1.6 percent, in 1889-1890, 1.3 percent, and in 1889-1890 0.51 percent. Only Sweden and Denmark had lower incidences of illiteracy than Germany in the 1880s" (p. 169).

16. The first rotary press in Germany was constructed in 1873, zinc lithographic plates replaced stone plates shortly after, and type-casting machines replaced hand labor in the early 1880s [Fullerton, 2015].

17. Fullerton [2015] describes this time as "the great divide in the history of the mass book market...Before 1870 or 1871 the market's growth had been steady; after, it was explosive" (p. 168).

18. There is substantial evidence of this. Groups of authors such as *The Inklings* (J.R.R. Tolkien and C.S. Lewis, among others) and the *Bloomsbury Group* (Virginia Woolf, E.M. Forster, Lytton Strachey, and others) met regularly to share ideas, read one another works, provide feedback. See Farrell [2003] for a detailed discussion of these and and a number of other collaborative circles of authors, how these groups interacted, and how authors benefited from these groups.

19. This persistence in location is not only observed for existing economic activity but also to new developments. As Duranton and Puga [2004] observe, "only 1.9% of the land area of the United States was built-up or paved by 1992. Yet, despite the wide availability of open space, almost all recent development is less than one kilometre away from earlier development" (p. 2065).

20. For example, Simone Wegge's extensive work on the economic history of migration and demography in 19th century Germany is primarily limited to Hesse-Cassel (see, for example, Wegge, 2021; Wegge, 2002; Wegge, 1998.

21. For example, a city may have a higher number of author-year observations if two authors lived there for the

duration of their lives or if a dozen authors spent their university years there but attended at different times.

## Acknowledgments

To be added after the review process to maintain the anonymity of the authors.

## Data Availability

The data underlying this article are available in the technical appendix, with the exception of the author panel dataset. The final dataset and the code used to generate the results will be made available on a GitHub repository prior to publication.

# A Appendix

## A.1 Tables

Table A.1: Summary statistics of three birth cohorts

| | born 1700-1785 | born 1786-1830 | born 1831-1899 |
|---|---|---|---|
| No. Authors | 53 | 43 | 57 |
| No. Observations | 2960 | 2381 | 2805 (2418) |
| Mean lifespan | 61 | 62 | 69 (53) |
| Mean publications (18-65) | 11.3 | 11.9 | 13.5 (10.4) |
| Mean no. relocations | 7.4 | 9.2 | 8.6 (6.5) |

*Notes:* Table only includes observations within the geographic extent of 1910 Germany. The last column lists numbers over the whole life and up to 1932 in brackets. Our analysis ends in 1932 before the Nazi party comes into power.

Table A.2: Top 10 author locations

| 1700-1805 | | 1806-1870 | | 1871-1932 | |
|---|---|---|---|---|---|
| Berlin | 8.28% | Berlin | 14.69% | Berlin | 28.28% |
| Weimar | 8.21% | Weimar | 4.22% | Munich | 12.35% |
| Hanover | 5.53% | Tübingen | 3.72% | Paris | 2.81% |
| Leipzig | 4.23% | Stuttgart | 3.66% | Wroclaw | 2.33% |
| Göttingen | 3.60% | Munich | 3.63% | Frankfurt | 2.23% |
| Kaliningrad | 3.19% | Dresden | 3.47% | Wiesbaden | 2.16% |
| Osnabruck | 2.78% | Bonn | 2.71% | Dresden | 1.92% |
| Frankfurt | 2.75% | Wroclaw | 2.46% | Hamburg | 1.78% |
| Jena | 2.45% | Hamburg | 1.86% | Lübeck | 1.75% |
| Darmstadt | 2.27% | Hanover | 1.83% | Dithmarschen | 1.65% |
| No. author-year obs. | 2752 | | 3233 | | 3070 |

*Notes*: The table presents the percent of author-year observations per city per period. We include all ages from birth to death. We only include observations with known locations. Wroclaw was formerly known as Breslau. Kaliningrad was formerly known as Königsberg.

Table A.3: Probability of being located in various types of locations by age cohort

| | Independent City | Capital City | Large City | University City | Center of book trade | Important City |
|---|---|---|---|---|---|---|
| Under 18 | 4.14*** | −1.72*** | −1.45*** | 0.38 | −1.67*** | −1.63*** |
| | (1.01) | (0.11) | (0.10) | (0.19) | (0.10) | (0.11) |
| Over 40 | −11.69*** | −0.79*** | −0.97*** | 0.40 | −1.04*** | −0.96*** |
| | (1.00) | (0.11) | (0.11) | (0.21) | (0.11) | (0.11) |
| Born 1700-1785 | 4.33*** | −1.10*** | −2.40*** | 2.01*** | −1.97*** | −3.43*** |
| | (1.01) | (0.09) | (0.11) | (0.16) | (0.10) | (0.17) |
| Born 1786-1830 | 3.21** | −1.28*** | −1.49*** | 1.56*** | −1.23*** | −1.63*** |
| | (1.03) | (0.10) | (0.10) | (0.17) | (0.10) | (0.11) |
| Under 18*Born 1700-1785 | −3.44*** | 1.32*** | 1.30*** | −2.12*** | 0.88*** | −13.86*** |
| | (1.02) | (0.14) | (0.15) | (0.24) | (0.15) | (0.19) |
| Under 18*Born 1786-1830 | −3.74*** | 1.66*** | 0.42** | −1.42*** | 0.52*** | 0.77*** |
| | (1.06) | (0.15) | (0.16) | (0.25) | (0.15) | (0.18) |
| Over 40*Born 1700-1785 | 11.83*** | 1.28*** | 1.66*** | −0.77*** | 1.41*** | 2.14*** |
| | (1.02) | (0.14) | (0.15) | (0.23) | (0.14) | (0.21) |
| Over 40*Born 1786-1830 | 11.46*** | 1.47*** | 1.19*** | −1.56*** | 1.08*** | 1.66*** |
| | (1.06) | (0.15) | (0.15) | (0.26) | (0.15) | (0.16) |
| Constant | −6.87*** | 0.54*** | 1.04*** | −2.89*** | 1.11*** | 0.35*** |
| | (1.00) | (0.07) | (0.07) | (0.14) | (0.07) | (0.07) |
| | 2836.12 | 9943.32 | 9235.49 | 5632.94 | 9494.80 | 6593.62 |
| N | 7759 | 7759 | 7759 | 7759 | 7759 | 7759 |

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Notes: The table presents regression estimates with standard errors in parentheses. Standard errors are clusters at the author level. The observations are restricted to 1700-1932 and all observations with missing location are removed.

See Section 4 for definitions of location types. The reference categories are authors aged 18-40 and authors born 1831-1900. The results in all columns were estimated using logistic regression. Details to calculations are given in Appendix B.2.

Table A.4: Attraction of Berlin, Munich, and other large cities, by birth cohort

| | Probability of being located in: | | | Distance |
|---|---|---|---|---|
| | Berlin | Munich | Other Large | to Berlin |
| Under 18 | $-1.54^{***}$ | $-1.04^{***}$ | $0.57^{***}$ | $0.27^{***}$ |
| | (0.12) | (0.15) | (0.14) | (0.04) |
| Over 40 | $-0.55^{***}$ | $-0.79^{***}$ | 0.02 | 0.05 |
| | (0.11) | (0.16) | (0.16) | (0.05) |
| Born 1700-1785 | $-1.68^{***}$ | $-3.88^{***}$ | $-0.41^{**}$ | $0.17^{***}$ |
| | (0.12) | (0.46) | (0.16) | (0.04) |
| Born 1786-1830 | $-1.31^{***}$ | $-1.56^{***}$ | $0.51^{***}$ | $0.26^{***}$ |
| | (0.12) | (0.18) | (0.14) | (0.04) |
| Under 18*Born 1700-1785 | $0.93^{***}$ | $-12.26^{***}$ | $-0.17$ | $-0.15^{***}$ |
| | (0.20) | (0.47) | (0.21) | (0.05) |
| Under 18*Born 1786-1830 | $0.57^{**}$ | $-0.24$ | $-1.19^{***}$ | $-0.16^{**}$ |
| | (0.21) | (0.38) | (0.21) | (0.05) |
| Over 40*Born 1700-1785 | $1.03^{***}$ | $2.48^{***}$ | $0.54^{*}$ | $-0.10$ |
| | (0.17) | (0.51) | (0.22) | (0.05) |
| Over 40*Born 1786-1830 | $1.19^{***}$ | $1.30^{***}$ | $-0.85^{***}$ | $-0.13^{*}$ |
| | (0.16) | (0.26) | (0.23) | (0.06) |
| Constant | $-0.29^{***}$ | $-1.38^{***}$ | $-2.07^{***}$ | $5.42^{***}$ |
| | (0.07) | (0.08) | (0.10) | (0.03) |
| Deviance | 7042.49 | 3072.88 | 5585.26 | 1300894.50 |
| Num. obs. | 7759 | 7759 | 7759 | 7759 |

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Notes: The table presents regression estimates with standard errors in parentheses. Standard errors are clusters at the author level. The observations are restricted to 1700-1932 and all observations with missing location are removed. See Section 4 for definitions of location types. The reference categories are authors aged 18-40 and authors born 1831-1900. Columns 1-3 were estimated using logistic regression. Column 4 was estimated using a Poisson model. Details to calculations are given in Appendix B.2.

Table A.5: Attraction of other authors by birth cohort

| | 1+ authors within 30km | 2+ authors within 10km | Distance to nearest author | # authors within 10km |
|---|---|---|---|---|
| Under 18 | $-1.18^{***}$ | $-1.34^{***}$ | $0.57^{***}$ | $-1.00^{***}$ |
| | (0.10) | (0.10) | (0.08) | (0.07) |
| Over 40 | $-0.60^{***}$ | $-0.85^{***}$ | $0.69^{***}$ | $-0.29^{***}$ |
| | (0.11) | (0.10) | (0.10) | (0.06) |
| Born 1700-1785 | $-0.47^{***}$ | $-1.33^{***}$ | 0.17 | $-1.56^{***}$ |
| | (0.10) | (0.10) | (0.09) | (0.06) |
| Born 1786-1830 | $-0.67^{***}$ | $-1.32^{***}$ | $0.21^{*}$ | $-1.35^{***}$ |
| | (0.11) | (0.10) | (0.10) | (0.06) |
| Born 1700-1785*Under 18 | $0.97^{***}$ | $1.19^{***}$ | $-0.47^{***}$ | $0.81^{***}$ |
| | (0.14) | (0.15) | (0.11) | (0.10) |
| Born 1700-1785*Over 40 | $0.91^{***}$ | $1.15^{***}$ | $-1.02^{***}$ | $0.58^{***}$ |
| | (0.15) | (0.14) | (0.13) | (0.09) |
| Born 1786-1830*Under 18 | $0.84^{***}$ | 0.05 | $-0.23$ | 0.10 |
| | (0.15) | (0.17) | (0.12) | (0.12) |
| Born 1786-1830*Over 40 | $1.14^{***}$ | $1.52^{***}$ | $-1.12^{***}$ | $0.87^{***}$ |
| | (0.16) | (0.15) | (0.13) | (0.09) |
| log # authors per year | $0.78^{***}$ | $0.81^{***}$ | $-0.61^{***}$ | $0.72^{***}$ |
| | (0.06) | (0.07) | (0.03) | (0.05) |
| Constant | $-1.46^{***}$ | $-2.08^{***}$ | $12.17^{***}$ | $-0.42^{*}$ |
| | (0.22) | (0.25) | (0.13) | (0.19) |
| Deviance | 9774.98 | 9229.26 | 541310964.90 | 34546.23 |
| Num. obs. | 7744 | 7759 | 7744 | 7744 |

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Notes: The table presents regression estimates with standard errors in parentheses. Standard errors are clusters at the author level. The observations are restricted to 1700-1932 and all observations with missing location are removed. The reference categories are authors aged 18-40 and authors born 1831-1900. Columns 1 and 2 are estimated using logistic regression. Column 3 and 4 are estimated using a Poisson model. Details to calculations are given in Appendix B.2.

Table A.6: Migration patterns by birth cohort

| | Probability of Move | Distance of Move | Distance to Birth Location | Probability of Publishing |
|---|---|---|---|---|
| Born 1700-1785 | 0.15 | $-0.75^{***}$ | 0.03 | $-0.26^{**}$ |
| | (0.09) | (0.20) | (0.07) | (0.08) |
| Born 1786-1830 | $0.45^{***}$ | $-0.64^{***}$ | $-0.44^{***}$ | $-0.13$ |
| | (0.09) | (0.19) | (0.03) | (0.09) |
| Age | $-7.58$ | 28.05 | $57.59^{***}$ | $161.58^{***}$ |
| | (4.63) | (15.44) | (5.25) | (7.25) |
| $\text{Age}^2$ | $-74.78^{***}$ | $-101.18^{***}$ | $-39.80^{***}$ | $-140.00^{***}$ |
| | (4.55) | (12.49) | (4.35) | (6.41) |
| $\text{Age}^3$ | $54.25^{***}$ | $43.92^{**}$ | $14.18^{***}$ | $59.07^{***}$ |
| | (4.05) | (15.17) | (3.00) | (4.34) |
| Constant | $-2.61^{***}$ | $4.13^{***}$ | $5.27^{***}$ | $-2.80^{***}$ |
| | (0.08) | (0.18) | (0.04) | (0.09) |
| Deviance | 4972.78 | 3226986.18 | 3378709.76 | 5363.47 |
| Num. obs. | 7759 | 7617 | 7759 | 7759 |

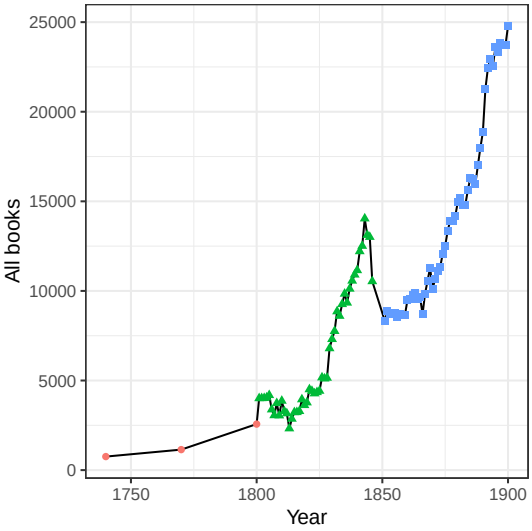
$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Notes: The table presents regression estimates with standard errors in parentheses. Standard errors are clusters at the author level. The observations are restricted to 1700-1932 and all observations with missing location are removed. The reference categories are authors aged 18-40 and authors born 1831-1900. Columns 1 and 3 were estimated using logistic regression. Columns 2 and 3 were estimated using a Poisson model. Details to calculations are given in Appendix B.2.

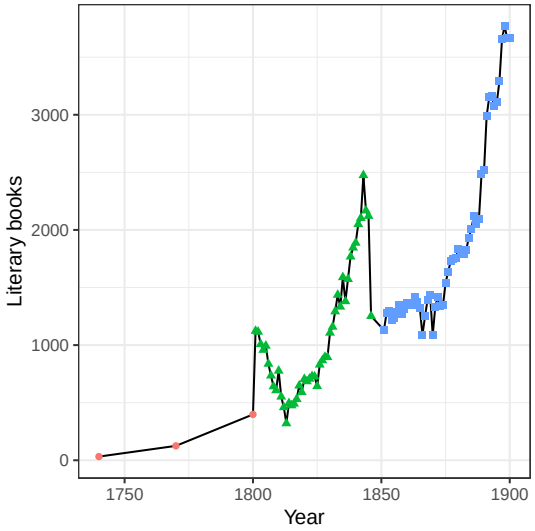
## A.2 Figures

Figure A.1: Growth in German publishing and book trade

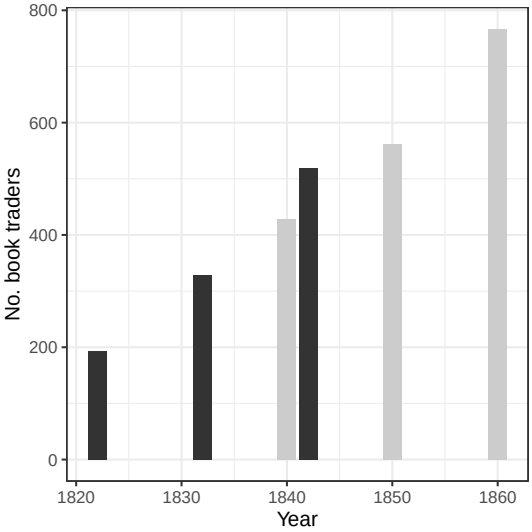
(a) No. book titles



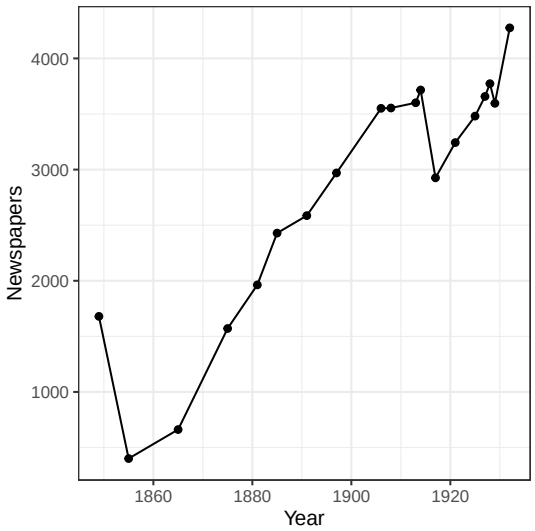

(b) No. literary book titles




(c) No. book traders in major cities




(d) No. newspaper titles

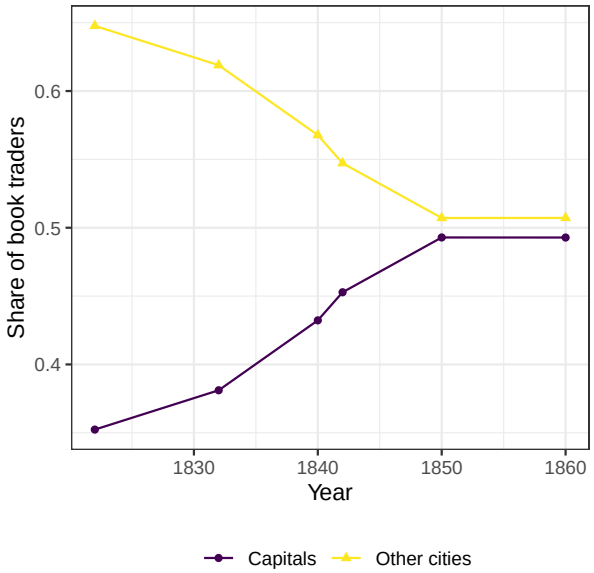



*Sources:* See Appendix B.3.2 for Figures (a) and (b); Appendix B.3.5 for Figure (c); Appendix B.3.3 for Figure (d).
*Notes:* Figures (a) and (b) shows the number of unique titles published within the area of the German book trade. Figure (a) includes all titles; Figure (b) only includes poetry, prose, and novels. Figure (c) shows the number of book traders in major cities. Figure (d) shows the number of daily newspaper titles published within the German Confederation / German Empire.

Figure A.2: Centers of book trade and publishing


*Source:* See Appendix B.3.5 for Figure (a) and Appendix B.3.6 for Figure (b).
*Notes:* Figure (a) shows the number of book traders by city (main cities in color). Figure (b) shows the city and year of establishment of literary book publishers active in the German empire (after 1870).

Figure A.3: Book traders by capital status

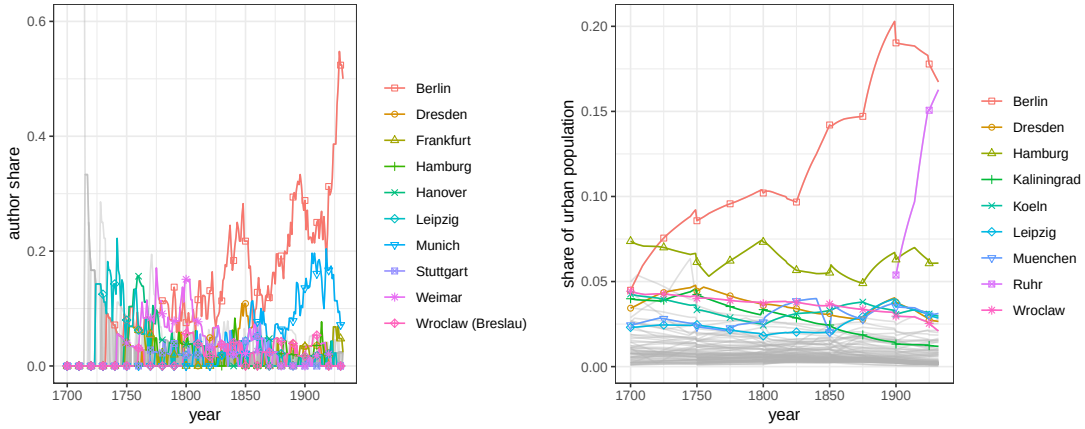



*Source:* See Appendices B.3.5 and B.3.9.
*Note:* The figure shows the number of book traders by city by capital status (i.e., whether a city was the capital of a state at the time).

Figure A.4: Share of authors and urban population in major cities

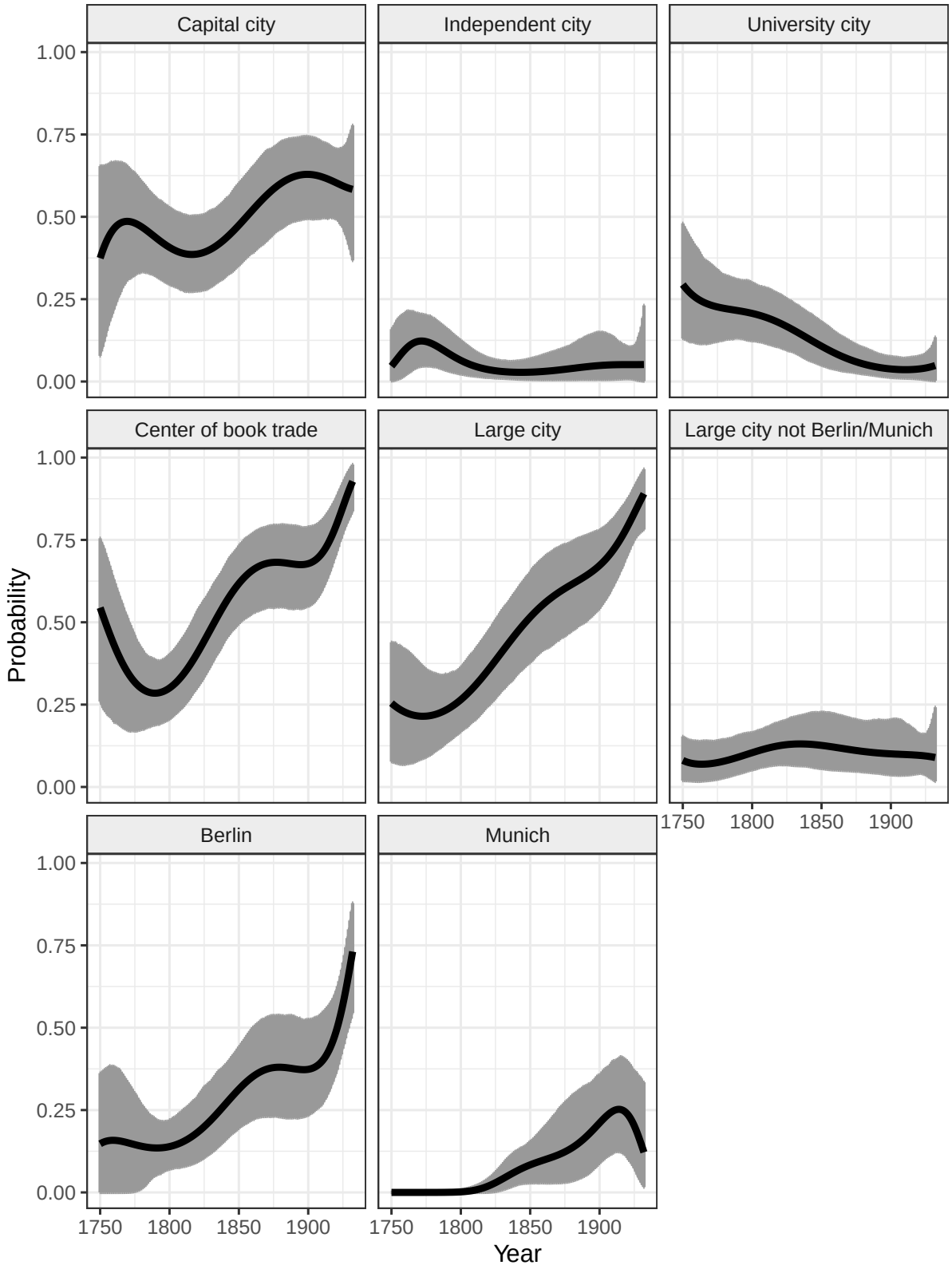
(a) Share of authors

(b) Share of urban population



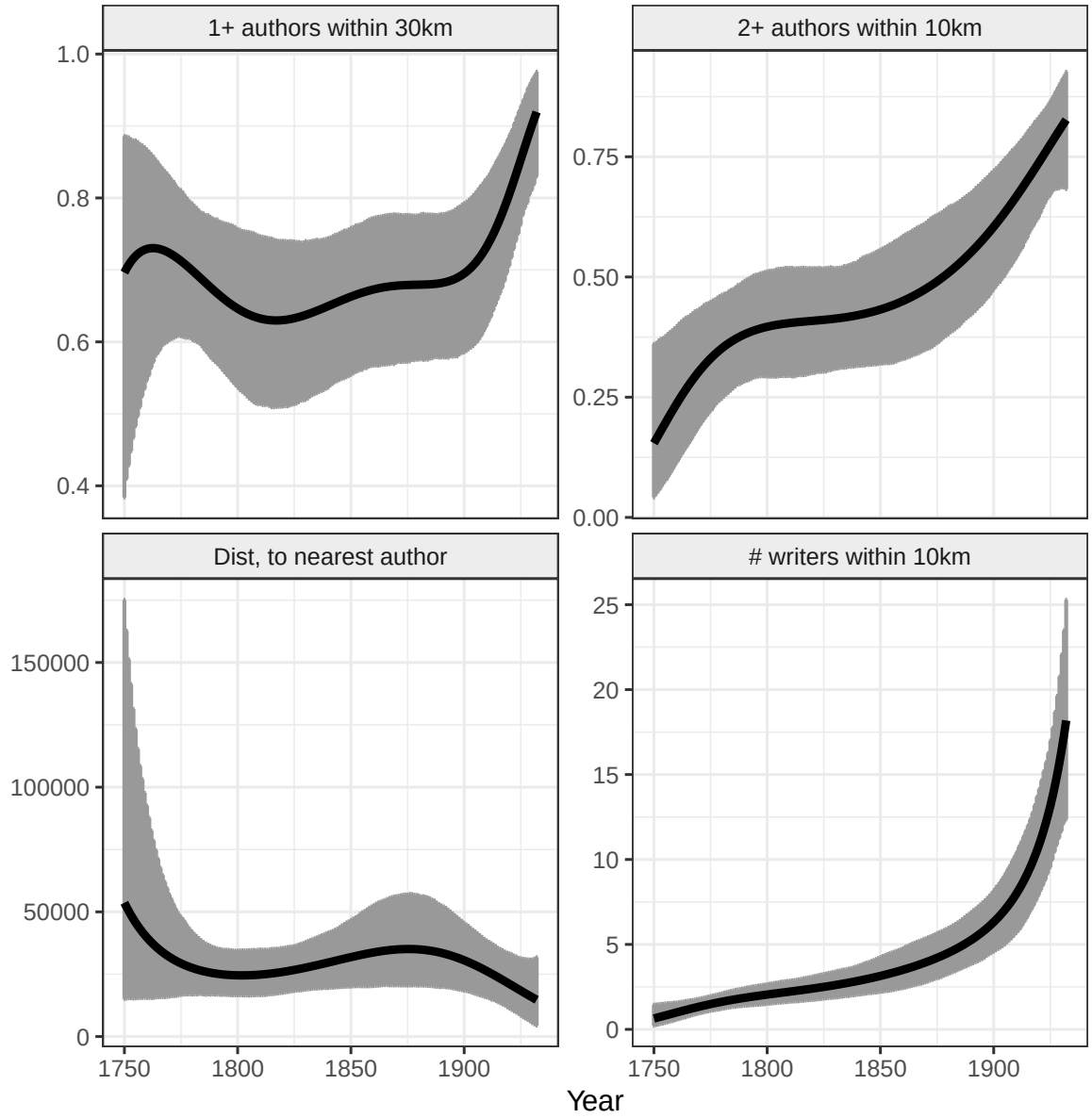

*Source:* See Appendices B.3.1 and B.3.7.
*Notes:* General population data is not available, so we calculate the share of the urban population.

Figure A.5: Attraction of various location types over time



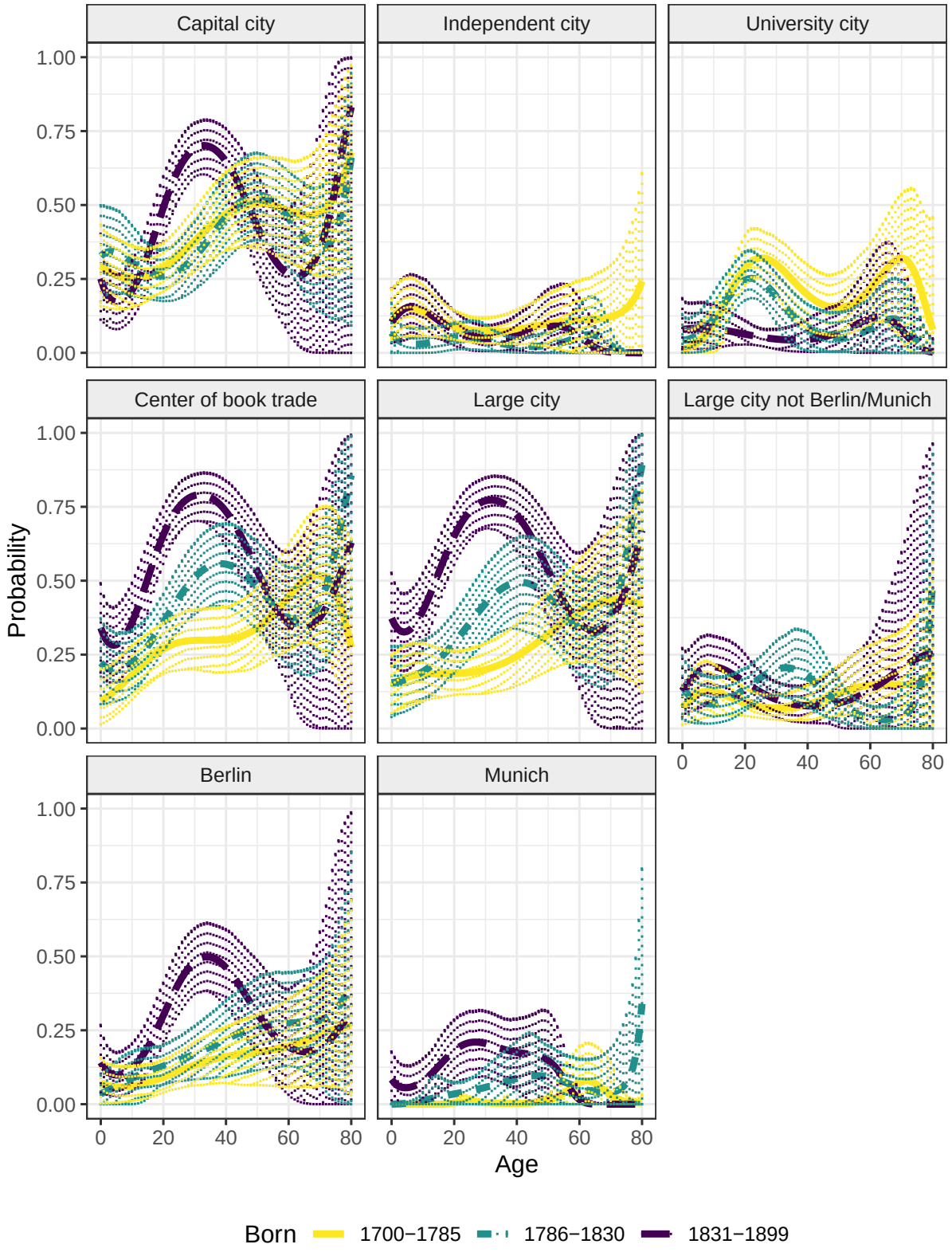

*Notes:* Probability is estimated for a 30-year old writer. Lighter colored lines represent 95 percent confidence intervals. Details to calculations are given in Section B.2.

Figure A.6: Proximity to other authors over time



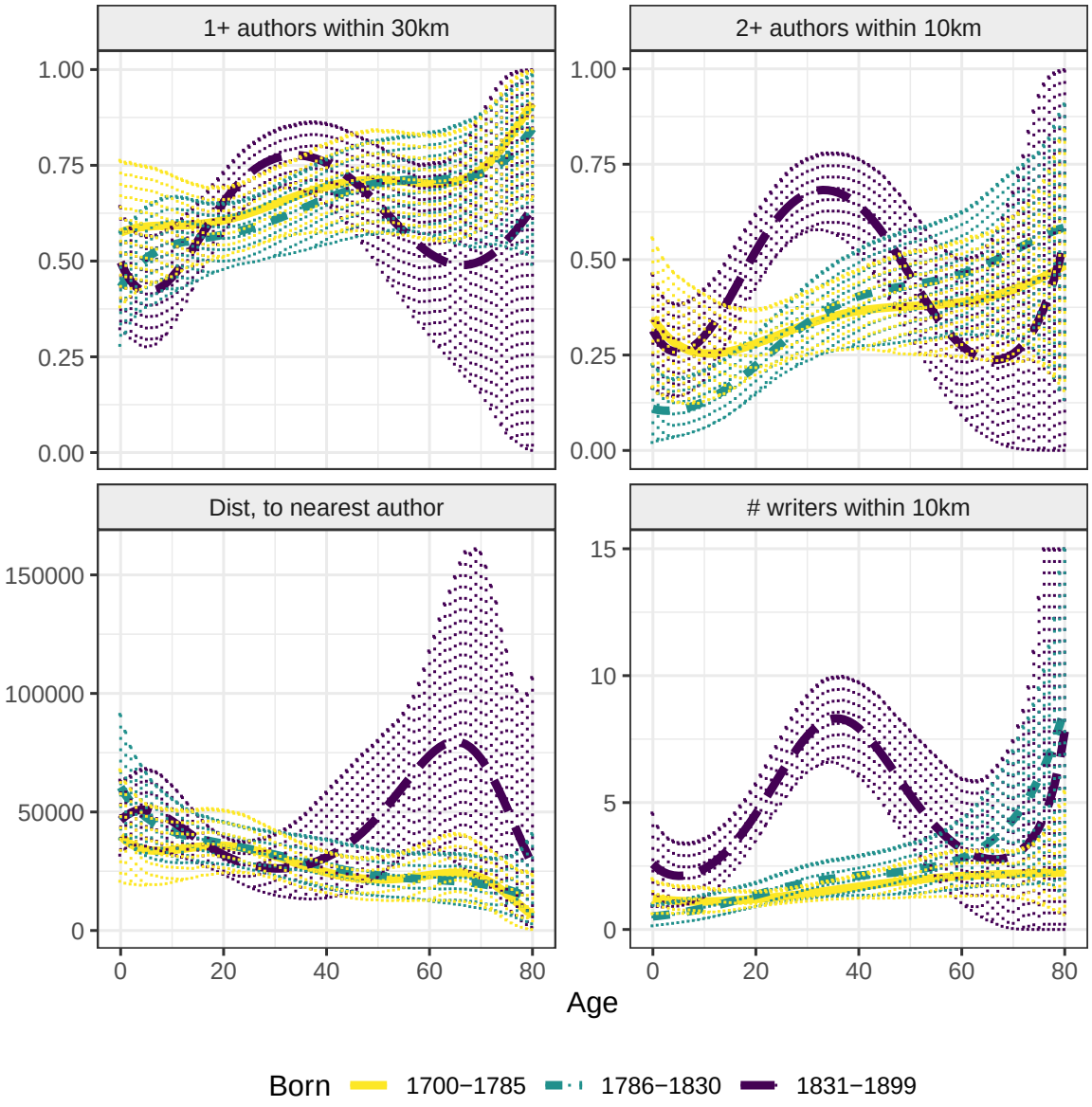

*Notes:* Probability is estimated for a 30-year old writer. Distances are in meters. Lighter colored lines represent 95 percent confidence intervals. Details to calculations are given in Section B.2.

Figure A.7: Attraction of various location types by birth cohort



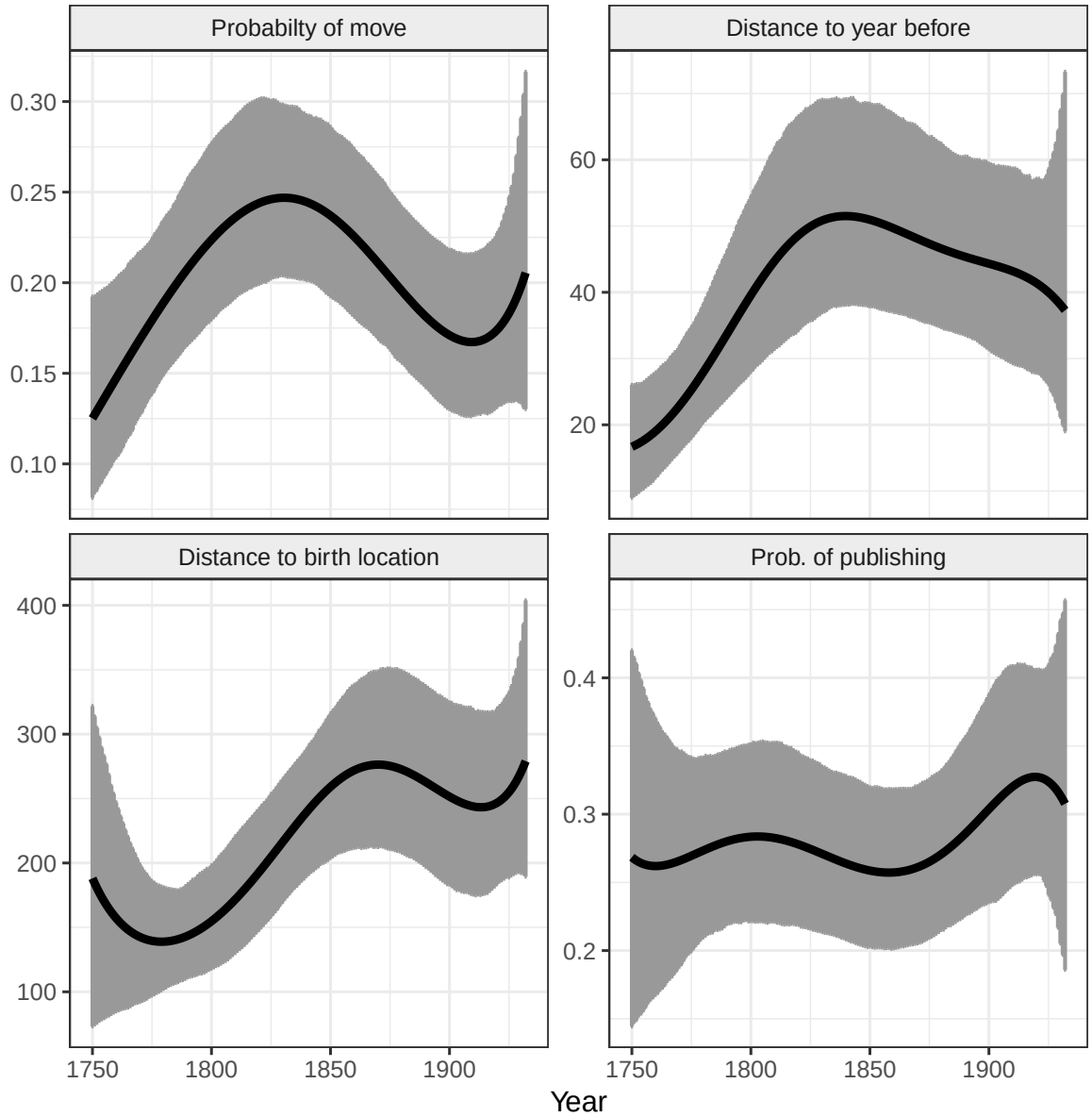

*Notes:* Lighter colored lines represent 95 percent confidence intervals. Details to calculations are given in Section B.2.

Figure A.8: Proximity to other authors over life-cycle



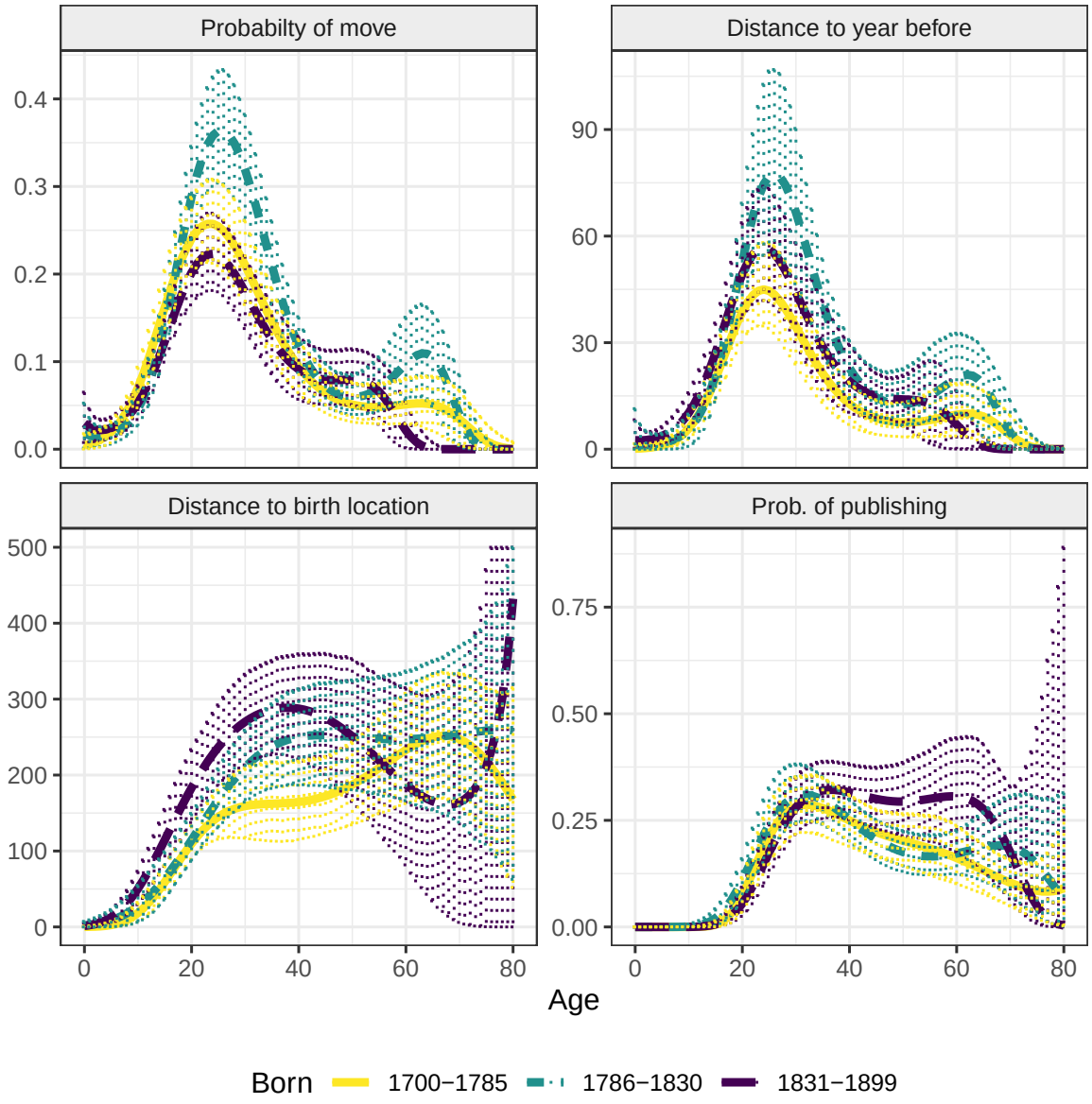

*Notes:* Distances are in meters. Lighter colored lines represent 95 percent confidence intervals. Details to calculations are given in Section B.2.

Figure A.9: Migration and publication trends over time




*Notes:* Probability is estimated for a 30-year old writer. All distances are in kilometers. Lighter shaded area represent 95 percent confidence intervals. Details to calculations are given in Section B.2.

Figure A.10: Migration and publication trends over life-cycle



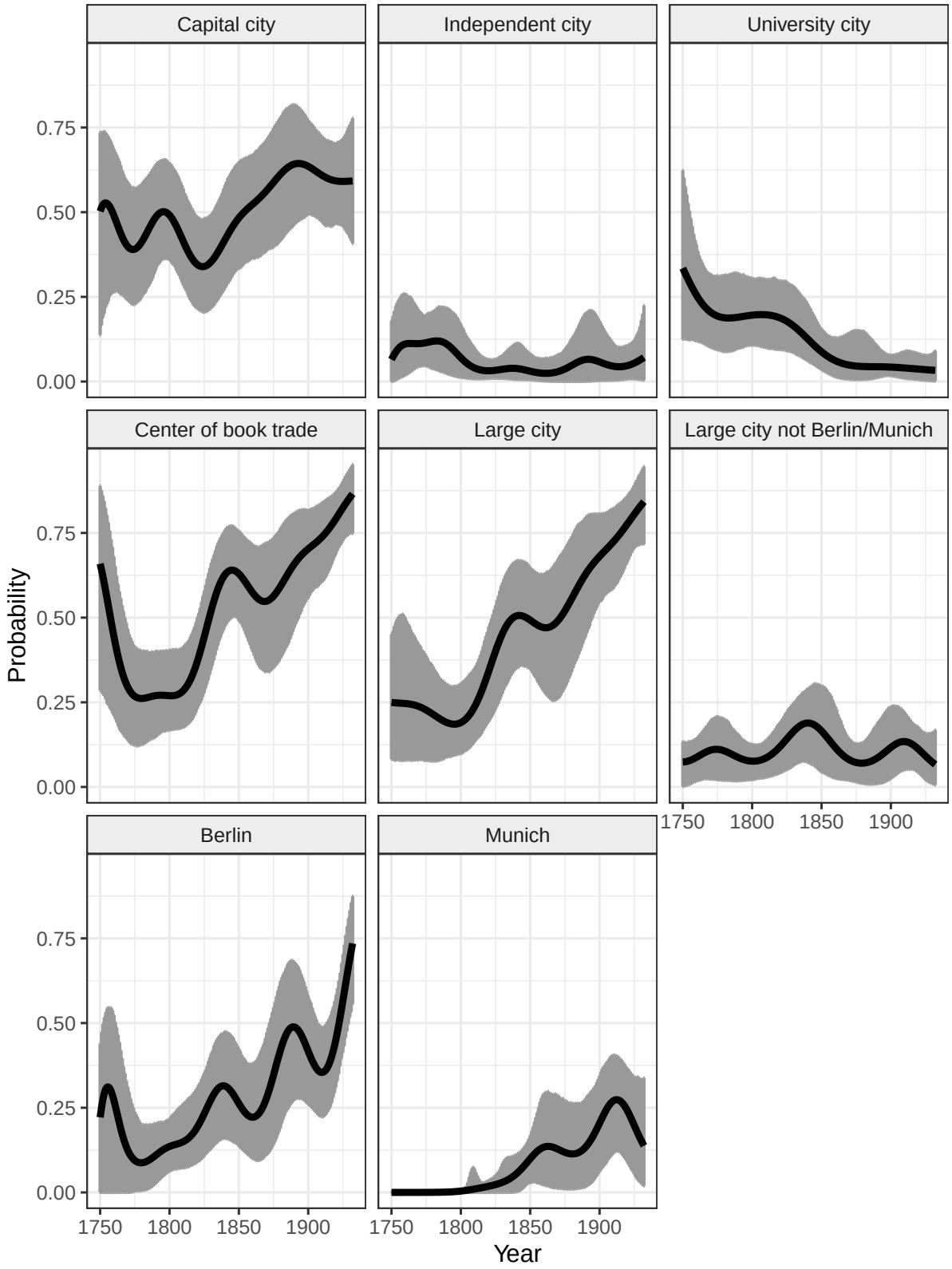

*Notes:* Distance is in kilometers. Lighter colored lines represent 95 percent confidence intervals. Details to calculations are given in Section B.2.

# B Supplementary Appendix

## B.1 Additional Results

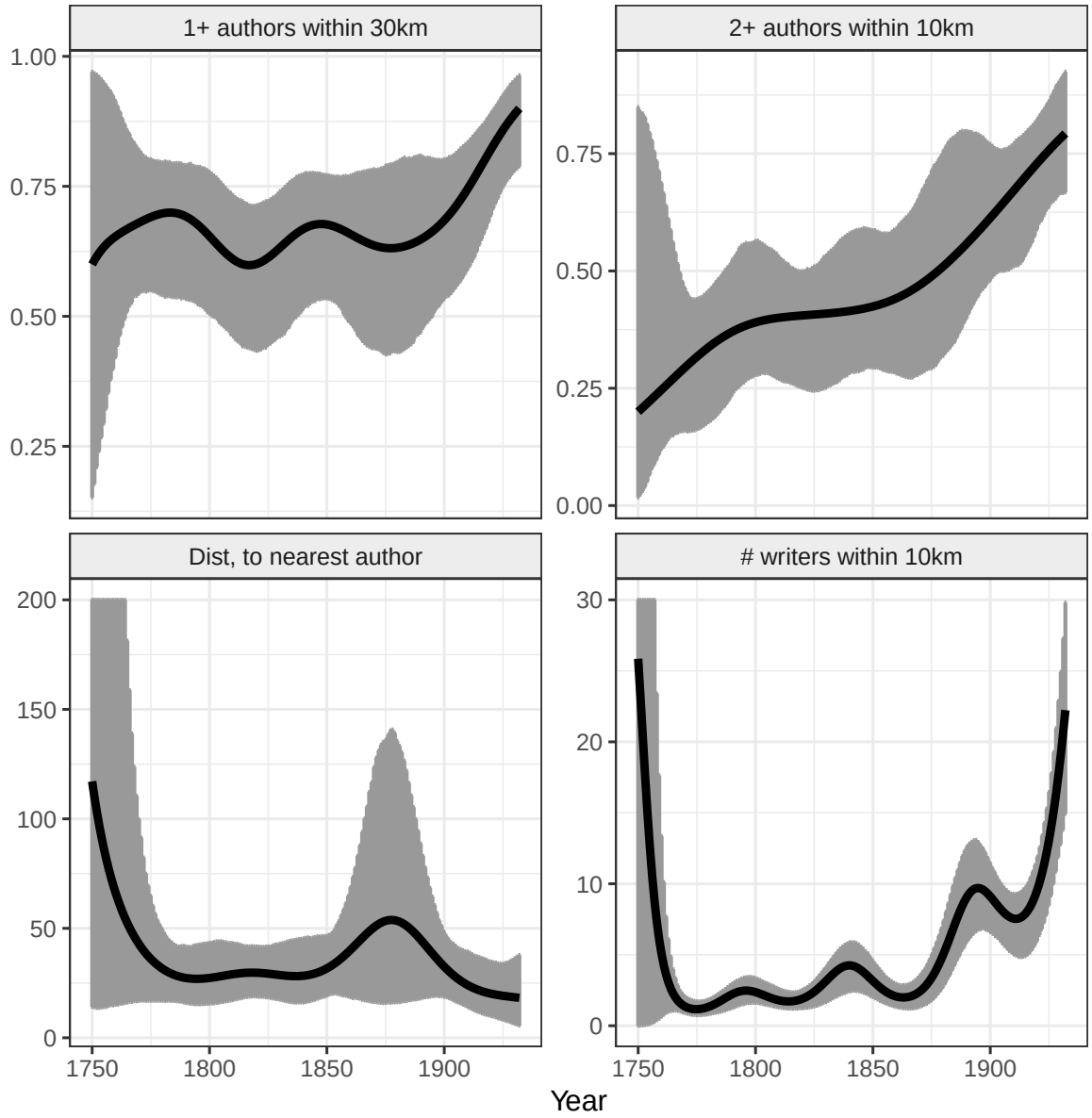
### B.1.1 GAM estimates

Figure B.1: Attraction of various location types over time, GAM estimates



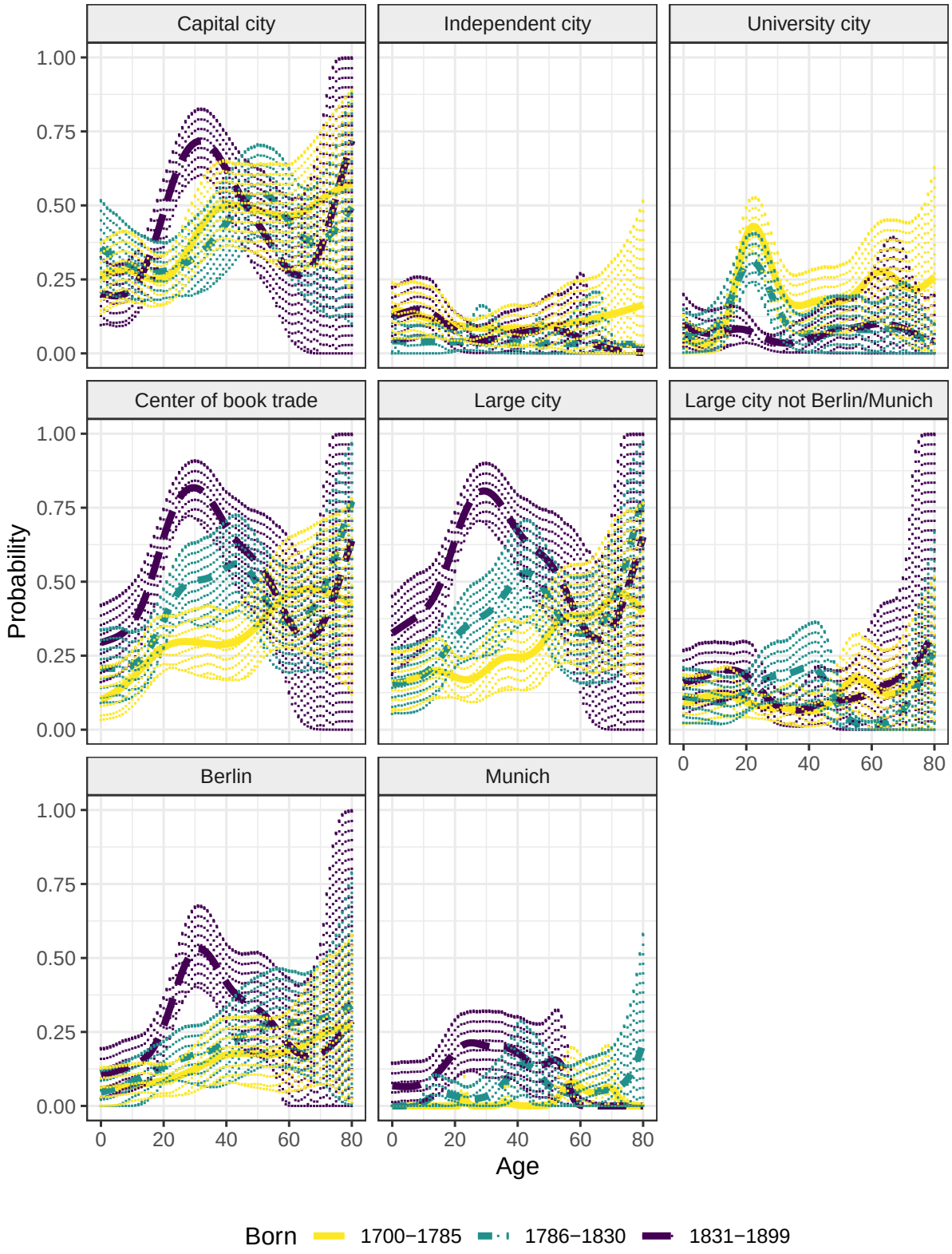

*Notes:* Probability is estimated for a 30-year old writer. Lighter colored lines represent 95 percent confidence intervals. Details to calculations are given in Section B.2.

Figure B.2: Proximity to other authors over time, GAM estimates



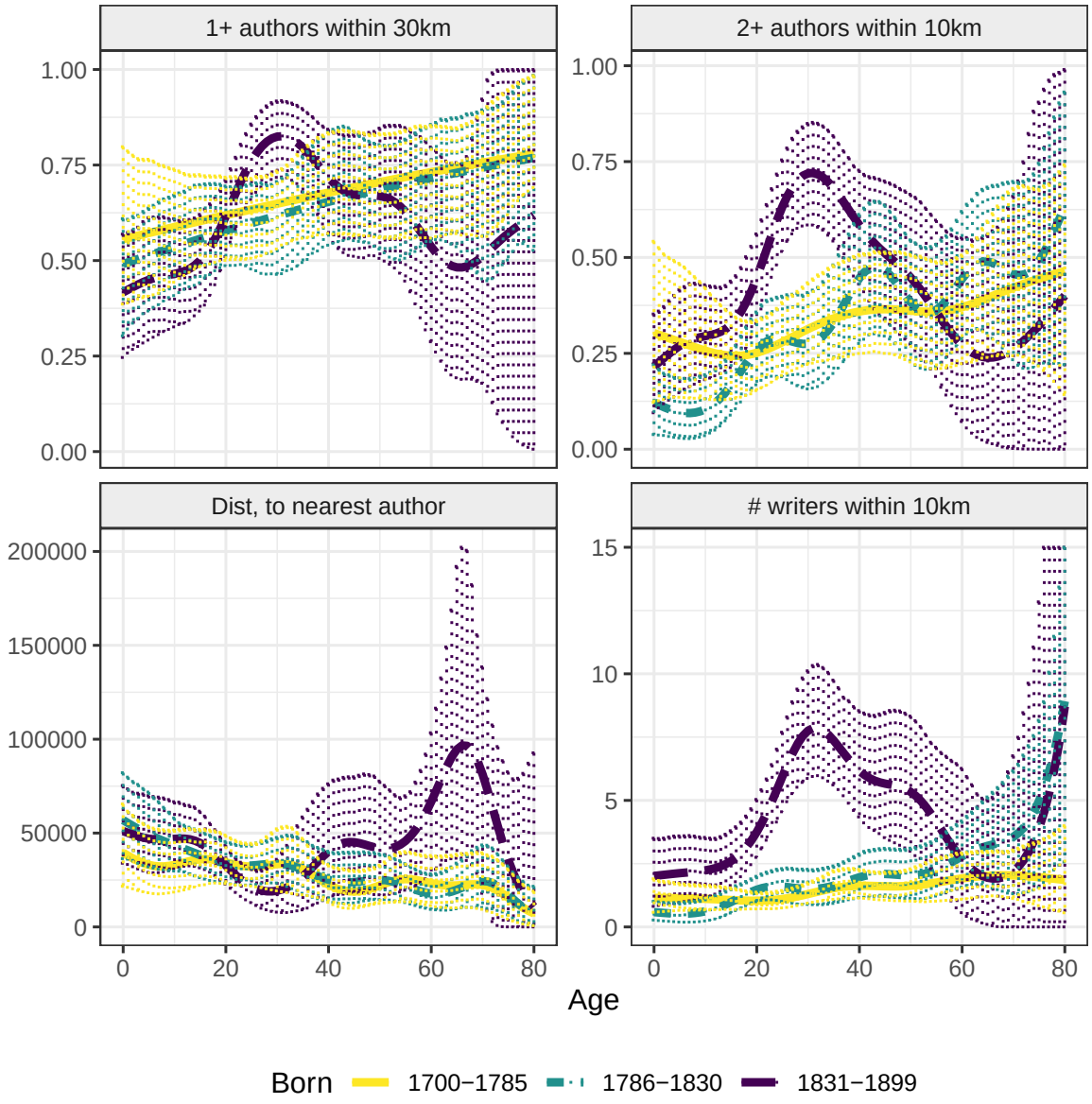

Probability is estimated for a 30-year old writer. Distances are in kilometers. Lighter colored lines represent 95 percent confidence intervals. Details to calculations are given in Section B.2.

Figure B.3: Attraction of various location types by birth cohort, GAM estimates



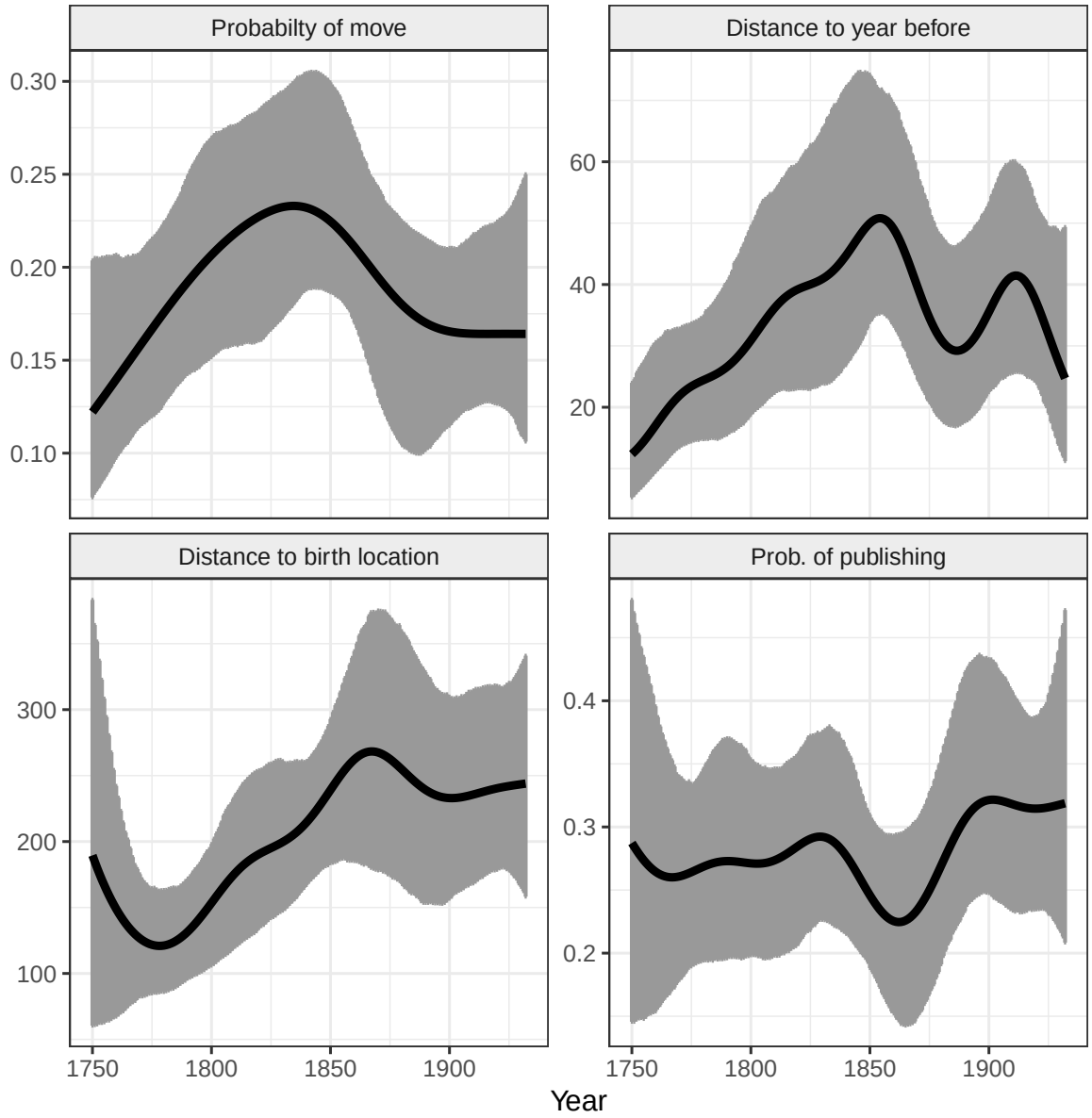

*Notes:* Lighter colored lines represent 95 percent confidence intervals. Details to calculations are given in Section B.2.

Figure B.4: Proximity to other authors over life-cycle, GAM estimates



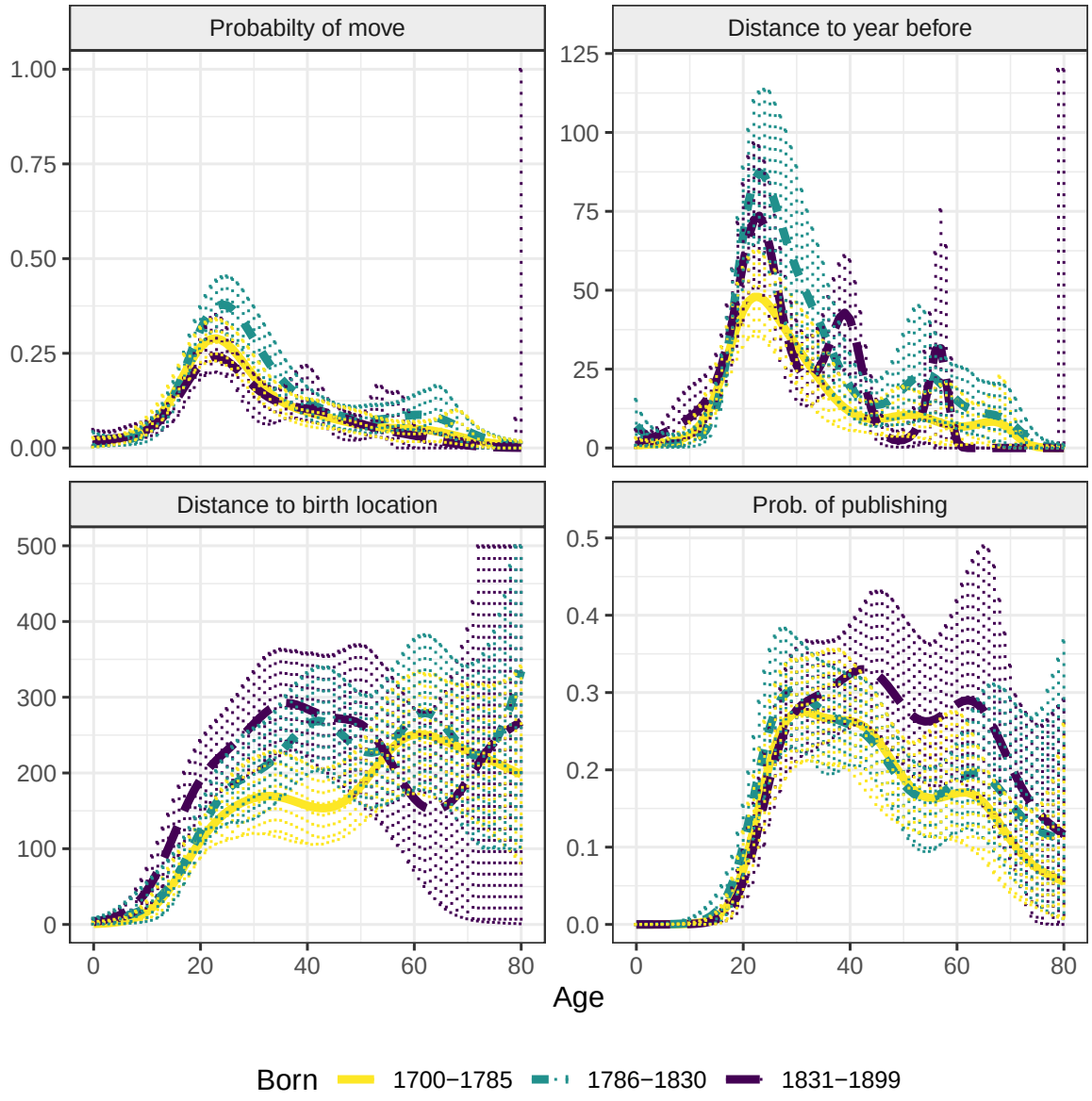

*Notes:* Distances are in meters. Lighter colored lines represent 95 percent confidence intervals. Details to calculations are given in Section B.2.

Figure B.5: Migration and publication trends over time, GAM estimates




*Notes:* Distances are in kilometers. Lighter colored lines represent 95 percent confidence intervals. Details to calculations are given in Section B.2.

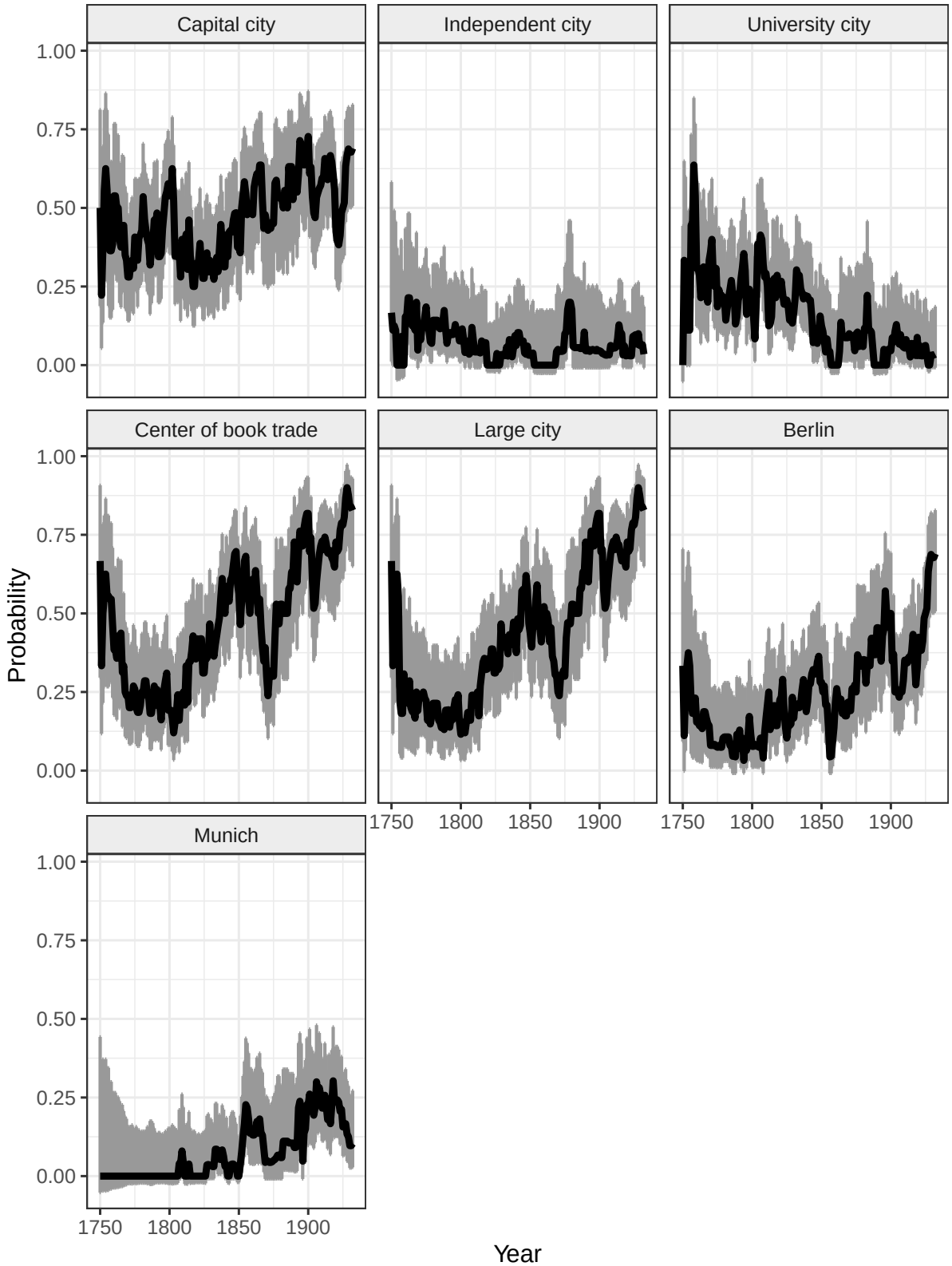
Figure B.6: Migration and publication trends over life-cycle, GAM estimates




*Notes:* Distance is in kilometers. Lighter colored lines represent 95 percent confidence intervals. Details to calculations are given in Section B.2.

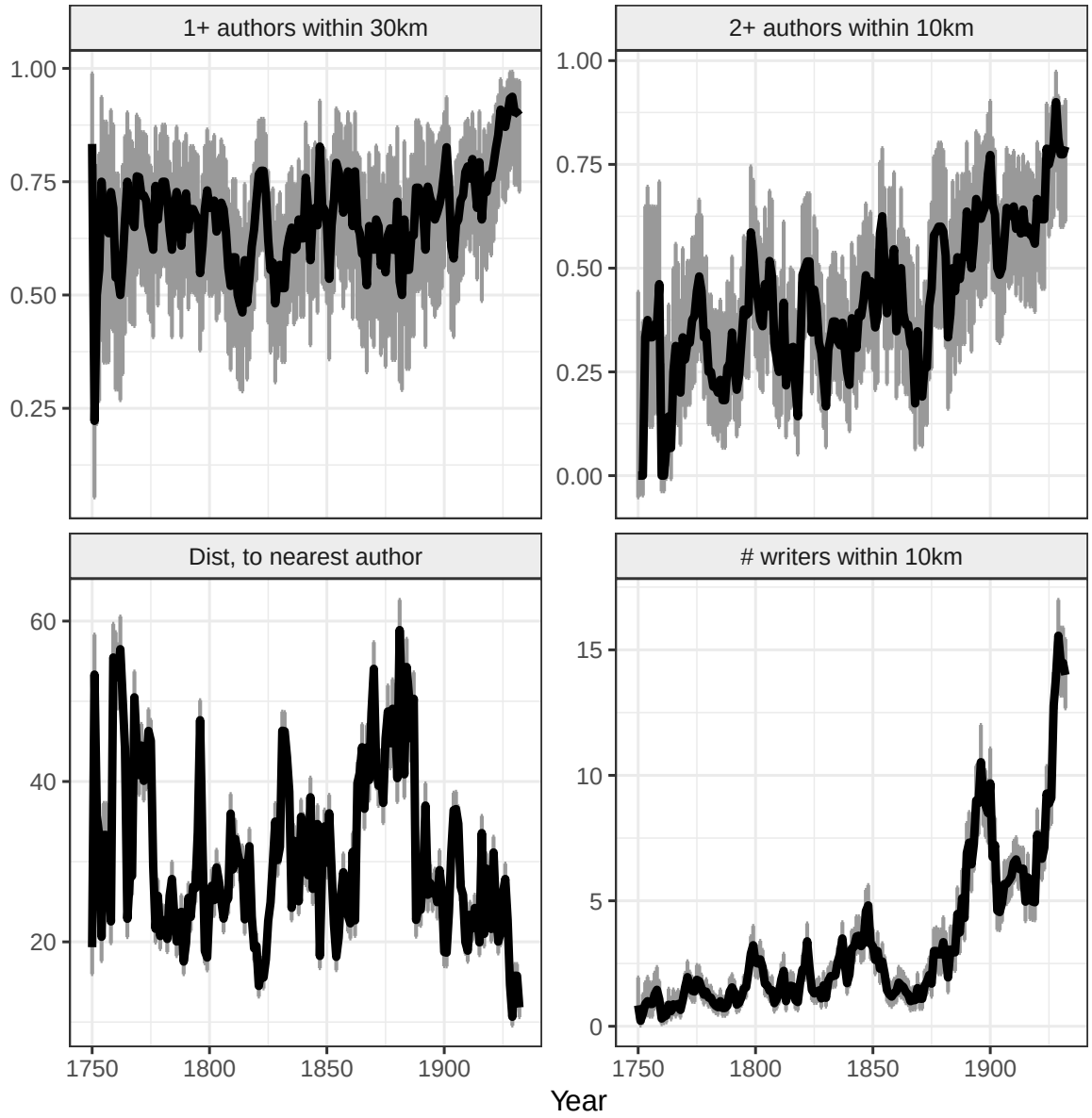
### B.1.2 Simple averages

Figure B.7: Attraction of various location types over time, yearly means



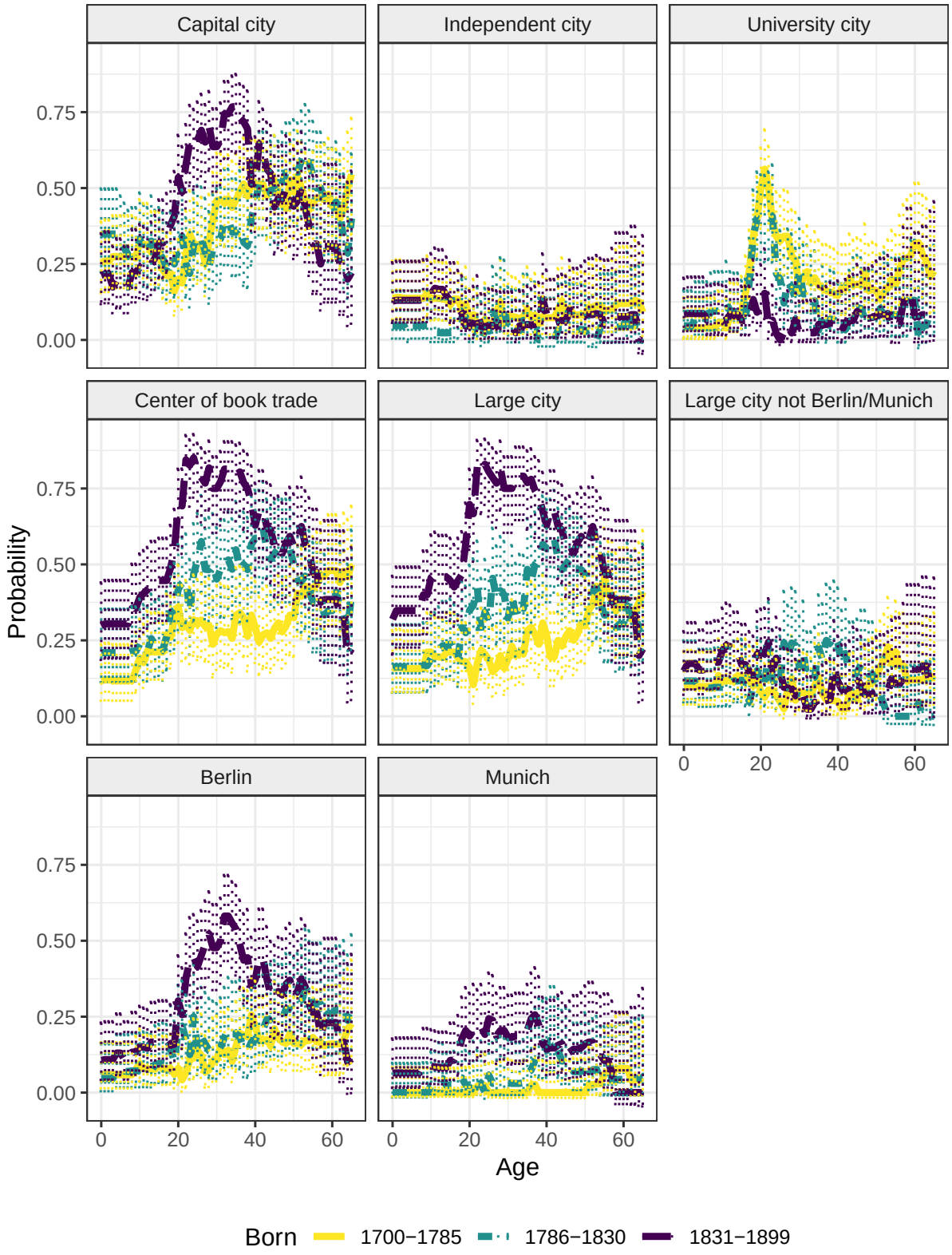

Yearly averages with 95 % confidence-intervals for authors aged 18 to 65 and living at a known location within the borders of 1910 Germany. Details to calculations are given in Section B.2.

Figure B.8: Proximity to other authors over time, yearly means



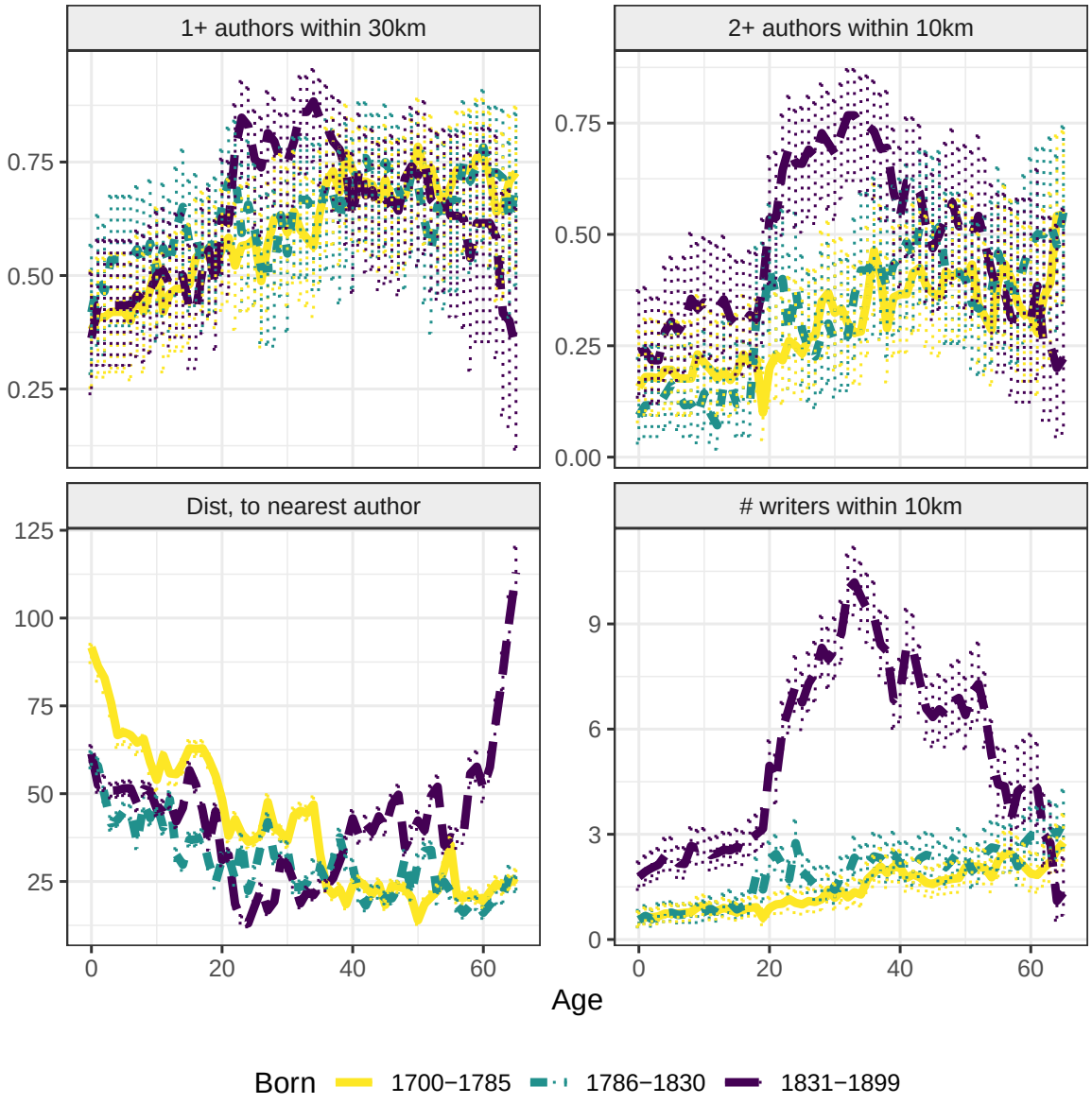

Yearly averages with 95 % confidence-intervals for authors aged 18 to 65 and living at a known location within the borders of 1910 Germany. Details to calculations are given in Section B.2.

Figure B.9: Attraction of various location types by birth cohort, age means



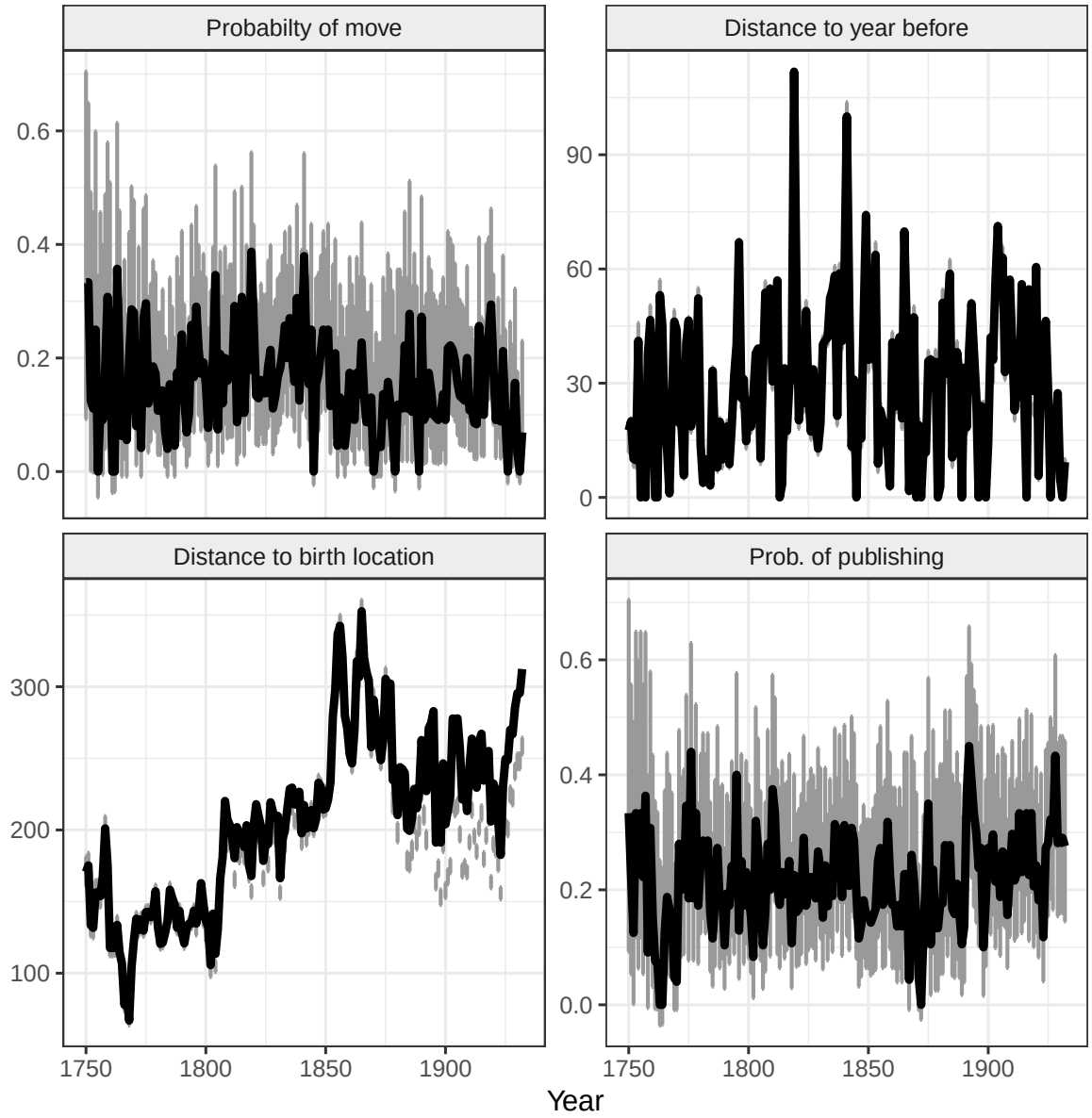

*Notes:* Yearly averages with 95 % confidence-intervals for authors living at a known location within the borders of 1910 Germany and between 1700 and 1932. Details to calculations are given in Section B.2.

Figure B.10: Proximity to other authors over life-cycle, age means



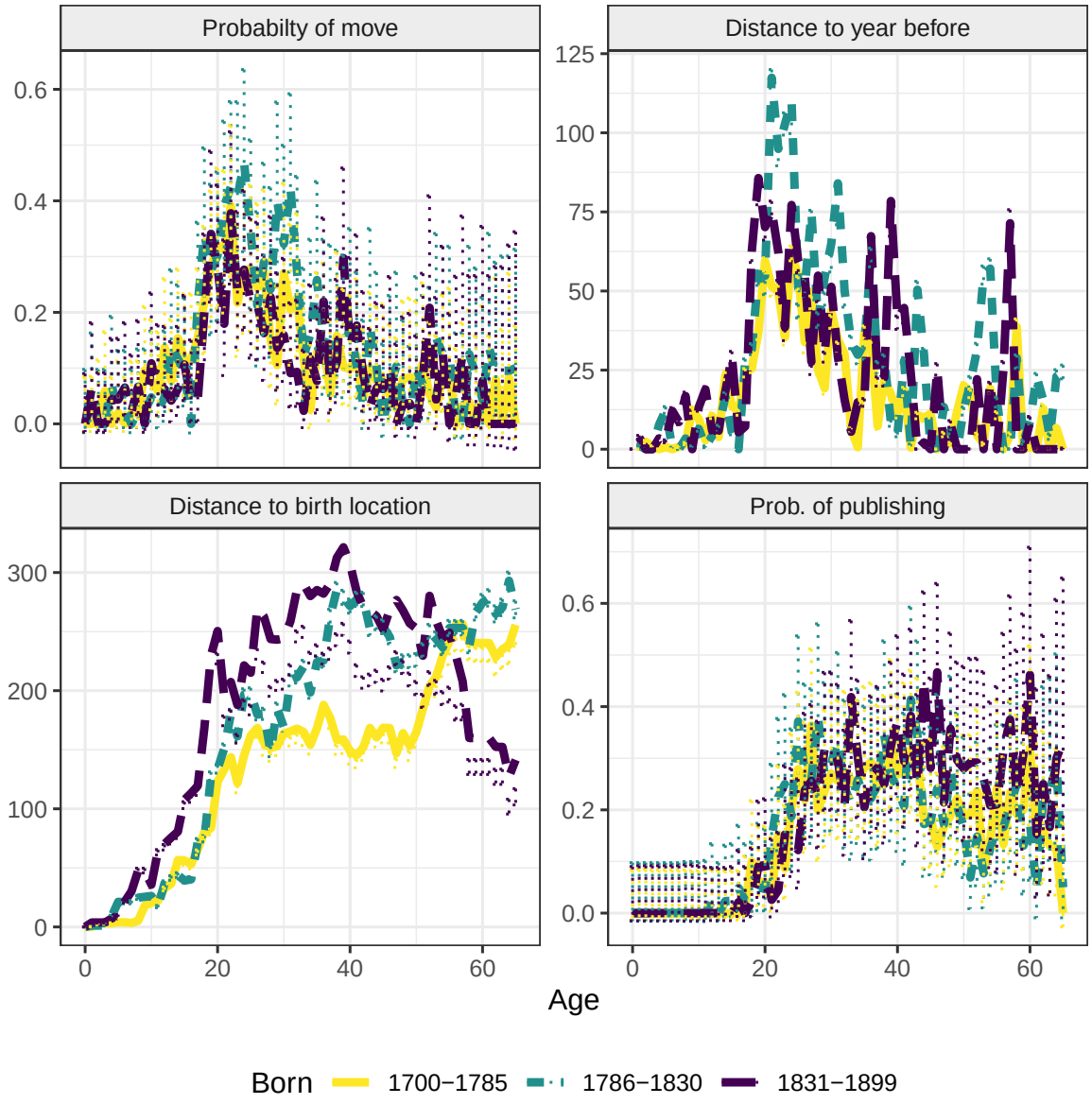

*Notes:* Yearly averages with 95 % confidence-intervals for authors living at a known location within the borders of 1910 Germany and between 1700 and 1932. Details to calculations are given in Section B.2.

Figure B.11: Migration and publication trends over time, yearly means




*Notes:* Distances are in kilometers. Yearly averages with 95 % confidence-intervals for authors aged 18 to 65 and living at a known location within the borders of 1910 Germany. Details to calculations are given in Section B.2.

Figure B.12: Migration and publication trends over life-cycle, age means




*Notes:* Distance is in kilometers. Yearly averages with 95 % confidence-intervals for authors living at a known location within the borders of 1910 Germany and between 1700 and 1932. Details to calculations are given in Section B.2.

## B.2 Technical Appendix

### B.2.1 Estimating location trends over time

We predict patterns in location type and proximity to other authors over the time using logit and Poisson regressions. All estimations in this section are limited to observation of authors between 18 and 65 and living at a known address within the borders of 1910 Germany. We derive the predictions by estimating the following relationship with $g$ as the link function:

$$g\left(\mathbb{E}\left(y_{it}\right)\right) = \sum_{k=1}^{5} \beta_{k,year} t^k + \sum_{k=1}^{5} \beta_{k,age} a_{it}^k + \beta_3 log(n_t + 1)\gamma \tag{1}$$

in which we estimate quintic polynomials for year $t$ and age $a_{it}$ and control for the logarithm of the total number of writers per year $n_t$, if $y$ depends on proximity to other writers. Depending on $y$, we estimate this relationship using logit (for binomial outcomes) and Poisson regressions (for count outcomes).

We use the $\beta$ estimates to predict the dependent variables for a 30-year old author between 1750 and 1932. We add the age polynomial to address age imbalances in our sample. Therefore, these estimates are different from simple averages, which we present as a robustness test. We do not interact year and age effects even though the patterns over the life-cycle are not constant as we show in our cohort analysis. This is motivated by two considerations. First, it simplifies the interpretation of the results. The estimates for a 30-year old writer are representative for all age groups in the relative size of the estimated outcome, i.e. a different age would not change the ordering of outcomes. Second, we postulate that authors are similar in their ideal, if not realized, location choices. They do show similar age patterns in migration and publication intensity.

Specifically, we estimate this relationship for the following outcomes:

- To analyze the attraction of various location types and major cities over time, we estimate:
  - the probability of being located in a capital city;
  - the probability of being located in an independent city;
  - the probability of being located in a city with an active university, which is not a large or capital city;
  - the probability of being located in one of the ten largest cities in a given year;
  - the probability of being located in a center of book trade;

  - the probability of being located in Berlin;
  - the probability of being located in Munich;
  - the probability of being located in a large city other than Berlin and Munich; and

- To analyze the formation of author clusters over time, we estimate:

  - the probability of there being at least one other author within 30km;
  - the probability of there being at least two other authors within 10km;
  - the distance to the nearest author; and
  - the number of other authors within 10m.

- To analyze migration and publication patterns over time, we estimate:

  - the probability of moving;
  - the distance from the current location to the location in the year before;
  - the distance from the current location to the birth location; and
  - the probability of publishing.

The 95% confidence intervals for these yearly point estimates are obtained using block bootstrap. Namely, we repeat the estimation 2,000 times by drawing authors randomly, with replacement to keep the group size constant. Then, we, take the 0.025 and 0.975 quantiles as limits to obtain intervals that contain 95 percent of the yearly estimated outcomes. We resample the authors instead of yearly observations to account for the correlation over time for observations of the same author.

As a robustness check, we, first, use thin plate regression splines for age and year instead of polynomials to relax the assumptions on the functional form. The generalized additive model framework serves as a check against over-fitting as well as other miss-specification. Second, we provide direct yearly means for all authors. Standard errors for the confidence intervals around these sample means are calculated using Agresti-Coull standard errors for the binomial outcomes and Poisson standard errors for count outcomes.

We present the yearly point estimates and simple means in graphs to illustrate the predicted outcomes over time. We use all observations with a known location within the 1910 borders of Germany for authors aged 18 to 65 for all years from 1700 to 1932 in the estimation. However, we truncate the graphs at 1750 due to a limited number of observations for the early years.

### B.2.2 Estimating life-cycle location trends by cohort

We predict patterns in location type and proximity to other authors over the life-cycle (from birth to age 80, using observations from known locations within 1910 Germany between 1700 and 1932) for each birth cohort. We derive the predictions by estimating the following relationship with $g$ as the link function:

$$g\left(\mathbb{E}\left(y_{it}\right)\right) = \beta_1 b_c + \sum_{k=1}^{5} \beta_{kc} a_{it}^{k} b_c + \beta_3 log(n_t + 1)\gamma \tag{2}$$

in which we interact birth cohort $b_i$ with a quintic age polynomial $a_{it}$ and control for the logarithm of the total number of writers per year $n_t$, if $y$ depends on proximity to other writers. Depending on $y$, we estimate this relationship using logit (for binomial outcomes) and Poisson regressions (for count outcomes).

Specifically, we estimate this relationship for the following outcomes:

- To analyze the attraction of various location types and major cities over the life-cycle, we estimate:
  - the probability of being located in a capital city;
  - the probability of being located in an independent city;
  - the probability of being located in a city with an active university, which is not a large or capital city;
  - the probability of being located in one of the ten largest cities in a given year;
  - the probability of being located in a center of book trade;
  - the probability of being located in Berlin;
  - the probability of being located in Munich; and
  - the probability of being located in a large city other than Berlin and Munich.
- To analyze the formation of author clusters over the life-cycle, we estimate:
  - the probability of there being at least one other author within 30km;
  - the probability of there being at least two other authors within 10km;
  - the distance to the nearest author; and
  - the number of other authors within 10m.
- To analyze migration and publication patterns over time, we estimate:
  - the probability of moving;
  - the distance from the current location to the location in the year before;
  - the distance from the current location to the birth location; and

– the probability of publishing.

The 95% confidence intervals for these yearly point estimates are obtained using block bootstrap. Namely, we repeat the estimation 2,000 times by drawing authors randomly, with replacement to keep the group size constant. Then, we, take the 0.025 and 0.975 quantiles as limits to obtain intervals that contain 95 percent of the estimates. We resample the authors instead of yearly observations to account for the correlation over time for observations of the same author.

As a robustness check, we, first, use thin plate regression splines for age, separately estimated by cohort, instead of polynomials to relax the functional form assumptions introduced. The generalized additive model framework serves as a check against over-fitting as well as other miss-specification. Second, we provide direct yearly age means for all authors. Standard errors for the confidence intervals around these sample means are calculated using Agresti-Coull standard errors for the binomial outcomes and Poisson standard errors for count outcomes.

We present the yearly point estimates and simple means in graphs to illustrate the predicted outcomes over the life-cycle (ages 0-80) for each birth cohort. We use all observations with a known location within the 1910 borders of Germany for authors aged 0 to 80 for all years from 1700 to 1932 in the estimation.

### B.2.3 Regressions with age and birth cohort dummies

We validate our analysis of trends over time and life-cycle trends by estimating regressions with age and birth cohort dummies. For this analysis, we estimate the following relationship with $g$ as the link function.

$$g\left(\mathbb{E}\left(y_{it}\right)\right) = \beta_{1,c} b_c + \beta_2 a_{it} + \beta_{3,c} b_c a_{it} + \beta_4 log(n_t + 1) \tag{3}$$

where $a_{it}$ is a categorical variable indicating whether author $i$ is under age 18, between age 18 and 40, or over age 40 in year $t$. $b_i$ is a categorical variable indicating whether author $i$ was born between 1700 to 1785, between 1786 and 1830, or 1830 and 1900. Last, $n_t$ is the total number of writers per year $n_t$, which we include as a control if $y$ depends on proximity to other writers, e.g. the number of writers within 10km. Depending on $y$, we estimate this relationship using

logit (for binomial outcomes) and Poisson regressions (for count outcomes).

Specifically, we estimate this relationship for the following outcomes:

- To analyze the attraction of various location types, we estimate:
  - the probability of being located in a capital city;
  - the probability of being located in an independent city;
  - the probability of being located in a city with an active university, which is not a large or capital city;
  - the probability of being located in one of the ten largest cities in a given year;
  - the probability of being located in a center of book trade; and
  - the probability of being located in an important city.
- To analyze the attraction of specific cities, we estimate:
  - the probability of being located in Berlin;
  - the probability of being located in Munich;
  - the probability of being located in other large cities (excluding Berlin and Munich); and
  - the distance from current location to Berlin.
- To analyze the formation of author clusters, we estimate:
  - the probability of there being at least one other author within 30km;
  - the probability of there being at least two other authors within 10km;
  - the distance to the nearest author; and
  - the number of other authors within 10m.
- To analyze migration and publication patterns over time, we estimate:
  - the probability of moving;
  - the distance from the current location to the location in the year before;
  - the distance from the current location to the birth location; and
  - the probability of publishing.

We use all observations with a known location within the 1910 borders of Germany for authors for all years from 1700 to 1932 in the estimation. We use the age category 18-40 as the reference category and the birth cohort 1830-1900 as the reference cohort. Standard errors are clustered on the author level.

### B.2.4 Computational implementation

All statistical work has been conducted using R [R Core Team, 2022]. The generalized linear model regressions are implemented using the glm function from the base {stats} package. Standard errors are re-estimated using the {sandwich} package [Zeileis, 2006]. {doParallel} has been used for parallelization [Corporation and Weston, 2022]. The generalized additive model regression are implemented using the gam function in the {mgcv} package [Wood, 2011]. All figures are produced using {ggplot2} [Wickham, 2016]. {sf} and {geosphere} have been valuable for calculating distances (Pebesma, 2018 and Hijmans, 2021).

All codes files for the regressions and resulting plots, as well as the complete author panel data, are available in our public repository using the following link [redacted to maintain anonymity of the authors, to be provided prior to publication].

## B.3 Data Appendix

### B.3.1 Author data

We followed the Mitchell [2019] methodology for data collection. We gathered a list of authors associated with German literature in Encyclopaedia Britannica, Inc. [2016] must have made at least one unique contribution to poetry or prose, which eliminated individuals whose contributions were strictly limited to translations, textbooks, song-writing, literary criticism, or other forms of publication.[22] Data on the general population or the general population of authors is not available, and therefore we do not make inferences about the general population of German writers.

Biographic information, number of publications per year, and data on location of residence was collected from Encyclopaedia Britannica, Inc. [2016], Literature Online [2016], and Deutsche Biographie [2016]. We identify the location of each author in every year of her life. In general, places of residence are only mentioned in an encyclopaedia entry if an author moved to a new location. For example, if the encyclopaedia entry indicates that a writer moved to Berlin in a given year and then moved to Leipzig 10 years later, then we assume that the author lived in Berlin continuously during that 10 year period. All empirical work is based on writer-individual data aggregated to a yearly level. This implies that we take a location in which a writer spends more than six months in a given year as the location for the whole year.

The Encyclopaedia Britannica, Inc. [2016] classification as German for this time frame corresponds geographically to the area of the pre-World War I German Empire, and we use a map of the 1910 German Empire to classify German locations in the remainder of this paper. We use these borders as they correspond to the Britannica definition for German author in our sample. This leads to a distinction between Germany and Austria that makes little sense politically before the German unification in 1871. However, less than 1 percent of years are spent by the writers in the sample in modern-day Austria prior to unification.

We identify the longitude and latitude of each location using OSM data by OpenStreetMap contributors [2021]. We calculate all geographic distances based on locality centers. Several writers have no location for some years due to extended travels or missing data.[23] In this case, we do not include these years to study location or movement. We use the boundaries of the

1910 German Empire to classify German locations for all time periods to ensure consistency.

The final data set is composed of 9,741 observations for 153 writers. A full list of the authors is available in Kuld and O'Hagan [2019].

### B.3.2 Annual book title production, 1740-1900

We gathered data on the number of unique book titles produced each year from the Rarisch [1976] study on book production, the publishing industry, and the book trade in the German Empire during the 19th century. In particular, we use the following archived data tables from Section "B. The Development of issued and on the book-market distributed book-titles by subject classification" (*B. Die Entwicklung der aufgelegten und im Handel vertriebenen Titel nach Sachgebieten*):

- B.1 Development of book-production by number of published titles for the years 1740, 1770 and 1800. (*B.1 Die Entwicklung der Buchproduktion nach Anzahl der aufgelegten Titel 1740, 1770 und 1800.*) File "ZA8564_B-1.xls" (Table 1, p. 13).
- B.4 The Development of book- and map-production of the german book-selling-sector, listed by the "codex nundinarius", 1801-1846. (*B.4 Die Entwicklung der Buch- und Kartenproduktion im Gebiet des deutschen Buchhandels nach dem codex nundinarius, 1801-1846.*) File "ZA8564_B-4.xls" (Appendix).
- B.5 The development of book and map production of the German book-selling sector, listed by the "Börsenblatt-Statistics", the Gazette of the German Book Trade, 1851-1900. (*B.5 Die Entwicklung der Buch- und Kartenproduktion im Gebiet des deutschen Buchhandels nach der Börsenblatt-Statistik, 1851-1900.*) File "ZA8564_B-5.xls" (Appendix).

By combining these three data tables, we are able to provide an estimate for the total number of unique book titles produced annually in Germany for the period 1740-1900. The data book title production and the respective sources is provided in Table B.1 below.

Table B.1: Total title production by year

| Year | No. titles | No. literary titles | Source |
|---|---|---|---|
| 1740 | 755 | 32 | B1 Rarisch |
| 1770 | 1144 | 125 | B1 Rarisch |
| 1800 | 2569 | 398 | B1 Rarisch |
| 1801 | 4008 | 1123 | B4 codex nundinarius |
| 1802 | 4010 | 1115 | B4 codex nundinarius |
| 1803 | 4016 | 1008 | B4 codex nundinarius |
| 1804 | 4049 | 958 | B4 codex nundinarius |
| 1805 | 4181 | 994 | B4 codex nundinarius |
| 1806 | 3381 | 837 | B4 codex nundinarius |
| 1807 | 3057 | 735 | B4 codex nundinarius |
| 1808 | 3733 | 641 | B4 codex nundinarius |
| 1809 | 3045 | 607 | B4 codex nundinarius |
| 1810 | 3864 | 777 | B4 codex nundinarius |
| 1811 | 3287 | 552 | B4 codex nundinarius |
| 1812 | 3162 | 460 | B4 codex nundinarius |
| 1813 | 2323 | 320 | B4 codex nundinarius |
| 1814 | 2861 | 495 | B4 codex nundinarius |
| 1815 | 3225 | 476 | B4 codex nundinarius |
| 1816 | 3231 | 491 | B4 codex nundinarius |
| 1817 | 3291 | 531 | B4 codex nundinarius |
| 1818 | 3945 | 648 | B4 codex nundinarius |
| 1819 | 3622 | 591 | B4 codex nundinarius |
| 1820 | 3772 | 705 | B4 codex nundinarius |
| 1821 | 4505 | 685 | B4 codex nundinarius |
| 1822 | 4414 | 707 | B4 codex nundinarius |
| 1823 | 4275 | 729 | B4 codex nundinarius |
| 1824 | 4346 | 727 | B4 codex nundinarius |
| 1825 | 4421 | 643 | B4 codex nundinarius |

*Continued on next page...*

Table B.1 – Continued from previous page

| Year | No. titles | No. literary titles | Source |
|---|---|---|---|
| 1826 | 5168 | 830 | B4 codex nundinarius |
| 1827 | 5106 | 867 | B4 codex nundinarius |
| 1828 | 5148 | 897 | B4 codex nundinarius |
| 1829 | 6794 | 895 | B4 codex nundinarius |
| 1830 | 7308 | 1107 | B4 codex nundinarius |
| 1831 | 7757 | 1159 | B4 codex nundinarius |
| 1832 | 8855 | 1292 | B4 codex nundinarius |
| 1833 | 8603 | 1436 | B4 codex nundinarius |
| 1834 | 9258 | 1333 | B4 codex nundinarius |
| 1835 | 9840 | 1589 | B4 codex nundinarius |
| 1836 | 9341 | 1380 | B4 codex nundinarius |
| 1837 | 10118 | 1573 | B4 codex nundinarius |
| 1838 | 10567 | 1768 | B4 codex nundinarius |
| 1839 | 10907 | 1846 | B4 codex nundinarius |
| 1840 | 11151 | 1888 | B4 codex nundinarius |
| 1841 | 12209 | 2050 | B4 codex nundinarius |
| 1842 | 12509 | 2103 | B4 codex nundinarius |
| 1843 | 14039 | 2475 | B4 codex nundinarius |
| 1844 | 13119 | 2168 | B4 codex nundinarius |
| 1845 | 13008 | 2121 | B4 codex nundinarius |
| 1846 | 10536 | 1252 | B4 codex nundinarius |
| 1851 | 8326 | 1130 | B5 Börsenblatt-Statistik |
| 1852 | 8857 | 1275 | B5 Börsenblatt-Statistik |
| 1853 | 8750 | 1298 | B5 Börsenblatt-Statistik |
| 1854 | 8705 | 1222 | B5 Börsenblatt-Statistik |
| 1855 | 8794 | 1238 | B5 Börsenblatt-Statistik |
| 1856 | 8540 | 1287 | B5 Börsenblatt-Statistik |
| 1857 | 8699 | 1345 | B5 Börsenblatt-Statistik |
| 1858 | 8672 | 1271 | B5 Börsenblatt-Statistik |

*Continued on next page...*

Table B.1 – Continued from previous page

| Year | No. titles | No. literary titles | Source |
|---|---|---|---|
| 1859 | 8666 | 1315 | B5 Börsenblatt-Statistik |
| 1860 | 9496 | 1367 | B5 Börsenblatt-Statistik |
| 1861 | 9566 | 1357 | B5 Börsenblatt-Statistik |
| 1862 | 9779 | 1350 | B5 Börsenblatt-Statistik |
| 1863 | 9889 | 1414 | B5 Börsenblatt-Statistik |
| 1864 | 9564 | 1374 | B5 Börsenblatt-Statistik |
| 1865 | 9661 | 1320 | B5 Börsenblatt-Statistik |
| 1866 | 8699 | 1088 | B5 Börsenblatt-Statistik |
| 1867 | 9855 | 1249 | B5 Börsenblatt-Statistik |
| 1868 | 10563 | 1395 | B5 Börsenblatt-Statistik |
| 1869 | 11305 | 1434 | B5 Börsenblatt-Statistik |
| 1870 | 10108 | 1085 | B5 Börsenblatt-Statistik |
| 1871 | 10669 | 1335 | B5 Börsenblatt-Statistik |
| 1872 | 11127 | 1418 | B5 Börsenblatt-Statistik |
| 1873 | 11315 | 1339 | B5 Börsenblatt-Statistik |
| 1874 | 12070 | 1346 | B5 Börsenblatt-Statistik |
| 1875 | 12516 | 1539 | B5 Börsenblatt-Statistik |
| 1876 | 13356 | 1635 | B5 Börsenblatt-Statistik |
| 1877 | 13925 | 1728 | B5 Börsenblatt-Statistik |
| 1878 | 13912 | 1752 | B5 Börsenblatt-Statistik |
| 1879 | 14179 | 1754 | B5 Börsenblatt-Statistik |
| 1880 | 15341 | 1836 | B5 Börsenblatt-Statistik |
| 1881 | 15191 | 1807 | B5 Börsenblatt-Statistik |
| 1882 | 14794 | 1789 | B5 Börsenblatt-Statistik |
| 1883 | 14802 | 1822 | B5 Börsenblatt-Statistik |
| 1884 | 15607 | 1926 | B5 Börsenblatt-Statistik |
| 1885 | 16305 | 2005 | B5 Börsenblatt-Statistik |
| 1886 | 16253 | 2118 | B5 Börsenblatt-Statistik |
| 1887 | 15972 | 2050 | B5 Börsenblatt-Statistik |

*Continued on next page...*

Table B.1 – Continued from previous page

| Year | No. titles | No. literary titles | Source |
|---|---|---|---|
| 1888 | 17000 | 2093 | B5 Börsenblatt-Statistik |
| 1889 | 17986 | 2483 | B5 Börsenblatt-Statistik |
| 1890 | 18875 | 2518 | B5 Börsenblatt-Statistik |
| 1891 | 21279 | 2988 | B5 Börsenblatt-Statistik |
| 1892 | 22435 | 3152 | B5 Börsenblatt-Statistik |
| 1893 | 22946 | 3162 | B5 Börsenblatt-Statistik |
| 1894 | 22570 | 3078 | B5 Börsenblatt-Statistik |
| 1895 | 23607 | 3114 | B5 Börsenblatt-Statistik |
| 1896 | 23339 | 3293 | B5 Börsenblatt-Statistik |
| 1897 | 23861 | 3659 | B5 Börsenblatt-Statistik |
| 1898 | 23739 | 3772 | B5 Börsenblatt-Statistik |
| 1899 | 23715 | 3664 | B5 Börsenblatt-Statistik |
| 1900 | 24792 | 3670 | B5 Börsenblatt-Statistik |

### B.3.3 Number of daily newspaper titles, 1849-1932

We retrieved data on the number of unique titles of daily newspapers, including main and supplementary editions, published within the German Confederation and Empire from 1849 to 1932 from the Rahlf [2015] time series data for Germany, 1834-2012. Specifically, we use the variable *Titel Tageszeitungen Haupt und Nebenausgaben* in the file "K11_1_Kultur___Tourismus_und_Sport_-_Zeitungen_und_Zeitschriften.csv" from data file ZA8603 in the GESIS Data Archive.

We provide the number of daily newspaper titles per year in Table B.2 below.

Table B.2: Number of daily newspaper titles

| Year | No. titles |
|---|---|
| 1849 | 1680 |
| 1855 | 401 |
| 1865 | 662 |
| 1875 | 1571 |
| 1881 | 1963 |
| 1885 | 2429 |
| 1891 | 2586 |
| 1897 | 2970 |
| 1906 | 3551 |
| 1908 | 3554 |
| 1913 | 3601 |
| 1914 | 3716 |
| 1917 | 2926 |
| 1921 | 3243 |
| 1925 | 3481 |
| 1927 | 3658 |
| 1928 | 3773 |
| 1929 | 3596 |
| 1932 | 4275 |

### B.3.4 Number of book trade association members, 1834-1932

We retrieved data on the number of book trade association members from 1834-1932 from the Rahlf [2015] time series data for Germany, 1834-2012. Specifically, we use the variable *Mitglieder Börsenverein des Deutschen Buchhandels*in the file "K11_1_Kultur___Tourismus_und_Sport_-_Zeitungen_und_Zeitschriften.csv" from data file ZA8603 in the GESIS Data Archive.

We provide the number of daily newspaper titles per year in Table B.3 below.

Table B.3: No. book trade association members

| Year | No. members |
|---|---|
| 1834 | 454 |
| 1835 | 507 |
| 1836 | 570 |
| 1837 | 606 |
| 1838 | 618 |
| 1839 | 610 |
| 1840 | 708 |
| 1841 | 685 |
| 1842 | 690 |
| 1843 | 707 |
| 1844 | 718 |
| 1845 | 723 |
| 1846 | 726 |
| 1847 | 752 |
| 1848 | 742 |
| 1849 | 749 |
| 1850 | 673 |
| 1851 | 687 |
| 1852 | 682 |
| 1853 | 701 |
| 1854 | 703 |
| 1855 | 703 |
| 1856 | 734 |
| 1857 | 758 |
| 1858 | 755 |
| 1859 | 767 |
| 1860 | 775 |
| 1861 | 834 |

*Continued on next page...*

Table B.3 – Continued from previous page

| Year | No. members |
|---|---|
| 1862 | 831 |
| 1863 | 856 |
| 1864 | 885 |
| 1865 | 911 |
| 1866 | 918 |
| 1867 | 911 |
| 1868 | 940 |
| 1869 | 966 |
| 1870 | 1000 |
| 1871 | 1010 |
| 1872 | 1043 |
| 1873 | 1146 |
| 1874 | 1153 |
| 1875 | 1227 |
| 1876 | 1252 |
| 1877 | 1297 |
| 1878 | 1331 |
| 1879 | 1424 |
| 1880 | 1435 |
| 1881 | 1436 |
| 1882 | 1456 |
| 1883 | 1463 |
| 1884 | 1507 |
| 1885 | 1549 |
| 1886 | 1610 |
| 1887 | 1636 |
| 1888 | 1830 |
| 1889 | 2286 |
| 1890 | 2366 |

*Continued on next page...*

Table B.3 – Continued from previous page

| Year | No. members |
|---|---|
| 1891 | 2420 |
| 1892 | 2494 |
| 1893 | 2543 |
| 1894 | 2575 |
| 1895 | 2646 |
| 1896 | 2698 |
| 1897 | 2720 |
| 1898 | 2773 |
| 1899 | 2821 |
| 1900 | 2858 |
| 1901 | 2886 |
| 1902 | 3001 |
| 1903 | 3080 |
| 1904 | 3240 |
| 1905 | 3260 |
| 1906 | 3319 |
| 1907 | 3405 |
| 1908 | 3381 |
| 1909 | 3398 |
| 1910 | 3417 |
| 1911 | 3459 |
| 1912 | 3543 |
| 1913 | 3552 |
| 1914 | 3613 |
| 1915 | 3609 |
| 1916 | 3560 |
| 1917 | 3577 |
| 1918 | 3593 |
| 1919 | 3741 |

*Continued on next page...*

Table B.3 – Continued from previous page

| Year | No. members |
|---|---|
| 1920 | 4132 |
| 1921 | 4295 |
| 1922 | 4682 |
| 1923 | 4849 |
| 1924 | 4819 |
| 1925 | 4931 |
| 1926 | 4971 |
| 1927 | 5015 |
| 1928 | 5080 |
| 1929 | 4990 |
| 1930 | 4891 |
| 1931 | 4613 |
| 1932 | 4487 |

### B.3.5 Number of book traders in the mid-1800s

To provide a measure of the growth in book trade in the German Empire over the 19th century, we combine data on the number of book traders (including book dealers and book shops) from Fullerton [2015] and Rarisch [1976]. In particular, we manually transcribed data from the following tables in Fullerton [2015]:

- *Table 2.1 Book dealers by city 1822, 1832, 1842* (p. 46)
- *Table 2.2 Concentrations of book dealers by state [in 1843]* (p. 48)
- *Table 5.2 The density of book dealerships in 1855* (p. 98)

Fullerton [2015] created these tables using information from *Börsenblatt für den Deutschen Buchhandel* (*Journal for the German Book Trade*). Tables 2.1 and 2.2 was created using information

from *Börsenblatt für den Deutschen Buchhandel* (1843, p. 211), and Table 5.2 was created using information from *Börsenblatt für den Deutschen Buchhandel* (1855, p. 205).

We combine this with data from the following archived data table in Rarisch [1976], Section "A. The development of enterprises in the field of book production and trade" (*A. Die Entwicklung der Betriebe im Bereich der Buchproduktion und des Handels*):

- A.7 Number of book-shops in 14 cities important for the book-trade-sector for the years 1840, 1850, and 1860. (*A.7 Zahl der Buchhandlungen in 14 bedeutenden Städten im Gebiet des deutschen Buchhandels, 1840, 1850 und 1860.*) File "ZA8564_A-7.xls" (Tabelle 11, S. 57).

This final dataset provides an estimate of the number of book traders in major German cities for the years 1822, 1832, 1840, 1842, 1850, and 1960. The annual figures and their respective sources are provided in Table B.4 below.

Table B.4: Number of book dealers and traders

| Year | City | No. traders | Source |
|---|---|---|---|
| 1822 | Berlin | 51 | Fullerton |
| 1822 | Leipzig | 59 | Fullerton |
| 1822 | Stuttgart | 5 | Fullerton |
| 1822 | Dresden | 4 | Fullerton |
| 1822 | Frankfurt a.M. | 18 | Fullerton |
| 1822 | Hamburg | 7 | Fullerton |
| 1822 | München | 8 | Fullerton |
| 1822 | Nürnberg | 19 | Fullerton |
| 1822 | Köln | 5 | Fullerton |
| 1822 | Augsburg | 5 | Fullerton |
| 1822 | Halle | 12 | Fullerton |
| 1832 | Berlin | 81 | Fullerton |
| 1832 | Leipzig | 79 | Fullerton |
| 1832 | Stuttgart | 14 | Fullerton |

*Continued on next page...*

Table B.4 – Continued from previous page

| Year | City | No. traders | Source |
|---|---|---|---|
| 1832 | Dresden | 15 | Fullerton |
| 1832 | Frankfurt a.M. | 38 | Fullerton |
| 1832 | Hamburg | 17 | Fullerton |
| 1832 | München | 15 | Fullerton |
| 1832 | Nürnberg | 26 | Fullerton |
| 1832 | Köln | 9 | Fullerton |
| 1832 | Augsburg | 20 | Fullerton |
| 1832 | Halle | 14 | Fullerton |
| 1840 | Berlin | 108 | Rarisch |
| 1840 | Leipzig | 113 | Rarisch |
| 1840 | Stuttgart | 30 | Rarisch |
| 1840 | Dresden | 25 | Rarisch |
| 1840 | Frankfurt a.M. | 35 | Rarisch |
| 1840 | Hamburg | 22 | Rarisch |
| 1840 | München | 22 | Rarisch |
| 1840 | Nürnberg | 26 | Rarisch |
| 1840 | Köln | 19 | Rarisch |
| 1840 | Augsburg | 16 | Rarisch |
| 1840 | Halle | 12 | Rarisch |
| 1842 | Berlin | 122 | Fullerton |
| 1842 | Leipzig | 140 | Fullerton |
| 1842 | Stuttgart | 41 | Fullerton |
| 1842 | Dresden | 42 | Fullerton |
| 1842 | Frankfurt a.M. | 33 | Fullerton |
| 1842 | Hamburg | 27 | Fullerton |
| 1842 | München | 30 | Fullerton |
| 1842 | Nürnberg | 30 | Fullerton |
| 1842 | Köln | 22 | Fullerton |
| 1842 | Augsburg | 16 | Fullerton |

*Continued on next page...*

Table B.4 – Continued from previous page

| Year | City | No. traders | Source |
|---|---|---|---|
| 1842 | Halle | 16 | Fullerton |
| 1850 | Berlin | 172 | Rarisch |
| 1850 | Leipzig | 133 | Rarisch |
| 1850 | Stuttgart | 50 | Rarisch |
| 1850 | Dresden | 32 | Rarisch |
| 1850 | Frankfurt a.M. | 34 | Rarisch |
| 1850 | Hamburg | 32 | Rarisch |
| 1850 | München | 23 | Rarisch |
| 1850 | Nürnberg | 25 | Rarisch |
| 1850 | Köln | 21 | Rarisch |
| 1850 | Augsburg | 17 | Rarisch |
| 1850 | Halle | 23 | Rarisch |
| 1860 | Berlin | 229 | Rarisch |
| 1860 | Leipzig | 188 | Rarisch |
| 1860 | Stuttgart | 66 | Rarisch |
| 1860 | Dresden | 48 | Rarisch |
| 1860 | Frankfurt a.M. | 48 | Rarisch |
| 1860 | Hamburg | 46 | Rarisch |
| 1860 | München | 35 | Rarisch |
| 1860 | Nürnberg | 35 | Rarisch |
| 1860 | Köln | 29 | Rarisch |
| 1860 | Augsburg | 24 | Rarisch |
| 1860 | Halle | 19.00 | Rarisch |

### B.3.6 Publishing houses

In order to provide a measure of the growth in the publishing industry over the 19th century, we recorded data on the founding year and location of literary publishing houses that were active in

the German Empire around 1900. This data was manually transcribed by Lukas Kuld from a list of such publishers on pp. 169-171 of Estermann and Füssel [2013]. More comprehensive annual data on all publishing houses that were founded from 1700 to 1932 is not available. However, we believe this measure provides an indication of where the publishing industry was the strongest and where publishing houses remained active the longest. We provide the list of cities and years in which a publishing house was founded in Table B.5 below.

Table B.5: Publishing houses

| Year | City |
|---|---|
| 1659 | Stuttgart |
| 1659 | Tubingen |
| 1719 | Leipzig |
| 1789 | Berlin |
| 1794 | Leipzig |
| 1804 | Heidelberg |
| 1806 | Berlin |
| 1811 | Leipzig |
| 1822 | Heidelberg |
| 1828 | Leipzig |
| 1831 | Stuttgart |
| 1833 | Minden |
| 1834 | Leipzig |
| 1837 | Leipzig |
| 1838 | Braunschweig |
| 1844 | Frankfurt |
| 1845 | Breslau |
| 1847 | Berlin |
| 1848 | Leipzig |
| 1848 | Stuttgart |
| 1849 | Hamm |
| 1850 | Berlin |

*Continued on next page...*

Table B.5 – Continued from previous page

| Year | City |
|---|---|
| 1850 | Leipzig |
| 1854 | Leipzig |
| 1869 | Leipzig |
| 1871 | Berlin |
| 1872 | Dresden |
| 1872 | Leipzig |
| 1873 | Berlin |
| 1873 | Leipzig |
| 1873 | Stuttgart |
| 1876 | Stuttgart |
| 1877 | Berlin |
| 1877 | Berlin |
| 1877 | Leipzig |
| 1878 | Leipzig |
| 1880 | Dresden |
| 1880 | Leipzig |
| 1881 | Stuttgart |
| 1884 | Munich |
| 1886 | Berlin |
| 1886 | Berlin |
| 1886 | Berlin |
| 1886 | Breslau |
| 1888 | Berlin |
| 1890 | Stuttgart |
| 1891 | Leipzig |
| 1893 | Berlin |
| 1894 | Berlin |
| 1894 | Leipzig |
| 1894 | Leipzig |

*Continued on next page...*

Table B.5 – Continued from previous page

| Year | City |
|---|---|
| 1895 | Berlin |
| 1895 | Berlin |
| 1895 | Dresden |
| 1895 | Leipzig |
| 1896 | Berlin |
| 1898 | Berlin |
| 1899 | Leipzig |
| 1900 | Berlin |
| 1901 | Berlin |
| 1901 | Leipzig |
| 1903 | Berlin |
| 1903 | Munich |
| 1904 | Leipzig |
| 1904 | Munich |
| 1904 | Stuttgart |
| 1905 | Berlin |
| 1905 | Munich |
| 1909 | Berlin |
| 1910 | Leipzig |
| 1913 | Leipzig |

### B.3.7 Most populous cities, 1700-1932

We combine the Bairoch et al. [1988] and Reba et al. [2018] (from the Reba et al. [2016] paper) historic urban population data to determine the 10 largest cities in terms of general population on annual basis from 1700-1932. We first identified cities within the borders of the 1910 German Empire. Next, we averaged the annual city population estimates from these two sources as the lists do not fully overlap. Next, we identified the 10 largest cities for each year by interpolating

our observations. For the analysis, we created a dummy variable indicating whether an author lived in one of 10 largest cities in a given year. We summarize the cities and then years in which the respective city was one of the 10 largest cities in Table B.6 below.

Table B.6: Most populous cities

| City | Years in Top 10 |
|---|---|
| Bremen | 1700 |
| Strasbourg | 1700-1854 |
| Dresden | 1700-1932 |
| Frankfurt | 1700-1932 |
| Kaliningrad (Konigsberg) | 1700-1862 |
| Cologne | 1700-1932 |
| Nurnberg | 1700-1727, 1890, 1910-1920 |
| Wroclaw (Breslau) | 1700-1931 |
| Berlin | 1700-1932 |
| Gdańsk (Danzig) | 1700-1789, 1802-1841, 1850 |
| Hamburg | 1700-1932 |
| Munich | 1728-1748, 1790-1849, 1851-1932 |
| Leipzig | 1749-1755, 1855-1932 |
| Metz | 1756-1801 |
| Wuppertal | 1842-1902 |
| Magdeburg | 1863 |
| Hanover | 1864-1899, 1903-1909, 1914 |
| Ruhr | 1900-1932 |
| Dusseldorf | 1921-1924, 1932 |
| Katowice | 1925-1932 |

### B.3.8 University cities

We define a university as a university city if the city was home to a university that was active for at least some period of time during the sample period (1700-1932). We used a list of universities in Germany from Wikipedia [2021] to obtain a list of university names and hyperlinks to the respective university Wikipedia pages. Data on the active periods was only available in unstructured form, so the years in which the university was active were manually transcribed into a CSV file. This data transcription was completed by our research assistant Jens Kirsten under the supervision of Lukas Kuld in autumn 2021.

In addition, we searched all locations in our author data for a university. This is still not an exhaustive list of German cities with a university. We only identify cities either in modern-day Germany or in the author dataset as having been home to a university for some period of time. We exclude cities where a university opened after 1932 or closed before 1700. If, for example, a

university was active before 1700, closed temporarily, and reopened after 1700, we only include the active periods between 1700-1932 in the list. For the analyses, we limit our definition to those cities that have an active university but are not included in the list of large cities or capital cities. Therefore this category mostly includes to the traditional university towns such as Heidelberg, Tübingen, Jena, and Göttingen.

The final list of cities and the years in which the university was active is provided in Table B.7 below (includes large and capital cities).

Table B.7: List of university cities before 1932

| City | Years active |
|---|---|
| Aachen | 1870-2021 |
| Altdorf | 1622-1809 |
| Aschaffenburg | 1808-1818 |
| Bamberg | 1770-1803 |
| Berlin | 1810-2021 |
| Bonn | 1786-1798 & 1818-2021 |
| Bützow | 1760-1789 |
| Cologne | 1388-1798 & 1919-2021 |
| Dillingen | 1553-1803 |
| Duisburg | 1654-1818 |
| Erfurt | 1379-1816 |
| Erlangen | 1743-2021 |
| Frankfurt | 1914-2021 |
| Frankfurt (Oder) | 1538-1811 |
| Freiburg im Breisgau | 1684-1713 & 1715-2021 |
| Fulda | 1734-1756 & 1763-1805 |
| Giessen | 1607-2021 |
| Göttingen | 1737-2021 |
| Greifswald | 1456-2021 |
| Halle an der Saale | 1693-2021 |
| Hamburg | 1919-2021 |

*Continued on next page...*

Table B.7 – Continued from previous page

| City | Year(s) |
|---|---|
| Heidelberg | 1700-2021 |
| Helmstedt | 1575-1810 |
| Ingolstadt | 1459-1800 |
| Jena | 1557-2021 |
| Kaliningrad (Konigsberg) | 1544-1945 |
| Kiel | 1652-2021 |
| Konstanz | 1713-1715 |
| Landshut | 1800-1826 |
| Leipzig | 1409-2021 |
| Mainz | 1476-1798 |
| Marburg | 1527-2021 |
| Munich | 1826-2021 |
| Munster | 1780-1818 & 1902-2021 |
| Osnabruck | 1794-2021 |
| Paderborn | 1614-1818 |
| Poznan (Posen) | 1611-1773 & 1919-1939 |
| Rinteln | 1619-1810 |
| Rostock | 1488-2021 |
| Strasbourg | 1621-2021 |
| Stuttgart | 1781-1794 & 1900-2021 |
| Trier | 1454-1798 |
| Tubingen | 1476-2021 |
| Wittenberg | 1502-1814 |
| Wroclaw (Breslau) | 1702-2021 |
| Würzburg | 1582-2021 |

### B.3.9 List of capitals and independent cities

Data on capital cities, independent cities, and the years in which the city had the respective status was collected from Wikipedia. We identified the Wikipedia pages of all German cities in the author panel dataset, and then we recorded the years in which the respective city was a capital or an independent city (if relevant) based on the information in both the structured and unstructured sections of the Wikipedia entry. The data was manually transcribed by the research assistant Jens Kirsten in December 2021. We only include cities in which at least one author in the panel dataset resided for at least one year. Therefore, this is not an exhaustive list of capitals and independent cities in Germany. We distinguish capitals of territorial states from independent imperial cites and city states.

The final list of cities and the years in which the respective city was a capital or independent city is provided in Table B.8 below.

Table B.8: List of capitals and independent cities

| City | Years as capital | Years independent |
|---|---|---|
| Aalen | | 1700-1803 |
| Ansbach | 1700-1791 | |
| Arnstadt | 1709-1716 | |
| Aschaffenburg | 1797-1810 | |
| Augsburg | | 1700-1805 |
| Baden-Baden | 1700-1705 | |
| Bamberg | 1700-1802 | |
| Bayreuth | 1700-1791 | |
| Berlin* | 1701-1932 | |
| Biberach | | 1700-1807 |
| Bonn | 1700-1794 | |
| Bonnigheim | | 1700-1750 |
| Braunschweig* | 1753-1807 & 1814-1918 | |
| Bremen | | 1700-1806 |
| Bruchsal | 1723-1803 | |

*Continued on next page...*

Table B.8 – Continued from previous page

| City | Capital | Independent City |
|---|---|---|
| Buckeburg* | 1700-1918 | |
| Celle | 1700-1705 | |
| Cologne | | 1700-1794 |
| Darmstadt* | 1700-1919 | |
| Dessau* | 1700-1832 | |
| Detmold | 1700-1806 | |
| Donaueschingen | 1716-1801 | |
| Dresden | 1700-1806 | |
| Dusseldorf | 1700-1716 | |
| Eisenach | 1741-1757 | |
| Essen | | 1700-1803 |
| Eutin | 1700-1803 | |
| Frankfurt* | 1810-1813 | 1700-1806 & 1813-1866 |
| Freiburg im Breisgau | 1700-1803 | |
| Fulda | 1700-1803 | |
| Gdańsk (Danzig)* | | 1807-1814 & 1920-1939 |
| Goslar | | 1700-1802 |
| Graudenz | 1700-1772 | |
| Halle an der Saale | 1700-1714 | |
| Hamburg | | 1700-1806 |
| Hanover* | 1700-1714 & 1837-1866 | |
| Hechingen* | 1700-1850 | |
| Heidelberg | 1700-1720 | |
| Heilbronn | | 1700-1802 |
| Hildburghausen* | 1700-1826 | |
| Hoxter | 1792-1803 | |
| Karlsruhe* | 1718-1918 | |
| Kassel | 1700-1815 | |
| Kaufbeuren | | 1700-1803 |

*Continued on next page...*

Table B.8 – Continued from previous page

| City | Capital | Independent City |
|---|---|---|
| Koblenz | 1700-1794 | |
| Lübeck* | | 1700-1811 & 1813-1937 |
| Magdeburg | 1714-1807 | |
| Mainz | 1700-1792 | |
| Mannheim | 1720-1778 | |
| Meersburg | 1700-1803 | |
| Meiningen | 1700-1871 | |
| Munich* | 1700-1919 | |
| Munster | 1700-1803 | |
| Neubrandenburg | 1774-1794 | |
| Neuwied | 1700-1806 | |
| Nordhausen | | 1700-1802 |
| Nordlingen | | 1700-1803 |
| Nuremberg | | 1700-1806 |
| Nurnberg | | 1700-1806 |
| Obersontheim | 1700-1713 | |
| Osnabruck | 1700-1802 | |
| Passau | 1700-1802 | |
| Pirmasens | 1741-1790 | |
| Potsdam | 1700-1815 | |
| Ravensburg | | 1700-1803 |
| Regensburg | | 1700-1803 |
| Reutlingen | | 1700-1803 |
| Schleswig | 1700-1721 | |
| Schweinfurt | | 1700-1803 |
| Schwerin* | 1700-1918 | |
| Stuttgart* | 1700-1718 & 1733-1918 | |
| Überlingen | | 1700-1803 |
| Ulm | | 1700-1803 |

*Continued on next page...*

Table B.8 – Continued from previous page

| City | Capital | Independent City |
|---|---|---|
| Weimar* | 1700-1918 | |
| Wiesbaden* | 1734-1866 | |
| Wolfenbuttel | 1700-1753 | |
| Wroclaw (Breslau) | 1700-1810 | |
| Würzburg | 1700-1803 & 1806-1814 | |
| Zweibrücken | 1700-1793 | |

*Indicates that the respective city retained its status as a capital and/or independent city for at least one year after the defeat of Napoleon in 1815.