WORKING PAPER SERIES

Creativity over Time and Space

Michel Serafinelli, Guido Tabellini

Working Paper n. 608

This Version: November, 2017

Creativity over Time and Space *

Michel Serafinelli[†], Guido Tabellini [‡]

November 2017

Abstract

Creativity is often highly concentrated in time and space, and across different domains. What explains the formation and decay of clusters of creativity? In this paper we match data on thousands of notable individuals born in Europe between the XIth and the XIXth century with historical data on city institutions and population. After documenting several stylized facts, we show that the formation of creative clusters is not preceded by increases in city size or wages. Instead, the emergence of city institutions protecting economic and political freedoms facilitates the attraction and production of creative talent.

Keywords: innovation, agglomeration, political institutions, immigration, gravity. JEL: R10, O10, J61, N13

*We thank Gary Cox, Stefania Gerevini, Massimiliano Marcellino, Enrico Moretti, Marco Tabellini, Francesco Trebbi, Fernando Vega-Redondo, Joachim Voth and seminar participants at U of Edinburgh, U of Toronto, Bocconi, CEA 2016, Rotman UBC/U of T Summer Conference, RCEF 2016, Dalhousie U, U of Ottawa, Copenaghen Business School, McGill, ESEM 2016, SITE, Bristol, U of Guelph, AEA 2017 and CIFAR for helpful comments, and to Igor Cerasa, Tommaso Coen, Viola Corradini, Clemence Idoux and Carlo Medici for excellent research assistance. We thank David Schonholzer and Eric Weese for helping us in defining historical state borders, and Ralf Meisenzahl and Joel Mokyr for sharing with us the data on UK inventors and patents. We are grateful to Bocconi University for financial support.

†Department of Economics, University of Toronto. email: michel.serafinelli@utoronto.ca

‡Department of Economics and IGIER, Università Bocconi; CEPR; CES-Ifo; CIFAR. email: guido.tabellini@unibocconi.it

1 Introduction

Creativity is often highly concentrated in time and space, and across different domains. In the XVth century, Florence was home to an amazing number of groundbreaking innovators in literature, paintings, sculpture, philosophy and science. At the turn of the XIXth century, Vienna hosted pioneers in paintings, medicine, biology, psychology, philosophy, music, who interacted with each other. Baghdad in the IXth century, London in the late XVIth century, Paris in the early XIXth century, San Francisco and New York in the past few decades, are other examples of clusters of creativity and innovations in a number of seemingly unrelated domains (Banks, 1997; Kandel, 2012).

What explains the formation and decay of such clusters of creativity? Are they driven by wealth, by specific features of local institutions, or by mere chance? More generally, asides from these exceptional clusters, how concentrated are creative activities in time and space? Is there co-agglomeration of creative people in different fields? And most important of all, what general lessons can be drawn from the historical analysis of creative clusters, to foster innovation and the production of creative talent? Given the central role of creativity and innovation in human progress and economic development, knowing the answer to these questions is particularly important.

In this paper we analyze data on European creative elites born in the XIth-XIXth centuries. We exploit information on the dates and location of birth and death of notable individuals in different creative endeavours (arts, humanities, science and business) throughout Europe. The source for these data is *Freebase.com*, a large data base owned by Google and coded by Schich et al. (2014), that stores information from a variety of publicly editable sources, most notably Wikipedia. We then match these individual data with a historical data set on European cities and local institutions put together by Bairoch et al. (1988) and by Bosker et al. (2013). Our unit of observation is thus a city in a particular century between the XIth and XIXth centuries.

Notable individuals in creative endeavours are a measure of upper tail talent and human capital. They are more likely to capture radical innovations and creativity, compared to general indicators of human capital. Moreover, these data cover periods and domains where patents data are not available. They are thus suitable to study the formation and decay of creative clusters, something that takes place over the very long run.

Our main variable of interest is the number of famous people born in a city (per 1000 inhabitants) in any given century. Births are a measure of the local production of creative talent. We also look at the number of famous immigrants, i.e. the number of deaths of famous creatives born elsewhere. Unfortunately we don't have information on where these notable individuals did their most important work. Nevertheless, the data compensate this inaccuracy with their breadth in terms of time, geography and disciplines.

We start by exploring the temporal and spatial patterns of the data, documenting the following stylized facts. First, births of creative people and famous immigrants are *spatially concentrated*, generally more so than population. They are also *clustered across disciplines*, as in the prominent examples of Florence and Vienna cited above, and this is true also after conditioning on local observables. Hence spillover effects associated with local proximity and/or local factors (observed or unobserved) are important for creative activities, and operate across disciplines and not just within each field.

Second, births and immigration display *persistence*, but smaller than for population. Cities that are at the frontier of creativity in one period retain an advantage that persists for a while but not indefinitely. Estimating a transition matrix, we also find that persistence of creativity is higher at the bottom of the distribution than at the top. Most small and uncreative cities remain in that condition. But at the top of the distribution there is more reshuffling in creative clusters than for population: while most large cities keep growing and remain large, creative clusters exhibit more change over the centuries. The finding that births and famous immigrants are less persistent than population (a proxy for economic activity), despite being spatially more concentrated, is in line with findings on spatial movements of clusters of innovation (Saxenian, 1994; Duranton, 2007; Kerr, 2010).

Third, the overall spatial proximity of births and the distribution of birth-to-death distances *did not change much* over the centuries. This stability is somewhat surprising, in light of the consolidation of states and the improvements in the means of transportation throughout this period. It suggests that the agglomeration of creative activities is not very sensitive to the cost of transportation and communication, but reflects historically stable forces.

We then address a number of questions about causality. A first natural question is whether creative clusters are influenced by local prosperity, or viceversa. With these non-experimental data, we cannot provide a definitive answer. Nevertheless, we can exploit the temporal pattern in the data, using population as an indicator of local prosperity in a distributed lag model.[1] We fail to reject the null hypothesis of no effect of lagged population on famous births and on famous immigrants. On the other hand, higher births and immigrants predict an increase in city size in the following century. This suggests that local prosperity does not play an important role in the formation or decay of creative clusters, in line with historical evidence on some prominent examples.[2]

[1] Using urban population or city size as an index of prosperity is in line with recent empirical studies on long-term economic growth - see Bosker et al. (2013, p.1418) for references and a discussion.

[2] For instance, Banks (1997) writes: "Elizabethan London (the apogee of English Renaissance - ed.) suffered “dearness without scarcity” (inflation); this fell most heavily on the aristocracy and the very poor. Then the wool trade collapsed, England entered “the worst economic depression in history” (Wilson, 1965), and Parliament anxiously debated means of averting a Bellum Rusticum." Lopez and Miskimin (1962) describe a similar picture for Renaissance Florence, presenting four pieces of evidence for this period: (a) an increase in the fraction of the poor (b) a decline in the city’s population (c) the fact that the Medici bank had around half the capital of the Peruzzi bank in 1340 (d) a

On the contrary, becoming a hub of creative talent may contribute to subsequent economic growth.

Population size is an imperfect measure of economic prosperity, and it is only measured at about the turn of the century. We thus repeated the analysis with historical data on wages of skilled workers, expressed in grams of silver. Here the sample only includes about 30 major European cities, but time is measured in decades rather than centuries, and for most cities the period goes from 1400 to the mid XIXth century. There is no evidence that wages started increasing before famous birth or the arrival of famous immigrants, while the opposite is true: surges in local creativity are associated with subsequent wage increases.

If not local prosperity, what explains the agglomeration of creative individuals in specific localities at different points in time? The rest of the paper explores the effect of local political institutions, and in particular the institutions typical of free cities. We measure city institutions by the dummy variable "Commune" coded by Bosker et al. (2013) and described in the next section. This variable classifies cities based on their political and administrative institutions at about the beginning of each century - this timing is important for our analysis.

As described by Pirenne (2014) and Parker (2004), communal institutions evolved from within the city, and were guided by the aspiration of the urban middle classes to gain freedom and independence from external influence (primarily in opposition to the Church or an external Lord). This had several implications of great relevance for the agglomeration of creative activities. First, Communes guaranteed freedom of movement and from censorship and other personal freedoms; for this reason they often received exiles escaping religious or political persecutions from elsewhere - eg. Burke (2016). Second, Communes protected basic economic freedoms and promoted trade. As emphasized by Cox (2017), self-governing cities were often in competition to attract merchants as well as human and financial capital. This made free cities a dynamic social environment, in frequent contact with other trading centers and open to innovations. Third, to consolidate their autonomy and territorial expansions and to strengthen identification with the community, independent cities promoted works of art that could become symbols of the city and enhance its prestige - see for instance Paoletti and Radke (2005), Connell (2002) on Florence and Norman (1999) on Siena. All these features of Communes encouraged creative activities and radical innovations, compared to cities that did not have these special institutions.

We first study the production of creative talent, measured by births of creative individuals in each century (per 1000 inhabitants). Here we do an event study analysis, where the treatment again is becoming a Commune. The main result of the event study is illustrated in Figure 1. Becoming a Commune at the beginning of the century (date 0 in Figure 1) is associated with a 5 percentage point increase in the birth of creative individuals (per 1000 inhabitants) during the current century, with an additional increase in the subsequent century.

decrease in wool production (although Cipolla (1964) points out that silk production grew in the same period).

To address the possibility of endogenous transitions into Commune status, we rely on instrumental variable estimation, exploiting the fact that the diffusion of communal institutions occurred in regional waves. It started in Northern Italy between the XIth and XIIth centuries, then it spread to Southern France; nearly at the same time as in Italy, Communes also began to appear in the Flanders and in Northern France (Pirenne 1925). Independent cities emerged in Germany in the XIII century, also in association with migration to the East where imperial control was weaker (Parker 2004). According to historians, these patterns reflect regional factors that influenced the transition into Commune, such as contagion effects in the aspirations of cities, or a vacuum of regional powers (Parker 2004). We thus instrument Commune with the incidence of Commune in the remainder of the region (defined by current NUTS 1 administrative borders), adapting a strategy introduced by Persson and Tabellini (2009) and Acemoglu et al. (2014) in their analysis of democratic transitions in a panel of countries. The identifying assumption is that, conditional on time and city fixed effects and other covariates, the regional waves of institutional transitions influence city level creativity only through the city political institutions. To make this assumption less restrictive, we also control for regional waves of creativity (measured by the spatial lag of births of notable individuals in the region) and for regional urban potential (namely the potential for urban growth in the region, measured by the spatial lag of urban population in the region as in Cox 2017 and De Vries 1984). These IV estimates confirm the results of the event study and are very robust. On average, births of creative people increase by 10 percentage points or more during the century, upon a transition into Commune.

We also explore whether other local institutions, such as being the seat of religious authorities and (possibly endogenous) universities foster or hinder the formation and attraction of upper tail human capital. A priori, the effect of religious institutions could go either way: on the one hand the Church was an important sponsor of works of arts and a source of economic and political power; on the other hand, religious authorities also promoted censorship and opposed scientific discoveries. The evidence suggest that the negative effects prevail: being the seat of religious authority is negatively associated with births of creative individuals. Universities instead are positively associated with creative births. The positive effect of becoming a Commune on local creativity is not affected by controlling for these institutions.

Next, we turn to studying whether the emergence (or disappearance) of Communes and of other city institutions (having a university, being the seat of religious authorities or a state capital) is correlated with the flow of famous immigrants. Using bilateral data on city of birth and of death by century, we estimate a gravity model of migration, identifying the causal effects of becoming a Commune on immigrants through a diff-in-diff methodology. The main finding here is that economic and political freedoms make a city an attractive destination: becoming a Commune is associated with an increase in the inflow of notable individuals, that almost doubles in size. On the contrary cities that become the host of religious powers (Bishop or Archbishop) are less likely to attract them, consistently with the idea that hierarchy and censorship stifles creativity.

A possible concern is that our data suffer from large measurement error. Certainly we have better records of more recent individuals, so that intertemporal comparisons have to be interpreted with care. Moreover, it is possible that information on individuals that were closer to political and economic centers is more readily available. Finally, these data treat all these notable individuals equally, without weighting them by their achievements and visibility. To address these concerns, we show that our results are robust to using similar data that weight individuals by their influence (Yu et al., 2016). Moreover, as discussed in the next section, our indicators are correlated with measures of technological innovation collected by Meisenzahl and Mokyr (2011) for the UK in the period leading to the Industrial Revolution.

Our paper is related to a large existing literature. The link between democratic institutions and innovation has been studied in the context of economic growth. In particular, Acemoglu and Robinson (2012) discuss of how bad political institutions can block innovation, and Acemoglu (2008) argue that democracies may be better able to benefit from new technologies than oligarchic societies, by letting individuals with comparative advantage in the new technology enter entrepreneurship.

A related body of work investigates how specific societal attitudes or specific features of the local environment affect innovation, with findings that are very consistent with ours. In particular, Bénabou et al. (2015a) and Bénabou et al. (2015b) show that religiosity is negatively correlated with indicators of innovation (measured at the country and at the individual level). See also Saxenian (1994), Florida (2005), Falck et al. (2011), Acemoglu et al. (2014), Akcigit et al. (2017a) and Akcigit et al. (2017b). Most of these studies generally focus on recent periods, however, and they often exploit patent data; in contrast, we study a very long historical period with data on creative elites.

The idea that scientific innovations associated with the industrial revolution reflect the flow of ideas across Europe is at the heart of the important work by Mokyr (2016). Cox (2017) argues that local autonomy and economic freedoms were key to European growth in the period leading to the industrial revolution, because they set in motion competition between decentralized localities, that in turn favored commerce and economic growth. Our finding of a positive effect of Commune on the production and attraction of creative talent complements his results. Our study is also motivated by the literature on upper tail human capital and the industrial revolution (Mokyr, 2009; Meisenzahl and Mokyr, 2011; Squicciarini and Voigtländer, 2014). A similar historical perspective to ours is taken by a set of studies using microdata on upper tail human capital. In particular Gergaud et al. (2016) analyze a database of more than one million famous individuals and more than seven million places associated with them throughout human history (3000BCE-2015AD). Other studies on similar topics include Schich et al. (2014), De la Croix and Licandro (2015) and Dittmar and Meisenzahl (2016). Relative to this work, we focus more specifically on creative individuals, on spatial patterns (using the city as the unit of analysis), on the flow of ideas, and on the effects of local self-government institutions on the formation of creative clusters. Regarding this last aspect, a related paper is by Donges et al. (2016) who show that German counties whose institutions are more inclusive as a consequence of the French occupation after 1789 turn out to be more innovative (in terms of patents per capita).

Our paper is also related to a growing literature on innovation. The first relevant body of work, recently surveyed in Carlino and Kerr (2015), analyzes the connections between agglomeration and innovation (Audretsch and Feldman, 1996; Carlino et al., 2012). In addition, our study is related to the literature, recently reviewed Combes and Gobillon (2015), empirically examining agglomeration advantages, and in particular the microfoundations of such advantages based on knowledge spillovers (Jaffe et al., 1993; Rosenthal and Strange, 2003; Moretti, 2004; Arzaghi and Henderson, 2008; Serafinelli, 2017).

The outline of the paper is as follows. The next section defines the data and their sources. Section 3 describes the main patterns in the data. Section 4 describes the results on the determinants of the birth of famous creatives, and in particular on the role of local institutions. In Section 5 we study more in detail the determinants of the migration of famous people between European cities. Section 6 discusses future directions and concludes.

2 Data and Variable Construction.

The data used in this paper cover Europe between the XIth and the XIXth century. Our data set combines information on notable individuals with population and socio-institutional variables at the city level.

Notable Individuals The records on notable individuals come from *Freebase.com*, as coded by Schich et al. (2014). Freebase is a "large Google-owned knowledge base that is publicly editable and available under a Creative Commons Attribution (CC-BY) license, which allows for both sharing and remixing of the data" (Schich et al. Supplementary Material, p.2). It stores information from a variety of sources, most notably Wikipedia, and contains dates and locations of birth and death, as well as occupations, of notable people.[3] Location information in the records is geocoded, making good quality latitude and longitude data available. Notability of people is "simply defined as the curatorial decision of inclusion" in the (partly crowd-sourced) Freebase (Schich et al., 2014,

[3]See Schich et al. (2014) for more detail on the nature, acquisition and cleaning of the records on notable individuals.

p.558). Using these records, we identify 40,980 notable individuals who can be matched by city of birth and/or city of death to the Bairoch et al. (1988) sample (described below).[4] Of these, we retain 21,906 individuals that became famous thanks to their creative endeavours in the following domains: arts (performing and non-performing), humanities and science, and business. Table 1 provides a count of the individuals active in each domain. The last row reports the total number of creative individuals.[5] Famous creatives in performing arts include: actors, singers, musician, playwrights; in non-performing arts: writers, novelists, journalists, composers, authors, architects; in humanities and science: mathematicians, physicians, philosophers, scientists, physicists, chemists, historians; in business: entrepreneurs, engineers, business-persons, sailors, managers.[6] The vast majority (86.2%) of the individuals studied did all of their work in a single area (e.g. Galileo Galilei in the Humanities and Sciences). Another 12.6% of our famous creatives had achievements in two fields (e.g. William Shakespeare in both the Performing arts and Non-performing arts), while 1.1% of individuals achieved fame in three different categories (e.g. Voltaire in the Humanities and Sciences, Performing arts, and Non-performing arts). A few individuals, including Leonardo da Vinci, were prominent in all four fields. A further breakdown of polymaths is available in Table A.1.

Using this information, we define the variable $Births_{ct}$ as the number of famous creatives born in city c during century t, per 1000 inhabitants at the beginning of each century. This variable measures the city production of upper-tail creative human capital. As a measure of attraction of upper-tail human capital, we define the variable $Immigrants_{ct}$ as the number of deaths in city c of famous creatives born elsewhere during century t, also per 1000 inhabitants. For individuals who die in a century different from that of birth, we face a problem. Ideally, we would like to attribute these famous people to the century in which their migration decision was taken. Hence, using the century of death risks erroneously posticipating their migration decision, while using the century of birth risks erroneously anticipating their migration decision. One of our goals is to estimate the effect of a change in city institutions on migrations. Using the century of birth minimizes the risk of erroneously attributing to institutional changes outcomes that actually took place earlier. Thus, unless otherwise noted, we use the more conservative assumption and measure immigrants by *century of birth*.

Table 2 reports descriptive information on the number of famous creatives (unscaled by city

[4] If an individual is born or dies in a small city not included in the Bairoch et al. (1988) sample, we assign it to the closest city in the sample, within a threshold of 71 Km (corresponding to the 95th percentile of the distance distribution).

[5] The last row of Table 1 does not correspond to the sum of the rows above, because some famous creatives were active in more than one domain.

[6] The remaining individuals in the Freebase.com database but not included in our analysis were either unclassifiable, or operated in domains were creativity is less important, such as sports, politics, religion or the military. The number of unclassfied notable individuals from Freebase.com is actually 25,580. For about 10,500 of these unclassified entries, we were able to establish the main area of work by scraping the internet.

population) born and immigrated in all the cities and total population in our sample - the city sample by Bairoch et al. (1988) described below. The Table shows how we have better records of more recent individuals. It also shows that there is substantial mobility of famous creatives in each century: the number of immigrants is a large fraction of the number of births.

Below we also discuss the robustness of our results to a similar data set that weights individuals by citations. Specifically, Yu et al. (2016) collect records of individuals present in more than 25 languages in Wikipedia, which contain dates and locations of birth, as well as occupations. Compared to the dataset constructed by Schich et al (2014), this dataset is significantly smaller and does not include the place of death of notable individuals. However, it is manually verified, and is enriched with the Historical Popularity Index (HPI), a measure that integrates information on the number of languages in which a biography is present in Wikipedia, the time since birth, and the number of page-views between 2008 and 2013. Using these records, we identify 1583 notable individuals who (a) can be matched by city of birth to the Bairoch et al. (1988) sample and (b) became famous thanks to their creative endeavours in the arts (performing and non-performing), humanities and science, and business. We define the variable *Births, Yu et al.*$_{ct}$ as the HPI-weighted number of famous creatives born in city c during century t, per 1000 inhabitants at the beginning of each century. The correlation between *Births* and *Births, Yu et al.* is 0.48.

We have also compared *Births* with records from the *Index Bio-bibliographicus Notorum Hominum (IBN)*, which was compiled from around 3000 biographical sources (mainly dictionaries and encyclopedias) with year of publication between 1600 and 1980 - see De la Croix and Licandro (2015). We define the variable *Births, IBN*.$_{ct}$ as the number of famous creatives in IBN born in city c during century t, per 1000 inhabitants at the beginning of each century. Unfortunately the geocoding information available to us[7] covers only 59 cities in our estimation sample. We therefore cannot perform our main estimation analysis with these data. However, it is interesting to note that the correlation between *Births* and *Births, IBN* is 0.57.

Correlation with measures of technological innovation A possible concern with using *Births* as a proxy for local creativity is that it may be tainted by measurement error and reporting bias. We thus investigate whether the variable *Births* is correlated with available measures of local innovation. Meisenzahl and Mokyr (2011) collected data on mechanics and engineers born in UK between 1660 and 1830, and on the patents that they created.[8] Many of these individuals were not great inventors, but rather highly skilled and able craftsmen, who adapted new technologies and

[7]We downloaded the data from the *Journal of Economic Growth* website.

[8]Meisenzahl and Mokyr (2011, p.42) argue that Britain's industrial precocity owed a great deal to "the technical competence of the British mechanical elite that was able to tweak and implement the great ideas and turn them into economic realities".

provided micro innovations. Almost one third of these 747 innovators in Meisenzahl and Mokyr (2011) are also included in our sample of famous creatives - probably the individuals with greater accomplishments. But interestingly, our variable *Births* is also correlated with the birthplace of the remaining mechanics and engineers, included in Meisenzahl and Mokyr (2011) but not classified as notable individuals in our sample.

Specifically, from the Oxford Dictionary of National Biography we obtained the place of birth of the innovators in Meisenzahl and Mokyr (2011) not in our sample, and matched it with the cities in our data set. Let *Inventors* be the number of mechanics and engineers born in a city during a century (per 1000 inhabitants), and *Patents* be their number of patents (also per 1000 inhabitants), both variables restricted to the individuals in Meisenzahl and Mokyr (2011) that are *not* in our data set. We then regress *Log(1+Inventors)* and *Log(1+Patents)* on *Log (1+Births)* plus other city observables corresponding to all the other city covariates described below (namely, the variables *Large state, Bishop, Archbishop, Capital, Plundered, Commune, Population, University)*. As shown in Figures A.1 and A.2, that depict the added-variable plots, our variable *Births* is positively and significantly correlated with both dependent variables. Specifically, a 10% change in *Births* is associated with a 1.8% change in *Inventors* and a 1.6% change in *Patents.* Thus, for the UK between the XVII and XIX centuries, our proxy for production of creative talent predicts indicators of local innovation, and in particular of the technological micro-innovations that contributed to the industrial revolution.

European Cities City population is measured at about the beginning of each century. The source is Bairoch et al. (1988). This is a wide-ranging database with information on 2,200 European cities that reached 5,000 residents between 800 and 1,800. Given the scarcity of data on notable individuals in the very early part of the sample, we restrict the analysis to the period between the XI and the XIX centuries, interpolating population for the missing century 1100.

Information on socio-economic and institutional variables comes from Bosker et al. (2013), who, for a subset of the cities in Bairoch et al. (1988), assembled a large database covering an extensive array of institutional characteristics of European cities between the IXth and the XIXth. The sample covered by Bosker et al. (2013) includes all cities in Bairoch et al. (1988) that reached 10 000 inhabitants between 800 and 1,800. In our analysis on the effect of institutions, our sample is always that of Bosker et al. (2013), except in section 3 where we describe the main features of the data and we rely on the larger sample of cities by Bairoch et al. (1988). Note that the data coded by Bosker et al. (2013) seek to capture the institutions that were in place at the beginning of each century. Thus, institutional changes that took place during century t would generally show up in century $t + 1$. In other words, although undoubtedly measured with much timing error, our data are more likely to erroneously postpone institutional changes rather than viceversa, compared to our outcomes of interest (births and immigrations of famous people) - a conservative feature of the data definition, given our goal of estimating the effects of institutional changes.

Our main institutional variable of interest is whether a city had a form of self-governance that gave it some autonomy and could constrain the dominant role of the church, state or feudal lords. This is captured by the dummy variable $Commune_{ct}$ coded by Bosker et al. (2013). This variable measures the extent of local participatory government at the beginning of each century. Typically Communes had autonomy in the regulation of commerce, taxation and other administrative activities. Communal institutions also guaranteed economic and personal freedoms and enforced the rule of law, also through the evolution of civil and penal codes. Besides check and balances on executive authority, Communal political institutions also had forms of limited representative democracy (Pirenne 1925).

Forms of local participative government began to develop in the Xth and XIth centuries when Europe was politically fragmented, after the fall of the Carolingian Empire. In the power vacuum that ensued, cities could claim a kind of self-rule that was frequently recognized by the sovereign in return for taxes or loyalty (Jones, 2003). This form of self-government emerged in the XIth and XIIth centuries in Southern Europe and spread elsewhere in the following centuries (Figure A.6 in the Appendix). The status of Commune is not irreversible, and we observe transitions in both directions, though transitions into the status of Commune (Figure A.7 in the Appendix) are much more frequent that transitions out of it (Figure A.8 in the Appendix). The XIIth century is the period with the highest number of transitions into the status of Commune, and the highest incidence of Commune in our sample is observed at the beginning of the XVIth century.[9]

Notable people could also be attracted by (or be born in) cities that had universities or that were the location of political or religious power, although religion also led to persecutions and hence could also expel rather than attract innovative individuals. To capture these features, we rely on the following variables, also in the data set by Bosker et al. (2013): whether a city has a university *(University)*, and three variables indicating a city's status in the political and ecclesiastical hierarchy, namely a dummy variable indicating whether a city is the seat of a bishop *(Bishop)*, is the seat of an archbishop *(Archbishop)* and is a capital city *(Capital)*. We also make use of variables, constructed by Bosker et al. (2013), indicating the number of times a city was plundered in

[9]During the XVth century, Communes in Italy started to grant long-term authority to a Lord or Signore. As noted by Guiso et al. (2016):

> "In several cases the Signoria retained the fundamental institutions of the commune, including the principle that power originated from the people and was to be exercised in the people's name. In cities such as Florence and Genoa, the Signoria also preserved the political institutions and the personal liberties that had characterized the commune period."

In this regard the Signoria was an evolution of the Commune (Prezzolini, 1948; Chittolini, 1999). Nevertheless, Bosker et al. (2013) code the transitions into Signoria as a loss of status of Commune.

the previous century *(Plundered)*, whether it is ruled by a large state *(Large state)*, and whether it belongs to a non-absolutist state (*Non-absolutist state*). Table 3 reports summary statistics for all these variables including births and immigrants for the 2137 cities in our sample.

To study the time series properties and their correlation with economic prosperity, we also collected data on the average wage of skilled worker expressed in grams of silver per day from a variety of sources.[10] The main source is Allen (2001), who in turn relies also on other studies. In addition, we used the data gathered by Bennassar (1999), Boulton (1996), Cabourdin (1968), Feliu (1991), Gibson and Smout (1995), Rappaport (2002), and Scholliers and Avondts (1977). These data are only available for 28 major European cities for a long enough period, but they are yearly and they cover several centuries. Table A.2 in the appendix lists the city and years included in our sample. To reduce measurement error and minimize missing observations, we express the wage as a 10-year average (called *Wage*), and obtain an unbalanced panel that covers the period 1260-1890.

Finally, as discussed below, in some specifications we define region-specific or nation-specific variables. They all refer to current administrative borders, as defined by Eurostat. NUTS 1 refers to macro-regions, NUTS 2 to regions. Other geographic variables are defined in context below, when we introduce them.

3 Stylized Facts

In this section we explore the spatial and temporal patterns of the data, documenting several stylized facts. Our goal here is not to test specific hypothesis or establish causality, but to describe the main features of the data. Nevertheless, it is useful to be guided by a simple conceptual framework. Broadly speaking, we can think of three forces that contribute to the formation and decay of creative clusters. (i) Shocks to the *local supply* of talented and creative individuals, such as due to education, or to a political and social environment that encourages innovation, or wars and negative shocks (eg. persecutions) in neighboring localities that induce an inflow of talented immigrants. In particular, the production and attraction of creative talent and innovators may be positively affected by institutions that protect freedoms and basic human rights and allow open exchange of ideas, such as the variable Commune. (ii) Shocks to the *local demand* for the services of creative individuals, for instance coming from an enlightened local lord or a rich bourgeoisie. (iii) *Spillover effects* amongst creative elites. Being close to other creative individuals could increase individual productivity through learning, exchange of ideas, or role model effects. For instance, see Bell et al. (2015). But spillover effects could also be negative, for instance through pecuniary externalities, if the demand for creative services is predominantly local and creative individuals in similar domains are good substitutes for each other.

[10]Skilled workers are typically building craftsmen, carpenters, or, more in general, masters and journeymen.

Spatial agglomeration How concentrated are famous creatives in space? Figure 2 displays the spatial distribution of *Births* in the XVth century, the middle of our sample. Darker tones indicate a larger number of *Births*, while population size is captured by the circle diameter. Famous *Births* are shown to be concentrated in a subset of the cities, not always those with larger populations. Amongst the large cities, Florence, Nuremberg and Siena have the most births of creatives, per 1000 inhabitants. Figure 3 displays the spatial distribution of *Births* in the XIXth century, the end of our sample period. Now many more cities are included in the sample, and the darker tones have shifted to Northern Europe and the UK. The spatial distribution of *Immigrants* displays similar patterns - cf. Figures A.4 and A.5: Florence and Rome have the largest number of famous immigrants per capita in the XVth century. *Births* and *Immigrants* are positively correlated: their correlation coefficient is 0.22 in the XVth century, 0.5 in the XIXth century, and 0.51 for the whole sample period. Thus, cities that give birth to creative individuals also tend to attract famous immigrants - in this sense, one can speak of clusters of creativity. Finally. Figure A.3 in the Appendix illustrates the distribution of *Births* in the XIXth century (histograms relative to other centuries are similar), displaying a large dispersion.[11]

How did spatial concentration evolve over time? Figure 4 plots the coefficient of variation of *Births*, *Immigrants* and population between cities, in each century between 1300 and 1800. A higher coefficient of variation indicates more geographic concentration (a plot of Gini coefficients is very similar). Recall that *Births* and *Immigrants* are expressed per capita. The following facts stand out. First, *Immigrants* are always more spatially concentrated than both population and *Births*, presumably due to sorting. Second, until 1600 *Births* were also more spatially concentrated than population. These features suggest that local factors or spillovers associated with spatial proximity were particularly important for creative activities. Third, while early on the tendency for famous creatives has been towards less concentration (convergence) over time, the spatial concentration and the spatial patterns of famous creatives did not change much between 1500 and 1800, despite the consolidation of states and the improvements in the means of transportation, suggesting that the forces behind agglomeration of creative activities are historically stable. Nevertheless, *Immigrants* were more concentrated in 1700, a century of significant innovations in several domains.

Persistence Next, we analyze the persistence of clusters of famous creatives. A comparison of Figure 2 and 3 suggests that there is some spatial movement of clusters over time. To further explore the temporal patterns of the data, Table 4 displays Spearman's rank correlation coefficients for *Births*, *Immigrants* and Population over each consecutive century (each row reports Spearman's ρ between the same variable measured in t and in $t + 1$). The Spearman's ρ for *Births* and *Im-*

[11] See Table A.3 and A.4 in the Appendix for descriptive statistics for *Births, Immigrants* and population in the XVth and the XIXth century.

migrants increases over time but it remains below 0.5 until the last century. *Births* are generally less correlated over time than *Immigrants*, and both are much less auto-correlated than population - though this may also reflect larger measurement error in famous creatives than for population. Table 4 suggests that cities at the frontier of creativity have an advantage, but generally not strong enough to guarantee dominance in creativity in the next century.

To complement this analysis, we estimate Markov transition matrices for *Births*, *Immigrants* and population. Specifically, for each variable (*Births*, *Immigrants* and population) and each century, we partition cities in five groups: the first group includes cities that in a given century had a value of zero for that variable; the remaining groups correspond to the quartiles of the distribution in any given century, conditional on being positive. Table 5 displays the probability that a city transits from the row group to the column group in the next century, estimated by Maximum Likelihood (assuming that transition probabilities have remained constant over time). Thus, with regard to *Births*, the first row says that a city that had 0 Birth in century t has a 0.61 probability of retaining 0 *Births* in $t+1$, it has a 0.13 probability of being in the first quartile of cities with positive Births in $t+1$, and so on.

For all variables, the top left and bottom right cells in Table 5 are the largest, indicating that the probability of remaining in the bottom and top groups, conditional on being there, is the highest. For cities at the top of the distribution, there is strong evidence of more persistence in population than in famous creatives: a city that belongs to the fourth quartile in population has a probability of 0.79 of remaining there in the next century, while for *Immigrants* this probability is 0.49, and for *Births* it is 0.43. On the other hand, at the bottom of the distribution persistence is roughly similar for famous creatives and for population: cities that have just a few famous creatives and belong to the first quartile have a probability of remaining there or falling in group 0 of about 0.55 for *Births* and *Immigrants*; the corresponding value for population is 0.5. In other words, emerging as a large city or a creative cluster is an unlikely event for cities that start out at the bottom of the distribution: most small and uncreative cities remain in that condition. But at the top of the distribution there is more reshuffling in creative clusters than for population: while most large cities keep growing and remain large, creative clusters exhibit more change over the centuries. This pattern conforms with anecdotal evidence about the rise and decline of creativity in cities like Baghdad, Florence, Rome and Vienna.

Notwithstanding the spatial movement of clusters and the changes in the coefficients of variation displayed in Figure 4, the overall pattern of spatial proximity of creatives is quite stable over time, despite the consolidation of states and the improvements in the means of transportation throughout these centuries. This can be seen in Figure 5 and 6. The former displays the distribution of the distance between places of birth of every pair of famous creatives born in the same century, for different centuries. The latter displays the distribution of birth-to-death distances. Both distributions remained quite stable in different centuries, despite the changes in the cost of transportation and communication during this period.

Agglomeration across disciplines We then explore whether creative clusters tend to be specialized or diverse. We want to know whether spillover effects and local factors (observed or unobserved) operate across or within disciplines. We thus estimate the matrix of pairwise correlation coefficients of famous people by discipline, for both *Births* and *Immigrants.* Disciplines are disaggregated in performing arts, non-performing arts, humanities and science, and business. Table 6 reports the results, in Panel A for *Births* and in Panel B for *Immigrants*. In the first line of each matrix, we condition on century dummy variables, to make variables comparable over time; thus, we first regress famous people in each discipline (*Births* or *Immigrants*) on a full set of century dummy variables, and then estimate the correlation coefficients across disciplines of the estimated residuals. In the second line of each matrix we condition on both century dummy variables and the set of observable city characteristics described above.[12] Thus, the second line of each matrix is a measure of co-agglomeration across disciplines due to either unobserved common local factors or spillover effects, while the first raw also includes the effect of observables. All correlation coefficients are positive and significant and quantitatively large, indicating that creative elites are clustered across disciplines, as in the prominent examples of Florence and Vienna cited above. Correlation coefficients are very similar with or without conditioning on the full set of observables, implying that common observable shocks are not responsible for the correlations. For instance, based on Panel A of Table 6, if we compare two cities with the same observables, but city A has 10 p.p. more *Births* of non-performing artists per 1000 inhabitants than city B, then on average city A also has 5.2 p.p. more *Births* in humanities and sciences compared to city B. This suggests that spillover effects and/or unobserved local factors operate across disciplines and not just within each field. Note that correlations tend to be somehow stronger for *Immigrants* than for *Births*.

What moves first: population or famous creatives? Given the patterns discussed so far, a natural question is whether creative clusters are influenced by local prosperity. With these non-experimental data, we cannot provide a definitive answer. Nevertheless, we can exploit the temporal variation in the data, using city size as an indicator of local prosperity in a distributed lag model that only exploits within city correlations.

Specifically, we estimate the following specification:

$$Y_{ct} = \alpha_c + \delta_t + \pi_1 \log(Population)_{ct} + \pi_2 \log(Population)_{ct-1} + u_{ct} \quad (1)$$

[12] They are: *Large state, Bishop, Archbishop, Capital, Plundered, Population, University, Commune*.

where Y_{ct} is either $Log(1+\ Births)$ or Log $(1+Immigrants)$ - we use this transformation because almost two thirds of city-century observations have zero *Births* or *Immigrants* - and α_c and δ_t are city and century fixed effects. Standard errors are clustered at the level of NUTS 2 regions. We use NUTS 2 level rather NUTS 1 level because of the higher number of groups (269 versus 69 in the full sample) but all results in the paper are similar if NUTS 1 is used. A finding that π_1 and π_2 are not significantly different from 0 would imply that the formation of creative clusters cannot be predicted by contemporaneous or past levels of prosperity, as captured by city size (or by other indicators described below), casting doubts on the possibility of a causal effect.

The results are presented in Column 1 of Tables 7 and 8 for *Births* and *Immigrants* respectively. We cannot reject the null hypothesis that creative people (born or immigrated) are uncorrelated with current and previous population.

In columns 2-4 of Tables 7 and 8 we replace population with each of the three variables indicating a city's status in the political and ecclesiastical hierarchy, *Bishop*, *Archbishop* and *Capital*. Again, we fail to reject the null hypothesis of no positive correlation, except for being a state capital (current and lagged $Capital$ are both positive and significant in the *Immigrants* regression, and the cumulative effect of $Capital$, π_1 + $\pi_2,$ is significant at 10% in the *Births* regression). Lagged *Archbishop* has a negative and statistically significant coefficient, suggesting that perhaps being close to religious power may stifle creativity, though the F test for significance of current plus lagged $Archibishop$ is only statistically significant for *Births*.

In column 5 we estimate Equation (1) replacing *Population* with the dummy variable indicating whether a city has a form of self governance (*Commune*). Now the coefficient on the current and lagged values of $Commune$ are positive and statistically significant in both Tables 7 and 8, suggesting that the failure to reject the null hypothesis in Columns 1-4 is not due to lack of statistical power. The following section investigates more closely the relation between local institutions and creativity.

Finally, in column 6 we include all these variables together on the RHS. Again, besides $Commune$, only being a state capital has a positive and significant correlation with *Births* and *Immigrants*.

This evidence is only suggestive, because time is measured at 100 year intervals, but it is robust. We obtain very similar results when (a) using *Births* rather than $Log(1 + Births)$ as dependent variable, (b) entering two lags rather than one, (c) allowing the coefficient of population to vary before/after the middle of our sample period. Also, we have used the interaction between a dummy variable indicating whether the city is an Atlantic port and dummy variables from 1500 (Acemoglu et al., 2005) as an alternative proxy for economic success, and reached similar conclusions. Overall, these results suggest that local prosperity (other than being a state capital) does not play an important role in the formation or decay of creative clusters.

What about the reverse ? Is local creativity a leading indicator of subsequent city growth ? To answer, Table (9) shows estimates with $\log(Population)$ as dependent variable, and the contemporaneous and lagged values of $Log(1+\ Births)$ or Log $(1 + Immigrants)$ on the RHS (with and without the other covariates). Here lagged values of famous births and of famous immigrants always have a positive and highly significant estimated coefficient. The contemporaneous correlation is negative and sometimes statistically significant (possibly because *Births* and *Immigrants* are scaled by population), but it is much smaller than the lagged effect, and the sum of the two coefficients of interest is always positive and significantly different from 0. The estimated coefficient of 0.690 of lagged $Log(1+\ Births)$ in column (6) implies that a 10% increase in *Births* relative to its average of 0.316 is associated with a 2% increase in city population in the subsequent century; the estimated effect of *Immigrants* is of about the same magnitude.[13] Thus, local creativity is an important predictor of subsequent city growth. Finally, note that becoming a *Commune* and a state *Capital* also predict increases in city size, in the current and subsequent century, in line with the results of Bosker et al. (2013).

What moves first: wages or famous creatives? To further investigate the dynamics of the relationship between economic prosperity and famous creatives, we also rely on wages of skilled workers. Wages are a better measure of economic prosperity than population. Moreover, they are available at a higher frequency for a subset of European cities, and this is important to detect which variable moves first.

We begin our exploration by estimating the following specification:

$$Log(1+Unscaled\ Births)_{ct} = \alpha_c + \delta_t + \sum_{k=0}^{5} \beta_k \log(Wage)_{ct-k} + \sum_{k=1}^{5} \beta_k \log(1+Unscaled\ Births)_{ct-k} + u_{ct} \quad (2)$$

where t now represents a decade. Thus, our lag structure goes back 50 years. Because population is not available at the frequency of 10 years, we measure Births as the sum of famous births during the decade, not scaled by population ($Unscaled\ Births$). The sample is an unbalanced panel of 28 cities over the period 1260-1890. Note that here we can include the lagged dependent, despite estimating with a city fixed effect, because the panel is sufficiently long (on average about 30 periods per city).[14] Results are similar when estimating without the lagged dependent variable. Standard errors are clustered at the city level. We also test for first order serial correlation in the estimated residuals, and we can never reject absence of serial correlation.

The results are presented in Column 1 of Table 10: the F test cannot reject that the sum of estimated coefficients of current and lagged wages is equal to zero. Note that failure to reject is

[13] Note that $Log(1 + Births) \simeq Births$, the dependent variable is $Log(Population)$ and $Births$ is expressed per 1000 inhabitants.

[14] See Judson and Owen (1999).

not due to lack of statistical power, because the estimated coefficients on the lagged endogenous variable (lagged births) are always highly significant.

Column 2 shows estimates of Equation (2) with $\log(Wage)$ as the dependent variable, and the contemporaneous and lagged values of $Log(1 + Unscaled\ Births)$ on the RHS. Here the F test for significance of current plus lagged famous creatives born in city c is significant at the 1% level, and all estimated coefficients of births except one are positive, suggesting that famous births are associated with a subsequent acceleration of economic activity.

These two regressions suggest that famous births help predict wages, but not viceversa. To further test that predictability is unidirectional, and only goes from births to wages, we add to the wage equation the leads of famous births for 5 decades. We thus estimate:

$$Log(Wage)_{ct} = \alpha_c + \delta_t + \sum_{k=1}^{5} \beta_k \log(Wage)_{ct-k} + \sum_{k=-5}^{5} \beta_k \log(1 + Unscaled\ Births)_{ct-k} + u_{ct} \tag{3}$$

If the estimated coefficients on the leads of births were significantly different from zero, this would suggest that shocks to wages in period t were correlated with famous births in subsequent periods. As shown in column 3, this is not the case. The F-test of joint significance of the lead values of famous creatives born in city c is not significant. On the other hand, the F test for significance of current plus lagged famous creatives born in city c is significant at 1% level. This again suggests that famous births are a leading indicator of subsequent accelerations in economic activity as measured by wages, but not viceversa.

In Columns 4-6 we repeat the same set of regressions replacing births with the number of famous immigrants born in each decade (again, not scaled by population).[15] Although in Column 6 the F test for significance of current plus lagged famous immigrants is not significant, the patterns are very similar to Column 1-3.

Overall, the estimates in Table 10 are consistent with the notion that the formation of creative clusters cannot be predicted by contemporaneous or past levels of wages. On the other hand, local creativity is a leading indicator of subsequent wage growth.

4 Local institutions and birth of famous creatives

This section explores the determinants of births of creative individuals, addressing a specific question: what features of local institutions are associated with the production of creative talent, as measured by births of notable individuals? In particular, we study the role that transitions into and out of the status of Commune play in the births of famous creatives. As already discussed, we

[15]Results are similar if immigrants are meausered by year of death, rather than birth.

expect that more democratic and participatory forms of self-government, protecting economic and political freedoms, favor a more open and innovative social environment, and thus are positively associated with births of creative individuals. We also investigate the effect of other institutional changes, such as religious institutions and universities.

4.1 OLS Estimates

The regression equation that forms the basis of our empirical analysis in this subsection is:

$$Log(1 + Births_{ct}) = \beta_1 Commune_{ct} + \beta_2 X_{ct} + \beta_3 Spatial\ Lags_{ct} + \alpha_c + \delta_t + u_{ct}, \quad (4)$$

where X_{ct} are city-level covariates, α_c and δ_t are city and century fixed effects. The covariates X_{ct} belong to two groups: those less likely to be affected by the status of Commune (*Large state, Bishop, Archbishop, Capital, Plundered)*, and those more likely to be influenced by Commune and hence to be "bad controls" (*Population*, *University*). The spatial lag of any variable Z_{ct} is defined as $\sum_{d \neq c} \varpi_d Z_{dt}$, where the weights ϖ_d are the inverse of distance between cities d and c within the same NUTS 1 region, and 0 outside the NUTS1 region. We include the spatial lags of $Log(population_{ct})$ and of $Log(1 + Births_{ct})$. In the terminology of De Vries (1984), the spatial lag of $Log(population_{ct})$ measures the potential for urban growth ("urban potential") in the macro-region, namely the urban development in the vicinity of each city. Likewise, the spatial lag of $Log(1 + Births_{ct})$ measures the "creative potential", namely neighboring creativity within the macro-region. These two indicators capture possibly omitted variables correlated with prosperity and creativity in the vicinity of each city. Our instrumental variable approach (discussed below) exploits regional waves of institutional change, and these spatial lags make the IV exclusion restriction more credible. We thus include these variables also in the OLS estimates, to allow comparison with the IV estimates shown below.

Table 11 shows OLS estimates of equation (4). Standard errors are clustered at the region (current NUTS 2) level. In Column 1 we estimate a parsimonious version of the baseline specification where we include period dummies and city fixed effects only. In Column 2, we add the first set of controls. In Column 3 we add the remaining covariates in X_{ct}, in Column 4 we add the *Spatial Lag* of *Births*, and in Column 5 both *Spatial Lags*. All columns report a positive and significant coefficient on Commune. According to the estimate in column 3, becoming a Commune is associated with a 7 percentage point increase in *Births*, or an increase of 0.8 unscaled births (a 26% increase relative to the average).[16] Regarding the other city-level variables, we find a positive and signifi-

[16]Since $Births_{ct} < 1$ and $Log(1 + Births_{ct}) \approx Births_{ct}$, the coefficient on Commune can be interpreted as a percentage point increase after a transition into commune status. Multiplying the coefficient of Commune (0.07) by mean population in the sample (in thousand; 11.8) yields the change in unscaled births.

cant coefficient for (possibly endogenous) *University*, and for *Capital*, while religious institutions have a negative estimated coefficient, statistically significant only in some specifications.

Note that the inclusion of $Log(population)$ and of $University$ in column 3 does not affect the estimated coefficient of $Commune$, suggesting that the effect of local political institutions is unlikely to go through these two channels. On the other hand, the inclusion of the spatial lags reduces the estimated coefficient of Commune from 0.70 to 0.51 in column 5, and the estimated coefficient of the spatial lag of $Log(1 + Births_{ct})$ is large and highly significant, suggesting the possible presence of spatially correlated unobserved determinants of creativity.

4.2 Event Study

A concern for the estimation of the relationship between *Commune* and *Births* arises from the possibility of differential pre-trends. Moreover there may be some interesting post-transition dynamics which are not captured by the estimation procedure in the previous OLS Table. We therefore turn to evaluating the local impact of transition into the status of Commune using an "event-study" research design as in Kline (2011) and Autor (2003), whose exposition we follow here. Such a design allows us to test for the presence of differential pre-trends and recover any dynamics of the Commune effect. We compare changes in *Births* of treated cities (i.e. localities that experience the transition into Commune status) both to cities that have not yet been treated and cities that will never be treated during our sample period[17].

Specifically, the regression equation is:

$$\log(1 + Births_{ct}) = \sum_{\tau=-2}^{T} \beta_\tau D_{ct}^\tau + \alpha_c + \delta_t + u_{ct}, \tag{5}$$

where α_c and δ_t are city and century fixed effects, and D_{ct}^τ are a sequence of "event-time" dummies that equal one when the transition to Commune is τ years away in city c and T is the end of the sample period (expressed in event time). Therefore $\tau = 0$ is the year of transition to Commune and the β_τ coefficients characterize the time path of creativity relative to the date of transition for "treated" cities, conditional on the unobserved variance components $\alpha_c, \delta_t, u_{ct}$. We impose the restriction that $\beta_\tau = \overline{\beta}\ if\ \tau \geqslant 2$ - i.e. the dynamics fades out after two centuries. Moreover, we normalize β_{-1} to zero, so that all post-event coefficients can be interpreted as treatment effects. We estimate (5) by OLS.

Figure 1 displays the estimates. Clearly, there is no evidence of pre-existing trends in *Births* of famous people, before the transition into Commune. Becoming a Commune at the beginning

[17] Among the treated cities, in estimating equation (5) below we remove those for which the status of Commune does not persist until the end of the sample period.

of the century (date 0 in Figure 1) is associated with a 5 percentage point increase in the birth of creative individuals (per 1000 inhabitants) during the current century, with an additional increase in the subsequent century. Results are very similar when adding the same controls as in Table 11, including the spatial lags.

4.3 2SLS Estimates

An important limitation of the estimation framework in Figure 1 is the possibility that both local institutions and creativity may be influenced by time-varying omitted factors. To tackle this challenge, we adapt a strategy introduced by Persson and Tabellini (2009) and Acemoglu et al. (2014) in their analysis of democratic transitions in a panel of countries. Namely, we instrument Commune with the proportion of other cities with Commune status in the region (defined by the current NUTS 1 administrative borders), leaving out the own-city observation - we call our instrument *Regional Commune.* As argued above, transitions into or out of Commune occur in regional waves that reflect learning and spillover effects across neighboring localities, or power vacuum in the regional level of government.

Note that we also control for regional creativity (measured by the spatial lag of *Log(1+Births)*) and for urban potential (measured by the spatial lag of $Log(population)$). This should remove possible concerns that transitions into Commune are more frequent in more prosperous areas, and that regional prosperity (or other unobserved regional variables) could exert a direct influence on *Births,* contaminating the validity of the instrument. The identifying assumption is that, conditional on period and city fixed effects, regional creativity, urban potential and other covariates, the regional waves of institutional transitions influence city level creativity only through a city's own political institutions.

The 2SLS estimates are shown in Table 12, which also reports a summary of the first-stage and the reduced-form results. The sequence of specifications mirrors that in the OLS Table. *Regional Commune* is always highly significant both in the first stage and in the reduced form regressions (F-statistics for the excluded instrument range from 18 to 23). Table A.5 in the Appendix reports the full first stage estimates - note that the religious variables are generally not statistically significant in the first stage regressions.

On average, upon a transition into Commune, births of creative people (per 1000 inhabitants) increase by about 11 percentage points in the more inclusive specification with all spatial lags. This corresponds to about 1.2 more creative births per century, or about a 40% increase relative to average births. The estimated coefficient of *Commune* is about twice as large as in the corresponding OLS regression. The fact that our IV strategy produces larger effects of city institutions on creativity may reflect attenuation bias in the OLS estimates due to measurement error in *Commune*.

Another possibility is that the effect of political institutions is heterogeneous across cities. If so, then consistent OLS estimates the average effect of *Commune* on creativity across all cities. On the other hand, 2SLS estimates the average effect for the cities that are marginal in the transition, in the sense that they become communes if and only if there exists a regional wave of institutional change.[18] If the effect of *Commune* on creativity is larger for cities that are marginal in the transition, the 2SLS estimates exceed those of consistent OLS.

Also note that the results are not affected by inclusion of city size on the RHS, which is never statistically significant, although we know from Bosker et al. (2013) that becoming a Commune is also associated with an increase in city size. This reinforces our previous claim that the formation of creative clusters does not seem to operate through local economic prosperity. The coefficient estimates of the other institutional variables are very similar to the OLS estimates (not shown). Overall, the evidence is consistent with the idea that becoming a Commune, and enjoying the resulting autonomy and economic and political freedoms, spreads a culture of openness that encourages innovation and creativity in arts, sciences and business.

4.4 Sensitivity Analysis

The main empirical result in this Section is that city institutions protecting economic and political freedoms explain the production of creative elites. We now investigate the robustness of the estimates and explore some alternative explanations for the estimated effects.

Alternative measures of creativity. In our data, being famous is equivalent to being included in the database *Freebase.com.* Yu et al (2016) have created a similar database that weights individuals by their influence (see section 2). Table A.6 in the Appendix replicates the 2SLS estimates of Table 12, replacing the dependent variable *Births* with the corresponding weighted variable obtained from Yu et al (2016). The estimated coefficient of $Commune$ is positive and significant across all different IV specifications, although the size of the implied estimated effect of a transition into Commune is a bit smaller than with the unweighted data used in Table 12 (expressed in percent of the mean of the dependent variable).[19]

[18]See Imbens and Angrist (1994) for a discussion. For a recent example, see Eisensee and Strömberg (2007).

[19]The conclusions regarding the role of local institutions are also unchanged when obtaining the OLS estimates using this weighted dependent variable. Regarding the event study, while the visual pattern is similar to that in Figure 1, the individual β_τ coefficients are not estimated very precisely. We therefore performed a more formal test of the null hypothesis that the transition into *Commune* has no impact on local creativity: we tested the hypothesis about the average of the β_τ coefficients for $\tau = 0$ and $\tau = 1$. The estimated average increase over the two centuries starting with the year of the opening is 13 p.p. and significant at 10%.

Specification In results untabulated here we have explored the sensitivity of the estimates to alternative specifications or sample restrictions. First, we have included the interaction between a dummy variable indicating whether the city is an Atlantic port and dummies from 1500 onwards, as in Acemoglu et al. (2005). Second, we have eliminated city-year observations with unusually high values of *Births* (trimming observations above the 99% percentile). Third, following Cox (2017), we have also included the growth rate of the spatial lag of *Log(Population)* - besides its level. Results are largely unchanged.

Sample restrictions In the previous analysis we exploited all transitions into and out of Commune, estimating an average effect. We have also estimated the effect of transitions in the two directions separately. Thus, when estimating the effect of entry into $Commune$, we dropped the city-century observations following a negative transition (from $Commune = 1$ back to $Commune = 0$). And when studying the effect of exits, we dropped the city-year observations prior to a positive transition (from $Commune = 0$ to $Commune = 1$). The OLS estimates remain very similar to those reported in Table 11, with a p. value below 1%, suggesting that the effect of Commune on *Births* is symmetric for transitions on both directions. When estimating by 2SLS, the estimated coefficient of *Commune* remains similar to the IV estimates reported in Table 12 (and is larger for transitions into $Commune$ than out of it), but the standard errors increase and the estimated coefficient of $Commune$ lose significance. Results are available upon request.

5 Migration of famous creatives

In this section we study the determinants of the migration of famous people between European cities, in a gravity model. This section has one goal: to describe how migration is correlated with observable features of European cities, and in particular which institutional features make a city an attractive destination.

5.1 Microfoundations

Let m_{jit} denote the number of immigrants (unscaled by city size) who die in city i and were born in city j during century t (throughout the century refers to the date of birth, as explained above). Also let b_{jt} denote the number of famous individuals born in city j during century t. By definition, we have:

$$m_{jit} = p_{jit} b_{jt} \tag{6}$$

where p_{jit} is the share of individuals who move from j to i in century $t.$

The share p_{jit} is the result of a deliberate decision to migrate. We model it as in the standard Random Utility Model, following Beine et al. (2016) and McFadden (1974). Specifically, let subscript k denote individuals, and define U_{kjit} as the utility of individual k born in j if he moves to i in century t. We assume:

$$U_{kjit} = w_{it} - c_{jit} + \varepsilon_{kjit} \tag{7}$$

where w_{it} refers to a deterministic component of utility, such as income and other benefits from being in city i, c_{jit} denotes the cost of moving from j to i in century t and ε_{kjit} is an individual specific random component of his utility. Note, that, due to our data limitations, we assume that the deterministic component of utility from being in city i, w_{it} , only depends on time and on the destination city, for all individuals irrespective of their origin.

If we assume that ε_{kjit} is independently and identically distributed according to an Extreme Value Type-1 distribution, then (6) and (7) imply that the expected number of immigrants from j to i in century t can be written as - see Beine et al. (2016):

$$E(m_{jit}) = \phi_{jit} y_{it} b_{jt} / \Omega_{jt} \tag{8}$$

where $\phi_{jit} = \exp(-c_{jit})$ is decreasing in the cost of moving from j to i, $y_{it} = \exp(w_{it})$ is a measure of attractiveness of location i, and $\Omega_{jt} = \sum_l \phi_{jlt} y_{lt}$ is the expected utility from all possible alternatives available to an individual born in j - including also the decision to stay (corresponding to $j = l$). Thus, the flow of immigrants from j to i is higher if city i is more attractive relative to the average of all other cities (weighted by the cost of moving and including the city of origin), if the city of origin j has more famous natives (i.e. more potential migrants), and if the cost of moving from j to i is lower.

Adding a well behaved error term e_{jit} to Equation (8), such that $E(e_{jit}) = 1$, allows us to estimate the following gravity equation with dyadic data referring to cities of birth and of death for which $i \neq j$:

$$m_{jit} = \phi_{jit} y_{it} \frac{b_{jt}}{\Omega_{jt}} e_{jit} = \exp(w_{it} - c_{jit}) \frac{b_{jt}}{\Omega_{jt}} e_{jit} \tag{9}$$

5.2 Data and Estimation

Since the dependent variable is a count variable with a very large number of zeros, we estimate equation (9) by Poisson Maximum Likelihood in the sample of cities included in Bosker et al. (2013). To reduce the size of what otherwise would be an excessively large data set, we discard all dyadic observations where the city of death i never had a famous individual who died in that city (an immigrant or a native) in any century, irrespective of whether the city gave birth to some famous individuals or not. These observations are automatically discarded when we include destination city fixed effects.

Table A.8 in the Appendix summarizes the main features of the dependent variable. There is a very large number of zeros (more than 99% of all observations), and when positive most dyadic observations have only a few immigrants from the same origin city per century. Nevertheless, more than 83% of the cities in our restricted sample received at least one immigrant throughout the period, and several of them, such as Paris and London, received several hundredths overall. The number of destination cities included in the data set ranges from 77 in the XIth century to 358 in the XIXth century.

The variables on the right hand side of (9) have the following observable counterparts.

To measure the bilateral cost of moving, c_{jit}, we use geographic distance (expressed in 10 kilometers), a time varying variable measuring the fraction of each century in which cities i and j belonged to the same historical state, and a set of dummy variables that equal 1 if cities i and j belong to the same (modern) NUTS1 region, and to two (modern) countries that share the same first official language.[20]

The utility of being in the destination city i, w_{it}, is proxied by population size (that only in this section is measured in 100,000) as a proxy for economic development, and by a set of dummy variables that capture the most relevant institutional variables in the dataset by Bosker et al. (2013), namely *Commune*, *University*, *Capital, Bishop*, *Archbishop*, and *Plundered*. We expect that being a *Commune*, a state *Capital* and having a *University* all make a city a more attractive destination, while having been *Plundered* has the opposite effect. The two ecclesiastic variables have an ambiguous sign: on the one hand the Church was a sponsor of creative endeavours in artistic domains and a source of wealth, but on the other hand it was also a source of discrimination and censorship.

The number of famous births in the origin city j is measured in logs (to be consistent with the exponential functional form of the Poisson regression (see (9)) and includes all famous people born in city j during century t, irrespective of whether or not they died in a different city, since they were all at risk of migrating.

Finally, the so called "multilateral resistance" term Ω_{jt} refers to the attractiveness of all the alternative destinations, for an individual born in city j. This term has no easily observable counterpart. We thus incorporate it as follows. In a first and most restrictive specification, we assume that the only relevant alternative to moving from j to i is remaining in the origin city, and thus proxy Ω_{jt} with w_{jt}, namely with the same institutional variables described above but referring to the origin city. Here we also include a full set of century fixed effects, to capture possible symmetric changes in the cost of moving or in the available alternatives. We then relax this assumption by adding also origin and destination fixed effects, to capture possible time invariant omitted variables. Finally, we estimate with a full set of destination and origin-century fixed effects, with

[20] We are grateful to David Schonholzer and Eric Weese for sharing with us their data on historical state borders.

which we fully capture the multilateral resistance term Ω_{jt} and any other variable that varies by origin and century (as well as time invariant destination variable).

To reduce the number of fixed effects to be estimated, in the specification with century, destination and origin fixed effects we discard all dyadic observations where the city of birth i never sent out any migrants (in any century), despite having a positive number of births, and the city of destination never received an immigrant (in any century), despite having a positive number of deaths of famous natives in some centuries. And in the specification with destination and origin-century fixed effects we also discard all dyadic observations where the city of birth i never sent out any migrants in a given century. These sample restrictions do not affect the estimated parameters of interest, since the omitted observations would be fully explained by the fixed effects. Finally, standard errors are always clustered two ways, by origin and by destination, as suggested by Cameron et al. (2011).

The inclusion of destination (and origin) fixed effects implies that we are identifying the parameters of interest with a diff-in-diff methodology. Namely we assume that *changes* in the institutions of interest are randomly assigned to cities, after controlling for the remaining covariates. In particular, we must assume that there is no time-varying unobserved heterogeneity making cities that adopt specific institutions also more likely to attract or send out immigrants (for reasons unrelated to the institutional changes). Note that immigrants are measured by century of birth, to minimize the risk that the migration decision precedes the institutional change. Nevertheless, the identifying assumption remains restrictive.

5.3 Results

Table 13 reports the estimates for the three specifications: with only century fixed effects (column 1), with century, destination and origin fixed effects (column 2), and with destination and origin-year fixed effects (column 3). The more credible specification is the one reported in column (3), but the estimated coefficients of the destination variables remain very stable in columns (2) and (3). The main institutional variables of interest have the same expected sign. Being a *Commune* and a state capital is associated with an inflow of immigrants, but university cities are not more attractive. On the other hand, being the seat of a Bishop has a negative effect on creative immigration. The estimated coefficients on all the distance measures are highly significant and with the expected sign. Cities that gave rise to more births send out more migrants (the estimated coefficient less than 1 implies that some of famous births do not migrate, as we know from the presence of several natives who die in the city of origin). Finally, population size is not robustly associated with any migration patterns.

The estimated coefficient of 0.587 for *Commune* implies that the arrival rate of immigrants in each century from the same destination city is exp(0.587) = 1.8 times larger in cities that are *Commune*, compared to the others. On average a non-Commune city receives about 1.6 immigrants per century from all origin cities in this sample, implying that becoming a *Commune* is associated with an increase of about 1.3 famous immigrants per century. Becoming a bishop-city has a negative effect of about the same size on the immigration rate. Becoming a state capital is associated with an increase in the arrival rate of immigrants of 2.3 times, implying about 2 more immigrants per century.

6 Concluding Remarks

It is often argued that open and democratic institutions are a prerequisite to sustain innovation-based growth - see for instance Lewin et al. (2016). This argument is strongly supported by the historical evidence of European cities.

As of yet there is no systematic study of the spatial patterns of creativity over a long historical period. After describing the main features of the formation and decay of creative clusters, we study how changes in city institutions affect local creativity. We find that institutions promoting local autonomy and protecting economic and political freedoms encourage the production and attraction of creative talent. The effects are quantitatively large. Becoming a Commune is associated with an increase in the births of famous people of about 40% relative to the average, while the attraction of famous immigrants almost doubles in size upon becoming a Commune. Religious institutions, on the other hand, are negatively correlated with the attraction of creative individuals, in line with other findings in the literature that emphasize a negative role of religion on innovation (Bénabou et al., 2015a,b).

What are the mechanisms through which becoming a Commune fosters local creativity? We know from Bosker et al. (2013) that transitions into Commune are also associated with subsequent increases in city size. Could this be the mechanism, namely transitions into commune enhance economic prosperity, and this in turn induces an increase in local creativity? Our evidence is inconsistent with this conjecture: surges in famous births and immigrants precede increases in city size and, for a smaller sample of European cities, they also precede wage increases (measured at 10 year intervals). Thus, our findings help explain the results in Bosker et al. (2013), but not viceversa. Free cities become the hub of creative talents, and this in turn contributes to improve subsequent city prosperity.

This leaves open the question of what are the mechanisms through which Communal institutions favor the production and accumulation of creative talent. A priori, several answers are possible. First, the protection of personal and economic freedoms changed the local culture, making it more receptive to innovations and new ideas. Second, the new institutions also changed incentives, through a more meritocratic and inclusive social environment, but also by encouraging works of art and innovations that would enhance the prestige of the city. Third, free cities attracted talented and creative individuals who escaped censorship and persecution elsewhere, and this created role models and facilitated social learning, breeding new generations of innovators. These mechanisms are not mutually exclusive, and it is likely that they are all relevant. Discriminating the relative importance of these different channels is an important task of future research.

References

Acemoglu, D. (2008). Oligarchic versus democratic societies. *Journal of the European Economic Association 6*(1), 1–44.

Acemoglu, D., U. Akcigit, and M. A. Celik (2014). Young, restless and creative: Openness to disruption and creative innovations. Technical report, National Bureau of Economic Research.

Acemoglu, D., S. Johnson, and J. Robinson (2005). The rise of europe: Atlantic trade, institutional change, and economic growth. *The American economic review 95*(3), 546–579.

Acemoglu, D., S. Naidu, P. Restrepo, and J. A. Robinson (2014). Democracy does cause growth. Technical report, National Bureau of Economic Research.

Acemoglu, D. and J. A. Robinson (2012). Why nations fail: The origins of power, prosperity and poverty.

Akcigit, U., J. Grigsby, and T. Nicholas (2017a). Immigration and the rise of american ingenuity. Technical report, National Bureau of Economic Research.

Akcigit, U., J. Grigsby, and T. Nicholas (2017b). The rise of american ingenuity: Innovation and inventors of the golden age.

Allen, R. C. (2001). The great divergence in european wages and prices from the middle ages to the first world war. *Explorations in economic history 38*(4), 411–447.

Arzaghi, M. and J. V. Henderson (2008). Networking off madison avenue. *The Review of Economic Studies 75*(4), 1011–1038.

Audretsch, D. B. and M. P. Feldman (1996). R&d spillovers and the geography of innovation and production. *The American economic review 86*(3), 630–640.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of labor economics 21*(1), 1–42.

Bairoch, P., J. Batou, and P. Chèvre (1988). La population des villes europeennes: banque de donnees et analyse sommaire des resultats, 800-1850. *Publications du Centre d'Histoire Economique Internationale de l'Universite de Geneve (2)*.

Banks, D. (1997). Clusters of talent. Newsletter of the Classification Society of North America, 48, February 1997 available at http://www.pitt.edu/ csna/news/csna.news48.html.

Beine, M., S. Bertoli, and J. Fernández-Huertas Moraga (2016). A practitionersŠ guide to gravity models of international migration. *The World Economy 39*(4), 496–512.

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen (2015). The lifecycle of inventors. Technical report, mimeo Harvard.

Bénabou, R., D. Ticchi, and A. Vindigni (2015a). Forbidden fruits: the political economy of science, religion, and growth. Technical report, National Bureau of Economic Research.

Bénabou, R., D. Ticchi, and A. Vindigni (2015b). Religion and innovation. *The American Economic Review 105*(5), 346–351.

Bennassar, B. (1999). *Valladolid au siècle d'or*. Editions EHESS.

Bosker, M., E. Buringh, and J. L. van Zanden (2013). From baghdad to london: Unraveling urban development in europe, the middle east, and north africa, 800–1800. *Review of Economics and Statistics 95*(4), 1418–1437.

Boulton, J. (1996). Wage labour in seventeenth-century london. *The Economic History Review 49*(2), 268–290.

Burke, E. (2016). *The Greeks of Venice, 1498-1600: Immigration, Settlement, and Integration*. Cursor mundi. Brepols Publishers.

Cabourdin, G. (1968). Valladolid au siècle d'or, une ville de castille et sa campagne au xvie siècle.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Carlino, G. and W. R. Kerr (2015). Agglomeration and innovation. *Handbook of Regional and Urban Economics*, 349.

Carlino, G. A., R. M. Hunt, J. K. Carr, and T. E. Smith (2012). The agglomeration of r&d labs.

Chittolini, G. (1999). *Il Comune e la signoria*. G. Chittolini editore, 1999.

Cipolla, C. M. (1964). Economic depression of the renaissance?. i. *The Economic History Review 16*(3), 519–524.

Combes, P.-P. and L. Gobillon (2015). The empirics of agglomeration economies. *Handbook of Regional and Urban Economics*.

Connell, W. J. (2002). *Society and Individual in Renaissance Florence*. Univ of California Press.

Cox, G. W. (2017). Political institutions, economic liberty and the great divergence.

De la Croix, D. and O. Licandro (2015). The longevity of famous people from hammurabi to einstein. *Journal of Economic Growth 20*(3), 263–303.

De Vries, J. (1984). European urbanization: 1500–1800.

Dittmar, J. E. and R. R. Meisenzahl (2016). State capacity and public goods: Institutional change, human capital, and growth in early modern germany.

Donges, A., J.-M. A. Meier, and R. Silva (2016). The impact of institutions on innovation.

Duranton, G. (2007). Urban evolutions: The fast, the slow, and the still. *The American Economic Review 97*(1), 197–221.

Eisensee, T. and D. Strömberg (2007). News droughts, news floods, and u.s. disaster relief. *The Quarterly Journal of Economics 122*(2), 693–728.

Falck, O., M. Fritsch, and S. Heblich (2011). The phantom of the opera: Cultural amenities, human capital, and regional economic growth. *Labour Economics 18*(6), 755–766.

Feliu, G. (1991). Precios y salarios en la cataluña moderna.

Florida, R. (2005). *Cities and the creative class*. Routledge.

Gergaud, O., M. Laouénan, E. Wasmer, et al. (2016). A brief history of human time: Exploring a database of'notable people'. Technical report, Sciences Po Departement of Economics.

Gibson, A. J. and T. C. Smout (1995). *Prices, food and wages in Scotland, 1550-1780*. Cambridge University Press.

Guiso, L., P. Sapienza, and L. Zingales (2016). Long-term persistence. *Journal of the European Economic Association*.

Imbens, G. and J. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*(4), 467–476.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *the Quarterly journal of Economics*, 577–598.

Jones, E. (2003). *The European miracle: environments, economies and geopolitics in the history of Europe and Asia*. Cambridge University Press.

Judson, R. A. and A. L. Owen (1999). Estimating dynamic panel data models: a guide for macroeconomists. *Economics letters 65*(1), 9–15.

Kandel, E. (2012). *The age of insight: The quest to understand the unconscious in art, mind, and brain, from Vienna 1900 to the present*. Random House.

Kerr, W. R. (2010). Breakthrough inventions and migrating clusters of innovation. *Journal of Urban Economics 67*(1), 46–60.

Kline, P. (2011). The impact of juvenile curfew laws on arrests of youth and adults. *American Law and Economics Review*, ahr011.

Lewin, A. Y., M. Kenney, and J. P. Murmann (2016). *China's Innovation Challenge: Overcoming the Middle-Income Trap*. Cambridge University Press.

Lopez, R. S. and H. A. Miskimin (1962). The economic depression of the renaissance1. *The Economic History Review 14*(3), 408–426.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, 105–142.

Meisenzahl, R. R. and J. Mokyr (2011). The rate and direction of invention in the british industrial revolution: Incentives and institutions. In *The rate and direction of inventive activity revisited*, pp. 443–479. University of Chicago Press.

Mokyr, J. (2009). *The Enlightened Economy: An Economic History of Britain 1700-1850*. Yale University Press.

Mokyr, J. (2016). *A Culture of Growth: The Origins of the Modern Economy*. Princeton University Press.

Moretti, E. (2004). Estimating the external return to higher education: Evidence from cross-sectional and longitudinal data. *Journal of Econometrics 120*, 175–212.

Norman, D. (1999). *Siena and the Virgin: art and politics in a late medieval city state*. Yale University Press.

Paoletti, J. T. and G. M. Radke (2005). *Art in Renaissance Italy*. Laurence King Publishing.

Parker, G. (2004). *Sovereign city: the city-state through history*. Reaktion Books.

Persson, T. and G. Tabellini (2009). Democratic capital: The nexus of political and economic change. *American Economic Journal: Macroeconomics 1*(2), 88–126.

Pirenne, H. (2014). *Medieval cities: Their origins and the revival of trade*. Princeton University Press.

Prezzolini, G. (1948). *The legacy of Italy*. New York Vanni [1948].

Rappaport, S. (2002). *Worlds within worlds: structures of life in sixteenth-century London*, Volume 7. Cambridge University Press.

Rosenthal, S. and W. Strange (2003). Geography, industrial organization, and agglomeration. *Review of Economics and Statistics 85*(2), 377–393.

Saxenian, A. (1994). *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press.

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

Scholliers, P. and G. Avondts (1977). *De Gentse textielarbeiders in de 19e en 20e eeuw: herkomst, huisvesting, arbeids-en levensomstandigheden van de werkkrachten van het bedrijf A. Voortman-NV Texas*. Vrije Universiteit.

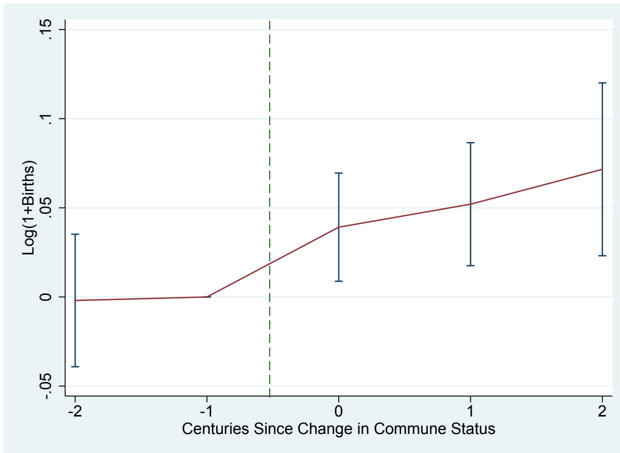
Serafinelli, M. (2017). Good firms, worker flows and local productivity. *Journal of Labor Economics, forthcoming*.

Squicciarini, M. P. and N. Voigtländer (2014). Human capital and industrialization: Evidence from the age of enlightenment. Technical report, National Bureau of Economic Research.

Yu, A. Z., S. Ronen, K. Hu, T. Lu, and C. A. Hidalgo (2016). Pantheon 1.0, a manually verified dataset of globally famous biographies. *Scientific data 3*.

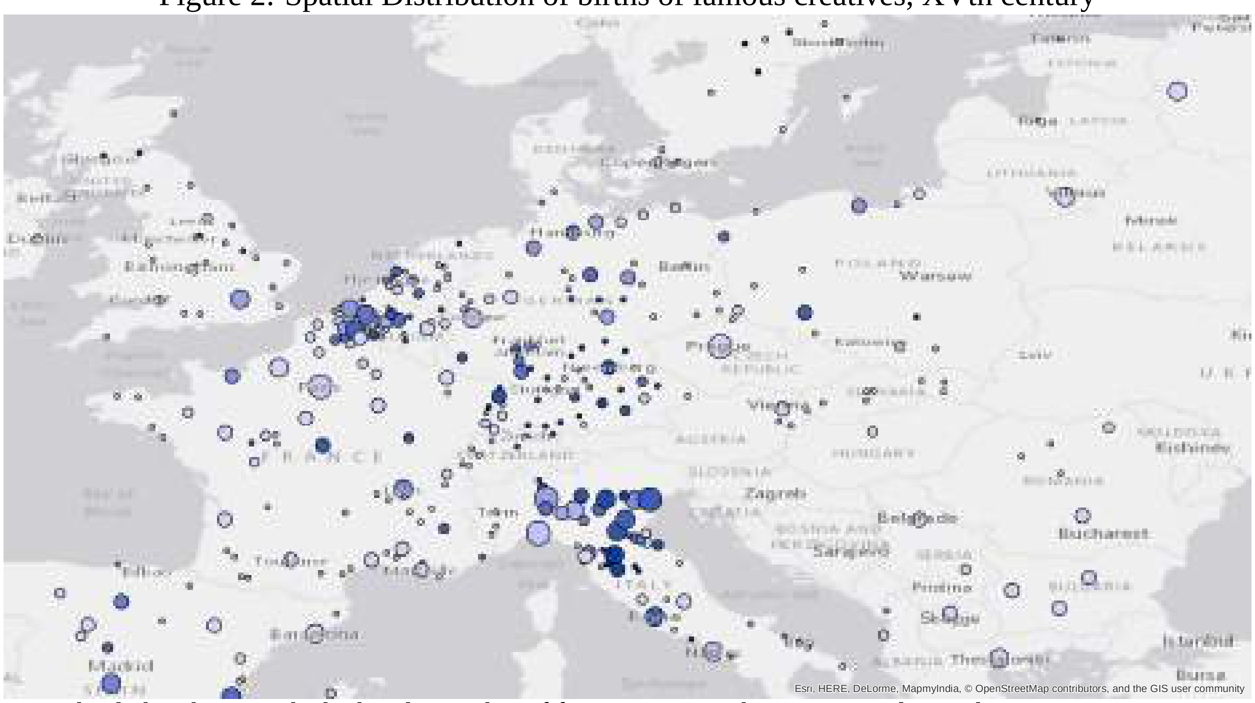
Figures and Tables

Figure 1: Local Institutions and Births of Famous Creatives



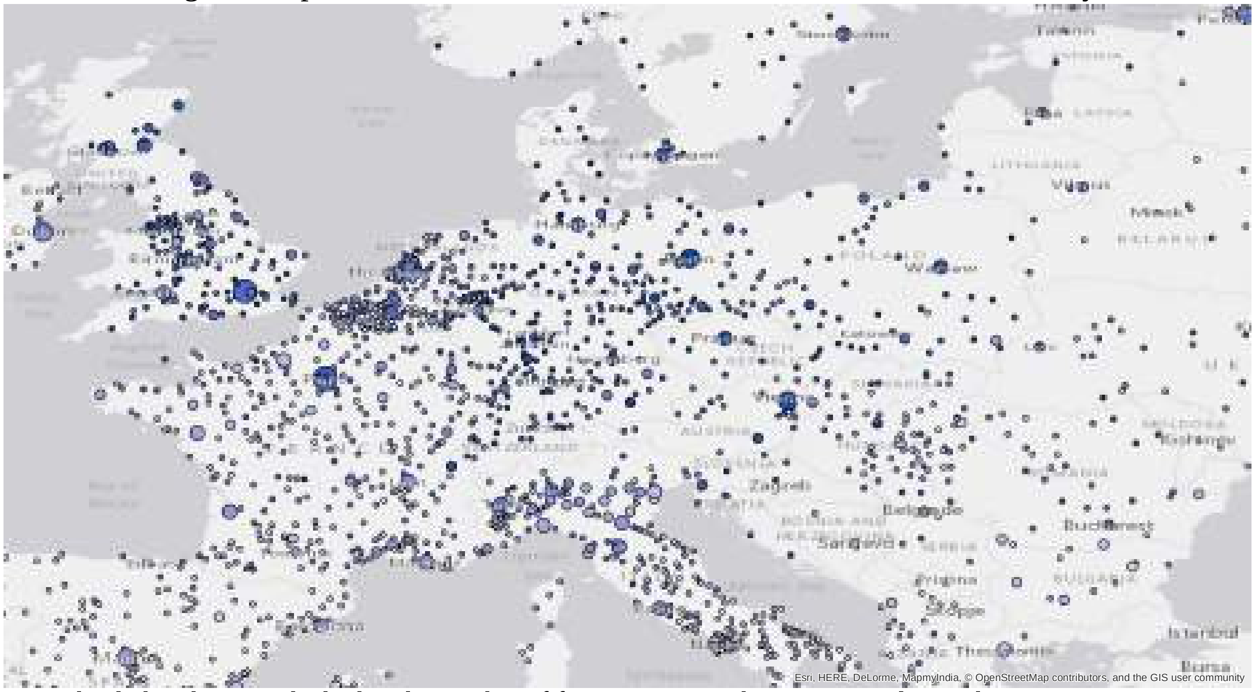
The figure plots point estimates for leading and lagging indicators for the change in commune status. Event time indicator "+2" set to 1 for all periods 2 periods after the event and 0 otherwise. The omitted category is one period prior to the change in commune status. The dependent variable is $Log(1 + Births)$. The variable $Births$ is equal to the number of famous creatives born in a city, per 1000 inhabitants. Vertical bars correspond to 95 percent confidence intervals with region-clustered standard errors.

Figure 2: Spatial Distribution of births of famous creatives, XVth century



The darker the tone, the higher the number of famous creatives born in a city during the century, per 1000 inhabitants. The larger circle, the larger the population of the city. See Table A.3 for descriptive statistics.

Figure 3: Spatial distribution of births of famous creatives, XIXth century



The darker the tone, the higher the number of famous creatives born in a city during the century, per 1000 inhabitants. The larger circle, the larger the population of the city. See Table A.4 for descriptive statistics.

Figure 4: Coefficient of Variation of Births, Immigrants, Population, XIVth to XIXth century

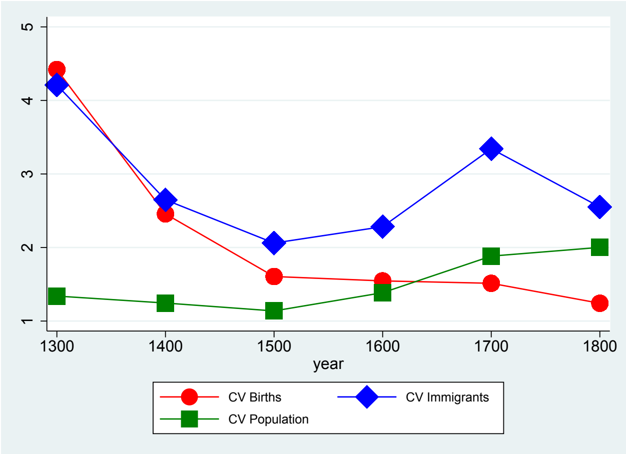


Figure 5: Distribution of distances between place of birth of any two famous creatives.

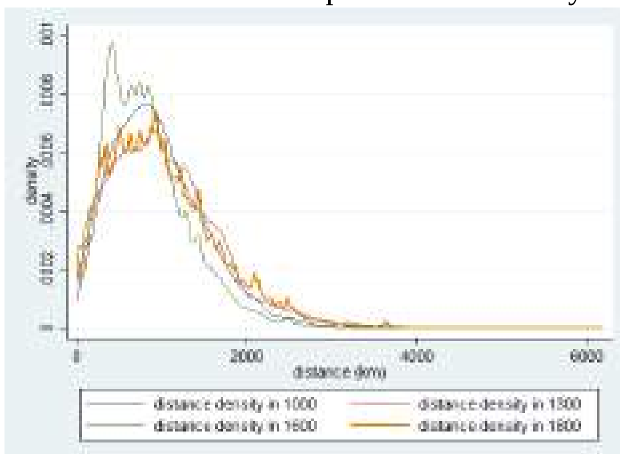


Figure 6: Distribution of birth-to-death distances over time

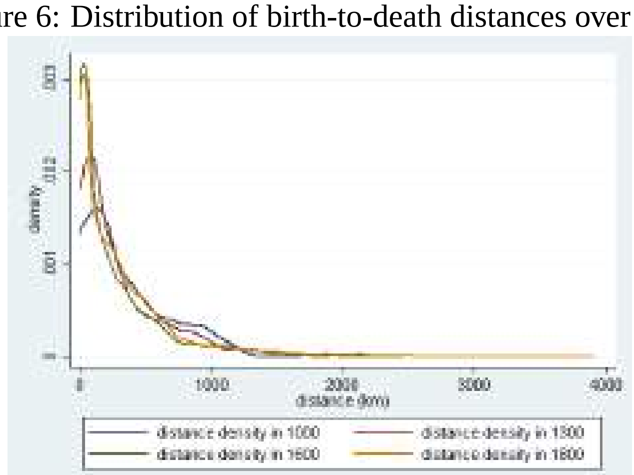


Table 1: Freebase Professional Categories

	Number of people
Performing arts	4692
Non performing arts	10002
Humanities and Sciences	8402
Business	2083
Total number of creative individuals	21906

The counts refer to the number of people for which a city of birth or death in Europe is observed during our sample period. The total number of creative individuals is lower than the sum across categories because some individuals are present in more than one category.

Table 2: Count of Famous Creatives and Population

Century	Unscaled Births	Unscaled Immigrants	Population (1000's)
1000	27	20	2004
1100	45	33	2089
1200	79	45	2993
1300	109	79	6092
1400	567	346	4867
1500	1243	746	6785
1600	1633	1068	10478
1700	3642	2036	12595
1800	15027	6774	25444

Table 3: Summary Statistics

Variable	**Mean**	**(Std. Dev.)**	**Min.**	**Max.**	**N**
Births	0.316	(0.809)	0	13.25	6226
Immigrants	0.195	(1.04)	0	31	6226
(Unscaled) Births	3.082	(16.713)	0	774	6226
(Unscaled) Immigrants	2.344	(20.362)	0	827	6226
Commune	0.504	(0.5)	0	1	3091
University	0.118	(0.323)	0	1	3091
Non-Absolutist State	0.352	(0.478)	0	1	3091
Population	11780.794	(25219.232)	316.228	948000	6226
Large state	0.627	(0.484)	0	1	3091
Bishop	0.382	(0.486)	0	1	3091
Archbishop	0.118	(0.323)	0	1	3091
Capital	0.07	(0.255)	0	1	3091
Plundered	0.022	(0.146)	0	1	3091
Births, Yu et al.	0.925	(3.607)	0	99.557	6226

Population is interpolated for year 1100

Table 4: Spearman's rank correlation coefficient for key variables over time

	(1)	(2)	(3)
Century	Births	Immigrants	Population
XIVth-XVth	0.164**	0.308***	0.838***
XVth-XVIth	0.331***	0.398***	0.870***
XVIth-XVIIth	0.424***	0.500***	0.839***
XVIIth-XVIIIth	0.423***	0.462***	0.791***
XVIIIth-XIXth	0.593***	0.516***	0.700***

Each row reports Spearman's rho for a given variable measured at t and t+1. For instance, the row "XVIth-XVIIth" reports the measure of statistical dependence between the variable in the column measured in XVIth century and the same variable measured in the XVIIth century.

Table 5: Markov transition matrices for key variables across cities

	Births					**Immigrants**					**Population**				
t/t+1	**0**	**q1**	**q2**	**q3**	**q4**	**0**	**q1**	**q2**	**q3**	**q4**	**0**	**q1**	**q2**	**q3**	**q4**
0	.61	.13	.12	.09	.05	.75	.09	.07	.06	.03	.61	.2	.09	.07	.03
q1	.17	.42	.24	.12	.04	.23	.47	.19	.07	.04	.25	.3	.23	.19	.03
q2	.22	.22	.24	.19	.12	.32	.19	.15	.18	.16	.1	.1	.21	.44	.15
q3	.12	.08	.19	.32	.29	.22	.12	.16	.23	.27	.05	.05	.13	.35	.41
q4	.19	.03	.1	.26	.43	.15	.03	.1	.23	.49	.01	.01	.05	.14	.79

This Table displays the probability of transition from each row to each column, estimated by Maximum Likelihood. For each variable (Births per 1000 inhabitants, Immigrants per 1000 inhabitants, and Population) and each century, we divide cities in five groups: the first group includes cities that in a given century featured a value of zero (the first row / column). The remaining groups correspond to the quartiles of the distribution in any given century conditional on being positive. The Table displays the probability of moving between each of the five categories from century t to century t+1, estimated by Maximum Likelihood.

Table 6: Coagglomeration

	Arts Non-Performing	Arts Performing	Humanities and Sciences
Panel A: Births			
Arts, Performing	0.4039 ***		
	0.3825 ***		
Humanities and Sciences	0.4988 ***	0.2757 ***	
	0.5229***	0.3165 ***	
Business	0.3327***	0.2789 ***	0.3386 ***
	0.3901***	0.2472***	0.3700***
Panel B: Immigrants			
Arts, Performing	0.4700 ***		
	0.4471 ***		
Humanities and Sciences	0.5261 ***	0.3434 ***	
	0.5566***	0.2817 ***	
Business	0.3673 ***	0.3944 ***	0.2864 ***
	0.3999 ***	0.3381 ***	0.3848 ***

Each entry represents a pairwise correlation of residuals from regressions of famous people per capita in each discipline by city of birth (Panel A) or death (Panel B) on period dummies (upper row) and period dummies plus controls (lower rows). The dependent variable is defined as Log(1+*Births in discipline i*) in Panel A and Log(1+*Immigrants in discipline i*) in Panel B. * p<0.1, . * p<0.1, ** p<0.05, *** p<0.01.

Table 7: Economic Success and Creativity, Births

	(1)	(2)	(3)	(4)	(5)	(6)
Log (Population)	-0.016					-0.017
	(0.013)					(0.013)
L.Log (Population)	0.016					0.015
	(0.013)					(0.011)
Bishop		0.004				-0.009
		(0.024)				(0.027)
L.Bishop		-0.037				-0.089***
		(0.026)				(0.031)
Archbishop			0.014			0.017
			(0.034)			(0.039)
L.Archbishop			-0.075**			-0.131***
			(0.033)			(0.037)
Capital				0.047**		0.041*
				(0.022)		(0.022)
L.Capital				0.024		0.032
				(0.023)		(0.026)
Commune					0.052***	0.054***
					(0.014)	(0.017)
L.Commune					0.040***	0.030*
					(0.015)	(0.017)
$\beta_X + \beta_{L.X} = 0$,pv	0.995	0.181	0.044	0.031	0.000	
Observations	3,434	3,045	3,045	3,045	3,045	2,191
Adjusted R-squared	0.244	0.291	0.292	0.293	0.301	0.287

The dependent variable is Log (1+Births). $\beta_X + \beta_{L.X} = 0$,pv is the p-value of equality to zero of the sum of coefficients of the variable measured in t and in t+1. Period dummies always included. Standard errors (clustered at the NUTS 2 region level) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 8: Economic Success and Creativity, Immigrants

	(1)	(2)	(3)	(4)	(5)	(6)
Log (Population)	-0.008					-0.013
	(0.012)					(0.013)
L.Log (Population)	0.016*					0.011
	(0.009)					(0.008)
Bishop		-0.022				-0.016
		(0.018)				(0.023)
L.Bishop		-0.013				-0.061***
		(0.018)				(0.022)
Archbishop			0.026			0.026
			(0.030)			(0.032)
L.Archbishop			-0.055**			-0.089***
			(0.023)			(0.026)
Capital				0.055**		0.035
				(0.027)		(0.024)
L.Capital				0.055*		0.051
				(0.030)		(0.034)
Commune					0.037***	0.034**
					(0.013)	(0.015)
L.Commune					0.037***	0.029**
					(0.012)	(0.014)
$\beta_X + \beta_{L.X} = 0$,pv	0.582	0.102	0.395	0.018	0.000	
Observations	3,434	3,045	3,045	3,045	3,045	2,191
Adjusted R-squared	0.141	0.138	0.138	0.144	0.146	0.147

The dependent variable is Log (1+Immigrants). $\beta_X + \beta_{L.X} = 0$,pv is the p-value of equality to zero of the sum of coefficients of the variable measured in t and in t+1. Period dummies always included. Standard errors (clustered at the NUTS 2 region level) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 9: Famous Creatives and Population

	(1)	(2)	(3)	(4)	(5)	(6)
Log(1+Births)	-0.075	-0.189			-0.138*	-0.243*
	(0.084)	(0.146)			(0.081)	(0.147)
L.Log(1+Births)	0.639***	1.000***			0.450***	0.690***
	(0.109)	(0.115)			(0.108)	(0.131)
Log(1+Immigrants)			-0.185*	-0.262**	-0.125	-0.119
			(0.104)	(0.133)	(0.103)	(0.121)
L.Log(1+Immigrants)			0.880***	1.226***	0.708***	0.898***
			(0.133)	(0.168)	(0.135)	(0.188)
Commune		0.266***		0.282***		0.268***
		(0.073)		(0.073)		(0.072)
L.Commune		0.126***		0.127***		0.115**
		(0.044)		(0.047)		(0.046)
Bishop		-0.214*		-0.203		-0.210*
		(0.127)		(0.126)		(0.127)
L.Bishop		0.019		0.007		0.029
		(0.104)		(0.102)		(0.103)
Archbishop		0.087		0.094		0.073
		(0.148)		(0.136)		(0.143)
L.Archbishop		-0.033		-0.044		-0.005
		(0.129)		(0.130)		(0.128)
Capital		0.354***		0.331***		0.328***
		(0.125)		(0.124)		(0.121)
L.Capital		0.436***		0.423***		0.415***
		(0.100)		(0.094)		(0.095)
$\beta_X^{Bir} + \beta_{L.X}^{Bir} = 0$,pv	0.000	0.000			0.003	0.002
$\beta_X^{Imm} = \beta_{L.X}^{Imm} = 0$,pv			0.000	0.000	0.059	0.048
Observations	3,434	2,191	3,434	2,191	3,434	2,191
Adjusted R-squared	0.290	0.337	0.294	0.337	0.308	0.354

The dependent variable is Log Population. $\beta_X^{Bir} + \beta_{L.X}^{Bir} = 0$,pv is the p-value of the F-test of joint significance of Log(1+Births) and L.Log(1+Births). $\beta_X^{Imm} = \beta_{L.X}^{Imm} = 0$,pv is the p-value of the F-test of joint significance of Log(1+Immigrants) and L.Log(1+Immigrants). Period dummies always included. Standard errors (clustered at the NUTS 2 region level) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 10: Famous Creatives and Wage Levels

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(Un.Births.+1)	Log(Wage)	Log(Wage)	Log(Un.Immig.+1)	Log(Wage)	Log(Wage)
F5.Log(Un.Births.+1)			0.019*			
			(0.010)			
F4.Log(Un.Births.+1)			-0.002			
			(0.013)			
F3.Log(Un.Births.+1)			-0.006			
			(0.010)			
F2.Log(Un.Births.+1)			0.001			
			(0.006)			
F1.Log(Un.Births.+1)			-0.014			
			(0.008)			
Log(Un.Births.+1)		0.006	0.006			
		(0.008)	(0.009)			
L1.Log(Un.Births.+1)	0.322***	-0.009	-0.004			
	(0.037)	(0.009)	(0.010)			
L2.Log(Un.Births.+1)	0.175***	0.010	0.007			
	(0.048)	(0.009)	(0.010)			
L3.Log(Un.Births.+1)	0.149**	0.010	0.014			
	(0.056)	(0.009)	(0.010)			
L4.Log(Un.Births.+1)	0.128**	0.006	-0.001			
	(0.054)	(0.011)	(0.014)			
L5.Log(Un.Births.+1)	0.078**	0.003	0.012			
	(0.038)	(0.009)	(0.009)			
Log(Wage)	0.093			-0.181		
	(0.129)			(0.164)		
L1.Log(Wage)	-0.348**	0.901***	0.842***	-0.011	0.907***	0.853***
	(0.134)	(0.062)	(0.068)	(0.173)	(0.059)	(0.064)
L2.Log(Wage)	0.264*	0.027	0.052	0.065	0.027	0.049
	(0.136)	(0.077)	(0.088)	(0.178)	(0.074)	(0.085)
L3.Log(Wage)	-0.001	-0.032	-0.011	0.193	-0.034	-0.011
	(0.171)	(0.053)	(0.056)	(0.291)	(0.050)	(0.055)
L4.Log(Wage)	0.134	0.009	0.016	-0.064	0.015	0.021
	(0.236)	(0.043)	(0.048)	(0.254)	(0.038)	(0.042)
L5.Log(Wage)	-0.126	-0.021	-0.018	0.058	-0.021	-0.020
	(0.178)	(0.052)	(0.053)	(0.139)	(0.050)	(0.049)
F5.Log(Un.Immig.+1)						0.005
						(0.010)
F4.Log(Un.Immig.+1)						0.005
						(0.010)
F3.Log(Un.Immig.+1)						0.005
						(0.016)
F2.Log(Un.Immig.+1)						-0.006
						(0.015)
F1.Log(Un.Immig.+1)						-0.009
						(0.008)
Log(Un.Immig.+1)					-0.011	-0.009
					(0.010)	(0.009)
L1.Log(Un.Immig.+1)				0.377***	-0.006	-0.002
				(0.050)	(0.009)	(0.011)
L2.Log(Un.Immig.+1)				0.188***	0.025***	0.023**
				(0.044)	(0.008)	(0.010)
L3.Log(Un.Immig.+1)				0.085*	-0.015	-0.021
				(0.043)	(0.015)	(0.017)
L4.Log(Un.Immig.+1)				0.121**	0.017*	0.019*
				(0.057)	(0.009)	(0.010)
L5.Log(Un.Immig.+1)				0.033	0.009	0.011
				(0.049)	(0.010)	(0.011)
$F^{L.Wage}$,pv	0.880	0.000	0.000	0.528	0.000	0.000
$F^{L.Births}$,pv	0.000	0.008	0.007			
$F^{L.Immigrants}$,pv				0.000	0.013	0.252
$F^{F.Births}$,pv			0.878			
$F^{F.Immigrants}$,pv						0.976
AR(1)z	-0.97	-1.33	-1.12	-0.49	-1.44	-1.14
Observations	815	815	692	815	815	692
Adjusted R-squared	0.665	0.929	0.881	0.633	0.929	0.881

The frequency of observation is a decade. The variable Wage represents the average nominal wage of skilled workers over the decade, measured in silver grams. The variable Un.Births is the (unscaled) number of Births. The variable Un.Immig. is the (unscaled) number of Immigrants, measured by year of birth. $F^{L.X}$,pv is the p-value of the F-test of joint significance of the lags of the variable X when X is the dependent variable, and of the contemporaneous and lag values when X is not the dependent variable. $F^{F.X}$,pv is the p-value of the F-test of joint significance of its leads. AR(1)z is the Arelanno and Bond (1999) test of first order serial correlation, distributed as N(0,1). Tests for correlation up to fifth order give similar results. Period dummies always included. Standard errors (clustered at city level) in parentheses.

Table 11: Commune and Births, OLS Regressions

	(1)	(2)	(3)	(4)	(5)
Commune	0.072***	0.067***	0.070***	0.058***	0.051***
	(0.012)	(0.011)	(0.012)	(0.012)	(0.011)
Large state		-0.033**	-0.035**	-0.008	-0.006
		(0.014)	(0.014)	(0.013)	(0.013)
Bishop		-0.032	-0.028	-0.009	-0.007
		(0.023)	(0.023)	(0.022)	(0.019)
Archbishop		-0.047	-0.058*	-0.056*	-0.050*
		(0.032)	(0.032)	(0.030)	(0.028)
Capital		0.057**	0.060**	0.064***	0.056**
		(0.025)	(0.025)	(0.025)	(0.024)
Plundered		-0.019	-0.016	-0.015	-0.021
		(0.027)	(0.026)	(0.024)	(0.023)
Log (Population)			-0.005	-0.014*	-0.011
			(0.009)	(0.008)	(0.008)
University			0.059**	0.059**	0.044**
			(0.024)	(0.023)	(0.022)
Spatial Lag of Log (1 + Births)				1.016***	1.145***
				(0.141)	(0.121)
Spatial Lag of Log Population					-0.026***
					(0.007)
Observations	3,045	3,045	3,045	3,045	3,045
Adjusted R-squared	0.300	0.305	0.308	0.377	0.409

Dependent Variable is Log (1 + Births). Standard errors (clustered at the NUTS 2 region level) in parentheses. Period Dummies and City FE always included.

Table 12: IV Estimates of the Effect of Commune on Births

	(1)	(2)	(3)	(4)	(5)
			Panel A: 2SLS estimates		
Commune	0.178***	0.168***	0.172***	0.124**	0.111**
	(0.052)	(0.052)	(0.057)	(0.056)	(0.052)
Spatial Lag of Log (1 + Births)				0.996***	1.122***
				(0.140)	(0.121)
Spatial Lag of Log Population					-0.025***
					(0.007)
Observations	2,961	2,961	2,961	2,961	2,961
Adjusted R-squared	0.116	0.127	0.129	0.225	0.268
Fstat, instrum., 1st stage	22.67	20.40	18.26	18.10	18.01
Baseline Controls	NO	YES	YES	YES	YES
Additional Controls	NO	NO	YES	YES	YES
			Panel B: First stage estimates		
Regional Commune	1.508***	1.498***	1.446***	1.437***	1.426***
	(0.316)	(0.331)	(0.338)	(0.336)	(0.336)
			Panel C: Reduced form estimates		
Regional Commune	0.256***	0.234***	0.231***	0.163***	0.163***
	(0.061)	(0.061)	(0.063)	(0.058)	(0.058)

Dependent Variable is Log (1 + Births). Standard errors (clustered at the NUTS 2 region level) in parentheses. Period Dummies and City FE always included.

Table 13: Determinants of the migration of famous creatives

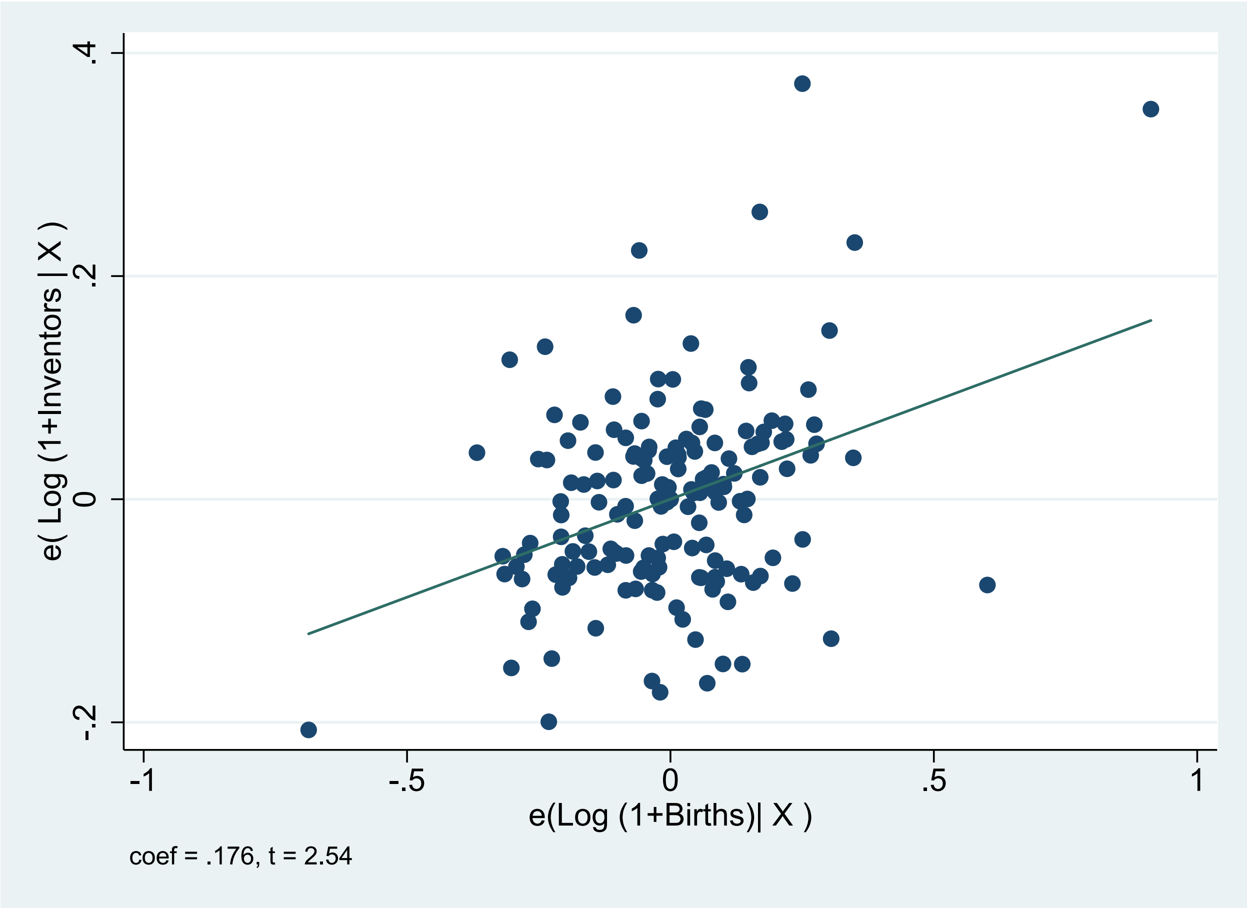
	(1)	(2)	(3)
To Population	0.220***	-0.074***	-0.053*
	(0.060)	(0.027)	(0.029)
To Commune	0.909***	0.576**	0.587**
	(0.169)	(0.232)	(0.240)
To Plundered	-0.540**	0.104	0.088
	(0.267)	(0.252)	(0.300)
To Capital	2.681***	0.945***	0.877***
	(0.289)	(0.203)	(0.206)
To Archbishop	-0.072	-0.310	-0.264
	(0.255)	(0.312)	(0.329)
To University	0.369*	-0.350	-0.423
	(0.211)	(0.263)	(0.328)
To Bishop	-0.438**	-0.618*	-0.580
	(0.211)	(0.355)	(0.358)
Log From (Unscaled) Births	0.889***	0.935***	
	(0.046)	(0.050)	
From Population	-0.102***	0.001	
	(0.029)	(0.044)	
From Commune	0.166*	0.056	
	(0.085)	(0.135)	
From Plundered	0.299**	0.122	
	(0.131)	(0.162)	
From Capital	0.062	0.028	
	(0.121)	(0.157)	
From Archbishop	0.143	0.065	
	(0.113)	(0.385)	
From University	-0.002	0.206	
	(0.069)	(0.152)	
From Bishop	0.017	-0.399***	
	(0.092)	(0.152)	
Distance	-0.163***	-0.190***	-0.190***
	(0.032)	(0.026)	(0.026)
Same Country	0.945***	1.048***	1.053***
	(0.254)	(0.149)	(0.155)
Same First Lang	2.303***	1.972***	1.936***
	(0.209)	(0.163)	(0.160)
Same NUTS 1	0.947***	0.855***	0.803***
	(0.156)	(0.136)	(0.133)
Same State	0.848***	1.968***	2.093***
	(0.259)	(0.202)	(0.197)
Dyadic Observations	495,003	394,368	360,380
Period Dummies	YES	YES	NO
Destination FE	NO	YES	YES
Origin FE	NO	YES	NO
Origin-Year FE	NO	NO	YES

Dependent Variable is Flow of Immigrants. Twoway clusterd standard errors (by origin and by destination) in parentheses. Population size measured in 100,000.

Appendix - For Online Publication

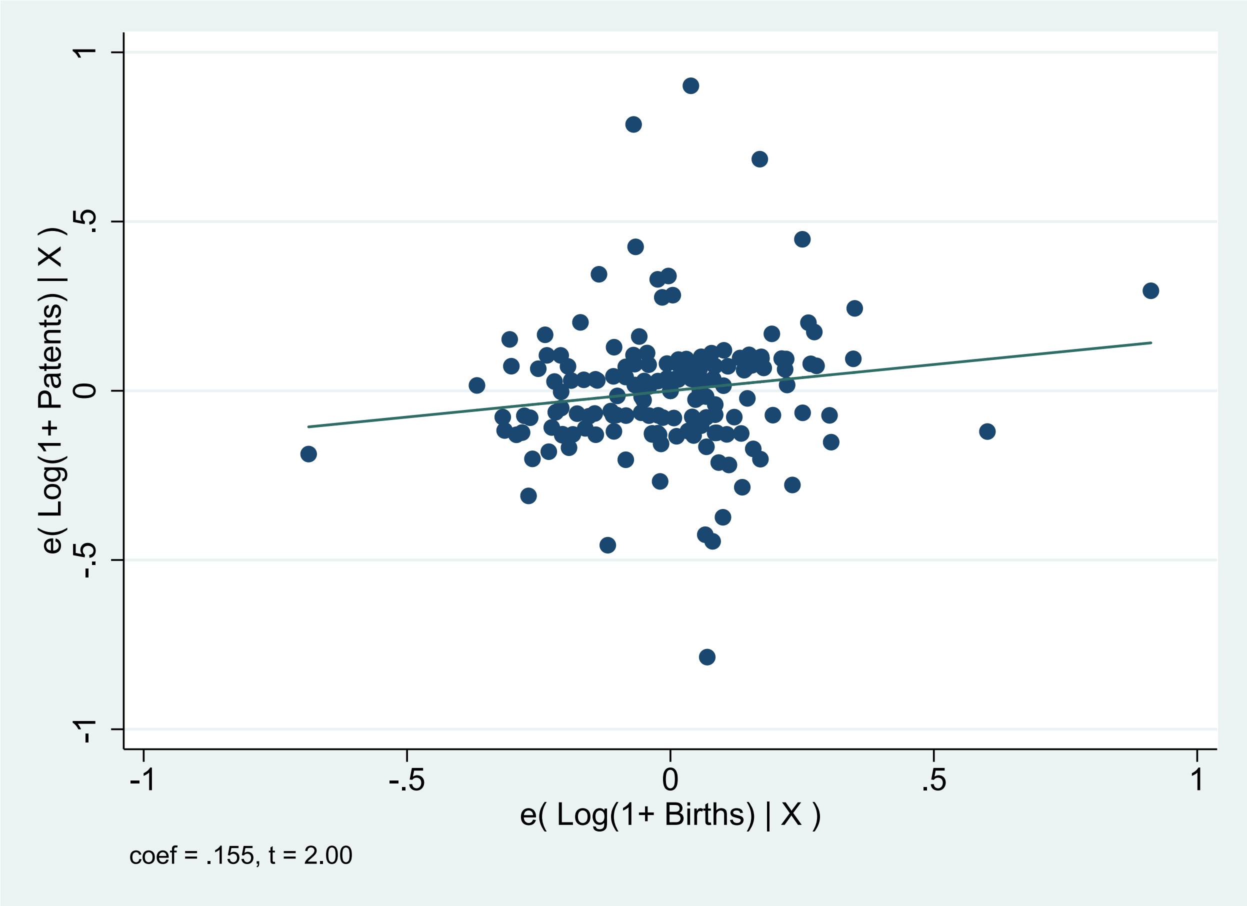
A.I Additional Figure and Tables

Figure A.1: Famous Creatives and Inventors, added-variable plot



We control for City FE, Century FE, Large state, Bishop, Archbishop, Capital, Plundered, Commune, Population, University. Standard Errors clustered by Region.

Figure A.2: Famous Creatives and Patents, added-variable plot



We control for City FE, Century FE, Large state, Bishop, Archbishop, Capital, Plundered, Commune, Population, University.Standard Errors clustered by Region.

Figure A.3: Births, XIXth century

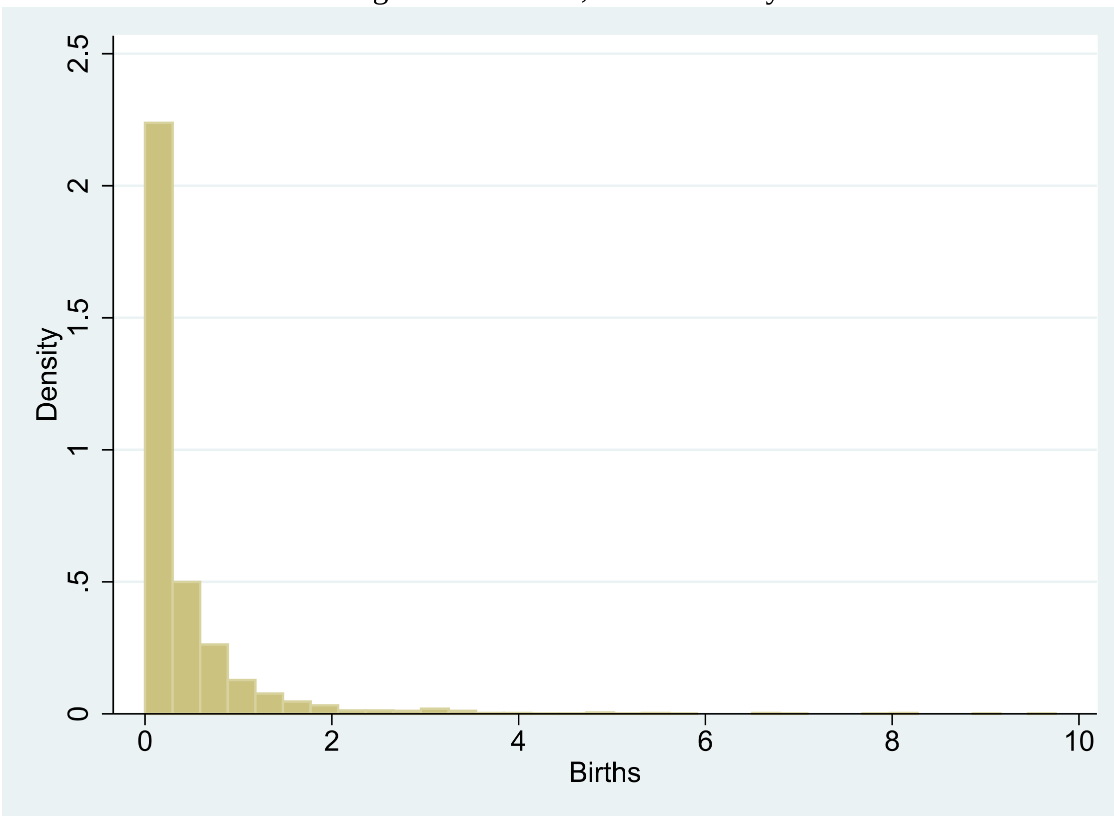
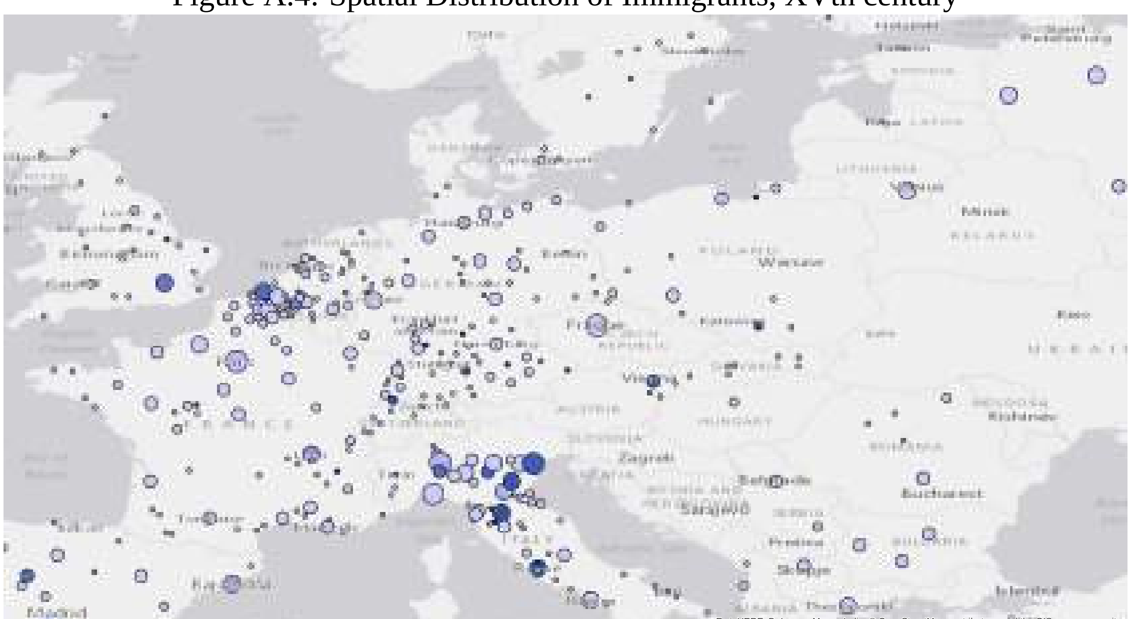
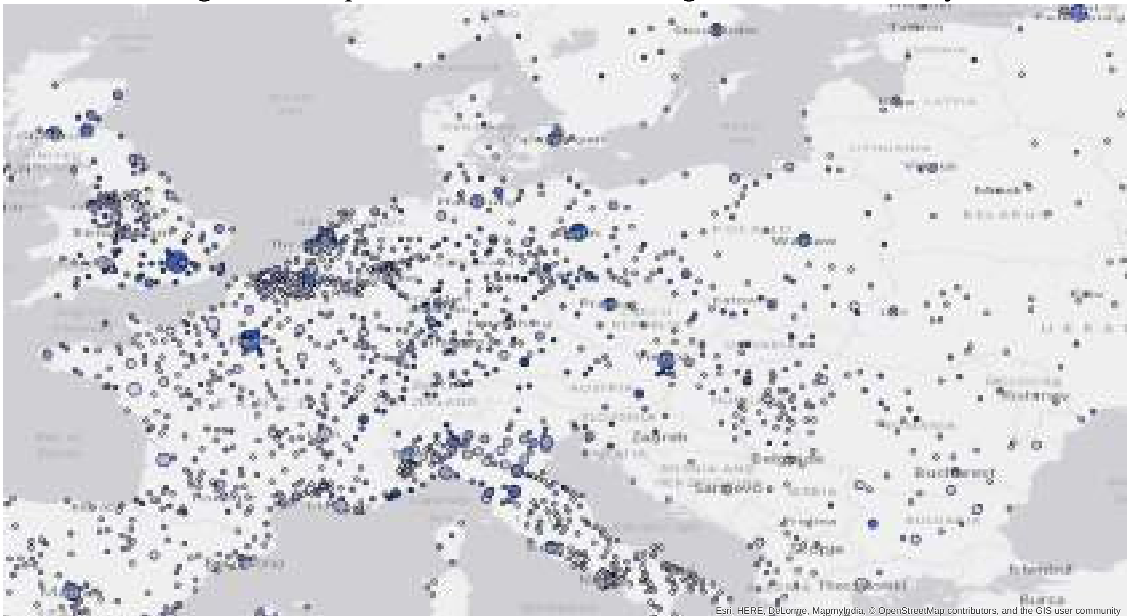


Figure A.4: Spatial Distribution of Immigrants, XVth century



the darker the tone, the higher the number of famous immigrants, per 1000 inhabitants. The larger circle, the larger the population of the city. See Table A.3 for descriptive statistics.

Figure A.5: Spatial distribution of Immigrants, XIXth century



the darker the tone, the higher the number of famous immigrants, per 1000 inhabitants. The larger circle, the larger the population of the city. See Table A.4 for descriptive statistics.

Figure A.6: Commune Status over Time

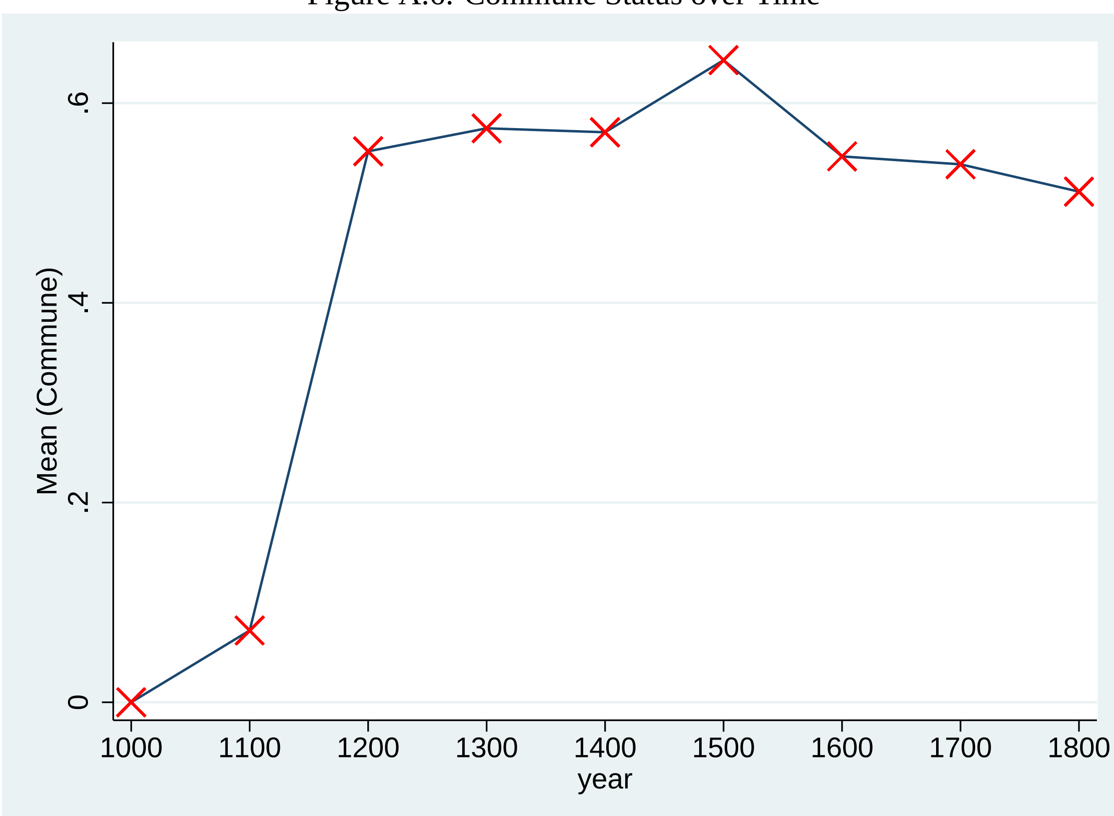


Figure A.7: Entry into Commune Status

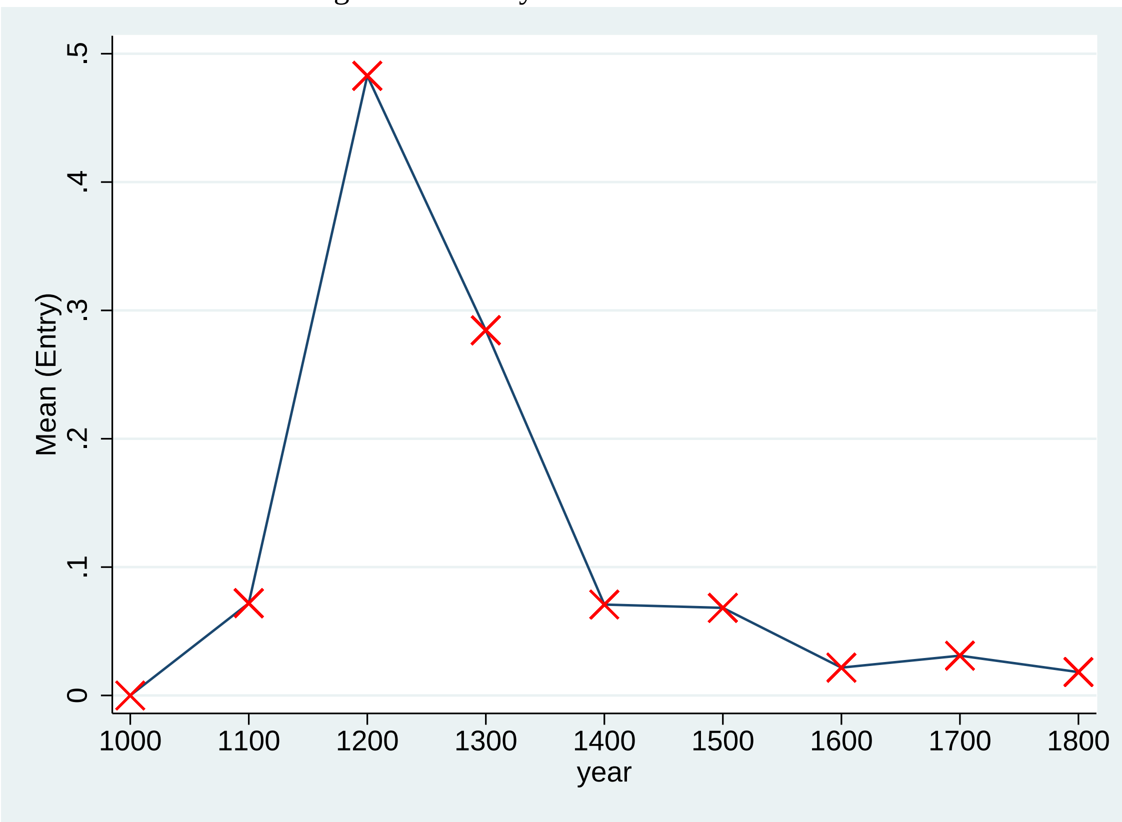


Figure A.8: Exit from Commune Status

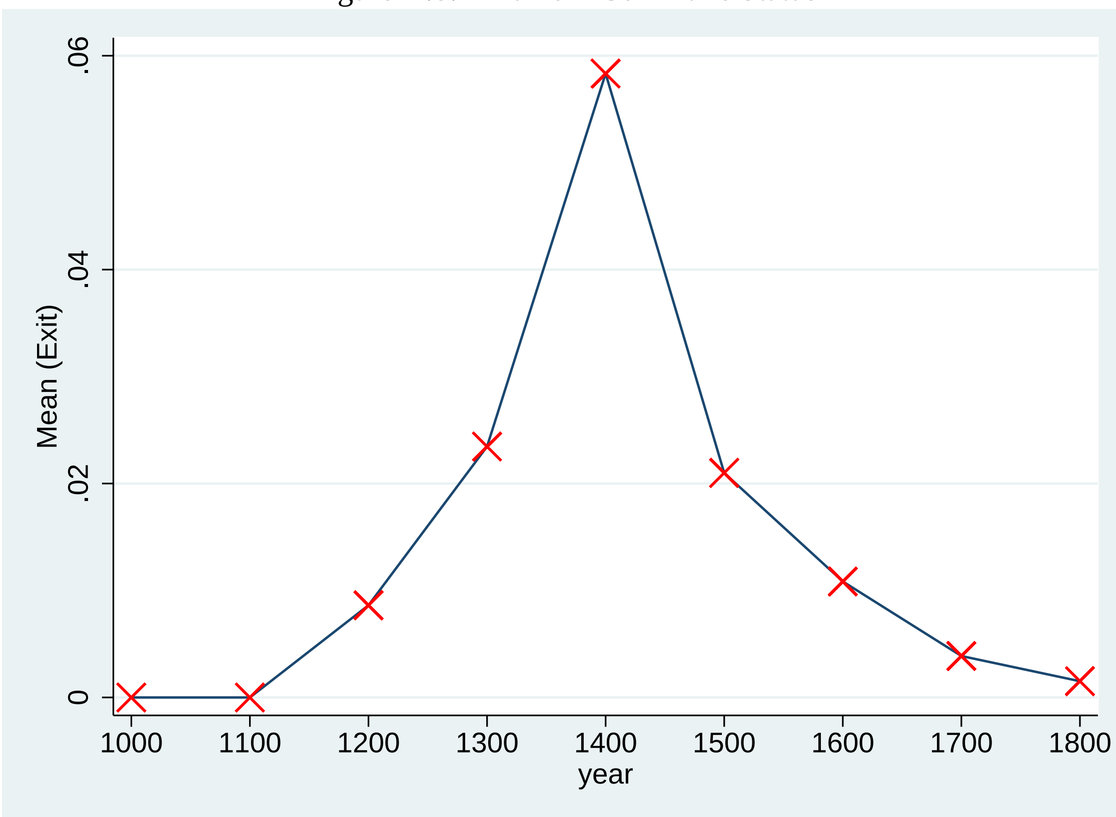


Table A.1: Polymaths

	Number of people
Performing arts + Non-performing arts	1072
Performing arts + Humanities and Sciences	40
Performing arts + Business	15
Non-performing arts + Humanities and Sciences	1245
Non-performing arts + Business	211
Humanities and Sciences + Business	181
Total two categories	2764
Performing arts + Non-performing arts + Humanities and Sciences	180
Performing arts + Non-performing arts + Business	10
Performing arts + Humanities and Sciences + Business	2
Non-performing arts + Humanities and Sciences + Business	55
Total three categories	247
All four categories	5
Total	21906

Note: Many of the creative individuals we considered achieved fame in multiple fields. For example, each of those listed under “Non-performing arts + Humanities and Sciences" achieved prominence both as a non-performing artist and in some field of the humanities or sciences, though not as a performing artist or in business.

Table A.2: Time Coverage of Data on Wages

City	Period
Amsterdam	1500-1910
Antwerp	1373-1913
Augsburg	1502-1803
Barcelona	1500-1804
Cambridge	1450-1700
Canterbury	1450-1700
Cavaillon	1600-1785
Dover	1450-1700
Edinburgh	1553-1642
Florence	1326-1913
Gdansk	1535-1814
Ghent	1835-1914
Krakow	1409-1910
Leipzig	1520-1913
London	1264-1913
Lviv	1520-1800
Lyon	1500-1592
Madrid	1520-1913
Milan	1520-1913
Munich	1427-1765
Naples	1514-1806
Oxford	1264-1913
Paris	1400-1911
Strasbourg	1395-1875
Valencia	1392-1785
Valladolid	1502-1560
Vienna	1440-1913
Warsaw	1558-1913

Note: Data do not necessarily cover the whole period as reported here; some gaps may be present. Sources: Allen (2001), Bennassar (1999), Boulton (1996), Cabourdin (1968), Feliu (1991), Gibson and Smout (1995), Rappaport (2002), and Scholliers and Avondts (1977)

Table A.3: Summary Statistics, XVth century

Variable	**Mean**	**(Std. Dev.)**	**Min.**	**Max.**	**N**
Population	11615.752	(17656.554)	1000	200000	419
Births	0.094	(0.251)	0	2	419
Immigrants	0.053	(0.157)	0	1.333	419

Table A.4: Summary Statistics, XIXth century

Variable	**Mean**	**(Std. Dev.)**	**Min.**	**Max.**	**N**
Population	12035.951	(31100.963)	1000	948000	2114
Births	0.617	(1.168)	0	13.25	2114
Immigrants	0.371	(1.51)	0	30	2114

Table A.5: Commune, First Stage

	(1)	(2)	(3)	(4)	(5)
Regional Commune	1.450***	1.381***	1.446***	1.434***	1.427***
	(0.207)	(0.205)	(0.338)	(0.337)	(0.336)
Large state		-0.053***	-0.020	-0.015	-0.014
		(0.019)	(0.032)	(0.032)	(0.032)
Bishop		0.126***	-0.016	-0.012	-0.012
		(0.037)	(0.054)	(0.054)	(0.054)
Archbishop		0.077	-0.109*	-0.107	-0.103
		(0.048)	(0.065)	(0.065)	(0.065)
Capital		0.080	-0.005	-0.004	-0.006
		(0.050)	(0.055)	(0.055)	(0.054)
Plundered		-0.032	-0.025	-0.024	-0.026
		(0.024)	(0.033)	(0.033)	(0.033)
Log (Population)			0.051***	0.049***	0.050***
			(0.012)	(0.012)	(0.012)
University			-0.012	-0.012	-0.017
			(0.054)	(0.054)	(0.054)
Spatial Lag of Log (1 + Births)				0.185**	0.226**
				(0.094)	(0.093)
Spatial Lag of Log Population					-0.009***
					(0.003)
Observations	7,227	7,227	3,110	3,110	3,110
Number of ID	657	657	657	657	657
Adjusted R-squared	0.446	0.454	0.427	0.427	0.429

Dependent Variable is Commune. Standard errors (clustered at the NUTS 2 region level) in parentheses. Period Dummies and City FE always included.

Table A.6: IV Estimates of the Effect of Commune on Births, using the Yu et al. data

	(1)	(2)	(3)	(4)	(5)
Commune	0.229**	0.259**	0.260**	0.237*	0.228*
	(0.113)	(0.116)	(0.123)	(0.122)	(0.122)
Spatial Lag of Log (1 + Births)				0.473**	0.538***
				(0.209)	(0.206)
Spatial Lag of Log Population					-0.019***
					(0.007)
Observations	2,961	2,961	2,961	2,961	2,961
Adjusted R-squared	-0.171	-0.170	-0.171	-0.164	-0.160
Fstat, instrum., 1st stage	21.60	19.04	17.03	16.96	16.86
Baseline Controls	NO	YES	YES	YES	YES
Additional Controls	NO	NO	YES	YES	YES

Dependent Variable is Log (1 + Births Yu et al.). Standard errors (clustered at the NUTS 2 region level) in parentheses. Period Dummies and City FE always included.

Table A.7: IV Estimates, Robustness to Inclusion of Country-Specific Linear Trends, and to Sample Composition

	(1)	(2)	(3)	(4)	(5)	(6)
Commune	0.148	0.181*	0.361***	0.361***	0.290**	0.297**
	(0.090)	(0.100)	(0.132)	(0.135)	(0.124)	(0.119)
Log (1+ Unscaled Immigrants)	0.132***	0.129***	0.163***	0.148***	0.169***	0.153***
	(0.024)	(0.024)	(0.028)	(0.027)	(0.028)	(0.026)
Spatial Lag of Log (1 + Births)	0.443***	0.607***	0.475***	0.720***	0.608***	0.846***
	(0.144)	(0.148)	(0.146)	(0.160)	(0.138)	(0.136)
Spatial Lag of Log Population		-0.020***		-0.035***		-0.033***
		(0.006)		(0.011)		(0.007)
Observations	1,625	1,625	1,456	1,456	1,338	1,338
Adjusted R-squared	0.428	0.425	0.263	0.298	0.294	0.337

Dependent Variable is Log (1 + Births). Columns 1-2 include linear country-specific time trends. Columns 3-4 drop the city-year observations after a negative transition (from Commune=1 back to Commune=0). Columns 5-6 drop the city-year observations prior to a positive transition (from Commune=0 to Commune=1). Full set of city-level controls always included. Standard errors (clustered at the NUTS 2 region level) in parentheses. Period Dummies and City FE always included. Commune is instrumented using Regional Commune.

Table A.8: Summary Statistics for the Dependent Variable in the Gravity Model

Variable	**Mean**	**(Std. Dev.)**	**Min.**	**Max.**	**Percent of Zeros**	**N**
Immigrants (dyadic obs.)	0.008	(0.13)	0	14	99.42	366763