

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## Regional Science and Urban Economics

journal homepage: [www.elsevier.com/locate/regec](http://www.elsevier.com/locate/regec)

# What makes an artist? The evolution and clustering of creative activity in the US since 1850<sup>☆</sup>



Karol Jan Borowiecki<sup>a,\*</sup>, Christian Møller Dahl<sup>b</sup>

<sup>a</sup> Historical Economics and Development Group (HEDG), Department of Business and Economics, University of Southern Denmark, Denmark

<sup>b</sup> Department of Business and Economics, University of Southern Denmark, Denmark

## ARTICLE INFO

JEL classification:

R1  
N33  
Z11

Keywords:

Creativity  
Artists  
Geographic clustering  
Agglomeration economies  
Urban history

## ABSTRACT

This research illuminates the historical development and clustering of creative activity in the United States. Census data is used to identify creative occupations (i.e., artists, musicians, authors, actors) and data on prominent creatives, as listed in a comprehensive biographical compendium. The analysis first sheds light on the socio-economic background of creative people and how it has changed since 1850. The results indicate that the proportion of female creatives is relatively high, time constraints can be a hindrance for taking up a creative occupation, racial inequality is present and tends to change only slowly, and access to financial resources within a family facilitates the uptake of an artistic occupation. Second, the study systematically documents and quantifies the geography of creative clusters in the United States and explains how these have evolved over time and across creative domains.

*“When I say artist I mean the one who is building things ... some with a brush – some with a shovel – some choose a pen.”* Jackson Pollock

## 1. Introduction

Industrial activity tends to cluster in certain locations, which leads to agglomeration economies (e.g., [Ellison and Glaeser, 1997](#)). Geographic clustering is particularly important for the cultural and creative sectors ([UNCTAD, 2018](#); [Borowiecki, 2015a](#)), among other reasons, because the production of creative output is particularly dependent on inter-human interactions and tacit knowledge sharing ([Borowiecki, 2013](#)). Nowadays, these interactions are facilitated in one of the many large artistic clusters in the United States (e.g., [O’Hagan and Borowiecki, 2010](#); [O’Hagan and Hellmanzik, 2008](#)). However, little is known on how or when the clustering of these activities has emerged

in the US and how is it interdependent on the different domains of the arts. This paper fills these knowledge gaps by drawing on novel historical data covering the periods from 1850 to 2010, and by focusing on several different creative domains. By doing so, we come close to documenting the *evolution* of artistic creativity.

Historical population growth in the United States was fast, from 23 million in 1850, 76 million in 1900, 151 million in 1950, to 309 million in 2010 ([Gibson and Jung, 2005](#)). Economic growth in terms of overall GDP was equally rapid and the resulting wealth increases benefited the whole society. This coincided with the emergence of a relatively rich group of geographically concentrated people who patronized the arts and supported the construction of concert halls, art galleries, and other types of cultural infrastructure ([Lovell, 2007](#); [Jovanovich and Renn, 2019](#)), as well as the less rich who are willing to pay taxes to support art teachers in schools. Along with rising income and improving education, the demand for art and culture shifted over the course of the late 19th and 20th centuries ([Galenson and Weinberg, 2001](#)). The change has

<sup>☆</sup> This paper has been disseminated previously under the title “The Origins of Creativity: The Case of the Arts in the United States since 1850”. The authors wish to thank Philipp Ager, David de la Croix, Walker Hanlon, Nathan Nunn, Casper Worm Hansen, and participants at the World Economic History Congress (Kyoto), Interdisciplinary Conference on Art Markets (Amsterdam), Workshop on Growth, History and Development (Odense), Economic History and Economic Policy Conference (Paris), FRESH Meeting (Odense), European Workshop on Applied Cultural Economics (Vienna), SOUND Economic History Workshop (Lund) and invited seminar at Syddansk Universitet Sønderborg for helpful consultations and insightful comments. Outstanding research assistance was provided by Martin Hørlyk Kristensen.

\* Corresponding author.

E-mail address: [kjb@sam.sdu.dk](mailto:kjb@sam.sdu.dk) (K.J. Borowiecki).

<https://doi.org/10.1016/j.regsciurbeco.2020.103614>

Received 5 October 2019; Received in revised form 27 October 2020; Accepted 30 October 2020

Available online 19 November 2020

0166-0462/© 2020 Elsevier B.V. All rights reserved.



**Fig. 1.** Creatives and Startup Densities (per 1000 people). Sources: IPUMS (2015), Fairlie et al. (2015). Note: Creatives include artists, musicians, authors, and actors.

been visible in numerous dimensions and in the exploration that follows we will document in particular a rapid increase in the number of artists and the rise in the diversity of artistic occupations.

High skilled people generate ideas (Glaeser, 2003) and their presence is correlated with city growth (e.g., Glaeser, 1994; Gergaud et al., 2016). Urban development comes from being an attractive “consumer city” for high skill people (Glaeser et al., 2001) and the presence of creative people, in particular of artists, may enhance this attractiveness (e.g., Florida, 2002). Creative and arts sectors are seen as “the key ingredient for job creation, innovation and trade” (UNCTAD, 2010) and are believed to constitute opportunities for developing countries to leapfrog into emerging high-growth areas of the world economy. Creativity is potentially “driving the economy, reshaping entire industries and stimulating inclusive growth” (OECD, 2014). As can be seen in Fig. 1, the correlation between the density of creative people and startup activity in US cities is surprisingly strong, also in the long run.<sup>1</sup> Finally, artistic creativity is needed to produce cultural goods and services, and the advantages of having a wealthy cultural supply and a meaningful cultural heritage nowadays are vast and non-negligible, ranging from economic gains from tourism inflows to non-monetary gains arising from a common identity.

Despite the remarkable importance of artistic creativity, economists have largely refrained from studying it. Creativity in the arts is explored even less by economic historians, especially since creativity is not a characteristic that has been priced highly on labor markets throughout most of history. In the past, employers have valued disciplined and hard-working workers (Crafts, 1985), as opposed to creative ones. Our focus is directed on the arts, which incorporate some of the earliest creative occupations. Therefore, this study may provide some unique insights on the long-term development, geographic spread, and individ-

<sup>1</sup> Even if simple scatter plots are subject to biases due to unobservable factors (e.g., educational attainment), the strength of the disclosed correlation is rather striking. Causal claims cannot be made, but improved opportunities to consume cultural goods such as theater (Glaeser et al., 2001) or the presence of cultural amenities (Falck et al., 2018) are an important factor in the location decision of high-skilled workers - and are likely related with the density of creatives. See also Glaeser et al. (2010) for a summary of the local causes and consequences of entrepreneurship.

ual motivations to engage in creative activity.<sup>2</sup>

This research is mostly descriptive and does not support any causal claims. Nonetheless, it makes three main contributions. First, we shed light on the determinants of a person engaging in an artistic occupation by using individual-level data. These analyses illuminate how the socio-economic profiles of artists have changed since 1850, thereby delivering unique insights into long-term patterns of artistic occupational choices. The focus is directed on individuals representing the visual arts, literary arts and performing arts, and music. The chosen categories are in line with the definition of traditional high art by Heilbrun and Gray (2001). Furthermore, by covering various types of creatives the data are suitable for identifying differences across artistic domains. Previously, the focus of economic history of the art has been directed at small samples of established artists, as these are typically observable today (e.g., Graddy, 2013; O'Hagan and Hellmanzik, 2008; Etro and Pagani, 2012; Borowiecki, 2017). The underlying study employs census data, which enables us to focus on the average artists instead.

The second contribution of the paper is the exploration of the dynamics associated with the historical development of US creative clusters. These investigations document and quantify the importance of specific geographic locations (cities) by including the census data along with data covering famous creatives who worked in the United States, as listed in the comprehensive biographical compendium - the Index Bio-bibliographicus Notorum Hominum (IBN). This allows important insights into the differences in the place and intensity of clustering between prominent and average creatives. Furthermore, it is possible to highlight differences in the geography across creative domains.

While the importance of certain cities within specific artistic domains is established, little is known about the clustering intensity and interrelations across different domains. Previous research built on heterogeneous samples, different data sources and time periods, which constrains the comparison of the relative importance and interplay across creative domains. Are some cities simply more cultured and attract artists from a range of different domains, or are there rather some forces in play that lead to specialization of artistic talent? This important question is not sufficiently well answered in the existing literature, and yet it is of relevance to our understanding of the clustering of creative activity and the benefits associated with agglomeration economies.

The third contribution is an analysis of how important artists affect local creative activity. In particular, it is explored how the presence of famous artists in a place are related to the probability of an individual becoming involved in a creative occupation. By making use of census data that reflect the creative involvement of average individuals, as opposed to prominent ones, it is possible to overcome the extreme non-random sample selection biases encountered in the fast growing related literature (e.g., O'Hagan and Borowiecki, 2010).

## 2. Literature review

This research primarily relates to four strands of the literature. First, this paper builds on literature on human capital theory of city growth by analysing how human capital accumulates across cities and over time. The focus on creative people only comes close to what Florida (2002) describes as the creative class, but the paper in its essence is about human capital formation in agglomerations (Glaeser, 2004). Natural advantage plays a role in the formation of agglomerations and firms show a higher clustering intensity than they would be if they were

<sup>2</sup> In the remainder of the paper by referring to creativity or creative people, what is meant is artistic creativity and artists, albeit artistic creativity is likely correlated with the same variables as is the sort of creativity that leads to innovation in production and economic growth. For example, within psychology, artistic, scientific or entrepreneurial creativity is studied along side each other (e.g., Ludwig, 1995), albeit in different contexts.

randomly located (Ellison and Glaeser, 1997; Duranton and Overman, 2005). New economic geography models suggest that transportation costs induce people and firms to cluster in cities (Fujita et al., 1999). Cities and industries that cluster are associated with higher wages and prices (Rosenthal and Strange, 2004; Puga, 2010), and hence clusters must experience some kind of advantages to be competitive (Ellison and Glaeser, 1999). These advantages take different forms in different times. For example, at the turn of the 20th century it may have mattered to be close to a coal mine or the harbor, but is not anymore nowadays. Similarly, advantageous climatic conditions have likely influenced the location for the movie industry in Hollywood, but possibly play a smaller role nowadays.

The second related strand focuses on the geographic concentration of artistic activity. Artists exhibit remarkable clustering patterns, both in terms of birth and migration. The predominant location for visual artists born in the first half of the 20th century is New York City, with all prominent American artists clustering there (O'Hagan and Hellmanzik, 2008). New York is also a major work location for music composers: It is the fifth most important city for composers born in the 19th century and, after Paris, the second most popular destination for 20th century composers (O'Hagan and Jan Borowiecki, 2010; Borowiecki and O'Hagan, 2012). Globally, Paris was the predominant music center over a remarkably long period of around four centuries and this was due to its large size (Borowiecki, 2015a). However, the incidence of the second World War caused a significant negative shock that led to massive relocations from the French capital to New York City. Borowiecki and Graddy (2018) find that US cities that experience larger immigrant artist inflows also see a greater inflow of native artists. In a more contemporary setting, Andersen (2014) show that location choices of artists are driven by access to other artists and by consumer demand.

Geographic clustering of creative activity is also associated with strong productivity gains (e.g., Tao, 2019). It has been shown that literary artists born between 1750 and 1925 experienced significant productivity gains when working in London, the predominant cluster within literary arts (Mitchell, 2019). Visual artists born 1850 to 1945 peaked earlier in the geographic clusters of Paris and New York (Hellmanzik, 2010). Music composers born in the late 18th and 19th centuries were more productive in the main hubs for music (Borowiecki, 2013) and the benefits increased with the peer group size at a decreasing rate (Borowiecki, 2015a). The focus in these studies is, however, on selected, specific artistic domains, and while interdisciplinary spillovers are sometimes acknowledged (e.g., Borowiecki, 2013, describes how composers in Paris have been in contact with literary and visual artists), little is known about how these domains interrelate or interlocate.

Third, this research relates to the literature on the attractiveness of cities. Urban development takes particularly place in cities that facilitate consumption for high skill people, and hence a "consumer city" is thriving more (Glaeser et al., 2001). The presence of cultural amenities facilitates such consumption and hence attracts high-skilled workers (e.g., Falck et al., 2018), which then in turn leads to various positive spillovers. Clustering of creative activity in cities may also have long-term effects on economic development, including increased innovation activity in certain areas (Knudsen et al., 2008) or persistent income differences across regions (Maloney and Caicedo (2014). The underlying research does not explicitly address the question of the consequences of cultural activity. Nonetheless, it outlines the dynamics and geography of emerging or established creative clusters, and hence provides important insights into the attractiveness of cities. These insights may have some relevance considering how cultural developments are correlated with start-up activity (see Fig. 1) and the need for further research on the local causes and consequences of entrepreneurship (Glaeser et al., 2010).

Fourth, the underlying study connects with research on the works and lives of artists. Influential overviews of artists' labor markets and research thereof are presented by Benhamou (2011) and Alper and Wassal (2006). Graddy (2013) explores the accuracy and persistency

of rankings of famous artists conducted by the 17th century art critic Roger de Piles. Etro and Pagani (2012) study contracts between patrons and successful artists in the 17th century market for figurative paintings in Italy. Few studies look explicitly at US artists. Galenson and Weinberg (2000) show how American artists born 1870 to 1940 introduced innovation into their art and how it influenced their careers. Alper and Wassal (1998) explore the determinants of persistence in artistic occupations, using the 1970 US Census. Alper and Wassal (2006) study employment and earnings of American artists using decennial US Census data from 1940 to 2000. Relative to this strand, we are able to explore data for a period of an unprecedented length covering 170 years and to shed light on the socio-economic background of average artists as opposed to the famous achiever. Blanchflower and Oswald (1998) have conducted an influential study on what determinants enhance people to become entrepreneurs, viewed contemporaneously as "popular heroes" (Blanchflower and Oswald, 1998, opening quote). About 20 years later the creative artist is increasingly often regarded as the new champion and even though the traits required by entrepreneurs and artists are in many regards similar, a study of what makes an artist becomes particularly meaningful.

### 3. Data

There are two main databases used here. First, we use US census data from the Integrated Public Use Microdata Series database IPUMS (2015). This comprehensive decennial population census was first undertaken in 1790, but only provides information from 1850 on occupational status (OCC1950), along with a wide array of background variables. The occupation variable is used in order to identify the following creatives: Artists and art teachers, authors, musicians and music teachers, and actors and actresses.<sup>3</sup> The occupation categories of artists and musicians also include the teachers within these domains. This is not necessarily a shortcoming, since also the presence of art teachers (be it in music or visual arts) serves as a proxy for creative activity. In places with more art teachers, more artists are educated and a higher population of artists can be expected, which in consequence leads to greater artistic activity and creativity (Ateca-Amestoy et al., 2017). The possible overrepresentation of teachers is nonetheless addressed econometrically in the results section, where we also control for the geographic spread of teachers.

The social setting or technical content of artistic occupations has likely changed over a period of one and a half centuries. Nonetheless, larger occupational groupings are claimed to constitute relatively accurate and consistent indicators of social status and general function (IPUMS, 2017). Hence, in our case at the very least the overall group of creative occupations should be useful in reflecting the trends studied. Furthermore, it is argued that the "specific occupational titles are less subject to meaningful change than common historical wisdom would suggest" (IPUMS, 2017). In any case, the harmonized occupation series used should work well as a means of locating individuals in the occupational structure as well as in geographic space and this as far back as the late nineteenth century.

The IPUMS provides 1%-samples for each of the available decades ("1-in-100 national random sample"), which are commonly used and also chosen for the underlying study. Furthermore, the focus of the analysis is on household heads. Focusing on household heads is important as they were typically the decision makers, and hence also took

<sup>3</sup> The records also make it possible to identify architects, dancers and dancing teachers, and editors and reporters. However, these occupations are very rare and deliver an insufficiently low number of observations, especially in the earlier periods. In analogy to creative occupations, we will refer to non-creative occupations when writing about occupations other than those of artists, musicians, authors, and actors. This is done merely for convenience, as opposed to the categorization of occupations into those that require or do not require creative talent.

**Table 1**  
Socio-economic background of creatives.

	<i>Dependent variable:</i>			
	Creatives LPM (1)	Creatives LPM (from 1950) (2)	Creatives LPM (3)	Creatives LPM (4)
female	0.231*** (0.014)	0.252*** (0.015)	-0.017 (0.027)	0.231*** (0.014)
age	0.031*** (0.001)	0.026*** (0.002)	0.035*** (0.003)	0.031*** (0.001)
age <sup>2</sup>	-0.0004*** (0.00001)	-0.0003*** (0.00002)	-0.0005*** (0.00003)	-0.0004*** (0.00001)
married	-0.302*** (0.019)	-0.298*** (0.020)	-0.517*** (0.034)	-0.302*** (0.019)
separated	-0.498*** (0.030)	-0.322*** (0.031)	-0.607*** (0.055)	-0.496*** (0.030)
divorced	-0.299*** (0.023)	-0.244*** (0.024)	-0.272*** (0.045)	-0.297*** (0.023)
widowed	-0.639*** (0.021)	-0.484*** (0.023)	-0.672*** (0.036)	-0.638*** (0.021)
family size	-0.091*** (0.005)	-0.069*** (0.007)	-0.094*** (0.007)	-0.092*** (0.005)
number children	-0.012** (0.006)	-0.018** (0.008)	-0.038*** (0.009)	-0.013** (0.006)
black	-0.347*** (0.013)	-0.185*** (0.014)	-0.606*** (0.021)	-0.343*** (0.013)
native	-0.127 (0.078)	0.089 (0.079)	-0.324** (0.163)	-0.127 (0.078)
asian	-0.365*** (0.029)	-0.428*** (0.030)	-0.744*** (0.054)	-0.377*** (0.029)
other	-0.635*** (0.026)	-0.117*** (0.026)	-1.114*** (0.045)	-0.650*** (0.026)
mixed	-0.040 (0.065)	0.129** (0.065)	-0.222 (0.142)	-0.053 (0.065)
earnings		-0.004*** (0.001)		
family income		0.008*** (0.003)		
education		0.193*** (0.002)		
migrant			0.245*** (0.017)	
log(city population)			0.109*** (0.005)	
Constant	0.251 (0.241)	-0.781*** (0.059)	-0.685*** (0.255)	0.450 (0.446)
Year FE?	Yes	Yes	Yes	Yes
State FE?	Yes	Yes	Yes	Yes
Year × State FE?	No	No	No	Yes
Observations	3,925,299	3,086,651	1,331,685	3,925,299
R <sup>2</sup>	0.003	0.006	0.005	0.003
Adjusted R <sup>2</sup>	0.003	0.006	0.005	0.003
Residual Std. Error	0.084 (df = 3925204)	0.087 (df = 3086563)	0.088 (df = 1331598)	0.084 (df = 3924576)
F Statistic	117.567*** (df = 94; 3925204)	203.027*** (df = 87; 3086563)	85.622*** (df = 86; 1331598)	16.215*** (df = 722; 3924576)

Note: All coefficients (and the associated s.e.) are multiplied by 10<sup>2</sup> to ease readability. \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01.

the decisions about location choice. This is crucial for an unbiased geographic analysis. Furthermore, narrowing down the analysis to household heads, who are typically the main breadwinners, filters out meaningful artistic occupations.

Given the particularly long time period covered here it is natural that some Census questions or universe have changed. It is important to note that these decisions about the design of each Census have been made quite certainly independently from developments in the labor market of creatives. Furthermore, the changes go in either direction (e.g., slight increases or decreases in the cut-off age). Of course, the census changes may still lead to biased estimates, however – given the long time-period covered – they should not be very meaningful on average. Any subjective judgments and data interference have been avoided

in the underlying paper as far as only possible and reasonable, however during pilot studies various data restrictions and censoring have been conducted, and no meaningful differences have arisen.

Appendix Fig. A1 visualizes the fast growth in the share of creatives across the domains covered. Over the second half of the 19th century the share of creative increases to about one in 1/000 respondents. The share of musicians is especially high and takes off during late 20th century until it reaches more than 0.5 percent. The other creative occupations fluctuate in a range between 0.1 and 0.15 percent. The higher uptake of these occupations by the population reflects the gradual increase in the popularity of creative occupations. Furthermore, from a technical point of view, it can be observed that the number of observations is sufficient for a quantitative analysis even for the earliest

census editions. However, one has to bear in mind a potentially higher volatility over those years.

The second database used is the comprehensive biographical compendium - the Index Bio-bibliographicus Notorum Hominum (IBN). IBN is aimed to facilitate the global research community easy access to existing biographical sources. The information in the IBN was compiled from around 3/000 biographical sources (mainly dictionaries and encyclopedias) covering almost all countries and historical periods.<sup>4</sup> The compendium lists 56/657 creatives (out of approx. 298/000 famous people) and provides for each person the name, birth and death places and a brief description of the individual's background. This allows us to filter out a set of 2/421 prominent creatives who were born or died in the US, which includes visual artists (1/060), authors (782), musicians (470), and actors (300). Appendix Fig. A2 shows the number of deaths of famous IBN creatives that occurred during the decade preceding a given census edition.

#### 4. Results

##### 4.1. The socio-economic background of creatives since 1850

This section presents the socio-economic background of the creatives covered by applying a simple regression model as summarized by Equation (1). In particular, a linear probability model (LPM) is estimated to explore how being involved in any of the four creative occupations covered is related to the background of an individual. The LPM is specified as follows:

$$y_{it} = \alpha + X_{it}\beta + \sum_{year} f_{year,i} + \sum_{state} f_{state,i} + u_{it}, \tag{1}$$

for  $t = 1, \dots, T$  and  $i = 1, \dots, N$ . The dependent variable in Equation (1) given by  $y_{it}$  is a binary variable indicating if individual  $i$  was in a creative occupation in census year  $t$ . Note that our data is unbalanced in the sense that individuals are only observed for a limited number of census years. The explanatory variables, summarized by  $X_{it}$  include a dummy for gender, a quadratic age polynomial, a set of indicator variables that identify the marital status (single is the baseline category), the number of own family members in household, the number of own children in the household and a set of dummy variables that identify the race (white is the baseline category). Finally, the model accounts for census year fixed effects and state fixed effects by inclusion of the terms  $\sum_{year} f_{year,i}$  and  $\sum_{state} f_{state,i}$ , respectively. Finally,  $u_{it}$  denotes an idiosyncratic error term.

The regression results are reported in Table 1. Most coefficients turn out to be statistically significant, but since this is co-driven by the high number of observations, the following interpretation will focus on the direction of the effects. The magnitudes of the coefficients are in general small, because the average proportion of people having one of the creative occupations is very low (about 0.0054). Being a woman increases the probability of having a creative occupation by 0.002, holding all else equal. Considering the low frequency of creatives, being female leads to a sizable increase by about more than one third. Controlling for all other variables, the probability of an individual having a creative occupation increases also with age, but at a decreasing rate. The results show also that those who are married, separated, divorced or widowed are less likely to take up a creative occupation than singles. Personal commitment, approximated by family size and number of children, are also negatively related with having an artistic occupation. On average, the black and Asian groups are less likely to engage in creative work than whites.

The baseline model (Equation (1)) is then extended by education and income variables. Educational attainment, which is available from

<sup>4</sup> The IBN data used in the underlying study has been generously provided by David de la Croix and its meticulous collection has been described in de la Croix and Licandro (2015).

1940, is measured on an ordinal scale between zero and 11, as described previously. Income is measured in two ways. First, the model includes labor only income (earnings), which is measured as the total pre-tax wage and salary income for the previous year and is available from 1940. This variable captures the effects of an artists' remuneration and hence is directly related to the financial incentives of choosing a creative occupation. Second, the model takes account of total family income, which is measured as the total pre-tax money income earned by one's family from all sources for the previous year, including non-labor income. This variable is available from 1950 and its inclusion is motivated by the fact that the participation decision depends likely not only on one's own earnings, but on the total income of the family. Both income variables have been adjusted for inflation and are logged.

Column (2) of Table 1 summarizes the results. It can be observed that better education increases the probability of a person having a creative occupation. On the other hand, earnings are negatively related with creative occupations — an association that is often found for creative workers, who typically earn less than the average (e.g., Alper and Wassal, 2006). Or from another perspective: The low earnings of artists are detrimental to taking up of an artistic occupations. Interestingly, total family income exhibits a positive relation with the uptake of a creative occupation. This is in line with the notion that potential access to family's financial support is a factor conducive in the participation

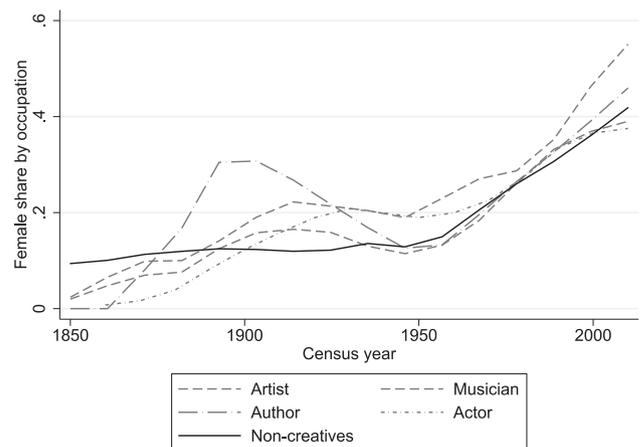


Fig. 2. Female share by occupation. Note: The “Non-creatives” category provides the average for all occupations other than the four creative occupations listed here.

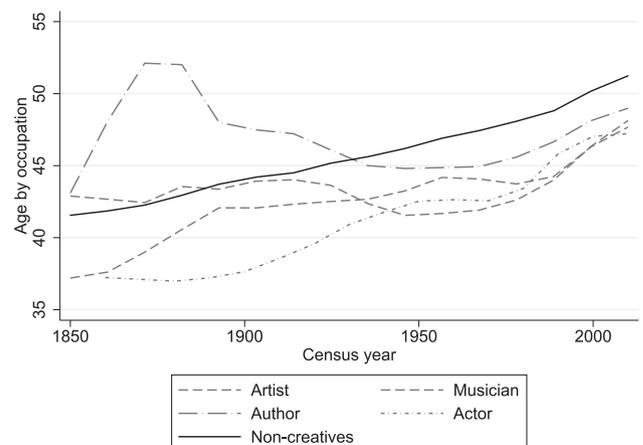


Fig. 3. Age by occupation. Note: See Fig. 2.

decision. Having access to financial resources facilitates the uptake of a creative occupation and the associated lower earnings and potentially higher uncertainties.

The baseline model is further extended by the inclusion of controls for migrants and logged city population. In column (3) it can be observed that these additional controls decrease the number of observations to about 1.3 million (from the initially almost 4 million observations). The newly added variables indicate that migrants are more likely to have a creative occupation, as are those who locate in larger agglomerations. The other previously presented results remain robust. Finally, the model includes in column (4) in addition year times state fixed effects.

4.1.1. The background of creatives over time

The regression approach can be complemented with and is supported by a simple graphical analysis depicting how creatives' background change over time and differ across domains. Fig. 2 shows the share of females for the four groups of creatives studied, along with

the share of females engaged in any other occupation (labeled non-creative occupations). The share of women in non-creative occupations is around 10% at the beginning of the observation window and gradually increases to around 40% by 2010. This corresponds with the overall labor force participation of women (e.g., Goldin, 2006, Fig. 1). During most of the second half of the 19th century relatively fewer female are involved in creative occupations than in other occupations. However, this changes from around 1890 when the share of females increases sharply and remains clearly above non-creative occupations before the two trends converge around 1980 for most domains.

These results challenge the conventional wisdom that the arts are predominantly a male only domain. For example, previous research - which is based on prominent creatives - shows that women are practically unobservable among famous artists (e.g., O'Hagan and Borowiecki, 2010; Hellmanzik, 2010). In contrast, by looking at the average artist, as recorded by the census data, it can be seen that women have often been involved in creative occupations and that their share in these occupations - relative to males - has typically been higher than in

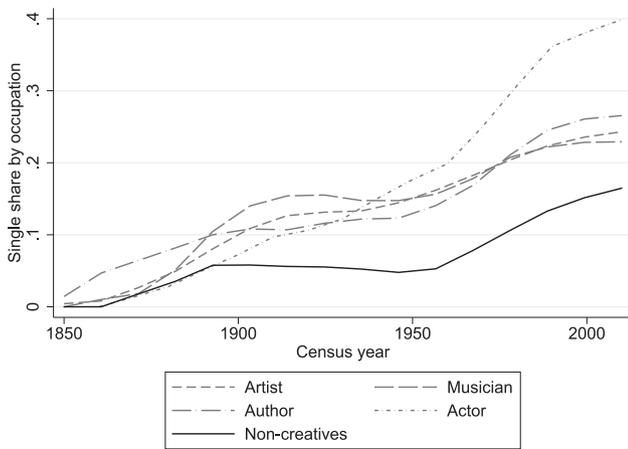


Fig. 4. Being single by occupation. Note: See Fig. 2.

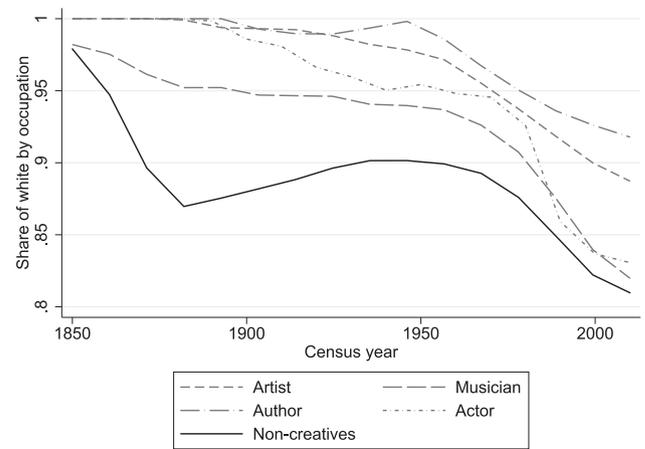


Fig. 5. White by occupation. Note: See Fig. 2.

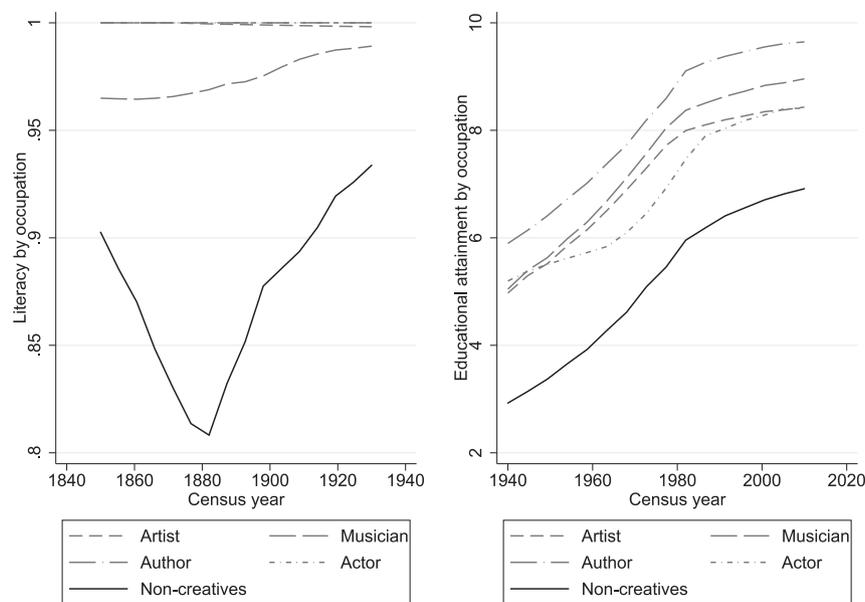


Fig. 6. Educational attainment by occupation. Note: See Fig. 2.

non-creative ones. The observed patterns are also reflected anecdotally in various events that occurred in the American arts education landscape. For example, the Art Students League, founded in 1875 in New York City, saw an increasing number of women artists from the early 1890s (Weber, 2012).

Turning next to age differences, we can observe in Fig. 3 that creative occupations are typically exercised by younger cohorts. One possible explanation for this is that older cohorts drop out from artistic occupations; however, the cross-section data used do not permit the investigation of the reasons behind this result in more depth. The exception is authors, who until the 1930s are on average up to ten years older than the average household head. The higher age of literary artists is possibly explained by their need to acquire a particular stock of cultural capital, before producing a literary artwork. These writers could also have been experimental innovators, who achieve success gradually and typically later in their careers (Galenson, 2007).

The results presented in Fig. 4 address family background and disclose how the share of singles in the American population gradually increases from levels below 5% up to about 18% by 2010. The rise in the share of single respondents among those involved in creative occupations is considerably steeper, and by the end of our observation window about one in four creatives is single, with actors reaching the highest proportion of 40%. The opposite is true for the respondents' family sizes in Fig. A3. Over the last one and a half centuries, a steady decrease from

about five family members to just above 2.5 can be observed. The family size of creatives is typically by at least one person smaller; however, this difference has decreased over the most recent 2-3 decades.

The observed overall higher share of singles and smaller family sizes among creatives is perhaps no surprise. Both variables are related to personal or time constraints and likely limit the individual's involvement in creative activities. These results are in line with the research on cultural participation, which finds very similar patterns and attributes them to the time constraints of a person (e.g., Ateca-Amestoy, 2008). Furthermore, as we will see later, creative occupations are also usually lower paid than non-creative ones, and hence perhaps their ability to afford to get married or have a family is limited.

Fig. 5 explores racial differences and shows that the share of whites decreases from 98% to around 80% over the time period covered. The figure indicates also that it takes almost a whole century before the first non-whites appear among artists or authors. Actors are somewhat less dominated by whites in the early 20th century and since 1980 the proportion of whites drops dramatically. Musicians are the most racially mixed group of creatives. This does not come perhaps as a surprise if one considers genres such as jazz, blues, or funk, all invented, mastered, and typically performed by blacks. These observations come though with a few shortcomings. First of all, the earliest two census editions do not include slaves, which means that the picture provided for 1850 and 1860 is incomplete. Second, non-whites who were involved in creative

**Table 2**  
Socio-economic background of creatives by domain.

	Dependent variable:			
	LPM Artist (1)	LPM Author (2)	LPM Musician (3)	LPM Actor (4)
female	-0.001 (0.006)	0.021*** (0.005)	0.218*** (0.011)	-0.007*** (0.003)
age	0.007*** (0.001)	0.007*** (0.0005)	0.017*** (0.001)	0.0001 (0.0003)
age <sup>2</sup>	-0.0001*** (0.00001)	-0.0001*** (0.00000)	-0.0002*** (0.00001)	-0.00001** (0.00000)
married	-0.080*** (0.008)	-0.059*** (0.007)	-0.138*** (0.015)	-0.025*** (0.004)
separated	-0.101*** (0.013)	-0.084*** (0.011)	-0.286*** (0.025)	-0.026*** (0.005)
divorced	-0.062*** (0.010)	-0.076*** (0.008)	-0.150*** (0.018)	-0.011** (0.005)
widowed	-0.131*** (0.009)	-0.113*** (0.007)	-0.371*** (0.017)	-0.025*** (0.004)
family size	-0.020*** (0.002)	-0.018*** (0.001)	-0.049*** (0.004)	-0.005*** (0.001)
number children	-0.004 (0.003)	0.006*** (0.002)	-0.011** (0.005)	-0.004*** (0.001)
black	-0.125*** (0.004)	-0.081*** (0.004)	-0.131*** (0.011)	-0.009*** (0.002)
native	0.020 (0.037)	-0.067*** (0.022)	-0.063 (0.064)	-0.017* (0.010)
asian	-0.025* (0.014)	-0.099*** (0.009)	-0.199*** (0.023)	-0.042*** (0.004)
other	-0.110*** (0.011)	-0.139*** (0.007)	-0.352*** (0.022)	-0.033*** (0.005)
mixed	-0.020 (0.025)	-0.035 (0.023)	-0.037 (0.052)	0.052*** (0.020)
Constant	0.117 (0.113)	-0.121** (0.057)	0.231 (0.206)	0.023*** (0.007)
Year FE?	Yes	Yes	Yes	Yes
State FE?	Yes	Yes	Yes	Yes
Observations	3,925,299	3,925,299	3,925,299	3,925,299
R <sup>2</sup>	0.001	0.001	0.002	0.001
Adjusted R <sup>2</sup>	0.001	0.001	0.002	0.001
Residual Std. Error (df = 3925204)	0.038	0.029	0.067	0.016
F Statistic (df = 94; 3925204)	26.045***	36.416***	74.964***	27.504***

Note: All coefficients (and the associated s.e.) are multiplied by 10<sup>2</sup> to ease readability. \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01.

artistic activity over the earlier part of the period studied may not have been counted as “artists” by historical census enumerators. This could be why it looks like there are no black artists or authors until the mid-20th century. Given the fact that some of the most important American art forms were created by African Americans, one needs to be careful in the interpretation of the census data.

Fig. 6 provides insights into the educational attainment of the creatives covered. The censuses until 1930 provide only a dummy indicator for literacy, as presented in the left panel of the figure. It can be observed that the vast majority of creatives are literate and clearly more so than those involved in non-creative occupations. From 1940 the census data provide a more sophisticated measure of educational attainment by indicating the level of school accomplishment or the number of college years completed. Based on this information an ordinal scale between zero and 11 has been compiled and is used in the right panel. The educational attainment is sharply increasing until about the 1990s when the increase becomes less marked. As in the pre-1940 period, the creatives have obtained significantly more education than the average non-creative worker. There are also interesting differences across the creative domains, with authors being the best educated, whereas actors are at the lower end of education attainment.

4.1.2. Disaggregating by the creative domain

Table 2 provides the baseline results disaggregated by the creative domain. Males are less likely to be authors or musicians, but somewhat more likely to work as actors. The other meaningful difference is that the probability of being an actor decreases with age. This supports the notion that actors, especially in the movie industry, are predomi-

nantly male and typically young. The online appendix presents further results disaggregated by creative domain for a model with education and income controls (Table A1) and with migrant and city size controls (Table A2).

4.2. The historical development of artistic clusters

This section provides historical insights into the geographic clustering patterns of creative activity in the United States. In the following depictions, the total number of creatives, as opposed to the share per population, is shown, and there are three reasons for this. First, it is established that the total number of artists (not the density) matters for benefits associated with peer effects: Whether artists are based in a small or large city, the experienced benefits are related to the size of the artist population (Borowiecki, 2015a). Second, it is more likely that the total number matters more for the attraction of high-skilled workers and possibly for spillover effects of creativity from the arts to the economic sectors. This is also supported by the observation that larger cities typically have cultural infrastructures that allow artists to reach greater audiences (e.g., a larger concert hall). Third, artists usually cluster in certain districts of a city, and hence considering the population size of a whole city as a denominator would be misleading, and an intra-city approach is not feasible in this research.

Initially, we analyze the geography of artistic talent by looking at the birthplaces and deathplaces of famous IBN creatives. Even though artists are highly mobile, there exists a very high correlation between their workplace and birthplace or deathplace. Furthermore, it is fairly established that the births of famous creatives typically occur in places

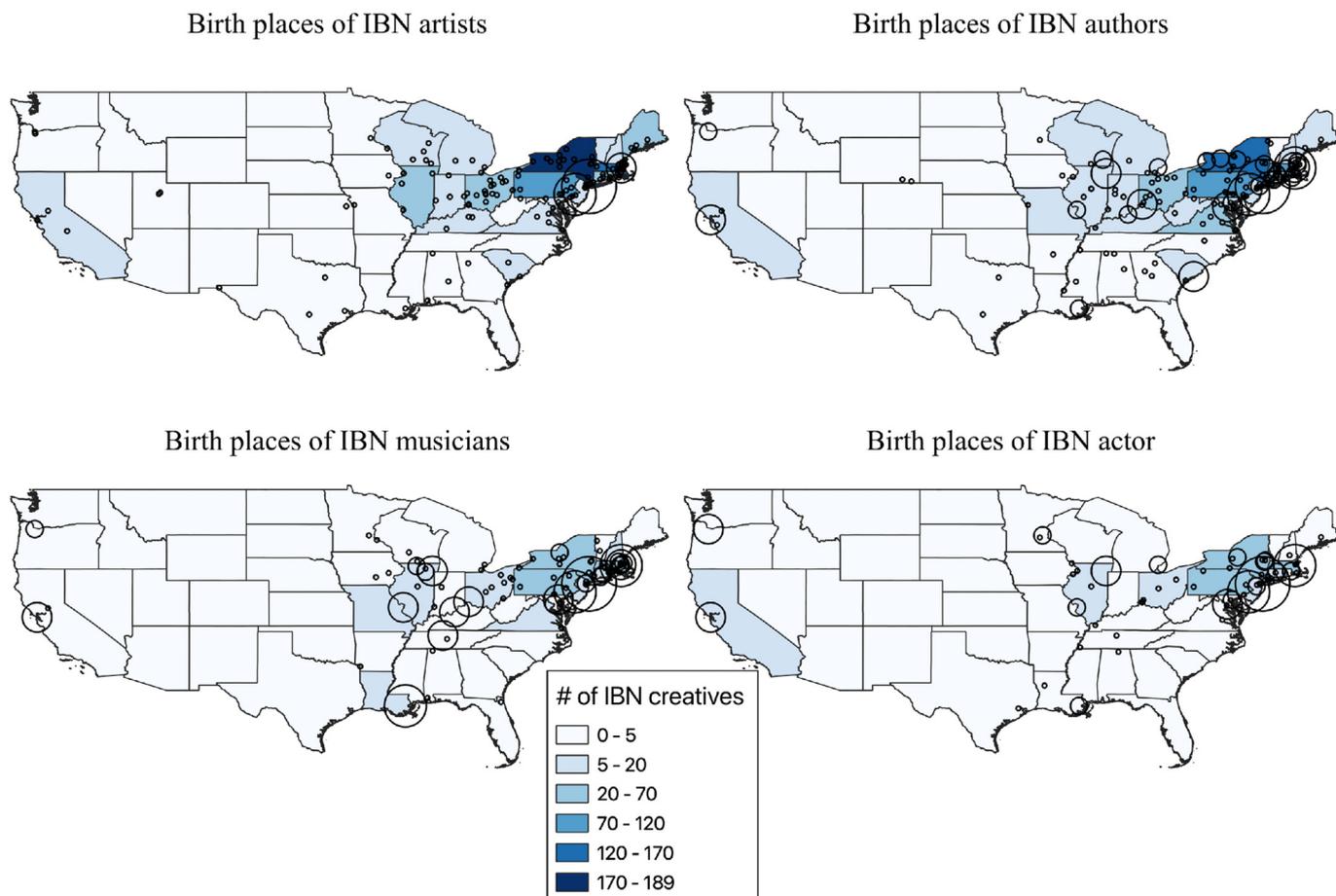


Fig. 7. Birthplace of IBN creatives.

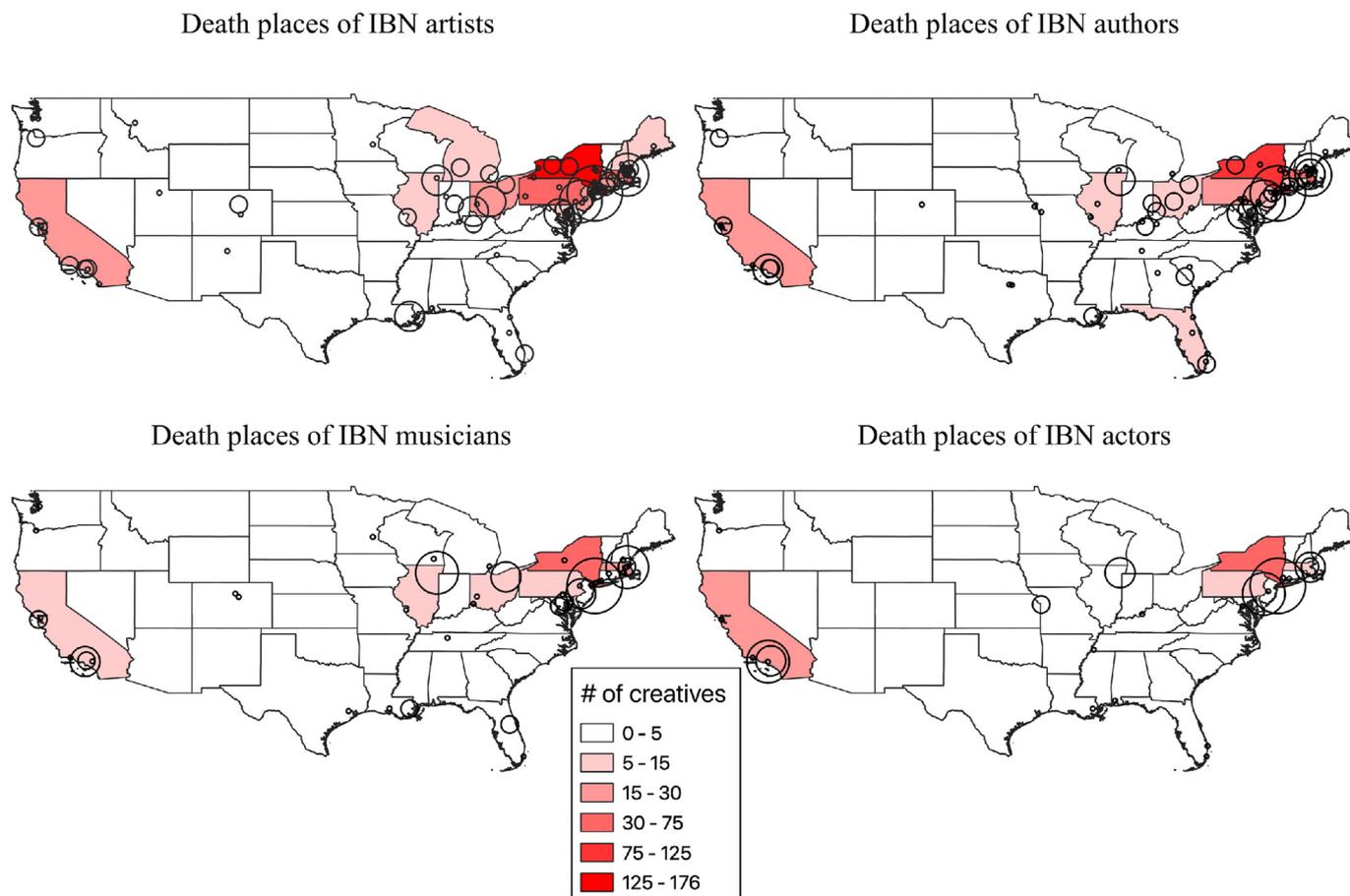


Fig. 8. Deathplace of IBN creatives.

where a given artistic domain has already been developed (for evidence and discussion see, for example, O'Hagan and Jan Borowiecki, 2010; Borowiecki and O'Hagan, 2012).

The maps depicting the birthplaces or deathplaces of IBN creatives are presented in Figs. 7 and 8, respectively. Each map indicates by a scaled point the importance of a city as a birthplace for a certain group of creatives and by shades the importance of a state.<sup>5</sup> Creative activity is primarily located in the Mid Atlantic, North Eastern, and a part of Mid Western regions, and along the West Coast. The geographic concentration is more intense for the deathplaces - this supports the previously posited high migration intensity (both internal migration and immigration). Across all creative domains studied, New York City emerges as the consistently largest cluster city, followed by Boston, Chicago, Los Angeles, and San Francisco.<sup>6</sup>

There are, however, also clear differences across the domains. For example, New Orleans is found to be a place with a very high concentration of births of musicians. In New Orleans funk was supposedly played for the first time ever but more importantly, it is the city where jazz originated. The insight that a significant number of famous musicians were born here lends support to the colloquial label assigned to the city as the *birthplace of jazz*. Another example is St. Louis, a city strongly associated with blues, but also jazz and ragtime. Interestingly though, while these two cities emerge as unusually important birthplaces, markedly fewer deaths are observed there. This indicates that

<sup>5</sup> For some few observations the exact city was not available, and only information on the county or state was provided.

<sup>6</sup> Miami also receives some prominence when it comes to deaths, but this is possibly more related to the fact that it is a popular destination for retirement.

many of the famous individuals born here migrated away. Perhaps the most famous example is Louis Armstrong, who was born in New Orleans, but died in New York City, where he also spent a significant part of his career.

The concentration intensity of the census creatives (the "average" creatives) is shown in Fig. 9. The geographic spread of the census creatives is considerably greater, while the clustering intensity appears to be somewhat lower - albeit still very noticeable - in comparison with the famous creatives. The greater spread is partly caused by the higher number of observations available for the census sample. However, the findings also indicate that extraordinary talent concentrates more than average individuals. Furthermore, it is interesting to observe that several dominant clusters emerge, and these are very similar as in the case of famous creatives.

The long time series enables us also to analyze changes in locations over time. The famous IBN creatives are observed for a period from before the census; hence, the earliest period covers the years before 1850. Fig. 10 depicts for this period the deaths of artists in the top-left panel, followed by three additional time intervals: 1850–1890, 1900–1940, and 1950–1980. The earliest period covers locations restricted to the East Coast: New York City, Boston, Philadelphia, and Charleston. Over time, the spread extends to the eastern part of the Mid West and later also to the West Coast, in particular California. A comparable development is observed if one looks at the location of census respondents in Fig. 11, now beginning with the period 1850 to 1890, and extending in the last period to the years from 1980 to 2010. Qualitatively the story of geographic spread is comparable with the IBN creatives; however, as previously noted, the clustering intensity is less marked. These maps more closely reflect the geographic concen-

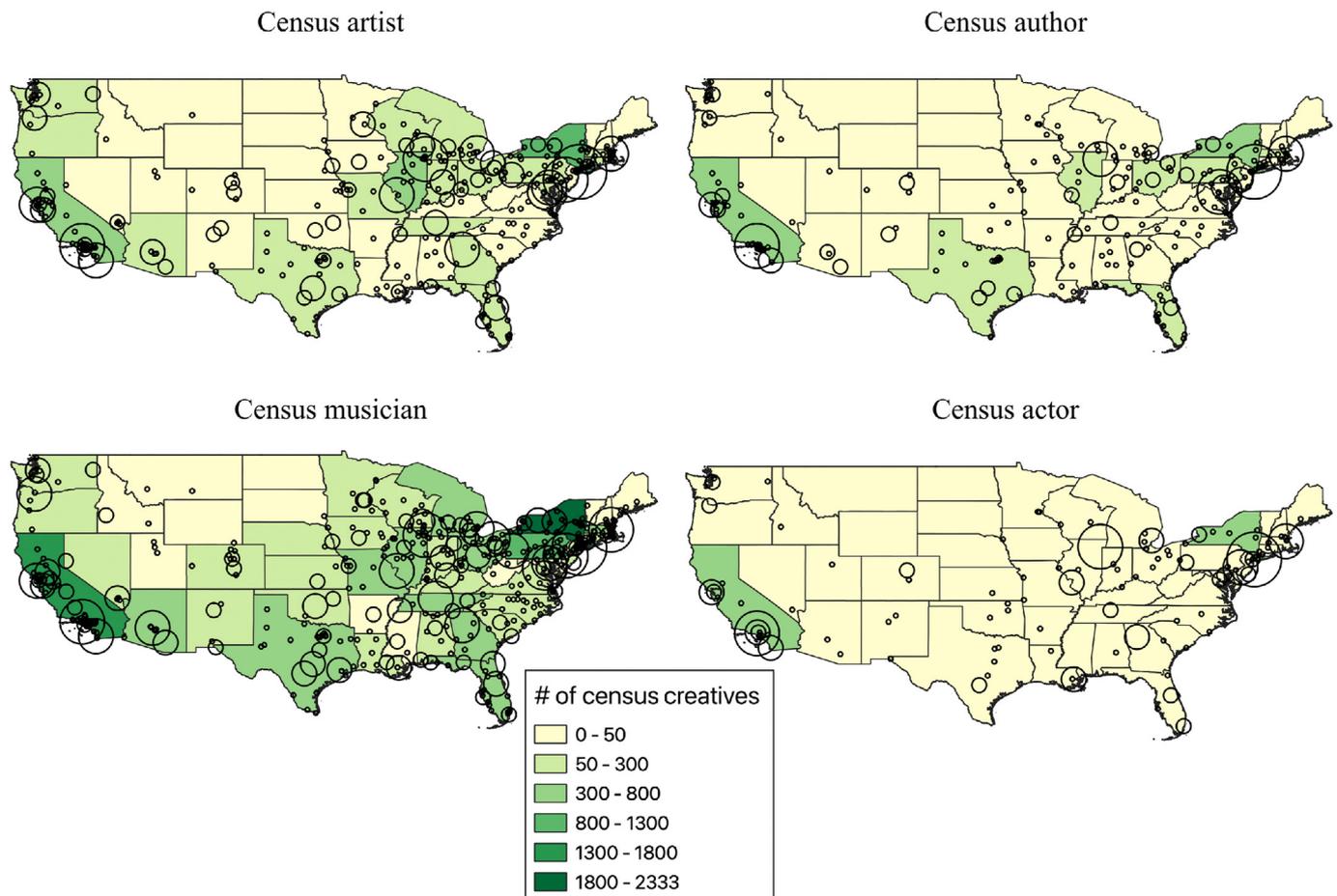


Fig. 9. Location of census creatives.

tration of the population across the United States; however, the clustering patterns go beyond demographic factors. Some of the cities have a clearly overrepresented share of artists (e.g., Chicago), whereas other large cities are characterized by a relatively insignificant population of artists (e.g., Houston).

The presented trends resonate with the wider economic context in several regards. The early economic development in the New England region has been stimulated by a range of subsidies to, for example, improve the infrastructure or by the introduction of early improved institutions, such as a legal system conducive to business activity (Newell, 2000). The decreasing cost of transportation, related to the introduction of railroads but not only, connected the North and Midwest, which in turn stimulated not only migration, but also economic growth (Fogel, 1965). The rapid expansion of settlements to the West opened up vast frontier lands, which became connected by rail already in 1869, and with the advent of the automobile decreased further the travel cost. Improved connectivity and the resulting high inflow of workers contributed to the spread of people, goods, and also ideas.

Fig. 12 depicts the clustering patterns of actors, which are particularly fascinating. The deaths of all famous IBN actors before 1840 occurred in New York City, later also in other cities of the Mid West, and eventually on the West Coast, quite as for artists. However, the period after 1950 shows a remarkable concentration in just two cities: New York and Los Angeles. The dominance of Los Angeles is related to the rapid growth of Hollywood and is in line with other economic history accounts of the development and dominance of the US movie

industry (e.g., Bakker, 2005; Sedgwick and Pokorny, 2010).<sup>7</sup> The forces that have created this geographic cluster include California's natural advantages (Ellison and Glaeser, 1997) in the form of sunlight, climate and the variety of landscape, which are particularly important for film-making. Similar results emerge if one looks at the more numerous observations from the census data in Fig. 13: The states of New York and California remarkably dominate the landscape.

#### 4.2.1. Measuring the geographic concentration of creatives

In an attempt to introduce a measure of the extent of geographic concentration, we compile Gini concentration indices for each of the creative occupations since 1850. The series are shown in Fig. 14, along with the Gini concentration index for teachers to provide some context. As one would expect, teachers are the most evenly distributed group of workers. The Gini index for teachers is almost constant and below 0.1 since the 1960s, which is the result of overall access to education (Goldin and Katz, 2010). At the other extreme are actors, who exhibit the greatest concentration. This is in line with the previously observed and discussed extraordinary clustering of actors (see Fig. 9). Authors and artists exhibit a very high concentration in the earliest decades, which is followed by a gradual decrease in concentration from about the 1940s. The Gini coefficient for authors and artists converges at the value of about 0.6. The patterns and extent of concentration seem very similar for authors and artists, which is likely a result of analogies in

<sup>7</sup> Changes over time in clustering patterns for IBN and census musicians are depicted in Appendix Figs. A6 and A7, and for authors in Figs. A8 and A9.

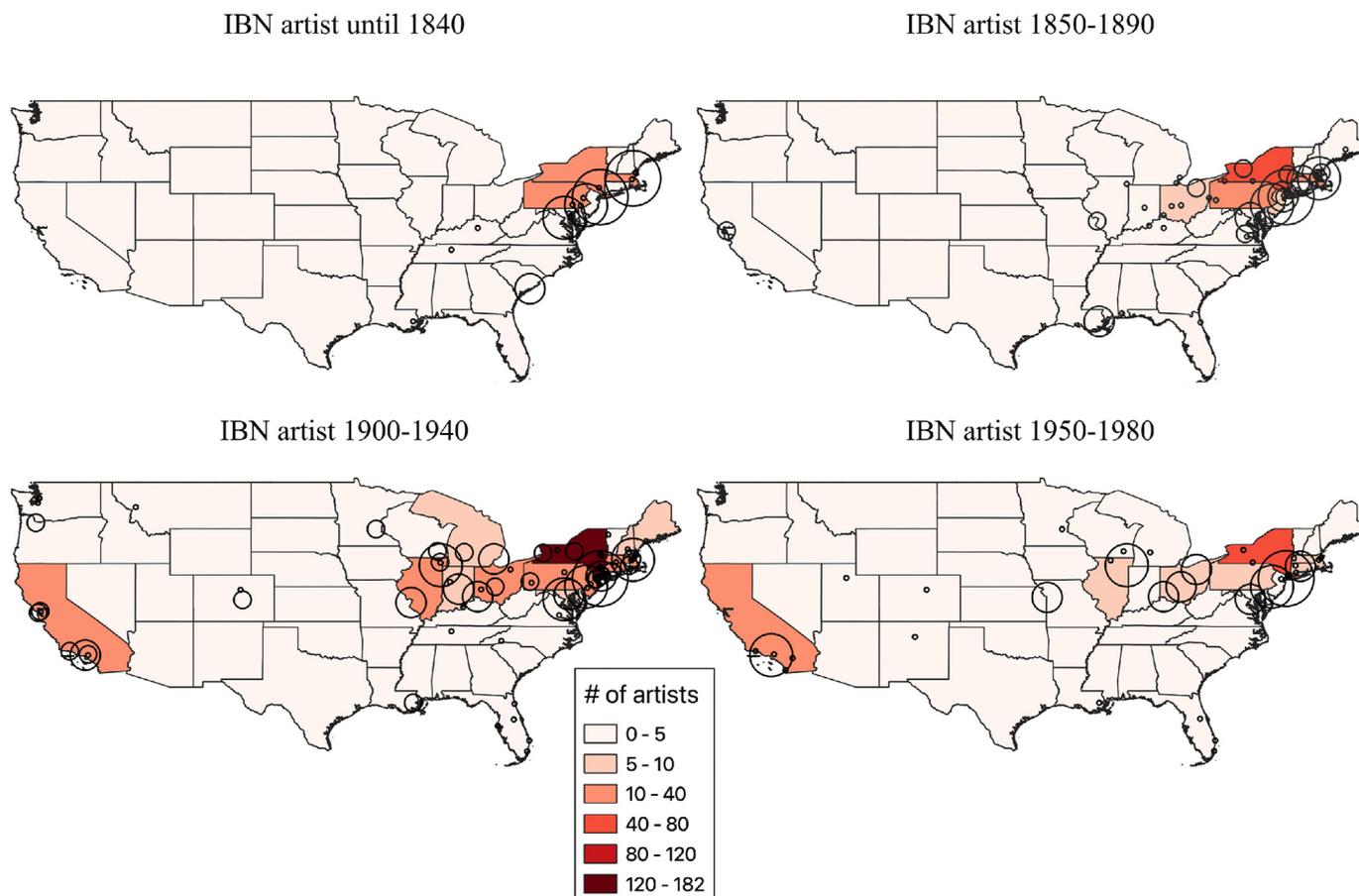


Fig. 10. Geographic distribution of deaths of artists (IBN).

how the two groups produce and disseminate their output: the production can occur remotely, whether writing a book or painting, but both require important access to supply-related infrastructure (publishers or gallerists and art museums). Moreover, the spillovers from interaction with other creatives are typically centered in specialized locations (e.g., Borowiecki, 2013). In some contrast to this, musicians concentrate less than artists and authors, albeit much higher than teachers. This could be due to the necessity to locate close to one audiences, especially considering live performances, but also appearances on the radio, which were initially more localized.

It is also interesting to observe the long-term decreases in the concentration of artists, authors and musicians over the course of the second half of the 20th century. These trends relate possibly to the various improvements in the transport infrastructure and communication technologies, which have decentred some of the creative activity away from artistic clusters towards the individual consumers (Borowiecki et al., 2016).

#### 4.2.2. Interdependencies across creative clusters

The data allows us also to explore cluster interdependencies and how cluster sizes of various domains relate to each other. These relationships are shown for census creatives in Table 3, using OLS-models (uneven columns) and Poisson-models (even columns). The regressions now include year fixed effects to account for unobservable differences across time. To account for differences in city size, all models include further controls for the logged population of a city. Since two of the creative occupations include teachers (i.e., artists and musicians), we further include controls for the number of teachers, as recorded by the

census.<sup>8</sup>

The correlations are positive and typically estimated with high statistical precision. For example, in column (1) we can observe that a one percent increase in the number of authors is associated with a 0.33% increase in the population of artists. The relationships are in general in the range between 0.06% and 0.33%, and also qualitatively consistent when estimated with Poisson-models. The overwhelming insight is that creatives mix: cities that are the domicile for a certain type of creatives (e.g., visual artists) are typically also more popular among individuals from other creative domains (e.g., authors or musicians).

#### 4.2.3. The role of famous artists

Finally, we investigate the interplay between famous and average creatives. Here we explore the probability of a census respondent reporting a creative occupation and how this relates to the presence of famous creatives in the same city in the past. The estimated model is basically the same as reported previously in Equation (1), but now in addition accounts for the history of significant artistic activity, measured as the number of deaths of famous IBN creatives that occurred in the decade prior to the given census. The identification of deaths measured at residence of census respondent during ten years prior to the given census enables a certain time lag for the person inspired by a famous IBN creative to become a creative herself.

The results are summarized in Table 4 and imply that places with greater artistic activity (i.e., where more creatives within a domain have

<sup>8</sup> The results would remain qualitatively the same, if the teacher control variable was not included. Appendix Table A3 estimates how creative clusters interrelate for case of famous IBN creatives.

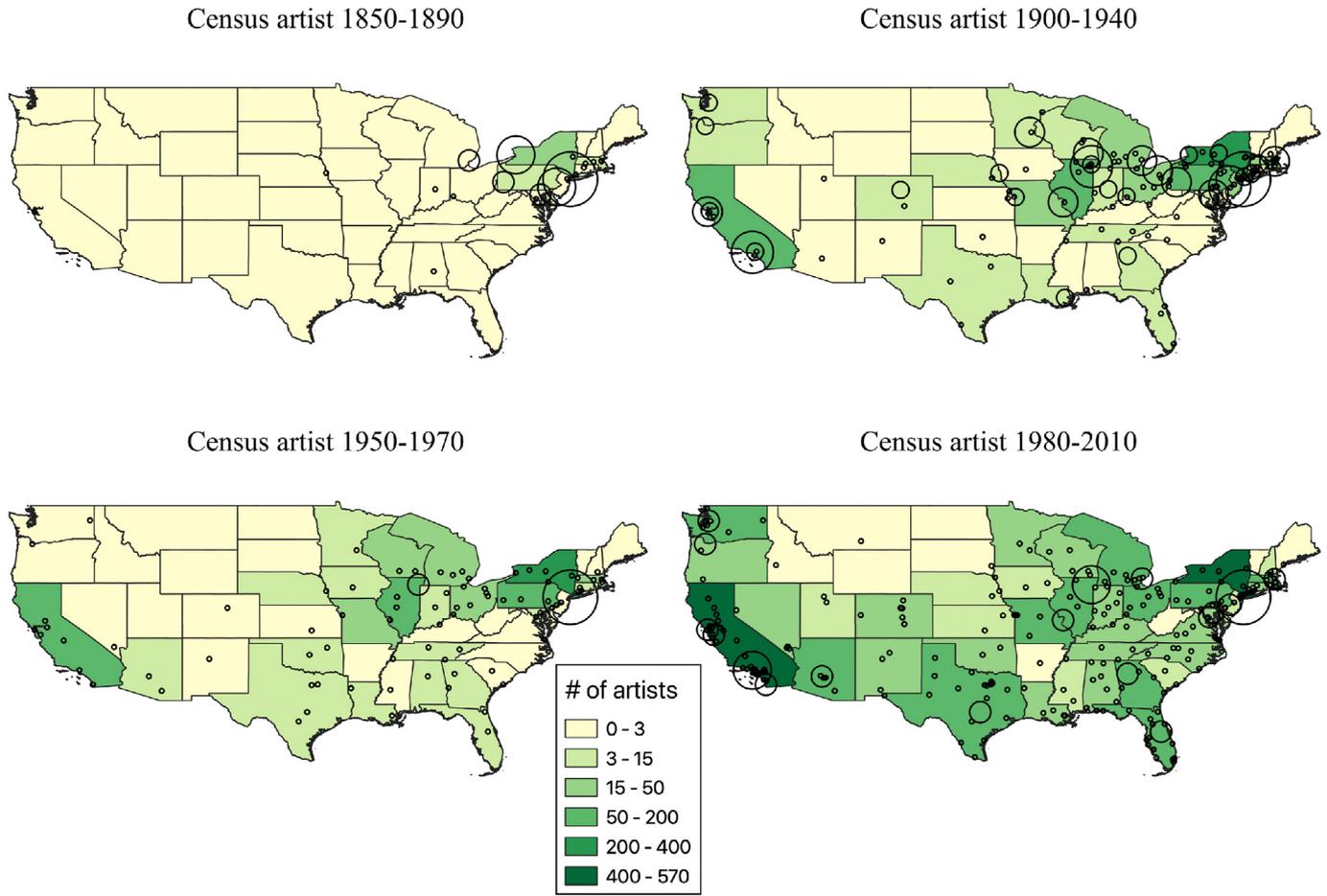


Fig. 11. Geographic distribution of artists (Census).

Table 3  
Interrelation of clusters across artistic domains.

	Dependent variable:					
	log(Artist) OLS (1)	Artist Poisson (2)	log(Author) OLS (3)	Author Poisson (4)	log(Musician) OLS (5)	Musician Poisson (6)
log(author)	0.33*** (0.03)	9.90*** (1.49)			0.13*** (0.02)	24.68*** (2.89)
log(artist)			0.21*** (0.02)	11.47*** (2.55)	0.23*** (0.02)	31.22*** (3.22)
log(actor)	0.21*** (0.04)	2.56** (1.21)	0.24*** (0.04)	1.48 (1.79)	0.19*** (0.03)	10.09*** (3.21)
log(musician)	0.17*** (0.02)	15.49*** (2.82)	0.06*** (0.01)	25.55*** (9.59)		
log(teacher)	0.09*** (0.01)	24.43*** (3.98)	0.02*** (0.01)	13.01 (9.05)	0.16*** (0.02)	115.58*** (9.17)
log(city population)	0.16*** (0.02)	58.69*** (5.46)	0.02* (0.01)	16.91* (9.04)	0.34*** (0.03)	101.22*** (13.50)
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3506	3506	3506	3506	3506	3506
R <sup>2</sup>	0.54		0.51		0.63	
Adjusted R <sup>2</sup>	0.54		0.50		0.62	
Log Likelihood		-110,523.70		-77,481.23		-185,492.60
Akaike Inf. Crit.		221,085.30		155,000.50		371,023.20
Residual Std. Error (df = 3487)	1.46		1.16		1.70	
F Statistic (df = 18; 3487)	228.88***		198.02***		323.31***	

Note: \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01.

died) tend to have more people involved in creative occupations within the same domain, albeit the results are statistically significant only for the case of authors and actors. This model is not suitable for making

any causal claims; however, it seems likely that the presence of famous creatives has at least some impact on the future development of creative clusters.

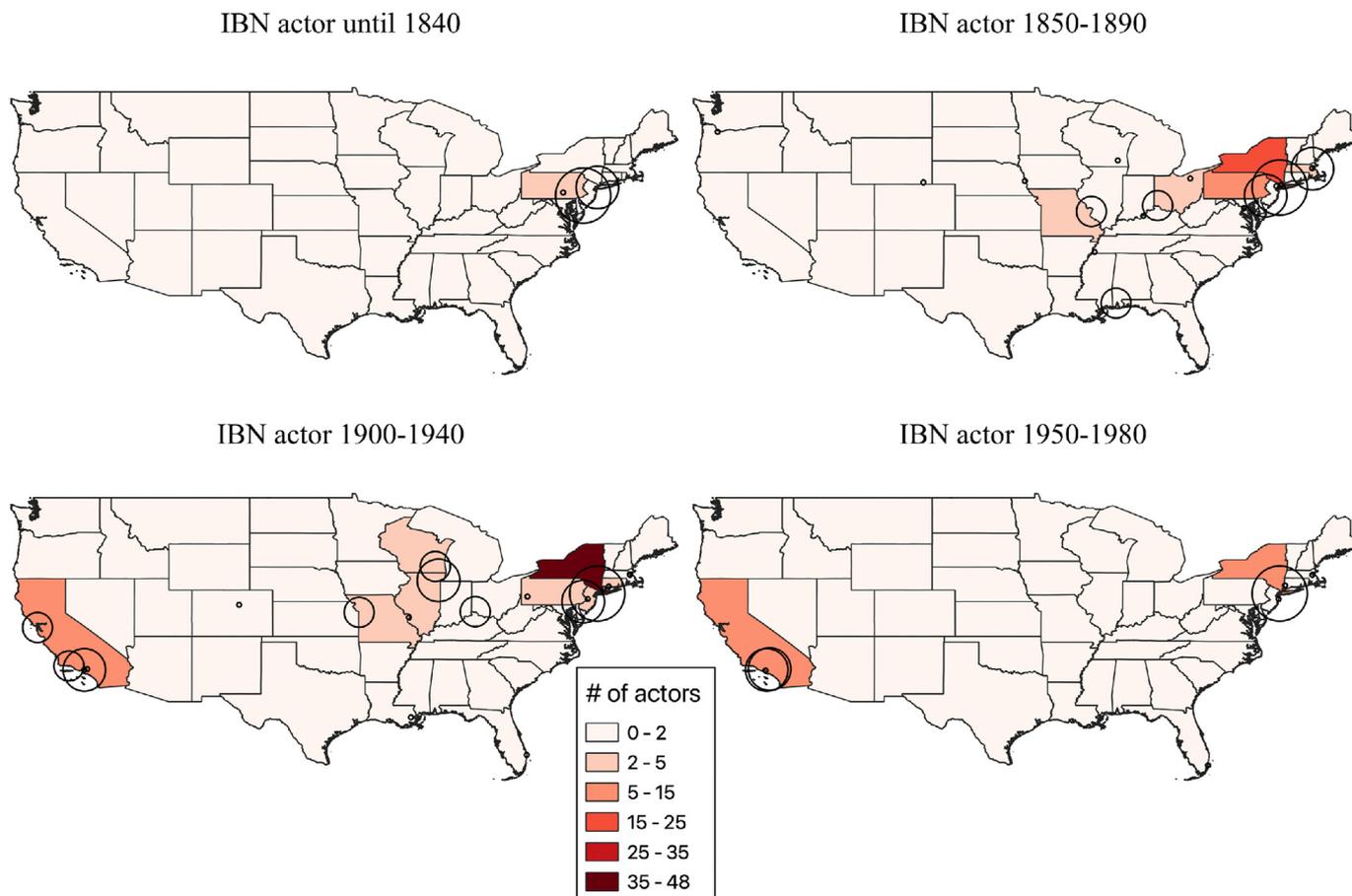


Fig. 12. Geographic distribution of deaths of actors (IBN).

Table 4  
Artistic occupation as a function of historical artistic activity.

	Dependent variable:			
	LPM Artist (1)	LPM Author (2)	LPM Musician (3)	LPM Actor (4)
log(IBN artist death)	0.047 (0.010)			
log(IBN author death)		0.022*** (0.005)		
log(IBN musician death)			0.009 (0.013)	
log(IBN actor death)				0.077*** (0.009)
Constant	0.125 (0.109)	-0.043 (0.026)	0.027 (0.136)	-0.036 (0.037)
Socio-economic controls?	Yes	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes	Yes
State FE?	Yes	Yes	Yes	Yes
Observations	504,508	504,508	504,508	504,508
R <sup>2</sup>	0.001	0.001	0.001	0.002
Adjusted R <sup>2</sup>	0.001	0.0005	0.001	0.002
Residual Std. Error (df = 504449)	0.046	0.021	0.056	0.026
F Statistic (df = 58; 504449)	8.844***	4.983***	9.515***	15.364***

Note: All coefficients (and the associated s.e.) are multiplied by 10<sup>2</sup> to ease readability. \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01.

### 5. Conclusions

Today, the role of creativity and the presence of creative people are arguably of immense importance for economic growth and the welfare

of societies. And yet, research on these topics is typically limited to contemporary approaches and is usually conducted outside the field of economics. This study adds new insights into the *historical* development of creative activity and clusters of creativity. This is achieved by looking

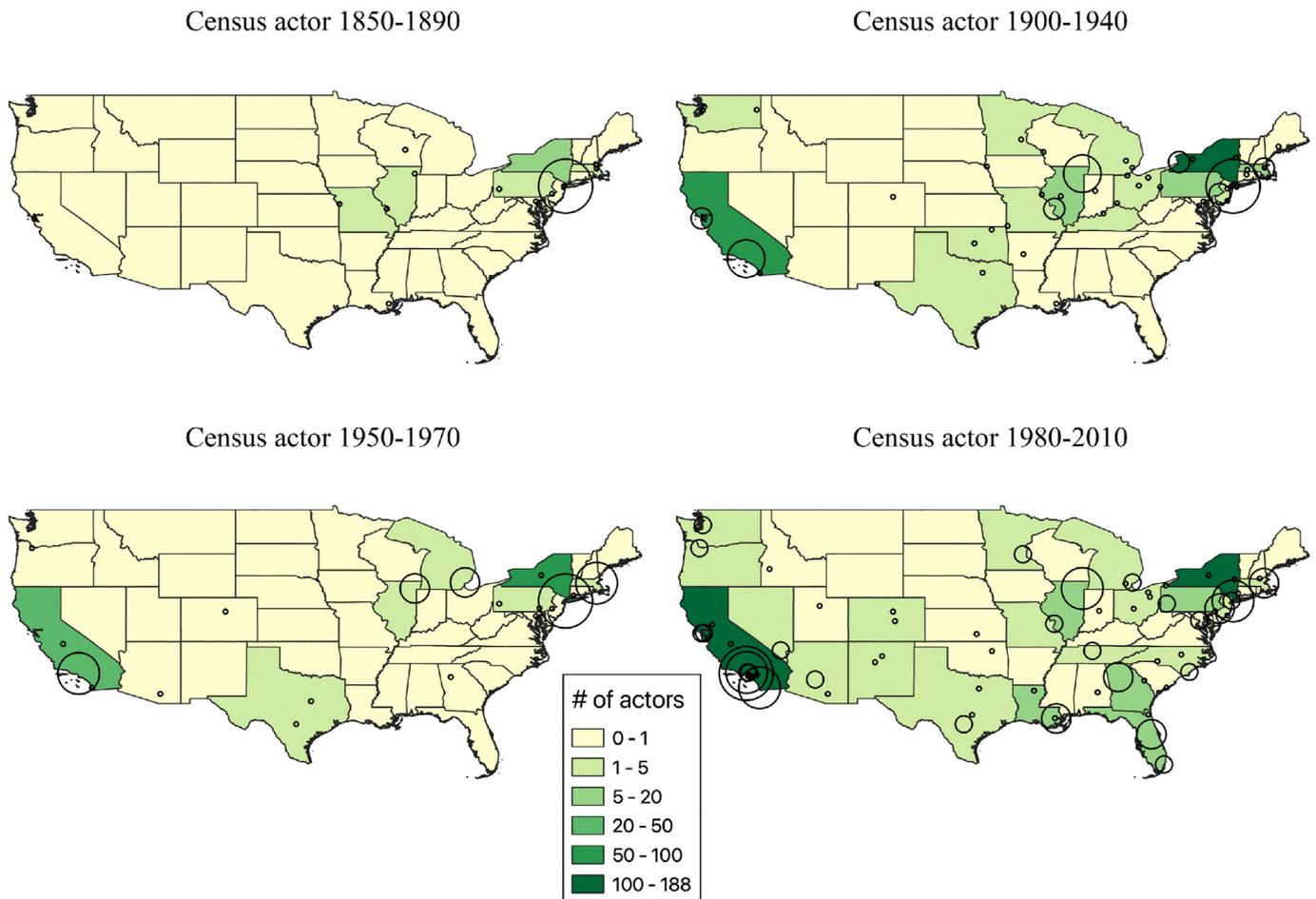


Fig. 13. Geographic distribution of actors (Census).

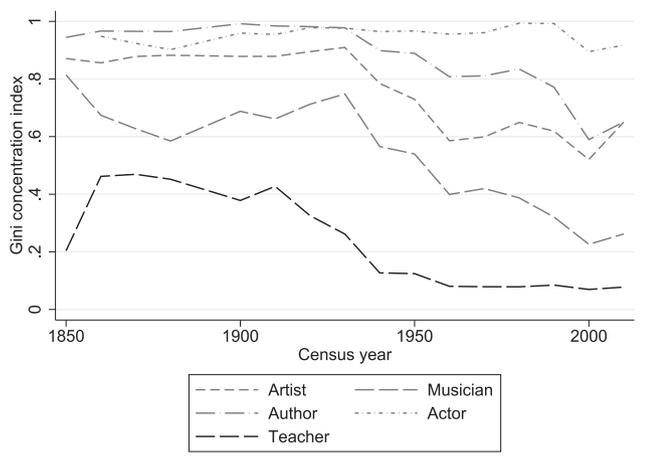


Fig. 14. Gini concentration index across creative occupations.

at the case of the arts, where the earliest creative achievements can be observed in a consistent and comparable way.

The underlying research documents the background of those involved in a creative occupation and furthermore illuminates how it has changed over the course of 170 years. Some of the disclosed patterns echo the overall socio-demographic trends of the period covered, but there are several novel and interesting insights: The proportion of female creatives is relatively high, time constraints can be a hindrance

for taking up a creative occupation, racial inequality is present and tends to change only slowly, and access to financial resources within a family facilitates the uptake of an artistic occupation.

Furthermore, we shed light on the geography of creative clusters in the United States and explain how these have evolved over time and across various domains. Even though it may seem that some of the patterns are already known — for example, that New York City is a very significant center for the arts — the extent of the dominance has not yet been quantified before, nor has it been compared across creative domains. Finally, by linking the census records with data on famous IBN creatives, the role of famous individuals for the growth of local clusters and creative employment is explored. Famous creatives have some influence on people taking up creative occupations and are related to the size of creative centers. Typically, superstar economies (Rosen, 1981) are criticized by the public mainly due to the extreme earnings received by a small group of individuals at the very top of the income scale. The insights presented here suggest the possibility of a positive externality of superstars in the form of a potentially long-lasting heritage that some famous creatives leave behind.

This research gives rise to several new questions. In particular, the question of how the presence of a famous creative impacts others: Does she introduce new knowledge, practices, networks, or infrastructure, or a multiple of these factors, which then potentially persist over time? Or is it perhaps the case that her presence stimulates the demand within a creative domain due to factors related to local identity and heritage (for a related discussion, see Borowiecki, 2015b)?

Of interest for contemporary policy makers and the public is also whether and how the historical development of creative activity is nowadays related to creativity. Anecdotally, there seems to be a very

high overlap between the creative clusters historically and the startup landscape in the United States these days. We have motivated this research by disclosing a strong and long-term correlation between artistic activity and entrepreneurial outcomes. According to [Compass \(2015\)](#), six US cities are listed among the global top-10 startup ecosystems, and the list begins with San Francisco (Silicon Valley), New York City, Los Angeles and Boston; that is, cities that have been identified in the underlying research as significant creative clusters in history. These cities are obviously also centers of higher education with some of the top universities in the country. Certainly the role of education cannot be overlooked, as it is likely at least as important as artistic activity in explaining why these areas are and have been centers of entrepreneurial outcomes (for a more detailed discussion see [Goldin and Katz, 2010](#)).

## Appendix

### A. Additional robustness tests

The novel long-term approach pursued in the underlying paper comes at a cost – the measurement of some of the variables has (usually slightly) changed across the 16 decades covered, while other variables are available only over limited time periods.

It is important to note that the changes across the Census waves in the definition or measurement of some of the variables covered, have been made quite certainly independently from changes in the labor market of creatives. Furthermore, the changes go sometimes in either direction (for example, we observe both, increases and decreases in the cut-off point of age). Therefore, while the census changes may still lead to biased estimates, these biases – given the long time-period covered – should not be very meaningful on average. Nonetheless, a series of robustness tests has been conducted to check on the consistency of the models estimated.

One particular change in the cut-off of a variable concerns the occupation variable (*occ1950*), which has been obtained in the earliest two census waves covered individuals aged 15+, but in later editions also respondents aged 14+ and 16+ have been surveyed. The volatility of the cut-off point regarding age is rather small, and concerns primarily individuals who – in many cases - have not reached yet the age to become involved in a creative occupation (or perhaps even in most occupations). Nonetheless, in an attempt to check on this potential bias, one may want to drop all individuals below the age of 16 to ensure that the same age cohort is covered throughout the time period. This has been done and the results are indistinguishable from the baseline specification, which is encouraging.

However, while it is beyond the scope of this study to contribute to the debate on how artistic creativity is related to startup activity, it becomes clear that these concepts are related and very persistent over time, perhaps even more so than previously thought.

### Declaration of competing interest

We declare that there exists no conflict of interest.

### Acknowledgement

This research has been co-funded by the Danish Research Council under the project “Origins of Creativity”.

## B. Additional Tables and Figures

## B.1. Socio-economic background of creatives: Additional models

**Table A1**  
Socio-economic background of creatives: Education and income

	Dependent variable:			
	LPM Artist (1)	LPM Author (2)	LPM Musician (3)	LPM Actor (4)
female	-0.0001 (0.006)	0.022*** (0.006)	0.233*** (0.012)	-0.010*** (0.003)
age	0.007*** (0.001)	0.007*** (0.001)	0.011*** (0.001)	0.0005* (0.0003)
age <sup>2</sup>	-0.0001*** (0.00001)	-0.0001*** (0.00001)	-0.0001*** (0.00001)	-0.00001** (0.00000)
married	-0.081*** (0.009)	-0.073*** (0.008)	-0.122*** (0.016)	-0.029*** (0.004)
separated	-0.081*** (0.013)	-0.066*** (0.011)	-0.163*** (0.025)	-0.027*** (0.005)
divorced	-0.055*** (0.010)	-0.075*** (0.009)	-0.097*** (0.019)	-0.018*** (0.004)
widowed	-0.108*** (0.009)	-0.102*** (0.008)	-0.262*** (0.018)	-0.026*** (0.004)
family size	-0.017*** (0.002)	-0.011*** (0.002)	-0.028*** (0.005)	-0.006*** (0.001)
number children	-0.004 (0.003)	0.003 (0.002)	-0.017*** (0.006)	0.001 (0.001)
black	-0.105*** (0.004)	-0.062*** (0.004)	-0.023* (0.012)	-0.005** (0.002)
native	0.047 (0.038)	-0.035 (0.022)	0.069 (0.064)	-0.011 (0.010)
asian	-0.032** (0.014)	-0.112*** (0.010)	-0.247*** (0.024)	-0.039*** (0.004)
other	-0.042*** (0.011)	-0.060*** (0.007)	-0.030 (0.022)	-0.021*** (0.005)
mixed	0.002 (0.025)	-0.011 (0.023)	0.067 (0.052)	0.058*** (0.020)
earnings	-0.009*** (0.001)	-0.005*** (0.001)	0.007*** (0.001)	-0.0005* (0.0002)
family income	0.009*** (0.001)	0.008*** (0.002)	-0.007** (0.003)	-0.0005 (0.001)
education	0.026*** (0.001)	0.032*** (0.001)	0.123*** (0.002)	0.003*** (0.0003)
Constant	-0.139*** (0.022)	-0.246*** (0.018)	-0.304*** (0.044)	0.032*** (0.008)
Year FE?	Yes	Yes	Yes	Yes
State FE?	Yes	Yes	Yes	Yes
Observations	3,260,641	3,260,641	3,260,641	3,260,641
R <sup>2</sup>	0.001	0.002	0.004	0.001
Adjusted R <sup>2</sup>	0.001	0.002	0.004	0.001
Residual Std. Error (df = 3260553)	0.038	0.031	0.070	0.014
F Statistic (df = 87; 3260553)	37.644***	58.088***	137.419***	23.266***

Note: All coefficients (and the associated s.e.) are multiplied by 10<sup>2</sup> to ease readability. \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01

**Table A2**  
Socio-economic background of creatives: City size and migrants

	<i>Dependent variable:</i>			
	LPM Artist (1)	LPM Author (2)	LPM Musician (3)	LPM Actor (4)
female	-0.038*** (0.012)	-0.003 (0.011)	0.047** (0.021)	-0.022*** (0.006)
age	0.007*** (0.001)	0.010*** (0.001)	0.018*** (0.002)	-0.0002 (0.001)
age <sup>2</sup>	-0.0001*** (0.00001)	-0.0001*** (0.00001)	-0.0002*** (0.00002)	-0.00001 (0.00001)
married	-0.146*** (0.016)	-0.085*** (0.013)	-0.240*** (0.026)	-0.047*** (0.009)
separated	-0.155*** (0.024)	-0.121*** (0.020)	-0.287*** (0.045)	-0.045*** (0.010)
divorced	-0.073*** (0.021)	-0.086*** (0.017)	-0.108*** (0.035)	-0.004 (0.011)
widowed	-0.163*** (0.017)	-0.114*** (0.013)	-0.365*** (0.028)	-0.030*** (0.009)
family size	-0.023*** (0.004)	-0.021*** (0.002)	-0.044*** (0.006)	-0.006*** (0.002)
number children	-0.009** (0.004)	0.001 (0.002)	-0.022*** (0.007)	-0.008*** (0.003)
black	-0.199*** (0.008)	-0.150*** (0.007)	-0.227*** (0.018)	-0.030*** (0.005)
native	-0.094 (0.066)	-0.129*** (0.047)	-0.060 (0.138)	-0.040 (0.027)
asian	-0.078*** (0.027)	-0.170*** (0.022)	-0.386*** (0.040)	-0.110*** (0.010)
other	-0.219*** (0.020)	-0.270*** (0.013)	-0.520*** (0.037)	-0.105*** (0.009)
mixed	-0.048 (0.059)	-0.194*** (0.044)	-0.074 (0.109)	0.094* (0.056)
migrant	0.039*** (0.008)	0.051*** (0.006)	0.131*** (0.013)	0.023*** (0.004)
log(city population)	0.029*** (0.002)	0.019*** (0.002)	0.038*** (0.004)	0.023*** (0.002)
Constant	-0.089 (0.121)	-0.366*** (0.062)	-0.138 (0.215)	-0.093*** (0.018)
Year FE?	Yes	Yes	Yes	Yes
State FE?	Yes	Yes	Yes	Yes
Observations	1,331,685	1,331,685	1,331,685	1,331,685
R <sup>2</sup>	0.001	0.002	0.002	0.001
Adjusted R <sup>2</sup>	0.001	0.002	0.002	0.001
Residual Std. Error (df = 1331598)	0.041	0.031	0.068	0.022
F Statistic (df = 86; 1331598)	20.235***	32.156***	38.384***	23.083***

Note: All coefficients (and the associated s.e.) are multiplied by 10<sup>2</sup> to ease readability. \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01

**Table A3**  
Interrelation of clusters across artistic domains of famous creatives

	Dependent variable:					
	log(Artist death)	Artist death	log(Author death)	Author death	log(Musician death)	Musician death
	OLS (1)	Poisson (2)	OLS (3)	Poisson (4)	OLS (5)	Poisson (6)
log(IBN author death)	0.19** (0.09)	0.10 (0.07)			0.26*** (0.06)	0.35** (0.14)
log(IBN artist death)			0.13** (0.06)	0.08 (0.14)	0.18*** (0.05)	-0.01 (0.11)
log(IBN actor death)	0.30*** (0.12)	0.08 (0.07)	0.44*** (0.09)	0.52*** (0.15)	0.27*** (0.09)	0.20 (0.14)
log(IBN musician death)	0.28*** (0.09)	0.12* (0.07)	0.28*** (0.07)	0.41** (0.18)		
log(city population)	0.07** (0.03)	0.14*** (0.04)	0.03 (0.03)	0.25** (0.11)	0.08*** (0.02)	0.61*** (0.11)
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes
constant?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	270	270	270	270	270	270
R <sup>2</sup>	0.44		0.50		0.55	
Adjusted R <sup>2</sup>	0.41		0.47		0.52	
Log Likelihood		-273.69		-263.69		-219.37
Akaike Inf. Crit.		567.37		557.37		468.75
Residual Std. Error (df = 255)	0.52		0.43		0.41	
F Statistic (df = 14; 255)	14.45***		18.16***		22.03***	

Note: \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01

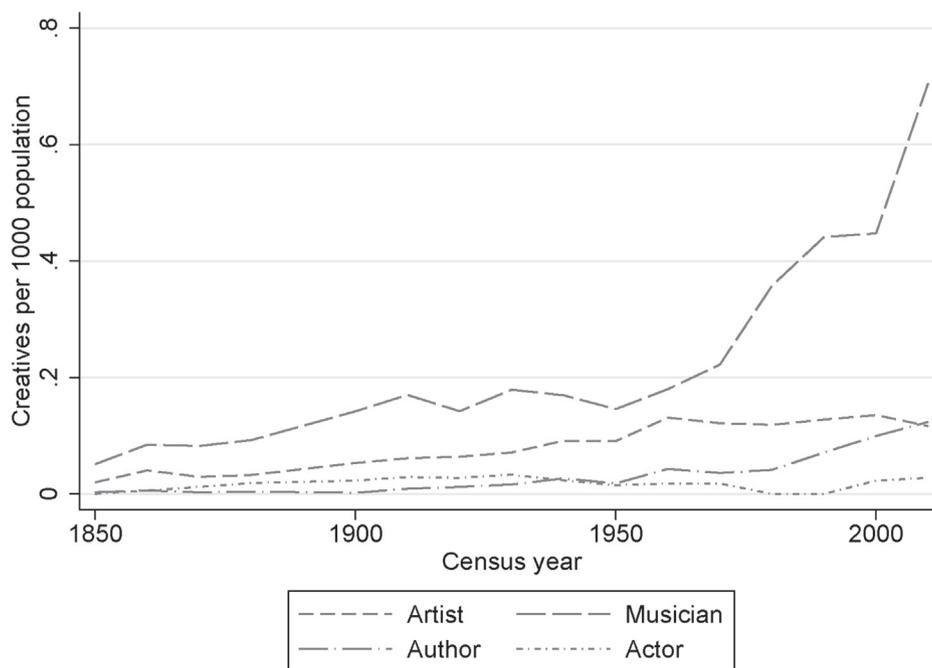


Figure A1 Share of household heads with artistic occupations.

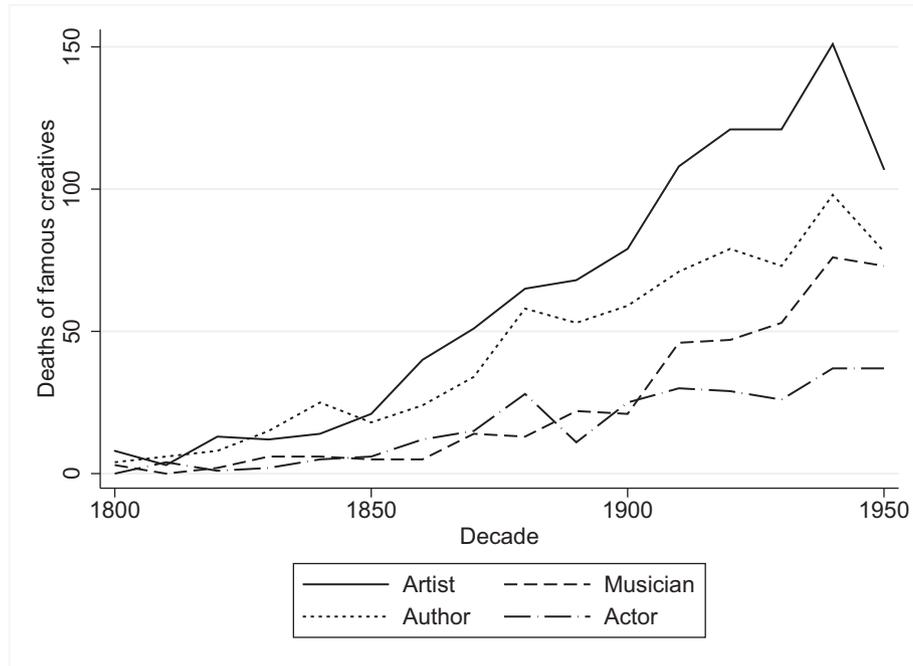


Figure A2 Deaths of famous creatives (IBN) by occupation.

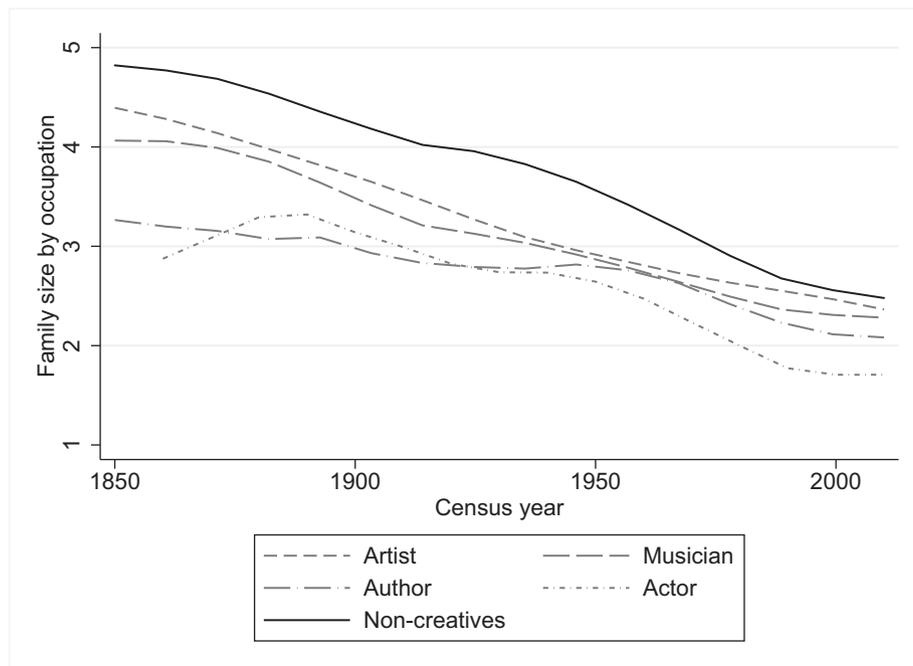


Figure A3 Family size by occupation Note: See Figure 2.

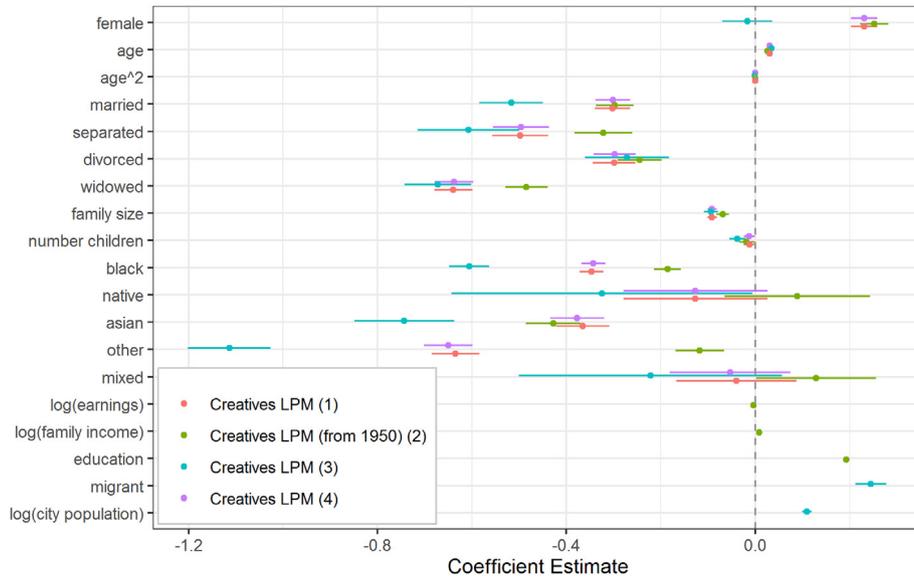


Figure A4 Socio-economic background of creatives. Note: This figure shows the point estimates (dots) and 95 per cent confidence bands based on robust s.e. (intervals) for the models presented in Table 1.

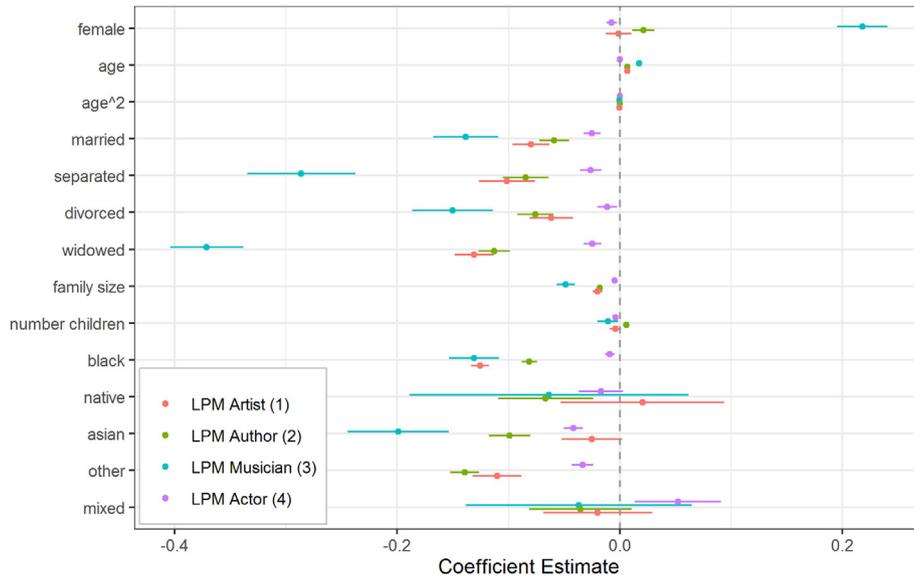


Figure A5 Socio-economic background of creatives by domain. Note: This figure shows the point estimates (dots) and 95 per cent confidence bands based on robust s.e. (intervals) for the models presented in Table 2.

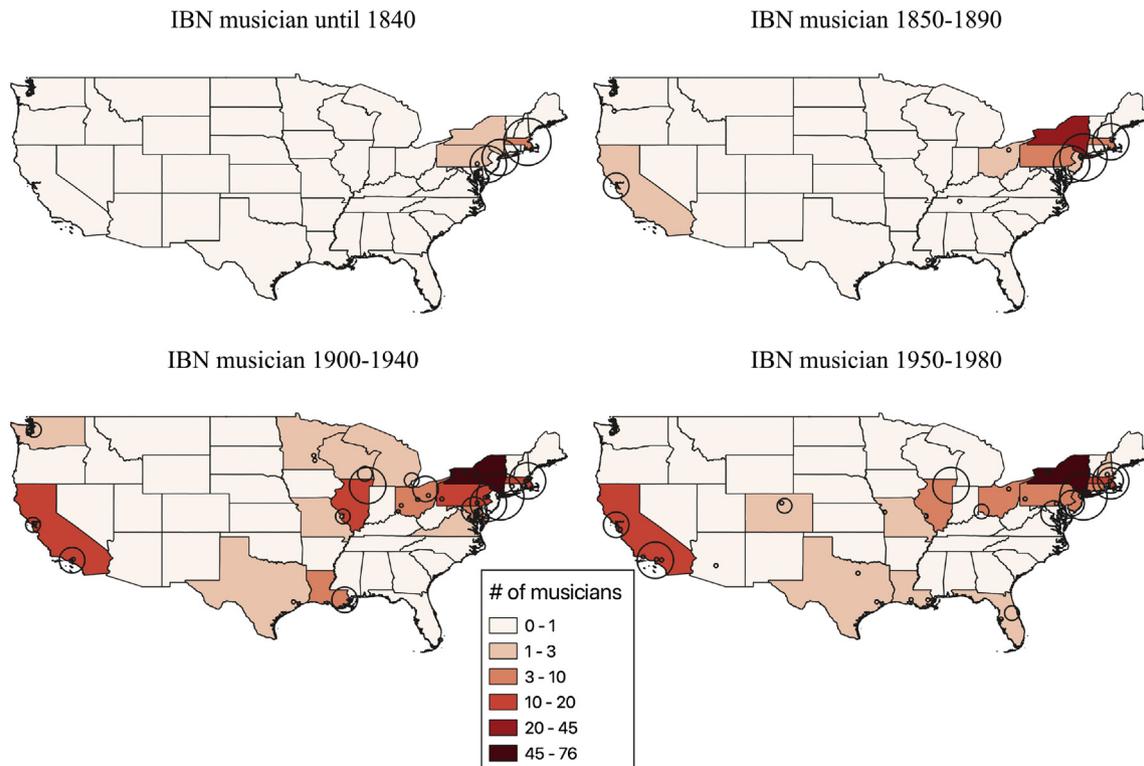


Figure A6 Geographic distribution of deaths of musicians (IBN).

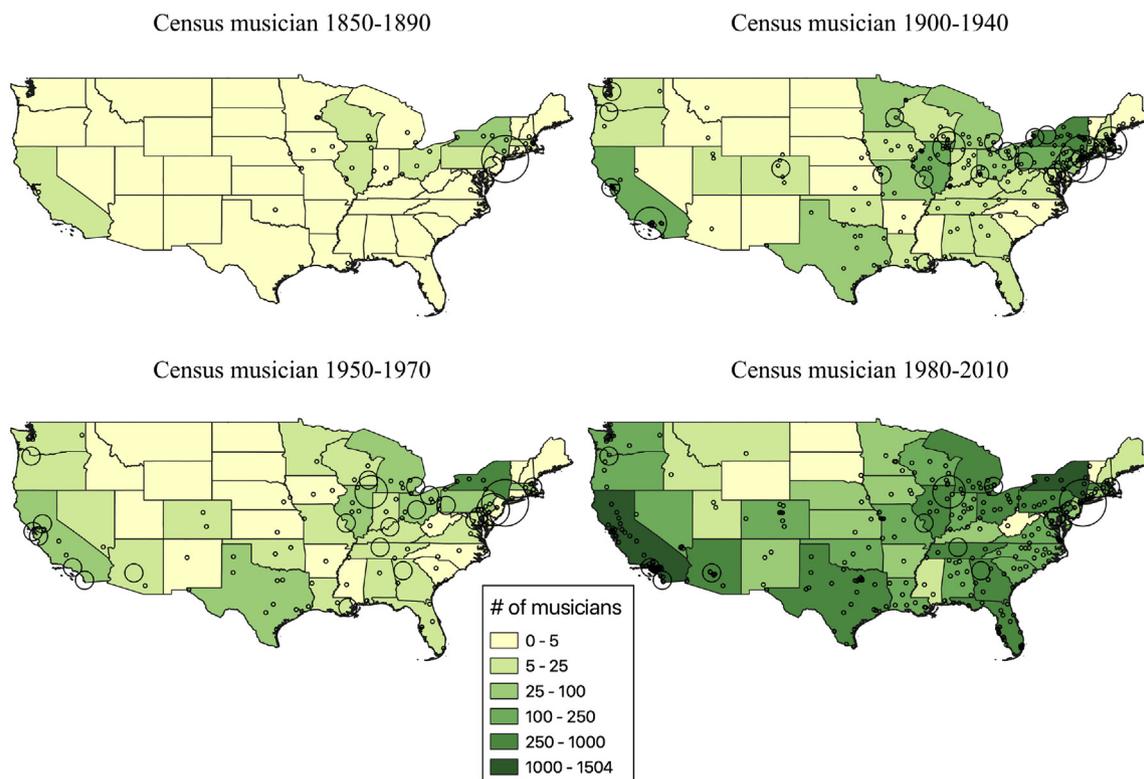


Figure A7 Geographic distribution of musicians (Census).

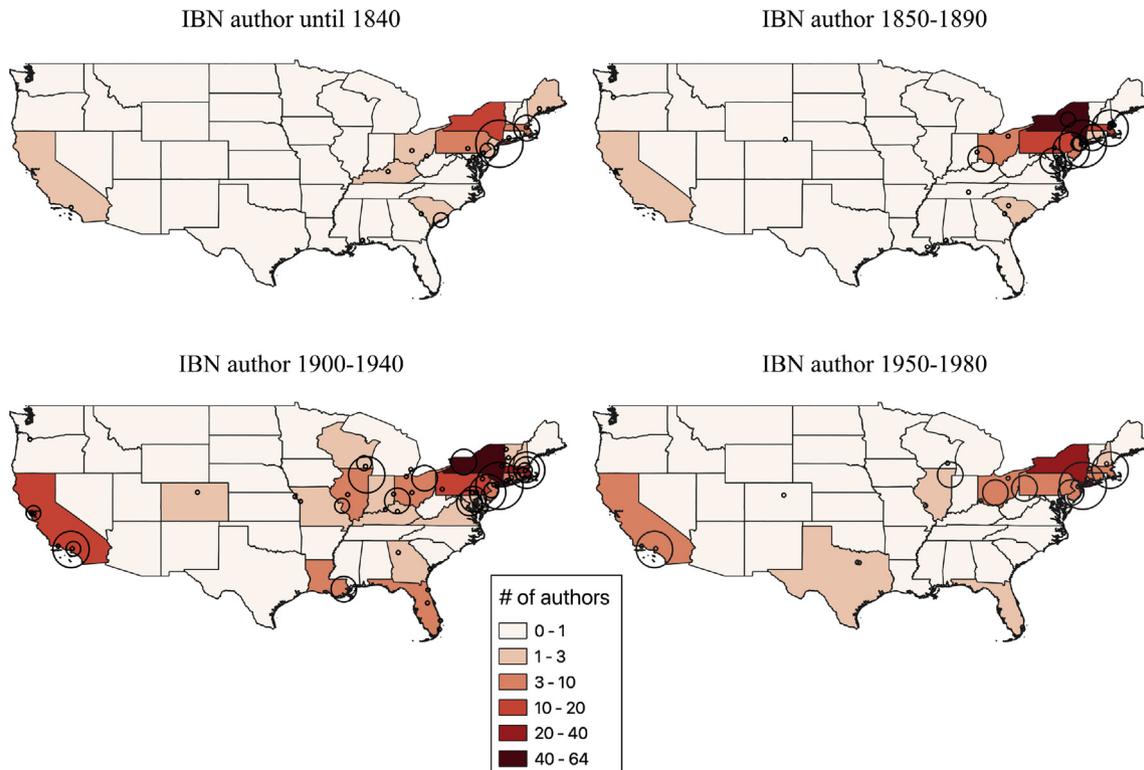


Figure A8 Geographic distribution of deaths of authors (IBN).

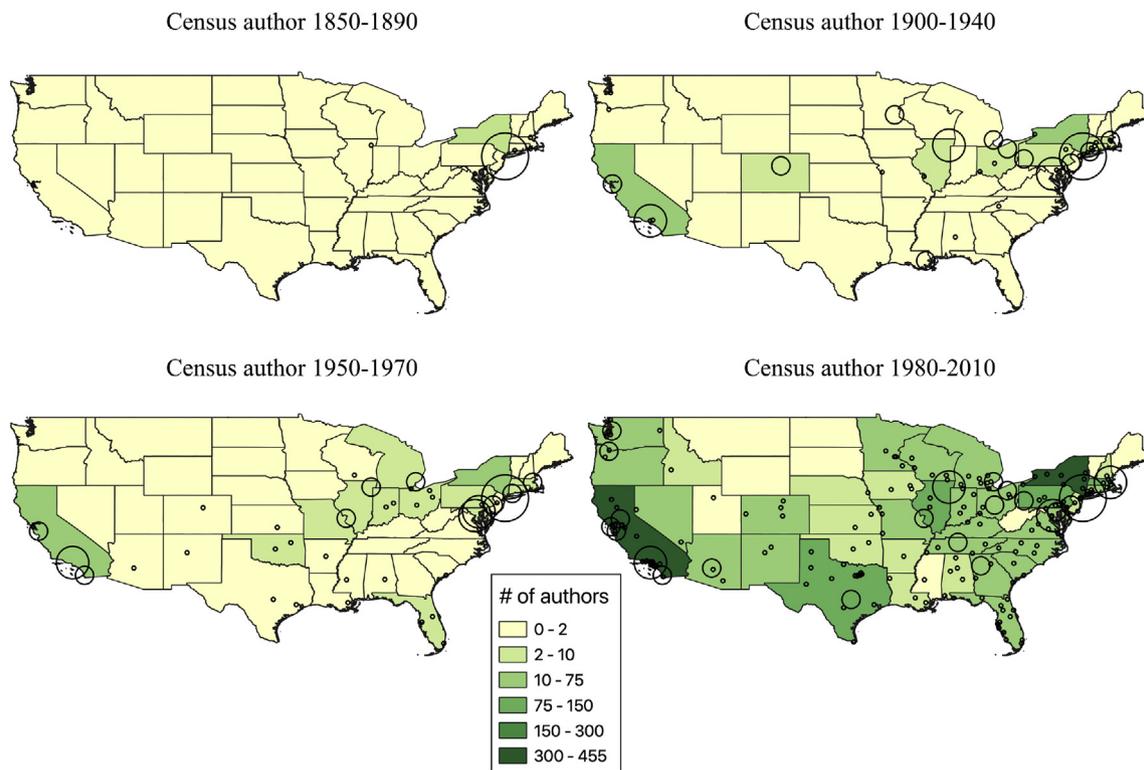


Figure A9 Geographic distribution of authors (Census).

## References

- Alper, Neil O., Wassal, Gregory H., 1998. Artists' labor market experiences: a preliminary analysis using longitudinal data. In: Heikkinen, M., Koshinen, T. (Eds.), *Economics of Artists and Arts Policy: A Selection of Papers*. The Arts Council of Finland, Helsinki.
- Alper, Neil O., Wassal, Gregory H., 2006. Artists' careers and their labor markets. In: Ginsburgh, V., Throsby, D. (Eds.), *Handbook of the Economics of Art and Culture*. Elsevier, Amsterdam.
- Andersen, 2014. Location and spatial clustering of artists. *Reg. Sci. Urban Econ.* 47, 128–137.
- Ateca-Amestoy, Victoria, 2008. Determining heterogeneous behavior for theater attendance. *J. Cult. Econ.* 32 (2), 127–151.
- Ateca-Amestoy, Victoria Maria, Ginsburgh, Victor, Mazza, Isidoro, O'Hagan, John, Prieto-Rodríguez, Juan, 2017. *Enhancing Participation in the Arts in the EU*. Springer, Mannheim.
- Bakker, Gerben, 2005. The decline and fall of the European film industry: sunk costs, market size, and market structure, 1890–1927. *Econ. Hist. Rev.* 58 (2), 310–351.
- Benhamou, Françoise, 2011. Artist's labour markets. In: Ruth, Towse (Ed.), *A Handbook of Cultural Economics*. Edward Elgar Publishing, Cheltenham, pp. 53–58 (chapter 7).
- Blanchflower, David G., Oswald, Andrew J., 1998. What makes an entrepreneur? *J. Labor Econ.* 16 (1), 26–60.
- Borowiecki, Karol Jan, 2013. Geographic clustering and productivity: an instrumental variable approach for classical composers. *J. Urban Econ.* 73 (1), 94–110.
- Borowiecki, Karol Jan, 2015a. Agglomeration economies in classical music. *Pap. Reg. Sci.* 94 (3), 443–468.
- Borowiecki, Karol Jan, 2015b. Historical origins of cultural supply in Italy. *Oxf. Econ. Pap.* 67 (3), 781–805.
- Borowiecki, Karol Jan, 2017. How are you, my dearest mozart? Well-being and creativity of three famous composers based on their letters. *Rev. Econ. Stat.* 99 (4), 591–605.
- Borowiecki, Karol Jan, O'Hagan, John W., 2012. Historical patterns based on automatically extracted data: the case of classical composers. *Hist. Soc. Res.* 37 (2), 298–314.
- Borowiecki, Karol Jan, Graddy, Kathryn, July 2018. Immigrant Artists: Enrichment or Displacement? CEPR Discussion Papers 13070, C.E.P.R. Discussion Papers.
- Borowiecki, Karol Jan, Forbes, Neil, Fresa, Antonella, 2016. *Cultural Heritage in a Changing World*. Springer.
- Compass, 2015. *The Global Startup Ecosystem Ranking 2015*. "The Startup Ecosystem Report Series".
- Crafts, Nicholas, 1985. *British Economic Growth during the Industrial Revolution*. Clarendon Press, Oxford.
- de la Croix, David, Licandro, Omar, 2015. The longevity of famous people from Hammurabi to Einstein. *J. Econ. Growth* 20 (3), 263–303.
- Durant, Gilles, Overman, Henry G., 2005. Testing for localization using micro-geographic data. *Rev. Econ. Stud.* 72 (4), 1077–1106.
- UNCTAD, 2010. *Creative Economy: A Feasible Development Option*. UNCTAD, Geneva.
- Ellison, Glenn, Glaeser, Edward L., 1997. Geographic concentration in U.S. Manufacturing industries: a dashboard approach. *J. Polit. Econ.* 105 (5), 889–927.
- Ellison, Glenn, Glaeser, Edward L., May 1999. The geographic concentration of industry: does natural advantage explain agglomeration? *Am. Econ. Rev.* 89 (2), 311–316.
- Etro, Federico, Pagani, Laura, 2012. The market for paintings in Italy during the seventeenth century. *J. Econ. Hist.* 72, 423–447.
- Fairlie, Robert W., Arnobio Morelix, E.J. Reedy, Russell, Joshua, 2015. *The Kauffman Index 2015: Startup Activity*. National Trends. Kauffman Foundation, Kansas City.
- Falck, Oliver, Fritsch, Michael, Heblich, Stephan, Otto, Anne, 2018. Music in the air: estimating the social return to cultural amenities. *J. Cult. Econ.* 42, 365–391.
- Florida, Richard, 2002. *The Rise of the Creative Class: and How It's Transforming Work, Leisure, Community, and Everyday Life*. Basic Books, New York.
- Fogel, Robert William, 1965. *Railroads and American Economic Growth: Essays in Economic History*. Johns Hopkins Press, Baltimore.
- Fujita, Masahisa, Krugman, Paul R., Venables, Anthony, 1999. *The Spatial Economy: Cities, Regions, and International Trade*. MIT Press, Boston.
- Galenson, David W., 2007. Old Masters and Young Geniuses: the Two Life Cycles of Artistic Creativity. Princeton University Press, Princeton.
- Galenson, David W., Weinberg, Bruce A., 2001. Creating modern art: the changing careers of painters in France from impressionism to cubism. *Am. Econ. Rev.* 91 (4), 1063–1071.
- Galenson, David W., Weinberg, Bruce A., August 2000. Age and the quality of work: the case of modern American painters. *J. Polit. Econ.* 108 (4), 761–777.
- Gergaud, Olivier, Laouenan, Morgane, Wasmer, Etienne, 2016. *A Brief History of Human Time: Exploring a Database of 'notable People'*. *Sciences Po Economics Discussion Papers No 2016-03*.
- Gibson, Campbell, Jung, Kay, 2005. Historical Census Statistics on Population Totals by Race, 1790 to 1990, and by Hispanic Origin, 1970 to 1990, for Large Cities and Other Urban Places in the United States. U.S. Census Bureau Population Division Working Paper No. 76.
- Glaeser, Edward L., 1994. Cities, information and economic growth. *Citiescape* 1 (1), 9–47.
- Glaeser, Edward L., 2003. The new economics of urban and regional growth. In: Clark, M. Feldman G., Gertler, M. (Eds.), *The Oxford Handbook of Economic Geography*. Oxford University Press, Oxford, pp. 83–98.
- Glaeser, Edward L., 2004. Book Review of Richard Florida's "The Rise of the Creative Class".
- Glaeser, Edward L., Kolko, Jed, Saiz, Albert, 2001. Consumer city. *J. Econ. Geogr.* 1 (1), 27–50.
- Glaeser, Edward L., Rosenthal, Stuart S., Strange, William C., 2010. Urban economics and entrepreneurship. *J. Urban Econ.* 67 (1), 1–14.
- Goldin, Claudia, 2006. The quiet revolution that transformed women's employment, education, and family. *Am. Econ. Rev.* 96 (2), 1–21.
- Goldin, Claudia, Katz, Lawrence F., 2010. *The Race between Education and Technology*. Harvard University Press, Boston.
- Graddy, Kathryn, 2013. Taste endures! The rankings of roger de Piles and three centuries of art prices. *J. Econ. Hist.* 73, 765–790.
- Heilbrun, James, Gray, Charles M., 2001. *The Economics of Art and Culture*. Cambridge University Press.
- Hellmanzik, Christiane, February 2010. Location matters: estimating cluster premiums for prominent modern artists. *Eur. Econ. Rev.* 54 (2), 199–218.
- IPUMS, 2015. *Integrated public use Microdata series*. Spring, <https://usa.ipums.org/2015>.
- IPUMS, 2017. *Integrated occupation and industry codes and occupational standing variables in the IPUMS. 2017-04-18* <https://usa.ipums.org/usa/chapter4/chapter4.shtml>.
- Jovanovich, Monica E., Renn, Melissa, 2019. *Corporate Patronage of Art and Architecture in the United States, Late 19th Century to the Present*. Bloomsbury Visual Arts, Bloomsbury.
- Knudsen, Brian, Florida, Richard, Stolarick, Kevin, Gates, Gary, 2008. Density and creativity in U.S. Regions. *Ann. Assoc. Am. Geogr.* 98 (2), 461–478.
- Lovell, Margaretta M., 2007. *Art in a Season of Revolution: Painters, Artisans, and Patrons in Early America*. University of Pennsylvania Press, Philadelphia.
- Ludwig, Arnold M., 1995. *The Price of Greatness: Resolving the Creativity and Madness Controversy*. The Guilford Press, New York.
- Maloney, William F., Valencia Caicedo, Felipe, 2014. Engineers, Innovative Capacity and Development in the Americas. IZA Discussion Paper No. 8271.
- Mitchell, Sara, 2019. London calling? Agglomeration economies in literature since 1700. *J. Urban Econ.* 112, 16–32.
- Newell, Margaret Alan, 2000. The birth of new England in the atlantic economy: from its beginning to 1770. In: Peter, Temin (Ed.), *Engines of Enterprise: an Economic History of New England*. Harvard University Press, Boston, pp. 11–68.
- OECD, 2014. *Resilient Economies for Inclusive Societies*. OECD, Paris.
- O'Hagan, John W., Hellmanzik, Christiane, 2008. Clustering and migration of important visual artist: broad historical evidence. *Hist. Methods* 40 (3), 121–136.
- O'Hagan, John W., Jan Borowiecki, Karol, 2010. Birth location, migration and clustering of important composers: historical patterns. *Hist. Methods* 43 (2), 81–91.
- Puga, Diego, 2010. The magnitude and causes of agglomeration economies. *J. Reg. Sci.* 50 (1), 203–219.
- Rosen, Sherwin, 1981. The economics of superstars. *Am. Econ. Rev.* 71 (5), 845–858.
- Rosenthal, S., Strange, W., 2004. Evidence on the nature and sources of agglomeration economies. In: Henderson, J.V., Thisse, J.F. (Eds.), *Handbook of Regional and Urban Economics*. Elsevier, Amsterdam.
- Sedgwick, John, Pokorny, Michael, 2010. Profitability trends in Hollywood, 1929 to 1999: somebody must know something. *Econ. Hist. Rev.* 63 (1), 56–84.
- Tao, 2019. Agglomeration economies in creative industries. *Reg. Sci. Urban Econ.* 77, 141–154.
- UNCTAD, 2018. *Creative Economy Outlook: Trends in International Trade in Creative Industries*. UNCTAD, Geneva.
- Weber, Clint B., 2012. *The Treasured Collection of Golden Heart Farm: A Biography of Wilhelmina Weber Furlong*. The Weber Furlong Publishing Company, New York.

## Further reading

- Towse, Ruth, 2010. Creativity, copyright and the creative industries paradigm. *Kyklos* 63 (3).
- Throsby, David, 2001. *Economics and Culture*. Cambridge University Press, Cambridge.
- Menger, Pierre-Michel, 2014. *The Economics of Creativity: Art and Achievement under Uncertainty*. Harvard University Press, Boston.
- Grove Music Online, 2013. *Oxford Music Online*. Oxford University Press.
- Burger-Helmchen, Thierry, 2013. *The Economics of Creativity: Ideas, Firms and Markets*. Routledge, London.