

# The geography of innovation in Italy, 1861–1913: evidence from patent data

ALESSANDRO NUVOLARI\* AND  
MICHELANGELO VASTA\*\*

\**Sant’Anna School of Advanced Studies, Institute of Economics, Pisa, Italy, [alessandro.nuvolari@santannapisa.it](mailto:alessandro.nuvolari@santannapisa.it)*

\*\**Department of Economics and Statistics, University of Siena, Siena, Italy, [michelangelo.vasta@unisi.it](mailto:michelangelo.vasta@unisi.it)*

In this paper we provide a systematic appraisal of the spatial patterns of inventive activity in Italy in the period 1861–1913 using patent data. First, we introduce a characterization of the spatial distribution of patents and of its evolution over time. Second, we assess the connection between different forms of human capital and patent intensity. We establish a robust correlation between secondary technical education and science and engineering university studies and patent activity. Third, we study the connection between patents and industrialization. Our main finding is that inventive activities were an important element of the industrialization process, even in a latecomer country such as Italy.

## 1. Introduction

The origins and the nature of the divide in economic performance between Northern and Southern Italy is an issue that has vexed more than one generation of economic historians. At this stage, the debate is still ongoing without sign of imminent closure.<sup>1</sup> In comparison with the traditional literature on the *questione meridionale* (the “Southern question”), the most recent contributions are characterized by attempts to provide more sophisticated quantitative and geographically disaggregated assessments of the historical dynamics of economic modernization. In particular, considerable progress has been obtained concerning indicators of economic performance, so that two recent and somewhat conflicting estimates for GDP per-capita are now available (Daniele and Malanima, 2011; Felice, 2013). Relatedly, also our knowledge of the spread of industrialization is much improved thanks to the estimates of industrial value added and of industrial labor force constructed at the fine-grained provincial level by Ciccarelli and Fenoaltea (2013) and by Ciccarelli and Missiaia (2013). In addition, the most recent literature has also produced both new regional estimates in broader dimensions of living standards using proxies such as heights, nutrition, education and human capital, infant mortality, life expectancy, etc. (Vecchi, 2011) and aggregate indicators of human development (Felice and Vasta, 2015). Finally, one should also point to the construction of regional indicators of social capital (Felice, 2012; Cappelli, 2017) following the lead of Putnam, Leonardi and Nanetti (1993) and Helliwell and Putnam (1995).

<sup>1</sup> See, for example, the recent controversy between Daniele and Malanima (2014a, 2014b) and Felice (2014).

Notwithstanding, this burgeoning literature, there has been so far only very limited research on the regional patterns of inventive activities. This is probably related to the dominance, in the Italian economic historiography, of a narrow interpretation of the Gerschenkronian framework for the study of European industrialization which has overlooked the crucial role of technology gaps (see [Federico and Toniolo, 1991](#) for a revealing example).<sup>2</sup>

This is even more surprising taking into account that, over the last twenty years or so, the geographical dimensions of inventive activities and the emergence, consolidation and demise of regional innovation systems are themes that have featured prominently in the economics literature. Indeed, this recent “innovation studies” literature suggests that differences in regional innovative activities are one of the primary candidates for explaining convergence and divergence in economic performance at regional level ([Feldman and Kogler, 2010](#)). This paper aims to fill this research gap by providing a preliminary examination of the geography of innovation in Italy during the Liberal Age (1861–1913). It is worth noting that, although Italy was a latecomer country characterized by a weak innovation system ([Nuvolari and Vasta, 2015a](#)), it was able to develop independently a number of significant breakthrough inventions including “macroinventions” à la [Mokyr \(1990\)](#): Ascanio Sobrero (1812–1888), an academic chemist working in Turin, in 1847 discovered nitroglycerine; Antonio Pacinotti (1841–1912) in 1860 made a major contribution to the development of the electric dynamo; finally, most notably, Guglielmo Marconi (1874–1937) pioneered the first successful long-distance radio-transmission equipment.

Our inquiry is based on a newly developed data-set comprising all patents granted in Italy in five benchmark years. In particular, we examine the relationship between human capital and inventive activities. Our focus on the innovation-human capital nexus is easily motivated by noting that the different trajectories of human capital accumulation are considered a crucial determinant of the economic divide between Northern and Southern Italy ([Felice, 2012](#); [Cappelli, 2015](#)). In this perspective, the connection between human capital and geographical patterns of innovative activities is a research issue that, so far, has not been explicitly tackled. Furthermore, it is worth noting that the literature dealing with the role of human capital in nineteenth century economic growth has been so far mostly concerned with “basic” (primary) education and literacy rates. In fact, it is possible to point to a stream of literature suggesting that during the period 1861–1913 the gap between North and South in technical education and in science and engineering university studies may have been even more critical than the gap in basic education ([Vasta, 1999](#); [Lacaita and Poggio, 2011](#)).<sup>3</sup> Interestingly enough, this line of research resonates with the recent view put forward by [Mokyr \(2005a\)](#) arguing that, in the early phases of industrialization, a key-role is played by the upper tail of the human capital distribution, rather than by the human capital of the average worker. In this paper, accordingly, we will consider human capital formation in comprehensive manner, taking into account both “basic” education, secondary technical education and university engineering studies. The rest of the paper is organized as follows. Section 2 contains a description of the data and the sources. Section 3 is devoted to the reconstruction of the (changing) geography of innovation in Italy. In Section 4, we assess the factors accounting for the concentration of patenting activities at provincial level.

<sup>2</sup> For a more elaborated discussion of this point, see [Vasta \(1996, pp. 27–32\)](#). Given these premises, it is not surprising that, in Italy, patents have remained for a long time a relatively neglected historical source.

<sup>3</sup> [Fenoaltea \(2011, pp. 243–244\)](#) also suggests that the inability of developing a sound system of secondary technical education and of scientific and engineering university studies with a genuine national coverage is one of the main “failures” of the Italian state in the Liberal Age.

Section 5 speculates on the possible relation between the geography of innovation and the process of industrialization. Section 6 concludes.

## 2. Sources and data

Our fundamental geographical unit is the “provincia”, which was an administrative unit of the time. Figure 1 contains a map of the provincial borders in 1911. The main advantage of using provincial data instead of regional ones is that they provide a relatively



Figure 1. Map of Italian provinces in 1911.

fine-grained picture moving from 16 to 69 geographical units and, therefore, provincial data may reveal patterns of spatial heterogeneity that may remain concealed in more aggregate data.

Following an established tradition both in economics of innovation and in the historical literature, we measure inventive activities at provincial level using patents.<sup>4</sup> We employ a data-set of Italian patents granted during the Liberal Age constructed by [Nuvolari and Vasta \(2015b\)](#).<sup>5</sup> This data-set contains all the 10,124 patents granted in Italy in five benchmark years. The data-set contains 3,892 “Italian” patents (that is patents whose inventors had stated an Italian residence) whose distribution across benchmark years is as follows: 1864–65 (254), 1881 (338), 1891 (486), 1902 (1,021), and 1911 (1,793).<sup>6</sup> The historical sources of these data are the Italian official serial publications of *Ministero di Agricoltura, Industria e Commercio* ([MAIC 1864–1885](#); [1886–](#), [1894–1901](#), [1902–1923](#)). The Nuvolari and Vasta patent data-set contains also information on the sectoral composition of the patents by classifying them in 14 industrial classes. In this paper, following [Nuvolari and Vasta \(2015b, pp. 871–872\)](#), we make use of the distinction between high-tech and low-tech patents, with high tech representing patents related with the technological paradigms of the First and the Second Industrial Revolution.<sup>7</sup>

Moreover, we have also retrieved information on the 7,578 granted in Italy in the same benchmark years to “foreign” inventors (that is patents whose inventors have stated a foreign residence).<sup>8</sup> We shall use the information on the sectors of activity of foreign patentees to assess the sectoral proximity of inventive activities at provincial level with foreign patterns of technological innovation. Several scholars have noted that one prominent feature of the patterns of technological accumulation of latecomer countries such as Italy is the absorption of international technological spillovers through the “creative” adaptation of foreign technology, rather than the development of autonomous innovations ([Nadiri and Kim, 1996](#); for some discussion on Italy in this period, see [Giannetti, 1998](#)). In this paper, we consider foreign patenting in Italy as a proxy of the flows from the international technological frontier.

To assess the role of absorptive capabilities of foreign technology, we construct a measure of the “technological distance” of each province to the sectoral distribution of foreign patents. We use a slightly modified version of the indicator of technological proximity proposed by [Bar and Leiponen \(2012\)](#):

$$B\&Ltech\_proximity = 1 - \left[ 1 - \sum_{k=1}^K \min(p_{ik}, p_{jk}) \right] \quad (1)$$

<sup>4</sup> For a recent comprehensive survey on this issue, see [Nagaoka, Motohashi and Goto \(2010\)](#).

<sup>5</sup> For an account of the evolution of the Italian patent systems in the preunitary states of the Restoration period to the legislation implemented after the political unification of the country, see [Nuvolari and Vasta \(2017\)](#).

<sup>6</sup> Given the small number of patents registered in the early years after the unification, we have decided to have an initial benchmark of two years, 1864 and 1865. For a full description of the database, see [Nuvolari and Vasta \(2015b\)](#).

<sup>7</sup> The high-tech sectors are chemicals, electricity, machine tools and machinery, steam engines, and weapons.

<sup>8</sup> In order to have a more comprehensive representation of the activities of foreign inventors in Italy, we added to the original data-set, which contains 6,227 foreign patents, also 1,351 patents granted on the basis of the existence of a former patent in a foreign country following the 1883 Paris convention (“imported” or “priority patents”) in three benchmark years (1891, 1902, and 1911).

in the formula,  $p_{ik}$  and  $p_{jk}$  indicate the shares of patents in sector  $k$ , respectively, of the province  $i$  and of the foreign patents ( $j$ ). The indicator is equal to 1 when the sectoral distributions of the province and of foreign patents are perfectly coincident and equal to 0 when there is no overlap in the sectoral distribution of patents.

In the construction of the data-set, patents were assigned to provinces using fractional counting.<sup>9</sup> This means that if a patent was granted to two inventors, one living in Milan and another living in Florence, at each province was assigned 0.5 of that patent.

One of the main limitations of patent data is that they typically comprise inventions of exceedingly different quality ranging from minor improvements to genuine technological breakthroughs (Schmookler, 1966). In order to address this issue, we have constructed an indicator of patent quality using duration data. Our quality indicator is the “scheduled” duration of the patent (measured in years).<sup>10</sup> The intuition is straightforward: patents taken for longer durations are seen, in the eyes of the patentees, as covering more important inventions than patents of shorter duration. Therefore, this indicator may be interpreted as representing an *ex ante* assessment of the value of the patent, with some possible revisions due to the extensions. For the benchmark years 1881, 1891, and 1902, we have also retrieved information on the fees that each patentee paid throughout the life time of the patent. This allows us to establish the “real” duration of the patent, that is the number of years for which the patent was actually in force. This indicator may be regarded as providing an *ex post* assessment of the value of the patent and as such it can represent a useful integration to “scheduled” duration. (Nuvolari and Vasta, 2015b, pp. 872–876).<sup>11</sup>

We construct a number of variables describing human capital formation at provincial level. The first variable of this type is the literacy rate of the province computed as number of people able to read and write on total population (MAIC, 1865, 1883, 1903, 1914).<sup>12</sup> Furthermore, we use a completely new set of estimates of the level of technical education at provincial level. We measure technical education as the number of students attending both *Scuole tecniche* and *Istituti tecnici*, which were respectively the lower and upper level of technical courses in the secondary education curriculum (Cives, 1990). The sources used for the construction of this variable are *Annuario Statistico Italiano* (MAIC, 1881, 1893, 1913) and *Ministero della Pubblica Istruzione [MPI]* (1883, 1901, 1912). Finally, we take into account the tertiary level of education using the number of students attending science and

<sup>9</sup> The article 24 of the *Regolamento* related to the (patent) Law no.1674 of 31 January 1864 established that patents applications were to be submitted to the *Ministero di Agricoltura, Industria e Commercio* (MAIC) via the local *Prefettura* (which was the local office representing the central government at provincial level). This means that, from an institutional point of view, the accessibility of the patent system was evenly distributed on the national territory. This also suggests that data on the residence of patentees can be reliably employed to characterize the geographical distribution of patenting.

<sup>10</sup> In Italy inventors could apply for a patent duration ranging from one to fifteen years (with increasing fees for longer duration). “Scheduled” duration comprises the initial duration plus “*prolungamenti*” (i.e., extension of the duration of the patent). We have checked the existence of extensions by looking at the *Bollettino* (MAIC 1864–1885, 1886–1893, 1894–1901, 1902–1923) for the following fourteen years after the patent was granted. It is worth noticing that for the 1911 cohort, we have followed the life of the patent up to 1923 because, in that year, a new Law (no. 1970) abolished the extensions and established a fixed patent duration to fifteen years. For a more detailed discussion of the computation of the “scheduled” duration, see Nuvolari and Vasta (2015b).

<sup>11</sup> The “real” duration has been computed for the benchmarks 1881, 1891, and 1902 since only for these periods the expiration of patents was published systematically in the *Gazzetta Ufficiale* (Nuvolari and Vasta 2015b, p. 867).

<sup>12</sup> For the year 1891 (when the Census was canceled due to the financial difficulties of the Kingdom), we have interpolated the data from the population censuses of 1881 and 1901.

engineering at universities in each province.<sup>13</sup> The sources of these variables are MAIC (1881, 1893, 1913) and MPI (1883, 1901, 1912).<sup>14</sup>

In order to account for the size of the provinces, we use population data collected from the population censuses (MAIC, 1865, 1883, 1903, 1914).<sup>15</sup> Besides human capital and technical skills, a rich stream of literature has highlighted the critical importance of agglomeration and urbanization on inventive activities (Mokyr, 1995). Accordingly, we have constructed a variable that measure urbanization by considering at the number of people living in cities with more than 30,000 inhabitants. It is worth noting that this variable has been thoroughly reconstructed for each benchmark year paying particular attention to the cities that were moving above or below the threshold.

In addition, we also assess the possible role played by the “intensity” of “access to information” available in each province. The potential role of this variable for innovation activities has been noted by Mokyr (2005b). We measure this factor by looking at the number of newspapers and periodicals published in a province. The source for this variable is the *Annuario Statistico Italiano* (MAIC, 1881, 1893, 1908, 1913).<sup>16</sup> Of course, this is a crude proxy, but it is difficult to construct more sophisticated indicators at the provincial level. Another possible determinant of inventive activities is the transport infrastructure. In this period, railways were clearly the fundamental invention that revolutionized the structure of transport systems. It is worth noting that at the Unification the railway network of the country was largely incomplete and it had an eminently local nature reflecting the borders of the preunitary states. Instead, at the end of the period considered, the railway network became possibly the most important transport infrastructure of the country. We use the number of kilometers of railways over the surface of the province to measure the density of railway infrastructure (MLP, 1878; *Ferrovie dello Stato*, 1911; Ciccarelli and Groote, 2017). Finally, we consider the possible role played by the density of manufacturing activities in the province. This effect is measured by using value added in manufacturing per male worker (Ciccarelli and Fenoaltea, 2013; Ciccarelli and Missiaia, 2013).<sup>17</sup> Table A.1 in Appendix contains the descriptive statistics of the variables presented above.

In the second part of the paper where we examine the relationship between inventive activities and industrialization (section 5), we make use of an additional set of variables which can be regarded as “proximate” factors of industrial localization. The first is the availability of water resources, which is measured by computing the average yearly discharge (flow) of rivers, canals, and streams in the province (measured in m<sup>3</sup>/s).<sup>18</sup> The second is

<sup>13</sup> We consider government universities, “Università libere” such as Bocconi University, “Politecnici” and institutions offering “certified” graduate courses of tertiary education.

<sup>14</sup> In this paper we consider only formalized processes of human capital formation. Of course, in this period a relevant part of human capital was also represented by skills accumulated by means of learning by doing on the shop floor. See Roses (1998) who highlights the role of this type of human capital in the context of a latecomer region (Catalonia).

<sup>15</sup> The population for 1891 has been interpolated using the observations of 1881 and 1901.

<sup>16</sup> Since the data contains information only on newly founded newspapers and periodicals, the observation for 1864 has been estimated assuming that in the period 1864–1880 there were only new newspapers and periodicals founded and no newspaper disappeared.

<sup>17</sup> We use this indicator of productivity because Italian population censuses tend to overestimate the female workforce of Southern provinces (Ciccarelli and Missiaia 2013). This variable is available only for the benchmarks 1881, 1901 and 1911.

<sup>18</sup> These data have been retrieved from the website [www.acq.isprambiente.it/pluter/](http://www.acq.isprambiente.it/pluter/) (constructed by the “Istituto Superiore per la Protezione e Ricerca Ambientale”) and they refer mostly to the period 1950–1970 (the assumption here is that this should be a still a reasonable proxy for the XIX century). Although the data are characterized

the level real wages. In this case we measure this variable by using a so far unexploited comprehensive survey of the hourly wages of unskilled workers (MAIC n.d.). In particular, the source reports the hourly wages of *terraiolo*, an unskilled worker in the construction sector which was employed for digging and transporting ground. This source takes into account if the workers received food or accommodation as part of their salary. We compute the real wages by dividing the nominal wages with the price of bread at provincial level.<sup>19</sup> These estimates of real wages refer to the 1874–1878 period.

Finally, in Section 5, we make use of the estimates of domestic market potential constructed by Missiaia (2016).

### 3. The geography of inventive activities: a preliminary snapshot

The maps reported in figure 2 show the geographical distribution of patents per million inhabitants. Overall, the figure suggests a clear pattern of regional differentiation in terms of the three major areas of the country (North, Center, and South). Besides the cities of the industrial triangle, the other provinces characterized by high densities of patenting activities are urbanized provinces with large populations (Roma, Palermo, Napoli, etc.). Initially (1864–1865 map), the distribution of patents is strongly concentrated in the North and in the northern provinces of the Center (especially in Tuscany). Notably, even in these regions, the distribution of patents is rather skewed with few provinces—Torino, Genova, Milano, Firenze, and Livorno holding the major bulk of patents. The industrial triangle (Torino, Genova, and Milano) is already clearly delineated, although in a somewhat embryonic shape. Subsequently, two main trends stand out. The first is an increasing spread of patenting activities: in 1864–1865 there are twenty seven provinces out of fifty nine with zero patents and fifty two out of fifty nine with less than five patents, whereas in 1911 there are only six provinces out of sixty nine with zero patents and thirty one out of sixty nine with less than five patents. Remarkably, this process of spatial diffusion is coupled with a process of growing concentration of the bulk of patenting activities in a few selected areas of the country. The main concentration is the industrial triangle which becomes clearly visible in all benchmark years since 1881. In 1911, Roma is also a province with a strong density of patenting activities, possibly reflecting its administrative role as the capital of the country. Overall the maps of figure 2 also suggest that regions may be rather heterogeneous as far as inventive activities are concerned, so that it is not uncommon to see provinces with high levels of patent per capita next to provinces with low levels.

The maps reported in figure 3 contain the geographical distribution of patents adjusted for their quality. More specifically, the figure reports the number of patents with duration of at least ten years per million inhabitants. The patterns are similar to the one emerging in figure 2. Also in this case, the industrial triangle is clearly the predominant area.

Figure 4 contains maps illustrating the geographical distribution of patents granted in high-tech sectors. Again we find a broadly similar pattern to that of the previous two figures.

by a relatively even distribution of gauging stations from a geographical point of view, we have decided to compute the average discharge only for the stations reporting a value higher than 2 m<sup>3</sup>/s, in order to limit the potential distortions caused by provinces with an over representations of stations on minor water flows. Furthermore, there are a few cases of provinces with missing data. In this case we have decided to attribute to the province an average value computed as the average value of all the neighboring provinces (Ancona, Bari, Cremona, Lecce, Napoli, Rovigo, Sondrio, Venezia and Verona). This approach to the measurement of water flows has been also used recently by Crafts and Wolf (2014) to study the location of the cotton industry in the UK.

<sup>19</sup> Data on bread prices at provincial level have been kindly provided to us by Giovanni Federico.

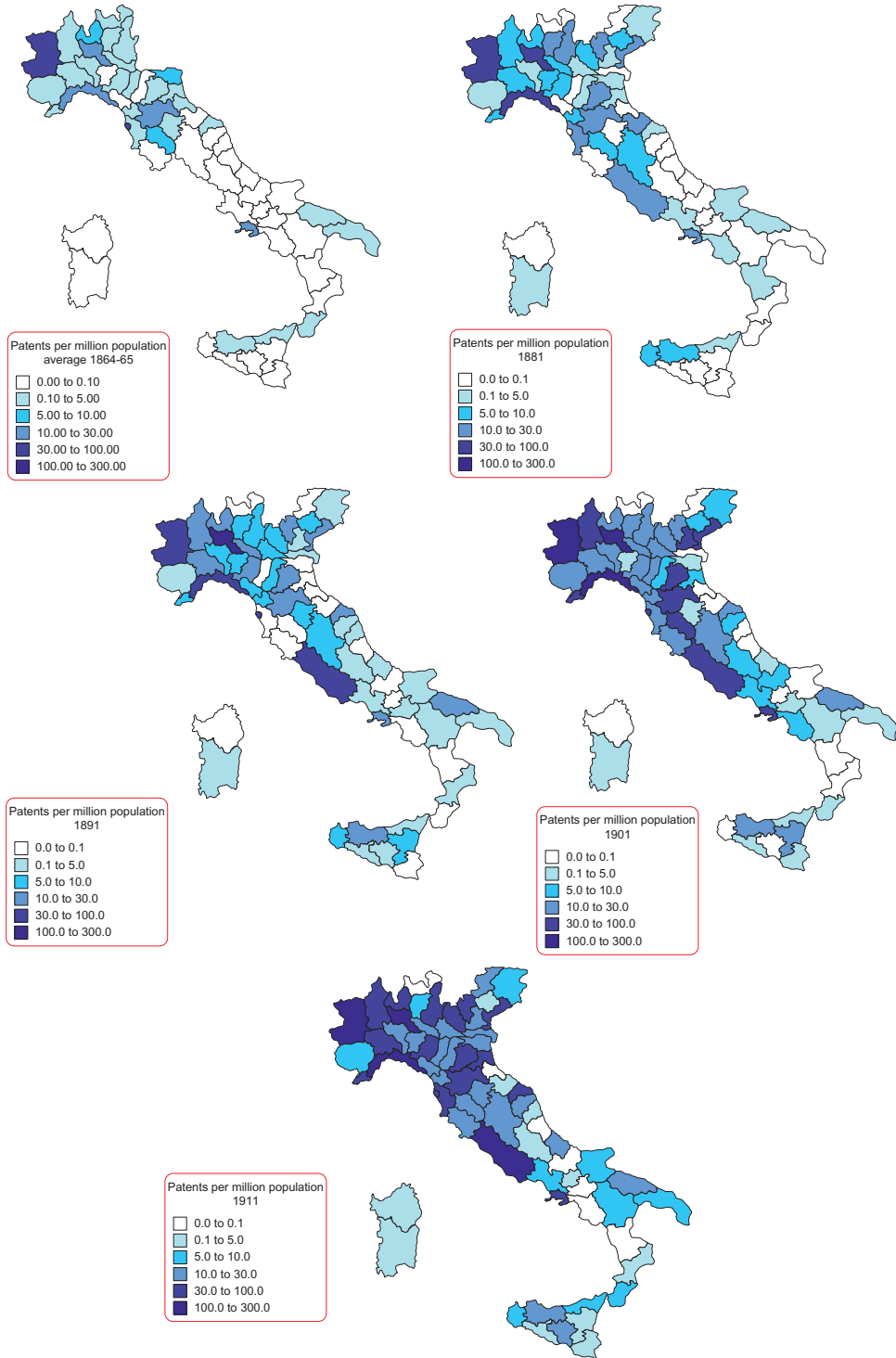


Figure 2. *Geographical distribution of patents per million population, 1864/65–1911.*



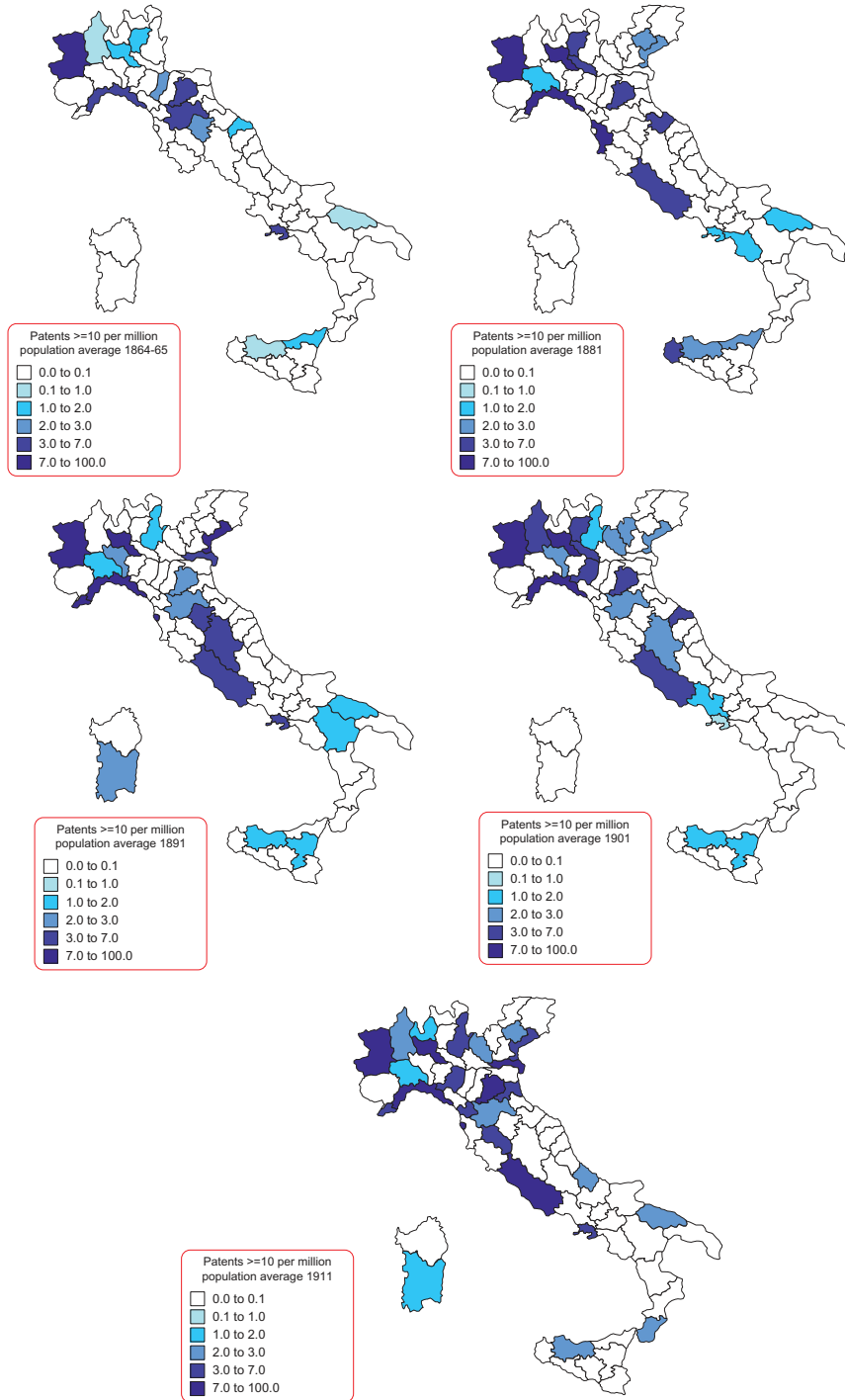


Figure 3. Geographical distribution of patents  $\geq 10$  per million population, 1864/65–1911.

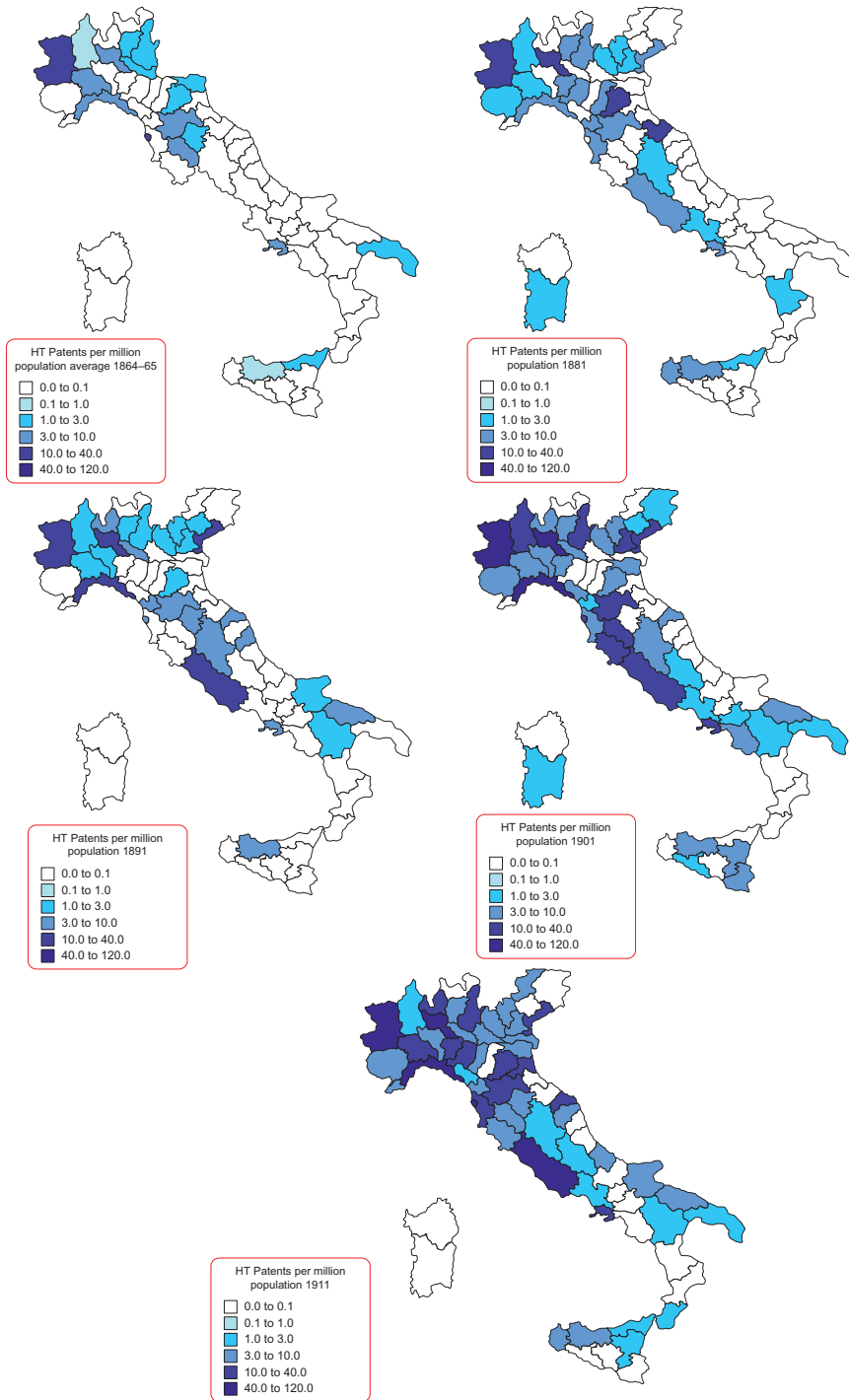


Figure 4. *Geographical distribution of HT patents per million population, 1864/65–1911.*

Interestingly enough, in terms of geographical concentration, the patterns of figure 4 seems an “intermediate” case between figure 2 and figure 3, which suggests that inventive activities in high-tech sectors were not characterized by a specific pattern of localization which substantially differed from that of regular patents.

Table 1 examines the changing correlation of patenting activities over time. Two points, in this case, merit attention. The first is that there is a clear path-dependent effect in the location of patenting activities, so that provinces with higher density of patents are characterized by higher densities also in the subsequent periods. The second point is that the strength of this path-dependent effect tends to increase over time.

In order to examine the possible historical legacy of pre-Unification institutional set-ups, table 2 contains the distribution of patents in the first two benchmark years (1864–65 and 1881) considering the geographical borders of the preunitary states. In this case, we consider the first benchmark as a proxy for patenting in the pre-Unification period. Table 2 indeed suggests that the early divide in innovative activities between the North, the Center and the South is likely to have its origins before the Unification.

#### 4. Inventive activities and human capital formation

In this section, we examine systematically the relationship between inventive activities and human capital formation. As dependent variables, we consider the number of different types of patents (all patents, quality adjusted patents, and patents in high-tech sectors)

Table 1. *Inter-temporal correlation patents per million population, 1864/65–1911*

1881–1864/65	0.6164***
1891–1881	0.8878***
1902–1891	0.9185***
1911–1902	0.9491***

Source: our own elaboration.

Note: \*\*\*indicate significance level of 1%.

Table 2. *Distribution of patents according to preunitary states borders (1864/65–1881)<sup>a</sup>*

	1864–65			1881		
	No. patents	Population	Patents per million population	No. patents	Population	Patents per million population
Regno di Sardegna	130.5	4,123,803	31.6	98.66	4,644,625	21.2
Regno Lombardo-Veneto	42.5	2,998,181	14.2	128	6,494,788	19.7
Ducato di Modena e Reggio	2	631,378	3.2	1.33	693,682	1.9
Ducato di Parma e Piacenza	1	474,598	2.1	4	494,023	8.1
Granducato di Toscana	34	1,826,334	18.6	18	2,039,400	8.8
Stato Pontificio	10	2,436,683	4.1	49	3,579,966	13.7
Regno delle Due Sicilie	34	9,179,703	3.7	39	10,513,144	3.7
Total	254	21,670,680	11.7	337.99	28,459,628	11.9

Source: our own elaboration.

<sup>a</sup>For 1864–65 Regno Lombardo-Veneto does not include the provinces of Veneto and Mantova and Stato Pontificio does not include Roma since both areas were not yet unified with the rest of the country.

granted in each province (fractional numbers have been rounded to the nearest integer). Since our dependent variables are count variables, we estimate the following Poisson regression model<sup>20</sup>:

$$E(\text{Patents}|X) = \exp(\alpha + \beta_1 \cdot \ln(\text{Literacy}) + \beta_2 \cdot \ln(\text{Tech\_Education} + \beta_3 \cdot \text{S\&T\_University} + \beta_4 \cdot \text{B\&L\_techproximity} + \sum_c \gamma_c Z_c + \sum_d \delta_d \text{year}_d + \ln \text{population})) \quad (2)$$

As determinants of patents, we consider three types of human capital. The first is the “basic” human capital endowment of the province (this is measured as the log of the literacy rate of the province in the year in question: *Literacy*); the second is the endowment of technical skills (this is measured as the log of the share of students attending secondary technical and vocational schools on the population of the province: *Tech\_Education*); the third is an indicator capturing the number of university students enrolled in science and technical fields on the population of the province (*S&T\_University*).<sup>21</sup> In addition to these variables, we also include the Bar and Leiponen measure of technical proximity between the province and foreign technology (*B&L\_techproximity*).

As control variables (indicated with *Z* in the formula), we consider urbanization, measured as an indicator ranging on scores 1, 2 and 3 (*Urban\_pop*)<sup>22</sup>; a proxy for the transport infrastructure, measured as the log of railways km per square kilometer (*Rail\_kmq*); a proxy of “access to information” measured as the log of the number of newspaper per million inhabitants (*Newspaper*); the level of labor productivity in manufacturing measured as the ratio between value added and male labor force (*Labor\_Productivity*). We also control for variations over time by including year fixed effects (indicated with *year* in the formula). Moreover, our specifications include the population of the province (in 1,000 inhabitants) as an “exposure variable”. This means that our results must interpreted in terms of rates with respect to the “exposure variable”, in this case patents per 1,000 inhabitants (Hilbe, 2011, pp. 134–135). Notably, the estimated coefficients of the co-variates expressed in logarithm form can be interpreted as elasticities.

Tables 3a–3c report Poisson regressions for three different types of patents as dependent variables. Table 3a considers the simple number of patents; table 3b considers the number of high-quality patents that is the number of patents with scheduled duration  $\geq 10$  years, and table 3c considers the number of patents in high-tech sectors. The coefficient size of the literacy rate (*Literacy*) is in general positive and significant, except in the specifications where labor productivity (*Labor\_Productivity*) is included amongst the controls. This is not unexpected since the two phenomena are likely to be strongly related. At the same time, it must be considered that the data on the manufacturing labor productivity is available only for a smaller sub-sample. Secondary technical education (*Tech\_Education*) has a similar

<sup>20</sup> Gourieroux, Monfort and Trognon (1984) have shown that the estimates of the Poisson model are consistent even if the count variable is not Poisson distributed and the data are characterized by over-dispersion. We have also carried out a similar set of negative binomial regressions obtaining fully consistent results in terms of size and significance of the coefficients.

<sup>21</sup> The indicator is constructed on 0, 1, and 2 scale, with 0 indicating provinces with no university students in scientific and technology fields, 1 an intermediate situation and 2 the provinces in the top decile of the distribution. Since most provinces have no university students in scientific and technology fields, the use of the log specification adopted for literacy and secondary technical education would have forced us to drop a much too large number of observations.

<sup>22</sup> The indicator is constructed assigning a score of 0 to provinces with no urban population, a score of 1 to provinces in an intermediate position and a score of 2 to provinces in the top decile of the distribution.

Table 3a. *Patenting activity (“simple” patents) and human capital formation (pooled Poisson regressions)*

Variables	(1) PATENTS	(2) PATENTS	(3) PATENTS	(4) PATENTS	(5) PATENTS
ln ( <i>Literacy</i> )	2.641*** (0.290)	1.654*** (0.184)	0.787*** (0.177)	0.672*** (0.253)	0.331 (0.246)
ln ( <i>Tech_Education</i> )		1.179*** (0.159)	0.373** (0.169)	0.00193 (0.154)	0.171 (0.142)
<i>S&amp;E_University</i>		0.517*** (0.0718)	0.126** (0.0624)	0.182*** (0.0636)	0.158*** (0.0581)
<i>B&amp;L_techproximity</i>			2.891*** (0.313)	3.073*** (0.411)	2.589*** (0.352)
<i>Urban_pop</i>			−0.0862 (0.110)	−0.0553 (0.126)	−0.0296 (0.114)
ln ( <i>Newspaper</i> )			0.383*** (0.0935)	0.384*** (0.113)	0.424*** (0.0951)
ln ( <i>Rail_kmq</i> )			0.132 (0.107)	0.124 (0.103)	0.141 (0.105)
ln ( <i>Labor_Productivity</i> )				0.668* (0.402)	2.021*** (0.497)
Year fixed effects	YES	YES	YES	NO	YES
Constant	−13.67*** (1.077)	−8.874*** (0.816)	−5.361*** (0.943)	−1.326 (3.124)	9.158** (3.841)
Observations	335	276	272	203	203

Notes: \*, \*\*, and \*\*\* indicate significance levels of 10%, 5%, and 1%, respectively. Exposure variable is population (in 000). Robust standard errors in parentheses.

Table 3b. *Patenting activity (patents with ten years duration) and human capital formation (pooled Poisson regressions)*

Variables	(1) PATENTS_10	(2) PATENTS_10	(3) PATENTS_10	(4) PATENTS_10	(5) PATENTS_10
ln ( <i>Literacy</i> )	2.450*** (0.350)	1.535*** (0.259)	0.443 (0.311)	0.450 (0.629)	-0.375 (0.526)
ln ( <i>Tech_Education</i> )		2.045*** (0.323)	1.638*** (0.394)	0.531 (0.387)	1.236*** (0.371)
<i>S&amp;E_University</i>		0.497*** (0.106)	0.196 (0.122)	0.258* (0.153)	0.240* (0.132)
<i>B&amp;L_techproximity</i>			3.515*** (0.724)	3.924*** (1.098)	3.035*** (0.920)
<i>Urban_pop</i>			-0.208 (0.190)	-0.143 (0.212)	-0.120 (0.204)
ln ( <i>Newspaper</i> )			-0.101 (0.189)	0.0104 (0.322)	-0.0214 (0.203)
ln ( <i>Rail_kmq</i> )			0.105 (0.110)	0.0293 (0.162)	0.142 (0.107)
ln ( <i>Labor_Productivity</i> )				0.358 (0.889)	3.589*** (1.071)
Year fixed effects	YES	YES	YES	NO	YES
Constant	-15.17*** (1.297)	-9.826*** (1.037)	-7.484*** (1.543)	-6.509 (6.666)	18.37** (7.997)
Observations	335	276	272	203	203

Notes: \*, \*\*, and \*\*\* indicate significance levels of 10%, 5%, and 1%, respectively. Exposure variable is population (in 000). Robust standard errors in parentheses.

Table 3c. *Patenting activity (high-tech patents) and human capital formation (pooled Poisson regressions)*

Variables	(1) PATENTS_HT	(2) PATENTS_HT	(3) PATENTS_HT	(4) PATENTS_HT	(5) PATENTS_HT
ln ( <i>Literacy</i> )	2.704*** (0.348)	1.549*** (0.232)	0.587** (0.262)	0.178 (0.379)	-0.237 (0.326)
ln ( <i>Tech_Education</i> )		1.318*** (0.205)	0.453* (0.244)	-0.0414 (0.201)	0.0891 (0.189)
<i>S&amp;E_University</i>		0.556*** (0.0841)	0.159* (0.0811)	0.255*** (0.0754)	0.222*** (0.0711)
<i>B&amp;L_techproximity</i>			3.747*** (0.497)	4.004*** (0.616)	3.160*** (0.563)
<i>Urban_pop</i>			-0.00774 (0.154)	0.0226 (0.180)	0.0974 (0.159)
ln ( <i>Newspaper</i> )			0.222* (0.133)	0.167 (0.152)	0.291** (0.132)
ln ( <i>Rail_kmq</i> )			0.0450 (0.152)	0.0448 (0.140)	0.0531 (0.144)
ln ( <i>Labor_Productivity</i> )				1.509*** (0.576)	3.363*** (0.590)
Year fixed effects	YES	YES	YES	NO	YES
Constant	-14.93*** (1.286)	-9.343*** (1.027)	-7.160*** (1.365)	2.903 (4.550)	17.24*** (4.636)
Observations	335	276	272	203	203

Notes: \*, \*\*, and \*\*\* indicate significance levels of 10%, 5%, and 1%, respectively. Exposure variable is population (in 000). Robust standard errors in parentheses.

behavior, but it is interesting to note that the size of this coefficient is higher in the regressions with high-quality (table 3b) or high-tech patents (table 3c) as dependent variables. The effect of the proxy for science and engineering university students (*S&T\_University*) is also positive and significant in all specifications except in column 3 of table 3b.

Finally, the effect of the Bar-Leiponen indicator of proximity to foreign patenting (*B&L\_techproximity*) is always positive and significant, which suggests that provinces whose patenting profile is characterized by a strong degree of sectoral overlap with foreign patenting are characterized by a more dynamic patenting activity. The effect of this variable is particularly evident in the case of high-quality (table 3b) and high-tech (table 3c) patents.

Table 4 reports Poisson regressions for patents with a “real” duration  $\geq 5$  years. In this case, as mentioned, we were forced by the availability of the “real” duration variable to restrict the sample to the benchmark years 1881, 1891, and 1902.<sup>23</sup> The results are substantially similar to those reported in tables 3a–3c. In particular, also in this case, the size of the coefficient of technical education is higher for high quality than for all patents (column 6 of table 4).

The evidence in tables 3a–3c and 4 suffers from possible biases due to endogeneity and to unobserved heterogeneity with potentially long historical roots. The Poisson regressions of tables 5a–5c, which is divided in three parts—one for each types of patents—try to address these econometric issues, and for this reason, they may be regarded as providing the most robust characterization of the relationship between human capital formation and inventive activities. In tables 5a–5c, all the co-variates have been lagged to the previous benchmark and this restricts our sample to the benchmarks 1891, 1902, and 1911. In this way, we deal with the possible endogeneity of the co-variates. Furthermore, in tables 5a–5c, we introduce progressively as controls a number of dummies in order to account for geographical “fixed effects”. After presenting the baseline, that is the pooled regression with lagged co-variates (column 1), we control, in the following columns: for Southern provinces (2), for macro-areas (North-West, North East, Center, South and Islands) in column (3), for the seven preunitary states in column (4), for regional effects (16 regions) in column (5). Finally, in column (6) we report a panel model with fixed-effect (FE) at provincial level, but without year fixed effects, and in column (7) a panel model controlling both for cross-sectional heterogeneity and time variation.

In line with the findings of tables 3a–3c and 4, in tables 5a–5c, we are also able to establish a systematic correlation between human capital formation and patenting. In general, we find positive and significant coefficients for technical education, science, and engineering university students and Bar & Leiponen technical proximity even when we introduce finer geographical fixed effects. The magnitude of these coefficients is also relatively stable across different specifications. The behavior of the literacy coefficients is somewhat more elusive, since, by virtue of the inclusion of geographical controls, the coefficient is not always positive and significant.

The results in columns (6) and (7) are clearly more puzzling, but in this case one must take into account that we are estimating the effects of human capital formation using a relatively narrow degree of within province variation with a panel of limited size and, in tables 5b and 5c, with a more restricted sample.

Overall, the results of the regression exercises reported in tables 3a–3c and 5a–5c suggest the existence of a significant correlation between human capital formation and inventive

<sup>23</sup> In this case, the models comprising labor productivity in manufacturing are restricted to two benchmark years (1881 and 1901).



Table 4. *Patenting activity and human capital formation (pooled Poisson regressions), 1881–1902*

Variables	(1) PATENTS_ REAL_5	(2) PATENTS_ REAL_5	(3) PATENTS_ REAL_5	(4) PATENTS_ REAL_5	(5) PATENTS_ REAL_5	(6) PATENTS	(7) PATENTS_ 10
ln ( <i>Literacy</i> )	2.629*** (0.413)	1.655*** (0.276)	0.525 (0.345)	0.0347 (0.524)	-0.270 (0.570)	0.370* (0.221)	-0.517 (0.408)
ln ( <i>Tech_Education</i> )		1.368*** (0.189)	1.117*** (0.199)	0.840*** (0.310)	0.940*** (0.288)	0.308** (0.128)	1.009*** (0.304)
<i>S&amp;E_University</i>		0.436*** (0.0957)	0.167 (0.132)	0.349** (0.141)	0.355** (0.143)	0.0939 (0.0648)	0.257* (0.143)
<i>B&amp;L_techproximity</i>			3.175*** (0.654)	4.235*** (1.059)	4.004*** (1.077)	2.515*** (0.362)	3.214*** (0.944)
<i>Urban_pop</i>			-0.394* (0.237)	-0.595** (0.259)	-0.653** (0.261)	0.0658 (0.0976)	-0.507** (0.243)
ln ( <i>Newspaper</i> )			0.0334 (0.170)	-0.0697 (0.222)	-0.0404 (0.199)	0.346*** (0.0914)	0.133 (0.200)
ln ( <i>Rail_kmq</i> )			-0.0536 (0.150)	-0.0512 (0.140)	0.0333 (0.138)	0.263*** (0.0636)	0.245** (0.123)
ln ( <i>Labor_Productivity</i> )				-0.0793 (0.879)	1.193 (1.157)	0.821* (0.439)	2.225** (1.029)
Year fixed effects	YES	YES	YES	NO	YES	YES	YES
Constant	-15.95*** (1.530)	-10.33*** (1.165)	-7.073*** (1.773)	-7.008 (6.939)	3.151 (9.167)	1.845 (3.435)	12.25 (7.765)
Observations	207	207	203	134	134	134	134

Notes: \*, \*\*, and \*\*\* indicate significance levels of 10%, 5%, and 1%, respectively. Exposure variable is population (in 000). Robust standard errors in parentheses.

Table 5a. *Patenting activity and human capital formation (panel Poisson regressions with lagged co-variables)*

Variables	(1) PATENTS None	(2) PATENTS South	(3) PATENTS Macroarea (5)	(4) PATENTS Preunitary State (7)	(5) PATENTS Regions (16)	(6) PATENTS FE	(7) PATENTS FE
<i>ln (Literacy)</i>	0.831*** (0.178)	0.312 (0.198)	0.157 (0.330)	0.400 (0.301)	0.567 (0.396)	3.318*** (0.450)	-1.521** (0.687)
<i>ln (Tech_Education)</i>	0.816*** (0.137)	0.733*** (0.141)	0.710*** (0.138)	0.799*** (0.138)	0.630*** (0.150)	0.362* (0.214)	-0.310** (0.122)
<i>S&amp;E_University</i>	0.207*** (0.0764)	0.260*** (0.0694)	0.275*** (0.0714)	0.282*** (0.0721)	0.331*** (0.0761)	0.104 (0.0965)	-0.0221 (0.0839)
<i>B&amp;L_techproximity</i>	2.327*** (0.292)	2.501*** (0.272)	2.446*** (0.313)	2.481*** (0.259)	2.358*** (0.279)	-0.766* (0.402)	-1.221*** (0.400)
Year fixed effects	YES	YES	YES	YES	YES	NO	YES
Constant	-6.314*** (0.741)	-6.314*** (0.741)	-4.451*** (0.814)	-4.737*** (1.071)	-4.819*** (1.179)		
Observations	207	207	207	207	207	198	198
Number of provincia	69	69	69	69	69	66	66

Notes: \*, \*\*, and \*\*\* indicate significance levels of 10%, 5%, and 1%, respectively. Exposure variable is population (in 000). Robust standard errors in parentheses.

Table 5b. *High-quality patenting activity and human capital formation (panel Poisson regressions with lagged co-variables)*

Variables	(1) PATENTS_10 None	(2) PATENTS_10 South	(3) PATENTS_10 Macroarea (5)	(4) PATENTS_10 Preunitary State (7)	(5) PATENTS_10 Regions (16)	(6) PATENTS_10 FE	(7) PATENTS_10 FE
ln ( <i>Literacy</i> )	0.761** (0.329)	0.609 (0.383)	-0.660 (0.723)	-0.399 (0.650)	1.276 (0.919)	4.152*** (1.485)	-2.229 (1.793)
ln ( <i>Tech_Education</i> )	1.473*** (0.383)	1.443*** (0.387)	1.320*** (0.370)	1.419*** (0.356)	0.954** (0.381)	-0.345 (0.492)	-0.793* (0.479)
<i>S&amp;E_University</i>	0.123 (0.162)	0.142 (0.159)	0.187 (0.143)	0.222 (0.146)	0.351*** (0.135)	0.428** (0.191)	-0.0988 (0.213)
<i>B&amp;L_techproximity</i>	2.395*** (0.564)	2.434*** (0.560)	2.745*** (0.656)	2.633*** (0.639)	2.371*** (0.574)	-1.463 (1.243)	-2.898** (1.150)
Year fixed effects	YES	YES	YES	YES	YES	NO	YES
Constant	-7.324*** (1.421)	-6.778*** (1.630)	-3.353 (2.338)	-2.881 (2.659)	-10.42*** (3.791)		
Observations	207	207	207	207	207	111	111
Number of provincia	69	69	69	69	69	37	37

Notes: \*, \*\*, and \*\*\* indicate significance levels of 10%, 5%, and 1%, respectively. Exposure variable is population (in 000). Robust standard errors in parentheses.

Table 5c. *High-tech patenting activity and human capital formation (panel Poisson regressions with lagged co-variates)*

Variables	(1) PATENTS_HT None	(2) PATENTS_HT South	(3) PATENTS_HT Macroarea (5)	(4) PATENTS_HT Preunitary States (7)	(5) PATENTS_HT Regions (16)	(6) PATENTS_HT FE	(7) PATENTS_HT FE
<i>ln (Literacy)</i>	0.735*** (0.184)	0.301 (0.227)	0.0341 (0.413)	0.238 (0.356)	0.770 (0.536)	2.769*** (0.693)	-2.176* (1.236)
<i>ln (Tech_Education)</i>	1.028*** (0.173)	0.962*** (0.178)	0.902*** (0.175)	1.047*** (0.160)	0.811*** (0.174)	0.780*** (0.215)	0.0206 (0.161)
<i>S&amp;E_University</i>	0.187** (0.0792)	0.237*** (0.0777)	0.280*** (0.0818)	0.297*** (0.0805)	0.392*** (0.0808)	0.0118 (0.109)	-0.0320 (0.123)
<i>B&amp;L_techproximity</i>	2.627*** (0.319)	2.753*** (0.315)	2.524*** (0.378)	2.629*** (0.309)	2.385*** (0.324)	-0.582 (0.554)	-0.861 (0.577)
Year fixed effects	YES	YES	YES	YES	YES	NO	YES
Constant	-6.787*** (0.794)	-5.219*** (0.917)	-5.199*** (1.295)	-6.030*** (1.657)	-7.486*** (2.190)		
Observations	207	207	207	207	207	174	174
Number of provincia	69	69	69	69	69	58	58

Notes: \*, \*\*, and \*\*\* indicate significance levels of 10%, 5%, and 1%, respectively. Exposure variable is population (in 000). Robust standard errors in parentheses.

activities measured using different types of patents. Furthermore, our findings point to two distinct channels of influence. The dominant channel is clearly represented by technical education. This variable affects all types of patenting both in the pooled (tables 3a–3c and table 4) and in the panel specifications (tables 5a–5c). The estimated elasticity in tables 5a–5c is ranging between 0.6 and 1.5.

The second channel of influence is represented by the literacy variable. Admittedly, the effect is more ambiguous: the literacy coefficients are mostly positive, although there are not always significant and, in few cases, they are even negative. All models, in general, suggest a connection between science and engineering university students and all types of patent activity.

To sum up, if we focus on inventive activities characterized by a higher degree of sophistication such as high-quality patents or patents in high-tech sectors, we find that secondary technical education and science and engineering university education are mostly relevant. This differential impact of various types of human capital on the quality of innovation is consistent with the notion that the upper tail of the human capital distribution was the key driver of technological breakthroughs in this historical phase (Mokyr, 2005a). In this perspective, our findings concur with those obtained by Khan (2015) and Squicciarini and Voigtlander (2015). Khan (2015) finds that, in the England in the second half of the XIX century, inventors that were creators of major technological breakthroughs were characterized by relatively high levels of science and technical education. Squicciarini and Voigtlander (2015), using the *Encyclopédie* subscriptions as indicator of the upper tail of skills, stress the critical role of this form of human capital for the rate of technological innovation in France during the first half of the XIX century.

Interestingly enough, these results pointing to the critical role of technical education for high-quality inventive activities are also in line with previous research on the Italian education system, emphasizing the importance of this type of human capital formation in fostering industrialization at local level (Zamagni, 1978, 1993, Vasta, 1999). Similar considerations hold on the connection between engineering university education and economic development (Vasta, 1996).<sup>24</sup>

Finally, it is worth noticing that in all models the coefficients of the Bar and Leiponen variable are mostly positive and significant. In our interpretation this finding suggests an important role played by the technological congruence *à la* Abramovitz (1986) and the pattern of technological development at provincial level.

## 5. The determinants of industrialization

After having provided an assessment of the relationship between different forms of human capital accumulation and patenting activities, in this Section, we examine the role played by innovative activities in shaping the broader patterns of industrialization at provincial level. So far, some recent studies have assessed the possible factors affecting the localization of industrial activities during this historical period both for manufacturing as whole (Fenoaltea, 2011, A’Hearn and Venables, 2013, Ciccarelli and Fenoaltea, 2013; Ciccarelli and Proietti, 2013, Ciccarelli and Fachin, 2016), but also for specific industries

<sup>24</sup> For a discussion of the connection between engineering and technical education and industrialization in a broader European perspective, see Fox and Guagnini (1993).

(A’Hearn, 1998).<sup>25</sup> However, none of these papers has explicitly considered technical progress. We tackle this gap by estimating a number of models in which we add technical change to the more “conventional” explanatory factors of the localization of industrial activities.<sup>26</sup>

The indicator of industrialization that we use is the growth rates of labor productivity, measured as value added per male worker, both in total manufacturing and in the engineering sector.<sup>27</sup> In particular, we estimate the following growth regression:

$$\hat{y} = \alpha + \sum_{i=1}^K \beta_i x_i + \gamma PATENTS + \delta y_0 + \varepsilon \quad (3)$$

where  $\hat{y}$  is the growth rate of labor productivity over the period 1881–1911,  $x_i$  are the control variables in circa 1880,  $PATENTS$  are different types of patents in 1881 normalized by population and  $y_0$  is the level of labor productivity in 1881. We consider four main determinants of industrialization as controls. The first is the human capital endowment of the provinces which is measured here by using both literacy rates as proxy for the “basic” education and technical education for “advanced” skills.

The second is the availability of water resources. This is in line with some recent contributions, highlighting the critical role of water resources in determining the localization of industrial hubs (A’Hearn and Venables, 2013 for Italy and Crafts and Wolf, 2014 for the UK).

The third determinant is the level of real wages. In this case, the literature has discussed two possible opposite effects of real wages on the industrialization process: on the one hand, it has been argued that low wages resulting in higher profits rates can stimulate a higher rate of investment in industrial plants (Mokyr, 1976); on the other hand, a more recent stream of literature maintains that high wages incentivize investments in capital goods and machinery (Allen, 2009).

The fourth determinant is domestic market potential which was originally introduced by Harris (1954) as a driver of industry localization. This theme has been recently recalled in Italian economic history by A’Hearn and Venables (2013) and Missiaia (2016).

Figure 5 contains maps illustrating the spatial distribution of all the variables we use in our analysis. The maps of the two panels of the first row represent the growth rates of labor productivity in manufacturing and in the engineering sector. They are both characterized by a North–South gradient. The panels in the second row show the distributions of our control variables. Literacy shows an evident North–South divide and provinces with relative high level of literacy in 1881 are indeed the most industrialized in 1911. Interestingly enough, technical education is characterized by a less clear-cut pattern of spatial distribution that, although reveals the existence of a North–South divide, shows also the existence of a number of hubs in the North, in the Center and, in some cases, also in the South, especially in Sicilia. Water resources are clearly clustered around the Po Valley, while real wages, instead, do not display a clear-cut geographical pattern. Indeed, it is possible to identify provinces with relative high real wages in the North, the Center and the South of the country. Finally, in the case of market potential we find again a sort of North–South

<sup>25</sup> For a recent example of this approach for England and France, see Kelly, Mokyr and O’Grada (2015).

<sup>26</sup> For other studies, which included patenting activities as possible factor for shaping industrialization, see Crafts and Wolf (2014) for cotton industry in the UK and Cinnirella and Streb (2017) for Prussia.

<sup>27</sup> Several contributions have pointed to the critical role played, in this historical period, by the engineering sector as a key driver of technical change (see, amongst others, Rosenberg 1976). For a case study of the Italian locomotive industry, see Ciccarelli and Nuvolari (2015).

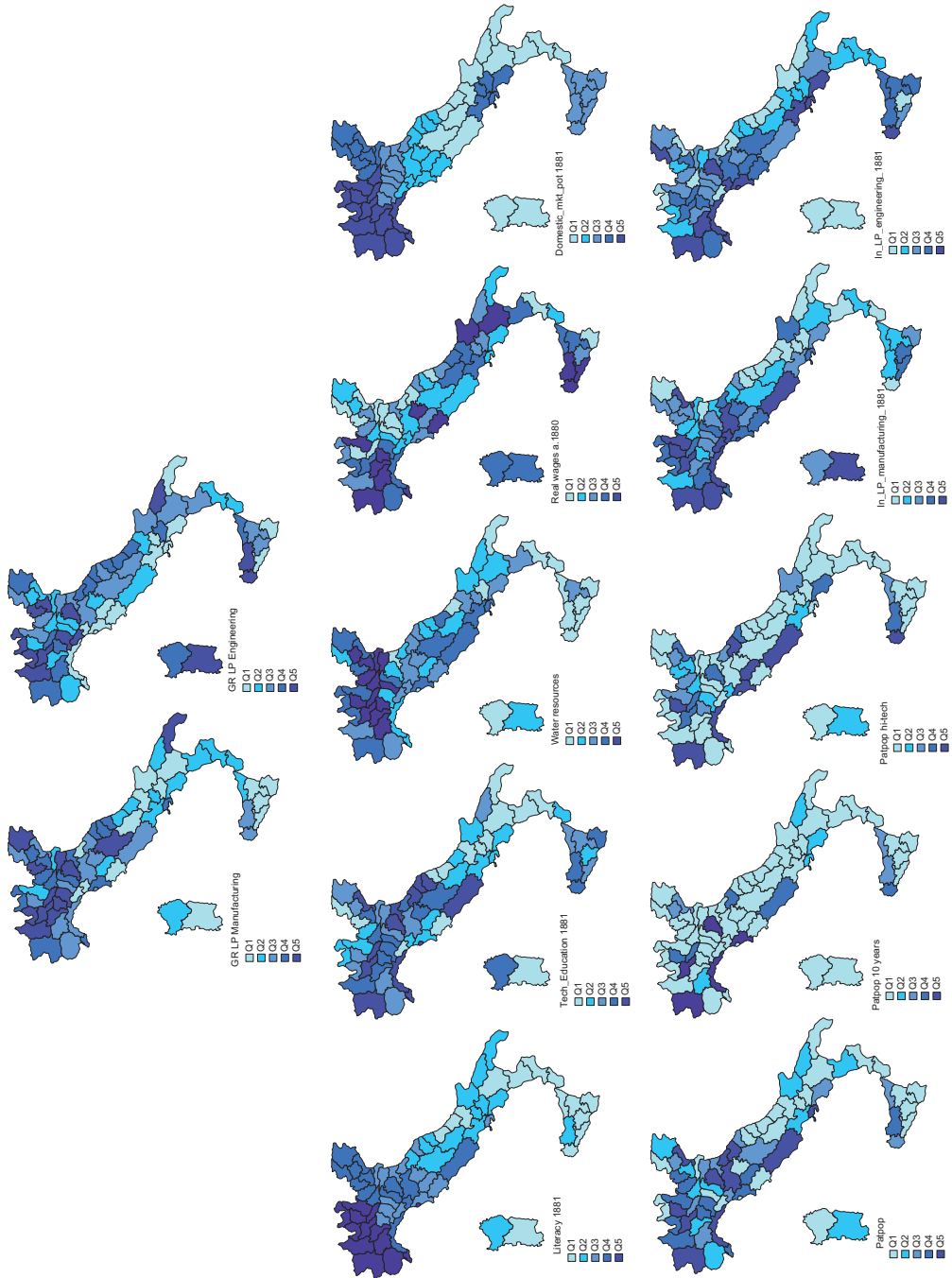


Figure 5. Maps of provinces industrialization and their determinants.

Table 6. *Determinants of industrialization 1911*

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	GR LP Manufacturing	GR LP Manufacturing	GR LP Manufacturing	GR LP Engineering	GR LP Engineering	GR LP Engineering
<i>Literacy 1881</i>	0.0252*** (0.00324)	0.0260*** (0.00317)	0.0259*** (0.00315)	0.000465 (0.00197)	0.00112 (0.00185)	0.00111 (0.00190)
<i>Tech_Education 1881</i>	1.816*** (0.592)	1.754*** (0.596)	1.773*** (0.590)	0.0271 (0.406)	-0.0303 (0.398)	-0.0335 (0.409)
<i>Water resources</i>	-8.25e-05 (0.000150)	-6.55e-05 (0.000150)	-7.50e-05 (0.000149)	-7.50e-05 (9.55e-05)	-5.39e-05 (9.38e-05)	-6.24e-05 (9.59e-05)
<i>Real wages a. 1880</i>	-0.414 (0.277)	-0.418 (0.276)	-0.369 (0.276)	0.294 (0.180)	0.293 (0.176)	0.327* (0.180)
<i>Domestic_mkt_pot 1881</i>	-0.000147 (0.000360)	-0.000178 (0.000361)	-0.000150 (0.000358)	0.000191 (0.000237)	0.000160 (0.000232)	0.000178 (0.000237)
<i>Patpop</i>	0.00463* (0.00271)			0.00482*** (0.00173)		
<i>Patpop 10 years</i>		0.0266* (0.0152)			0.0318*** (0.00956)	
<i>Patpop hi-tech</i>			0.0177* (0.00916)			0.0163*** (0.00582)
<i>ln_LP_manufacturing_1881</i>	-2.724*** (0.310)	-2.728*** (0.310)	-2.768*** (0.312)			
<i>ln_LP_engineering_1881</i>				-3.058*** (0.288)	-3.144*** (0.284)	-2.990*** (0.288)
<i>Constant</i>	-17.71*** (2.246)	-17.73*** (2.243)	-18.06*** (2.261)	-19.62*** (1.990)	-20.19*** (1.957)	-19.19*** (1.987)
<i>Observations</i>	69	69	69	69	69	69
<i>R-squared</i>	0.709	0.710	0.713	0.678	0.693	0.679

Notes: OLS regressions (dependent variable is the logarithm of 1911 manufacturing value added per capita for columns 1–3 and the logarithm of 1911 engineering value added per capita for columns 4–6), \*, \*\*, and \*\*\* indicate significance levels of 10%, 5%, and 1%, respectively. Standard errors in parentheses.



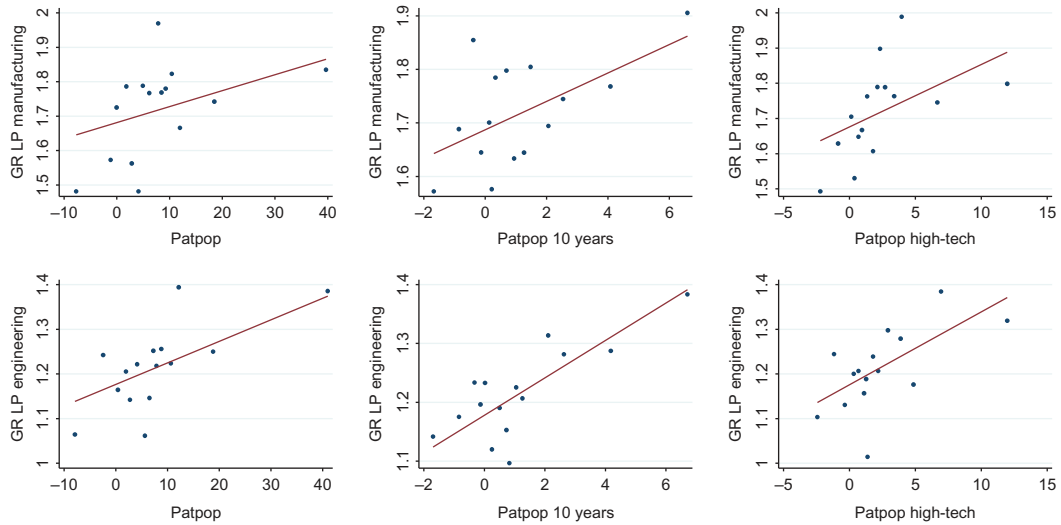


Figure 6. Binscatter plots of growth rates (1881–1911) of labor productivity (manufacturing and engineering) and patenting (1881).

Note: Binscatter plots constructed using fifteen equal sized bins and corresponding to the regressions reported in table 6 (controls include literacy, technical education, real wages, water resources, domestic market potential, labor productivity (in manufacturing and engineering in 1881)).

gradient, although it is worth noting the existence of few provinces in the South (in Campania and Sicilia) characterized by relatively high levels of this variable.

The panels in the third row contain three maps showing the spatial distribution of patenting activities in 1881 normalized by population and two maps of the initial levels of labor productivity. As for the maps on patents, it is possible to detect a certain degree of clustering around a North-West axis. Interestingly enough, in this case, the distribution of patents adjusted for quality is spatially more concentrated than that of standard patents.

Table 6 presents the results of the growth regressions. We find that literacy and technical education have a significant impact on the growth rates of labor productivity in manufacturing but not in the engineering sector.<sup>28</sup> Remarkably, the impact of technical education is considerably stronger than that of literacy. Moreover, it is worth noting that all regressions show a tendency toward convergence as indicated by the coefficients of initial level of labor productivity. However, the main finding of table 6 is the connection between various forms of patenting activities and the provincial levels of industrialization: all the three patent variables have a significant and positive impact both on manufacturing and on the engineering sector.

Figure 6 provides a further impressionistic outlook on our results by means of binscatter plots showing the partial correlation between the growth of labor productivity and patenting activities including, as controls, all the co-variates of the regressions in table 6.

Overall, the evidence presented in this section points to a potential connection between patenting and industrialization which deserves further investigation.

<sup>28</sup> This result is broadly consistent with Ciccarelli and Fachin (2016) who find that literacy and numeracy have a positive and significant effect on the growth of labor productivity in manufacturing for the period 1871–1911.

## 6. Conclusions

In this paper, we have made a first attempt to look at the geography of innovation in Italy during the Liberal age. This is a critical phase, since it represents the moment in which the industrialization process was launched on a national scale although in a preliminary fashion (Gerschenkron, 1955; Fenoaltea, 2011). Research on the origins of the Italian regional divides in economic performance has also focussed on the same period (Felice, 2013; Felice and Vasta, 2015). Using patent data, we were able to provide a comprehensive assessment of the innovative performance of Italian provinces. Our interpretation of the geography of innovation can be articulated in terms of two interconnected “tales”. The first “tale” is essentially a story of the factors affecting patenting activities; the second “tale” is instead a story of the determinants of industrialization in which innovative activities play a significant role.

As for the first “tale”, we can summarize our findings in four main points:

- (i) patenting activities are concentrated in the provinces of the so-called “industrial triangle” since a relatively early stage;
- (ii) the geography of patenting activities is also characterized by a rather clear-cut geographical divide pointing to a threefold partition of the country in North, Center, and South;
- (iii) there is a significant relationship between the localization of patenting activity and different types of human capital formation. In particular, we establish an important role for secondary technical education, and, in a somewhat weaker form, for science and engineering university education, in fostering innovation and, in particular, high-quality and high-tech patents. There is also a positive effect of literacy on patenting activity but it is more elusive.

In the second “tale”, using growth regressions of labor productivity, we have examined the nexus between the location of industrialization and innovation capacity, while controlling for other basic factors highlighted by the recent literature on Italian economic growth, namely: human capital endowment, the availability of water resources (which was critical factor for a country lacking major coal deposits), the comparative level of real wages and the domestic market potential.

Therefore, the general lesson of our paper is that, even for a latecomer country such as was Italy at the time, the understanding of the process of industrialization requires to pay proper attention to the characteristics of inventive activities and their determinants.

## Acknowledgments

We would like to thank Sara Pecchioli for outstanding research assistance and Michele Mannucci for granting us access to the historical library of Ufficio Tecnico Ing. A. Mannucci s.r.l. Firenze. We are particularly grateful to Gabriele Cappelli, Carlo Ciccarelli and Giovanni Federico for sharing their data with us. We thank Stefano Breschi, Stefano Falcone, Emanuele Felice, Andrea Giuntini, Stefano Maggi, Tom Nicholas, Chiara Tomasi and Patrick Wallis for helpful comments. The suggestions of the editors of this journal and of three anonymous referees are also gratefully acknowledged. The paper has benefited from the comments of participants at the Economic History Society Annual Conference (Warwick, 2014), the Economic History Workshop (Siena, 2014), the XVII<sup>th</sup> World Economic History Congress (Kyoto, 2015), the 11<sup>th</sup> European Historical Economics Society Conference (Pisa, 2015), at a seminar at the Department of Economics and Management of the University of Trento (2015) and at The Fourth Symposium on Quantitative History (Beijing, 2016).

*Conflict of interest statement.* None declared.

## References

- ABRAMOVITZ, M. (1986). Catching up, forging ahead, and falling behind. *Journal of Economic History* **46**, pp. 385–406.
- A'HEARN, B. (1998). Institutions, externalities, and economic growth in southern Italy: evidence from the cotton textile industry, 1861–1914. *Economic History Review* **51**, pp. 734–62.
- A'HEARN, B. and VENABLES, A. (2013). Regional disparities: internal geography and external trade. In G. TONIOLO (ed), *The Oxford Handbook of the Italian Economy since Unification*. Oxford: Oxford University Press, pp. 599–630.
- ALLEN, R. (2009). *The British Industrial Revolution in Global Perspective*. Cambridge: Cambridge University Press.
- BAR, T. and LEIPONEN, A. (2012). A measure of technological distance. *Economics Letters* **116**, pp. 457–9.
- CAPPELLI, G. (2015). One size that didn't fit all? Electoral franchise, fiscal capacity and the rise of mass schooling across Italy's provinces, 1870–1911. *Econometrica* **10**, pp. 311–43.
- CAPPELLI, G. (2017). The missing link? Trust, cooperative norms, and industrial growth in Italy. *Journal of Interdisciplinary History* **47**, pp. 333–58.
- CICCARELLI, C. and FACHIN (2016). Regional growth with spatial dependents: A case study on early Italian industrialization. *Paper in Regional Science*. DOI:10.1111/pirs.12217.
- CICCARELLI, C. and FENOALTEA, S. (2013). Through the magnifying glass: provincial aspects of industrial growth in post-unification Italy. *Economic History Review* **66**, pp. 57–85.
- CICCARELLI, C. and GROOTE, P. (2017). Railway endowment in Italy's provinces, 1839–1913. *Rivista di Storia Economica*, forthcoming.
- CICCARELLI, C. and MISSIAIA, A. (2013). The industrial labour force of Italy's provinces: estimates from population censuses, 1871–1911. *Rivista di Storia Economica* **29**, pp. 141–92.
- CICCARELLI, C. and NUVOLARI, A. (2015). Technical change, Non-Tariff Barriers and the Development of the Italian Locomotive Industry, 1850–1913. *Journal of Economic History* **75**, pp. 860–88.
- CICCARELLI, C. and PROIETTI, T. (2013). Patterns of industrial specialisation in post-Unification Italy. *Scandinavian Economic History Review* **61**, pp. 259–86.
- CINNIRELLA, F. and STREB J. (2017), The role of human capital and innovation in economic development: evidence from post-Malthusian Prussia. *Journal of Economic Growth*, forthcoming. doi:10.1007/s10887-017-9141-3.
- CIVES, G. (ed) (1990). *La scuola italiana dall'Unità ai nostri giorni*. Firenze: La nuova Italia Editrice.
- CRAFTS, N. and WOLF, N. (2014). The location of the UK cotton textiles industry in 1838: a quantitative analysis. *Journal of Economic History* **74**, pp. 1103–39.
- DANIELE, V. and MALANIMA, P. (2011). Il divario nord-sud in Italia, 1861–2011. Soveria Mannelli: Rubbettino.
- DANIELE, V. and MALANIMA, P. (2014a). Perché il Sud è rimasto indietro? Il Mezzogiorno fra storia e pubblicistica. *Rivista di Storia Economica* **30**, pp. 3–36.
- DANIELE, V. and MALANIMA, P. (2014b). Due commenti finali. *Rivista di Storia Economica* **30**, pp. 243–8.
- FELDMAN, M. and KOGLER, D. (2010). Stylized facts in the geography of innovation. In HALL B. and ROSENBERG N. (eds), *Handbook of Economics of Innovation vol. I*. Dordrecht: Elsevier, pp. 381–410.
- FEDERICO, G. and TONIOLO, G. (1991). Italy. In SYLLA, R. and TONIOLO, G. (eds), *Patterns of European industrialization. The Nineteenth Century*. London: Routledge, pp. 197–217.
- FELICE, E. (2012). Regional convergence in Italy, 1891–2011: testing human and social capital. *Econometrica* **6**, pp. 267–306.
- FELICE, E. (2013). *Perché il Sud è rimasto indietro*. Bologna: il Mulino.
- FELICE, E. (2014). Il Mezzogiorno fra storia e pubblicistica. Una replica a Daniele e Malanima. *Rivista di Storia Economica* **30**, pp. 187–242.
- FELICE, E. and VASTA, M. (2015). Passive modernization? The new Human Development Index and its components in Italy's regions (1871–2007). *European Review of Economic History* **19**, pp. 44–66.

- FENOALTEA, S. (2011). *The Reinterpretation of Italian Economic History: from Unification to the Great War*. New York: Cambridge University Press.
- FOX, R. and GUAGNINI, A. (eds) (1993). *Education, Technology and Industrial Performance in Europe, 1850-1939*. Cambridge-Paris: Cambridge University Press.
- GERSCHENKRON, A. (1955). Notes on the rate of industrial growth in Italy, 1881-1913. *Journal of Economic History* **15**, pp. 360-75.
- GIANNETTI, R. (1998). *Tecnologia e sviluppo economico italiano*. Bologna: Il Mulino.
- GOURIEROUX, C., MONFORT, A. and TROGNON, A. (1984). Pseudo maximum likelihood methods: Theory. *Econometrica : journal of the Econometric Society* **52**, pp. 681-700.
- HARRIS, C.D. (1954). The market as a factor in the localization of industry in the United States. *Annals of the Association of American Geographers* **44**, pp. 315-48.
- HELLIWELL, J. and PUTNAM, R. (1995). Economic growth and social capital in Italy. *Eastern Economic Journal* **21**, pp. 295-307.
- HILBE, J. (2011). *Negative Binomial Regression*. Cambridge: Cambridge University Press.
- KHAN, Z. (2015). 'Knowledge, human capital and economic development: Evidence from the British Industrial Revolution, 1750-1930', NBER Working paper, n. 20853.
- KELLY, M., MOKYR J. and O'GRADA C. (2015), 'Roots of the Industrial Revolution', UCD Centre for Economic Research Working Paper series, no. WP15/24.
- LACAITA, C. and POGGIO, P.P. (2011). *Scienza, Tecnica e Industria nei 150 anni di Unità d'Italia*. Milano: Jaca Book.
- MISSIAIA, A. (2016). Where do we go from here? Market access and regional development in Italy (1871-1911). *European Review of Economic History* **20**, pp. 215-41.
- MOKYR, J. (1976). *Industrialization in the Low Countries, 1795-1850*. New Haven: Yale University Press.
- MOKYR, J. (1990). *The Lever of Riches*. Oxford: Oxford University Press.
- MOKYR, J. (1995). Urbanization, technological progress, and economic history. In GIERSCHE H. (ed), *Urban Agglomeration and Economic Growth*. Berlin: Springer-Verlag, pp. 3-37.
- MOKYR, J. (2005a). Long-Term economic growth and the history of technology. In AGHION, P. and DURLAUF, S. (eds), *Handbook of Economic Growth*. Dordrecht: Elsevier, pp. 1113-80.
- MOKYR, J. (2005b). The intellectual origins of modern economic growth. *Journal of Economic History* **65**, pp. 285-351.
- NADIRI, M. and KIM, S. (1996), 'International R&D spillovers, trade and productivity in Major OECD Countries', NBER Working Paper n. 5801.
- NAGAOKA, S., MOTOHASHI, K. and GOTO, A. (2010). Patent statistics as an innovation indicator. In HALL B.H. and ROSENBERG N. (eds), *Handbook of Economics of Innovation* vol. 2. Dordrecht: Elsevier, pp. 1083-1127.
- NUVOLARI, A. and VASTA, M. (2015a). The ghost in the attic? The Italian national innovation system in historical perspective, 1861-2011. *Enterprise & Society* **16**, pp. 270-90.
- NUVOLARI, A. and VASTA, M. (2015b). Independent invention in Italy during the Liberal Age, 1861-1913. *Economic History Review* **68**, pp. 858-86.
- NUVOLARI, A. and VASTA, M. (2017), The Italian patent system(s) during the long nineteenth century (from the Congress of Vienna to World War I). In G. GOODAY and S. WILF (eds), *Fashioning Global Patent Cultures: Diversity and Harmonization in Historical Perspective*. Cambridge: Cambridge University Press, forthcoming.
- PUTNAM, R.D., LEONARDI, R. and NANETTI, R.Y. (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press.
- ROSENBERG, N. (1976). *Perspectives on Technology*. Cambridge: Cambridge University Press.
- ROSES, J.R. (1998). Measuring the contribution of human capital to the development of the Catalan factory system (1830-61). *European Review of Economic History* **2**, pp. 25-48.
- SCHMOOKLER, J. (1966). *Invention and Economic Growth*. Cambridge (MA): Harvard University Press.
- SQUICCIARINI, M. and VOIGTLANDER, N. (2015). Human capital and industrialization: evidence from the Age of Enlightenment. *Quarterly Journal of Economics* **130**, pp. 1825-83.

- VASTA, M. (1996), Innovative capabilities in the Italian take-off: the electrical and chemical industries (1880–1914), Ph.D Thesis, University of Oxford.
- VASTA, M. (1999). Capitale umano e ricerca scientifica e tecnologica. In F. AMATORI, D. BIGAZZI, R. GIANNETTI and L. SEGRETO (eds), *Storia d'Italia. Annali 15, L'industria*. Torino: Einaudi, pp. 1041–1124.
- VECCHI, G. (2011). In *Ricchezza e in Povertà. Il benessere degli Italiani dall'Unità ad oggi*. Bologna: Il Mulino.
- ZAMAGNI, V. (1978). Istruzione e sviluppo economico. Il caso italiano. 1861–1913. In G. TONIOLO (ed), *L'economia italiana 1861–1940*. Roma-Bari: Laterza, pp. 137–78.
- ZAMAGNI, V. (1993). *The Economic History of Italy 1860–1990*. Oxford: Oxford University Press.

### Official publications

- Ferrovie dello Stato, Ufficio Statistica (1911), *Ferrovie Italiane 1861–1909*. Riproduzione dei lavori grafici presentati all'Esposizione Internazionale di Torino del 1911, Roma.
- Ministero della Pubblica Istruzione [MPI] (1883), *Bollettino ufficiale*, Roma.
- Ministero della Pubblica Istruzione [MPI] (1901), *Bollettino ufficiale del Ministero della Istruzione Pubblica*, Roma.
- Ministero della Pubblica Istruzione [MPI] (1912), *Bollettino ufficiale del Ministero della Istruzione Pubblica*, Roma.
- Ministero di Agricoltura, Industria e Commercio – Direzione Generale di Statistica (n.d.), *Salari. Prezzi medi di un'ora di lavoro degli operai addetti alle opere di muratura ed ai trasporti di terra e mercedi medie giornaliere degli operai addetti alle miniere (1862–1878)*, Roma.
- Ministero di Agricoltura, Industria e Commercio [MAIC] (1864–1885), *Bollettino delle privative industriali del Regno d'Italia*, Torino.
- Ministero di Agricoltura, Industria e Commercio [MAIC] (1865). *Statistica del Regno d'Italia. Popolazione. Censimento generale vol. II*. Torino: Popolazione.
- Ministero di Agricoltura, Industria e Commercio [MAIC] (1881), *Annuario Statistico Italiano. Anno 1881*, Roma.
- Ministero di Agricoltura, Industria e Commercio [MAIC] (1883). *Censimento della popolazione del Regno d'Italia al 31 dicembre 1881 vol. II*. Roma: Popolazione classificata per età sesso, stato civile e istruzione elementare.
- Ministero di Agricoltura, Industria e Commercio [MAIC] (1886–1893), *Bollettino delle privative industriali del Regno d'Italia*, Roma.
- Ministero di Agricoltura, Industria e Commercio [MAIC] (1893), *Annuario Statistico Italiano. 1892*, Roma.
- Ministero di Agricoltura, Industria e Commercio [MAIC] (1894–1901), *Elenco degli attestati di privativa industriale, di prolungamento, completivi, di importazione e di riduzione*, Roma.
- Ministero di Agricoltura, Industria e Commercio [MAIC] (1902–1923), *Bollettino della proprietà intellettuale*, Roma.
- Ministero di Agricoltura, Industria e Commercio [MAIC] (1903), *Censimento della popolazione del Regno d'Italia al 10 febbraio 1901, vol. II. Numero delle famiglie e numero degli abitanti classificati secondo la qualità della dimora, il luogo di nascita, il sesso, l'età, lo stato civile e l'istruzione – ciechi e sordo-muti – stranieri – lingue parlate*, Roma.
- Ministero di Agricoltura, Industria e Commercio [MAIC] (1908), *Annuario Statistico Italiano 1905–1907*, Roma.
- Ministero di Agricoltura, Industria e Commercio [MAIC] (1913), *Annuario Statistico Italiano, vol. II-1912*, Roma.
- Ministero di Agricoltura, Industria e Commercio [MAIC] (1914). *Censimento della popolazione del Regno d'Italia al 10 giugno 1911 vol. II*. Roma: Popolazione presente classificata per sesso, età, stato civile ed istruzione.
- Ministero dei Lavori Pubblici. Direzione Generale delle Strade Ferrate [MLP] (1878), *Relazione Statistica sulle costruzioni e sull'esercizio delle strade ferrate italiane per l'anno 1877*, Roma.

**Appendix**Table A.1 *Descriptive statistics*

Variables	Years*	N	Minimum	Maximum	Mean	SD	Number of o
PATENTS (number)	1861	59	0	93	4.32	13.44	27
	1881	69	0	90	4.91	13.35	25
	1891	69	0	127	7.07	18.83	18
	1901	69	0	251	14.81	38.03	13
	1911	69	0	504	26.06	75.49	6
PATENTS_10 (number)	1861	59	0	22	1.02	3.41	45
	1881	69	0	9	0.68	1.74	51
	1891	69	0	17	1.00	2.74	48
	1901	69	0	22	1.06	3.25	48
	1911	69	0	54	2.64	8.80	42
PATENTS_HT (number)	1861	59	0	34	1.54	4.99	42
	1881	69	0	17	1.29	2.83	40
	1891	69	0	37	2.17	6.04	40
	1901	69	0	99	5.77	15.27	26
	1911	69	0	198	9.67	30.04	19
Literacy (% on population)	1861	59	7.20	51.10	19.88	12.06	0
	1881	69	11.73	65.04	30.94	15.03	0
	1891	69	17.58	91.16	49.29	21.62	0
	1901	69	17.42	76.52	42.22	17.35	0
	1911	69	24.82	82.45	51.43	16.47	0
Tech_Education (% on population)	1861	-	-	-	-	-	-
	1881	69	0.02	0.34	0.09	0.05	0
	1891	69	0.03	0.55	0.13	0.08	0
	1901	69	0.02	0.40	0.14	0.08	0
	1911	69	0.06	0.61	0.28	0.12	0
S&E_University (% on population)	1861	-	-	-	-	-	-
	1881	69	0.00	0.07	0.01	0.01	51
	1891	69	0.00	0.08	0.01	0.02	52
	1901	69	0.00	0.10	0.01	0.02	53
	1911	69	0.00	0.11	0.01	0.02	54
B&L_techproximity	1861	59	0.00	0.84	0.12	0.20	27
	1881	69	0.00	0.77	0.16	0.20	25
	1891	69	0.00	0.78	0.18	0.21	17
	1901	69	0.00	0.85	0.29	0.25	13
	1911	69	0.00	0.86	0.37	0.25	6
Urban_pop (% on population)	1861	59	0.00	82.59	13.15	16.61	26
	1881	69	0.00	79.71	14.94	16.44	23
	1891	69	0.00	78.81	16.11	16.67	19
	1901	69	0.00	77.92	17.32	16.98	18
	1911	69	0.00	76.40	18.64	17.34	16
Rail_kmq	1861	59	0.00	0.06	0.01	0.01	32
	1881	69	0.00	0.10	0.03	0.02	3
	1891	69	0.01	0.12	0.04	0.02	0
	1901	69	0.01	0.15	0.06	0.03	0

*(Continued)*

Table A.1 *Continued*

Variables	Years*	N	Minimum	Maximum	Mean	SD	Number of 0
Newspaper (number)	1911	69	0.01	0.18	0.06	0.03	0
	1861	59	0	36	2	5.48	34
	1881	69	0	216	21.1	35.8	1
	1891	69	3	243	25.78	41.18	0
	1901	69	3	273	27.55	45.43	0
Labor Productivity (1911 million Lira per 000 male workers)	1911	69	7	417	45.28	71.52	0
	1861	59	0.74	1.39	0.97	0.13	0
	1881	69	0.77	1.40	1.02	0.14	0
	1891	69	–	–	–	–	–
	1901	69	0.91	2.01	1.31	0.23	0
Population (in 000 inhabitants)	1911	69	1.24	2.38	1.72	0.28	0
	1861	59	100.63	948.32	367.30	199.65	–
	1881	69	114.30	1114.99	412.46	224.24	–
	1891	69	122.74	1268.08	440.31	246.10	–
	1901	69	123.88	1442.18	470.66	271.96	–
	1911	69	129.93	1726.55	502.48	305.06	–

*Note:* \*Patents data refer to 1864–65 instead of 1861 and 1902 instead of 1901; data for rail per kmq refer to 1876 instead of 1881, 1886 instead of 1891, and 1909 instead of 1911; data for newspaper refer to 1864 instead of 1861, 1895 instead of 1901, and 1905 instead of 1911; data for labor productivity refer to 1871 instead of 1861.