UNIVERSITÀ
DI SIENA
1240

QUADERNI DEL DIPARTIMENTO
DI ECONOMIA POLITICA E STATISTICA

Alessandro Nuvolari
Michelangelo Vasta

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

n. 724 – Dicembre 2015

THE GEOGRAPHY OF INNOVATION IN ITALY, 1861-1913: EVIDENCE FROM PATENT DATA

Alessandro Nuvolari
(Sant'Anna School of Advanced Studies, Pisa)

Michelangelo Vasta
(Department of Economics and Statistics, University of Siena)

ABSTRACT: In this paper we provide a systematic appraisal of the spatial patterns of inventive activity in Italy in the period 1861-1913. Our main source of evidence is a data-set containing all patents granted in Italy in five benchmark years (1864-65, 1881, 1891, 1902, 1911). Our geographical unit of analysis is the province, an administrative district of the time. First, using some simple descriptive statistics, we introduce a characterization of the spatial distribution of patents and of its evolution over time. Second, we perform an econometric exercise in which we assess the connection between different forms of human capital and patent intensity. We are able to establish a robust correlation between literacy and "basic" patent intensity and robust correlation between secondary technical education and scientific and engineering studies and "high quality" patent intensity. Third, we study the connection between patents and industrialization. Our exercise shows that patents exerted a significant role in accounting for the level of industrial production. Interestingly enough, in this context, the role of patents was possibly more relevant than that of the availability of water-power and of the level of real wages (two factors that were pointed out by the previous literature, mostly on the basis of rather impressionistic accounts of the evidence). Our study warrants two main conclusions. First, domestic inventive activities were an important element of the industrialization process, even in a late-comer country such as Italy. Second, at the time of the unification, Northern provinces were characterized by more effective innovation systems. This factor contributes to explain the growing divide in economic performance between the North and the South of the country during the Liberal age.

Acknowledgement: 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 and Carlo Ciccarelli for sharing their data with us. We also thank Chiara Tomasi, Emanuele Felice, Stefano Falcone, Andrea Giuntini and Stefano Maggi for helpful suggestions. The paper has also benefited from the comments of participants at the Economic History Society Annual Conference (Warwick, 2014), the Economic History Workshop (Siena, 2014), the XVIIth World Economic History Congress (Kyoto, 2015), the 11th European Historical Economics Society Conference (Pisa, 2015) and at a seminar at the Department of Economics and Management of the University of Trento.

December 2015

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.[1] In comparison with the traditional literature on the *questione meridionale* (the “Southern question”), the most recent contributions are characterized by an attempt to provide more and more sophisticated quantitative assessments of different dimensions of economic performance between the different areas of the country since the unification. In particular, considerable progress has been obtained as far as aggregate indicators of economic performance are concerned, so that two recent and somewhat conflicting sets of 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 constructed at the fine-grained provincial level by Ciccarelli and Fenoaltea (2013) and of industrial labour force constructed by Ciccarelli and Missiaia (2013). In addition, the most recent literature has also worked to the elaboration of 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 also to the construction of aggregate indicators of human development (Felice and Vasta 2015). Finally, one should also point to pioneering attempts to assess some potential cultural determinants of economic performance through the construction of regional indicators of social capital (Felice 2012).

Notwithstanding this burgeoning literature and the multiplication of research efforts in different directions, there has been so far only a very limited examination of the regional differences in the patterns of inventive activities and innovation performance . This is somewhat 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 or “localized” innovation systems are themes that have featured prominently in the economics of innovation 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 for a thorough recent survey). 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). Our inquiry is based on a newly developed data-set comprising all

[1] See, for example, the recent controversy between Daniele and Malanima (2014a, 2014b) and Felice (2014).

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 at local level is easily motivated by noting that the different trajectories of human capital accumulation are, in fact, very often considered as one of the primary factors accounting for the divide in economic performance between Northern and Southern Italy (Felice 2012, Cappelli 2015). In this perspective, the possible 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 scientific and engineering studies university may have been even more critical than the gap in basic education (Vasta 1999, Lacaita and Poggio 2011).[2] In this paper, accordingly, we will consider the geography of human capital formation in comprehensive manner, taking into account both "basic" education and secondary technical education and engineering studies. After our quantitative assessment of the relationship between different types of human capital and inventive activities, we shall put our findings in broader perspective by assessing the relative contribution of innovation and human capital to the industrialization process using a simple quantitative framework. The rest of the paper is organized as follows. Section 2 contains a description of the data and the sources used in the paper. Section 3 is devoted to our reconstruction of the (changing) geography of innovation in Italy during the Liberal Age. In section 4, we try to assess the potential factors accounting for the concentration of patenting activities at provincial level. Section 5 speculates on the possible relation between the geography of innovation and the process of industrialization. Section 6 draws conclusions.

2. Sources and Data

Our fundamental geographical unit is the "provincia", which was a local administrative unit of the time. Provinces vary significantly in terms of population and geographical extension, since their actual borders reflected historical and geographical accidents. The main advantage of using provincial data is that they provide a fine-grained picture (moving from

[2] Fenoaltea (2011, pp. 243-4) also suggests that the inability of developing a sound system of secondary technical education and scientific and engineering studies with a genuine national coverage is one of the main "failure" of the Italian state in the Liberal Age.

the regional to the provincial level means in this historical period means that moving from 16 to 69 geographical units) and, therefore, provincial level data may reveal patterns of spatial heterogeneity that can remain concealed when using regional data. This is in line with some recent contributions on the Italian regional divide such as Ciccarelli and Fenoaltea (2013), Ciccarelli and Missiaia (2013) and Ciccarelli and Proietti (2013). Interestingly enough, one of the main results emerging from the contribution of Ciccarelli and Fenoaltea (2013) is exactly that the patterns of economic activities at provincial level are somewhat more intricate than those visible at the broader regional level which were highlighted in some previous contributions (for example Daniele and Malanima (2011)).

Following an established tradition both in economics of innovation and in historical literature, we measure inventive activities at regional level using patents.[3] This is done by using a data-set of Italian patents granted during the Liberal Age constructed by Nuvolari and Vasta (2015a).[4] This dataset contains all the 10,124 patents granted in Italy in five benchmark years. In this paper we consider only the patents assigned to Italian residents in these benchmark years. The data-set contains 3,897 "Italian" patents whose distribution across benchmark years is as follows: 1864-65 (257), 1881 (338), 1891 (486), 1902 (1,022) and 1911 (1,794).[5] The choice of these benchmarks has been dictated by our concern of ensuring an even coverage of the entire liberal age period. The historical sources of these data are the Italian official serial publications of *Ministero di Agricoltura, Industria e Commercio* (MAIC 1864-1885, 1886-1893, 1894-1901, 1902-1916).

In the construction of the data-set, patents were assigned to provinces using fractional counting. 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 technical improvements to existing products and processes to genuine technological breakthroughs (Schmookler 1966). In order to address this issue, we have also constructed an indicator of patent quality using renewal data. Our quality indicator is the "scheduled" duration of the patent (measured in years).[6] This duration is computed by adding to the initial duration of the patent, that is the number of years for which the patent

[3] For a recent comprehensive survey on this issue, see Nagaoka, Motohashi and Goto (2010).
[4] For an account of the evolution of the Italian patent systems in the pre-unitary states of the Restoration period to the legislation implemented after the political unification of the country, see Nuvolari and Vasta (2015b).
[5] 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 (2015a).
[6] In Italy inventors could apply for a patent duration ranging from 1 to 15 years (with increasing fees for longer duration). For a more detailed discussion, see Nuvolari and Vasta (2015a).

was originally granted, all the years for which the patent was extended. The intuition is straightforward: patents taken or prolonged 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.

In order to account for the size of the provinces, we use population data collected from the censuses (MAIC 1865, 1883, 1903, 1914).[7]

We construct a number of variables describing the process of 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 the total population (MAIC 1865,1883,1893,1903,1914).[8] Another critical dimension of the process of human capital formation is the spread of technical education which could represent an important component of the innovation process. In this paper, 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 MAIC (1881, 1893, 1913) and BUPI (*ad annum*). Furthermore, we take also into account the tertiary level of education using the number of students enrolled at universities in each province.[9] We consider both the total number of students and the number of students attending scientific and engineering university courses. The sources of these variables are MAIC (1881, 1893, 1913) and BUPI (*ad annum*). This disaggregated perspective on human capital is in line with the recent argument of Mokyr (2005a) suggesting that in this historical phase the density of the upper-tail distribution of human capital may have been more than the basic level as a source of innovations.[10]

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 in the year in question.

---

[7] The population for 1891 (year in which the Census was cancelled due to the financial difficulties of the Kingdom) has been interpolated using the observations of 1881 and 1901.
[8] For the year 1891 (year in which the Census was not taken) we have used the share of spouses who were able to sign the marriage register.
[9] We consider government universities, "Università libere" such as Bocconi University, "Politecnici" and institutions offering "certified" graduate courses of tertiary education kind.
[10] It should be noted that 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).

This variable has been constructed on the basis of census data and other related sources. 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 MAIC (1881, 1893, 1908, 1913).[11] Of course, this is a rather rough 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 pre-unitary 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 Pompeo 2014).

In the second part of the paper where we examine the relationship between inventive activities and industrialization (section 5) and 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, which is measured by computing the average yearly discharge (flow) of rivers, canals and streams in the province (measured in m3/s).[12] The second is the level real wages. In this case we measure this variable by using a so far unexploited coeval source which is a comprehensive survey of the hourly wages of a wide number of unskilled workers in the construction sector (MAIC n.d.). We compute the real wages by dividing the nominal

[11] 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.

[12] These data have been retrieved from the website 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 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 m3/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.

wages with the price of bread at provincial level.[13] These estimates of real wages refer to the 1874-78 period.

Finally, in section 3 and 5, we make also use of the recent estimates of the spatial distribution of the labour force and of industrial value added constructed by Ciccarelli and Missiaia (2013) and Ciccarelli and Fenoaltea (2013).

3. The geography of inventive activities: a preliminary snapshot

The maps reported in Figure 1 show the geographical distribution of patents per million inhabitants. Initially, right after the Unification (1864-1865 map), the distribution of patents is strongly concentrated in the North and in the northern provinces of the Centre (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 27 provinces out of 59 with zero patents and 52 out of 59 with less than 5 patents, whereas in 1911 there are only 6 provinces out of 69 with zero patents and 31 out of 69 with less than 5 patents. Alongside with this process of gradual spatial diffusion, the second trend consists in the growing concentration 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, which reflects its administrative role as the capital of the Kingdom. Overall the maps of Figure 1 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 relatively high level of patent per capita next to provinces with low levels of patent per capita.

[Figure 1 around here]

The maps reported in Figure 2 contain the geographical distribution of patents adjusted for their quality. More specifically, the figure reports the number of duration-years of the patents of each province per million inhabitants. In other words, if in a province we have 5 patents with duration of 2 years and 3 patents with duration of 5 years, we have assigned a score of 25 duration-years (5x2 + 3x5 = 25) to the province in question. Then this score is

[13] Data on bread prices at provincial level have been kindly provided to us by Giovanni Federico.

divided by the resident population in millions. Overall, the dynamics is similar to the one emerging in Figure 1. Also in this case, the industrial triangle is clearly the predominant area.

[Figure 2 around here]

Figure 3 presents the density of patent activities across the three macro-regions of the country by means of box-plots.[14] Again, the critical role of the industrial triangle is visible in all benchmark years. The figure also suggests a very clear pattern of regional differentiation. Even if the distributions tend to overlap, the median for the North provinces is always above of the median of the Centre provinces and the median of the Centre provinces is always above the median of the South Provinces. . 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.). Figure 4 contains the distribution of the quality adjusted patents represented again by means of box-plots. Also in this case the profiles of the distribution seem to mirror closely those emerging from Figure 3.

[Figures 3 and 4 around here]

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.

[Table 1 around here]

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 pre-Unitary states. This means that we have considered each province as belonging to the pre-unitary state in which it was originally located. In this case, we consider the first benchmark as a proxy for the level of patenting in the pre-Unification period. The overall patterns are in line with the evidence discussed so far.

[14] Throughout this paper we adopt the following geographical partition of the country: the Northern area comprises Piedmont, Liguria, Lombardy, Veneto and Friuli; the Central-area comprises Emilia-Romagna, Tuscany, Marche, Umbria and Latium; the Southern and islands area comprises Campania, Abruzzo, Molise, Basilicata, Puglia, Calabria, Sicily and Sardinia.

Table 2 suggests that the divide between the North and the Centre and the South (at least for the largest states) is likely to have its origins before the Unification. Interestingly, in the initial phase, the Granducato di Toscana displays a remarkable performance overtaking that of Regno Lombardo-Veneto.[15] After the Unification, the provinces previously located in the Granducato di Toscana decline and they are overtaken by the growth of the provinces of the Stato Pontificio which is probably accounted for by the effect of the city of Roma (annexed in 1870 and becoming capital in 1871). We believe that the main message arising from Table 2 is that of a strong path-dependence in the location of inventive activities, in the sense that one cannot exclude that the predominance of the industrial triangle may have some long term roots in the history of the pre-unitary states. Figure 5 sets out the data of Table 2 by means of histograms.

[Table 2 around here]

[Figure 5 around here]

In Figure 6 we consider the connection between patenting activities and industrialization. In order to do so, we construct an Index of Relative Patenting (IRP) which is analogous to the Index of Relative Industrialization (IRI) estimated by Ciccarelli and Fenoaltea (2013).[16] IRP is computed as the ratio between the share of patents in a province on total patenting in Italy and the share of the population in a province on the total population. In mathematical terms, we have that:

$$IRP_{it} = \frac{pat_{it}}{\sum_{i=1}^{N_t} pat_{it}} \Bigg/ \frac{pop_{it}}{\sum_{i=1}^{N_t} pop_{it}} \qquad (1)$$

where i indicates the province, t indicates the year, pat_{it} the number of patents granted to inventors in province i in year t, N_t the number of provinces in the dataset in year t, pop_{it} the population resident in province i in year t.

[Figure 6 around here]

[15] It is worth noting that Granducato di Toscana was the only Italian pre-unitary state that did not have a formal patent legislation during the Restoration period.

[16] Ciccarelli and Fenoaltea (2013) computed their Index as the ratio between the share of industrial value added of a province and the share of the male population over age 15 of a province. Here, for both the IRI and the IRP, we use as denominator the share of the provinces in present population. We prefer this approach because it seems more appropriate to use the total population to normalize the patent indicator.

Again, the peculiar character of the industrial triangle is immediately visible for all benchmark years. In the first benchmark there are few other provinces which seem also to be characterized by concomitant high levels of relative patenting and industrialization (Firenze, Napoli and Livorno). It is interesting that the predominant role of the industrial triangle in terms of relative industrialization (IRI) is emerging at relatively late stage, approximately since 1891 (Ciccarelli and Fenoaltea 2013), whereas in terms of patenting activities is clearly visible since the Unification.

Table 3 examines the relationship between IRP and IRI by means of OLS regressions. In all the samples there is a clear significant relationship between industrialization and patenting. The coefficients range between 1.7 and 2.7 so that relative small changes in relative industrialization are actually related with larger variations in relative patenting which perhaps may be interpreted as pointing to the existence of scale economies in inventive activities. The behaviour of the relationship is also similar in the pooled regression. Finally, the R^2 coefficient is increasing over time suggesting that the connection between industrialization and patenting is becoming less noisy over time.

[Table 3 around here]

Table 4 shows the differences in the patterns of spatial correlation between relative industrialization and relative patenting by means of the Moran I statistic. Moran I statistic is a correlation coefficient which assesses the degree of spatial autocorrelation of a distributed variable in a given geographical space. In other words, Moran I statistic measures whether a variable displays a tendency to be systematically clustered in space, or, on the contrary, it is randomly spread. Higher values of Moran I statistic indicate a stronger degree of spatial autocorrelation. In other words, higher values of the Moran I statistic mean that provinces with relatively high values of IRI or IRP tend to be neighbouring to other provinces with relatively high values of IRI or IRP. Moran I statistic is computed as follows. Assume that the variable x is defined over a number locations n . We can construct a matrix W (spatial contiguity matrix) which indicates whether two provinces are bordering. The matrix is symmetric and each element w_{ij} is equal to 1 when the locations i and j are neighbouring and to 0 otherwise. The elements on the main diagonal of the contiguity matrix are equal to 0. In this case Moran I statistic is equal to

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} z_i w_{ij} z_j}{2 \cdot (\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}) \cdot \sum_{i=1}^{n} z_i^2} \qquad (2)$$

where is $z_i = x_i - \bar{x}$ (the deviation of x_i from the mean). Higher values of I indicate a stronger degree of (positive) spatial autocorrelation. Cliff and Ord (1981, pp. 42-46) illustrate how to compute significance intervals for Moran's I statistic under two different hypotheses: the first one is that the observations x are normally distributed (normality assumption) whereas the second one (randomised assumption) assumes that the realizations of x were extracted from one of the possible n ! permutations of the n values of the variable x over the n locations.

[Table 4 around here]

The main point emerging from Table 4 is that industrialization (IRI) exhibits a significant degree of spatial autocorrelation whereas this does not hold for patenting (IRP) which is characterized by relatively low levels of Moran I statistic that are always not significant. Our interpretation is that patenting is affected to a much higher degree than industrialization by some idiosyncratic factors that have a limited tendency to spread to neighbouring provinces. For example, the maps in Figures 1 and 2 suggest that the existence of a large urban centre inside the province, which is a characteristic only of a restricted number of them, may be a powerful driver of patenting. On the other hand, industrial production seems to have a stronger tendency to spread over space, possibly driven by geographical factors that are more widespread than those relevant for patenting (e.g. the availability of water resources, transport infrastructures, suitable locations for factories, etc.).

4. Inventive activities and human capital formation

In this Section we examine systematically the relationship between inventive activities and the process of human capital formation. Both the dependent variables and the covariates are measured as the provincial shares on the Italian total. In this way, the coefficients can be interpreted and compared in a straightforward way as elasticities. The dependent variables are the share of patents in each province over the national total (*s_patents*) and the share of patents with duration>=10 years over the national total (*s_patents10*).

We consider three types of human capital. The first is the level of “basic” human capital endowment of the province (this is measured as the share of literate population in the province over the national total: *s_literacy*); the second is endowment of technical skills (share of students attending technical and vocational schools over the national total: *s_tech_stud*); the third comprises the share of students attending university courses over the national total (*s_university_stud*) and the share of students attending scientific and engineering courses over the national total (*s_tech_university_stud*).

As additional control variables we employ the share of population of the province (*s_population*) which is a measure of its size; the urbanization level of the province (measured as the share of people living in towns with more of 30,000 in the province over the national total: *s_urban_population*); a proxy for the transport infrastructure, measured by the share of kilometres of railways in the provinces (*s_rail*); finally, we include a proxy of “access to information” measured as the share of newspapers and periodicals on the national total (*s_newspaper*);

Therefore our econometric specifications are the following:

$$\begin{aligned} s\_patents &= a + b \cdot s\_literacy_{it} + c \cdot s\_tech\_stud_{it} + d \cdot s\_university\_stud_{it} \\ &+ e \cdot s\_tech\_university\_stud_{it} + f \cdot s\_population_{it} + g \cdot s\_urban\_population_{it} \\ &+ h \cdot s\_rail_{it} + i \cdot s\_newspaper_{it} + \varepsilon_{it} \end{aligned} \quad (1)$$

$$\begin{aligned} s\_patents10 &= a + b \cdot s\_literacy_{it} + c \cdot s\_tech\_stud_{it} + d \cdot s\_university\_stud_{it} \\ &+ e \cdot s\_tech\_university\_stud_{it} + f \cdot s\_population_{it} + g \cdot s\_urban\_population_{it} \\ &+ h \cdot s\_rail_{it} + i \cdot s\_newspaper_{it} + \varepsilon_{it} \end{aligned} \quad (2)$$

Table 5 reports the results of regressions for the share of patents (*s_patents*) using pooled OLS (equation 1). The effects of literacy and technical education are both positive and significant in all specifications (only in column 7 the effect of technical education is not significant).[17] The size of these coefficients is also roughly similar. In particular, when both literacy and technical education are included in the regressions, the provincial share of patents is related to the shares of literacy and technical education in about a one-to-one ratio. Interestingly enough, the coefficient for the total university students is not significant. However, if we focus on scientific and engineering university education, we are able to find a positive and significant relationship, although its magnitude is considerably lower than the effects of literacy and technical education.

[17] In 1891 the literacy rate is computed using a different source (marriage registers instead of than census data). This may explain the lower significance of the variable.

We also assess the issue of the degree of spatial autocorrelation of the OLS residuals. For each time span of our sample, in the last rows of Table 5 we report the Moran's I statistic computed on the cross-sectional errors. In general, we do not find strong evidence of significant spatial autocorrelation (the size of the Moran's I statistic is relatively small and not significant but for 1911). To probe further into the autocorrelation effect of the 1911 year, we have included two additional regressions splitting the sample in two sub-periods. Column (6) reports the results of a regression that is limited to the 1881-1902 period. Reassuringly, the coefficients are fully in line with those estimated in column (5). Column (7) reports the results for a regression limited to the year 1911. In this case we find a similar degree of spatial autocorrelation in the OLS residuals to that of the cross-sectional errors of the other models for the year 1911.[18]

[Table 5 around here]

Table 6 reports the results of regressions estimating the connection between high quality patenting activity and human capital formation (equation 2). In this case we find a similar effect of the literacy variable to that of Table 5. Instead, in this case, the estimated effect of technical education is significantly higher (in the region of a two-to-one ratio). This finding suggests that technical education can indeed represent a critical ingredient for the generation of "significant" inventions. The effects of scientific and engineering university students is in line with the estimates of Table 5 (however the estimated effect is ambiguous in column 6 and 7 when the sample is split in different sub-periods). The last rows of Table 6 report the diagnostic test for the spatial autocorrelation of the OLS residuals. In this case the behaviour is analogous to that of Table 5.

[Table 6 around here]

In Table 7 we exploit the panel structure of the data and we re-estimate the model of Table 5 using fixed effects. The fixed effect model allows to control for unobserved heterogeneity across single provinces assuming this remains constant over time. In other words, the fixed effects specification is consistent with the assumption of a significant degree

[18] Moreover, we have examined several maps of the residuals for the year 1911. It appears rather clearly that the negative spatial autocorrelation for this year is due to the strong effect of some Northern provinces (in particular the three provinces of the industrial triangle) that are surrounded by neighbouring provinces characterized by OLS residuals of opposite "size" suggesting a possible sort of "hub" effect of the triangle provinces.

of micro-heterogeneity at the level of provinces in terms of socio-cultural characteristics with potentially long historical roots, a theme that has a long tradition in the literature on Italian economic development (Putnam, Leonardi and Nanetti 1993).[19]

[Table 7 around here]

The results reported in Table 7 confirm the positive effect of literacy and of scientific and engineering education (in the regressions where this variable is included). In this case, there is no significant effect of technical education. It is worth noticing that the estimates of Table 7 suggest that the connection between human capital formation and patent activities was probably varying over time. In particular, it is possible to find a significant a strong effect of the literacy variable in the models containing the 1864/65 benchmark (column 1 and column 7). On the other hand, in the models including the last benchmark year (1911), the effect is not significant. Interestingly enough, it is possible to still find significant effects for both literacy and scientific and engineering education in the model estimated in column 9 which covers the period 1881-1902. Our suggested interpretation is that the patenting activities were probably affected more strongly by the literacy gaps among provinces right after the political unification of the country. In the final benchmarks 1902 and 1911, the literacy gaps among provinces were somewhat reduced and this can account for the disappearance of a significant effect of this variable in the regression of Table 7.

Table 8 reports panel regressions estimates of the relationship between high quality patenting activities and human capital formation (equation 2). Technical education is playing a major role in affecting the generation of "high quality" patents. In all the specifications in which this variable is included, even changing the period observed, we find a significant and positive coefficient suggesting a connection in the region of a one-to-one ratio. It is also possible to find a positive and significant effect of the variable measuring scientific and engineering university students. This effect is about half of the effect observed for technical education. Finally, the literacy variable is not significant (except for the model estimated in column (7) when technical education is not included).

[Table 8 around here]

Overall the results of the regression exercises reported in Tables 5-8 suggest the existence of a significant connection between human capital formation and inventive

[19] The adoption of the fixed effect approach is also corroborated by the Hausman test.

activities measured using patents. Furthermore, we are able to discern two relative distinct channels of influence. Literacy seems to exert mostly a significant effect on “basic” inventive activities. However, when we focus on inventive activities characterized by a significantly higher degree of sophistication, we find that secondary technical education and scientific and engineering university education are mostly relevant. This differential impact of different types of human capital on the quality of innovation hints to the possible role played by the upper-tail of the human capital distribution suggested by Mokyr (2005a). In this perspective, our findings resonate with those obtained by Khan (2015) and Squcciarini and Voigtlander (2015). Khan finds that, in the English case, in the second half of the nineteenth century, inventors that were creators of major technological breakthroughs were characterized by relatively high levels of scientific and technical education. Squcciarini and Voigtlander, 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.

The results pointing to the critical role of technical education for high quality inventive activities are also in line with other studies which have emphasized for Italy the importance of this type of human capital formation in fostering the process of industrialization also at local level (Zamagni 1978 and 1993, Vasta 1999, Conti 2001). Similar considerations hold on the connection between engineering education and economic development (Lacaita 1993, Vasta 1996).[20]

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 possible role played by innovative activities in shaping the broader patterns of industrialization at provincial level. So far, a few recent studies have assessed the possible factors affecting the localization of industrial activities during this historical period both for manufacturing as whole (Fenoaltea 2011, Ciccarelli and Fenoaltea 2013, Ciccarelli and Proietti 2013, A’Hearn and Venables 2013), but also for individual industries (A’Hearn 1998). However, none of these papers has explicitly considered the role played by technical progress. We tackle this gap in this section estimating a number of models in which we add technical change to the

[20] For a discussion of engineering and technical education and industrialization in broader European perspective, see Fox and Guagnini (1993) and Wengenroth (2000).

more 'conventional' explanatory factors tested in the literature on the localization of industrial activities.

The indicator of industrialization that we use is the share of manufacturing employment on the provincial population. In particular, we consider as dependent variables both the level (share) of manufacturing employment in 1911 and the growth of this share in the 1881-1911 period. Following the literature just mentioned, we consider four main determinants of industrialization.[21]

The first is the "basic" human capital endowment of the provinces which is measured here by using literacy rates.

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, Crafts and Wolf 2014).

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 is 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 incentivizes investments in capital goods and machinery (Allen 2009).

Finally, we add to this set of variables, a measure of the technological performance of the province using the indicators constructed using the patent dataset described in section 2. We focus on the 1881-1911 period, considering the dependent variables in 1911 and the co-variates in circa 1880.

Figure 7 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 both the level of industrial employment in 1911 and its growth over the period 1881-1911 (our dependent variables). They are both characterized by a rather strong clustering in the North-West areas of the country around the industrial triangle (the growth seems to have a somewhat wider distribution, but still clustered around a clear North-West axis). The panels in the second row show the distributions of other co-variates. Provinces with relative high level of literacy in 1881 are indeed the most industrialized (both considering levels in 1911 and growth over the 1881-1911 period). Water resources are clearly clustered mostly around the Po Valley, while

[21] For a recent example of this type of approach, see Kelly, Mokyr and O'Grada (2015). However, in their study they do not explicitly consider the role played by patents. For other studies, which included patenting activities as possible factor for shaping industrialization, see Crafts and Wolf (2014) for cotton industry in the United Kingdom and Cinnirella and Streb (2013) for Prussia.

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 Centre and the South of the country.

[Figure 7 around here]

The panels in the third row contain three maps showing the spatial distribution of patenting activities in 1881. The first map contains the raw number of patents normalized by population; the second and third maps contain quality adjusted measurement of patenting activity based on their duration. The second map weighs each patent according to the year of "scheduled" duration, while the third one considers only patents with duration of 10 years or more normalized by population. Again, 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 9 presents the results of regressions relating the share of industrial employment in 1911 with the industrialization determinants we have just discussed. Overall, we find that literacy and patenting activity have a significant impact on industrialization. The size of the coefficients of literacy is quite stable across the three different specifications: an increase of 1 per cent in the literacy rate at provincial level would determine an increase of the industrial employment share of about 0.1 per cent. Patenting is also significant. All these three patent variables have a significant and positive impact. The effect of top quality patents is particularly remarkable: adding a top quality patent per million inhabitants would lead to an increase of industrial employment of 0.45 per cent.

[Table 9 around here]

In Table 10 we use the same framework to study the dynamics of industrialization, that is the rate of growth of industrial employment in the different provinces over the 1881-1911 period. In this case, the only variable which has a significant and positive impact is the literacy rate. All patenting proxies are not anymore significant. We have also included as covariate the initial (1881) level of industrial employment in order to assess a possible convergence effect. Indeed, this variable seems to play a role since the size of this coefficient is always negative and significant. In other words, provinces with initial low level of industrial employment are characterized by faster rate of industrialization in the subsequent 30 years period. However, the geographical spread of this convergence process must be better qualified. The specifications 4-6 show a negative and significant effect of a dummy variable for the South provinces. This means that the process of convergence in industrialization that we have detected is mostly concentrated in the peripheral provinces of the North and Center, which comprise several provinces without patents.[22] This piece of evidence is also in line with the results of Table 4 on the spatial correlation in the Index of Relative Industrialization (IRI).

[Table 10 around here]

Overall, the evidence presented in this section points to a possible connection between patenting and the industrialization process of the country which so far has not received adequate attention. However, a word of caution must be spent: the results of Tables 9 and 10 can of course be driven by the experience of a limited number of provinces rather than describing a generalized phenomenon. In particular, it is clear that the three provinces of the Industrial triangle may have a crucial role as shown in the maps presented in Figures 1, 2 and 7.

6. Conclusions

In this paper we have made a first attempt to look at the geography of innovation in Italy during the Liberal age. According to economic historians, this is a critical phase in the long run development of the Italian economy since it represents the moment in which the industrialization process was actually 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 in the same period (Felice 2013; Felice and Vasta 2015). Using patent data, we were able to provide what we believe is the first historically comprehensive assessment of the different innovative performance of Italian provinces. 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. This precocious concentration may reflect to some degree the legacy of the history of pre-unitary states.

[22] As a further robustness check, we have also estimated an additional set of regressions excluding the provinces with patenting activities. These results are consistent with the estimates of Tables 9 and 10. This is also visible in the second map of the first row of Figure 7.

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, Centre and South.

iii) We find a significant relationship between the localization of patenting activity and different types of human capital formation. In particular, we establish that literacy is related with standard patenting and that technical education and scientific and engineering university education are related with high-quality patenting.

iv) Finally, we were also able to reveal an interesting connection between patenting and industrialization.

We have further explored this latter point using a framework in which the location of industrialization is the outcome of the interaction of four types of factors: 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 innovation capacity. In this way, our interpretation of the geography of innovation can be seen 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. On this latter point, we have been able to highlight a broad correlation between literacy rates and both the levels and the growth rates of industrial employment. We have also identified a long term connection between innovative activities and levels of industrial employment. However, the impact of innovative activities on the growth rates of industrial employment is not particularly evident. Our interpretation is that innovative activities were an engine of industrialization only for a very restricted number of provinces. Therefore, the general lesson of our paper is that the Italian industrialization process during the Liberal Age it is not susceptible to be summarized in terms of overarching meta-narratives, but should be told in terms of a more nuanced and complex tale of localized and interacting development trajectories.

References

A'Hearn, B. (1998), 'Institutions, externalities, and economic growth in southern Italy: evidence from the cotton textile industry, 1861-1914', *The Economic History Review,* vol. 51, (4), pp. 734-762.

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.

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', *Cliometrica,* forthcoming.

Ciccarelli, C. and Fenoaltea, S. (2013), 'Through the magnifying glass: provincial aspects of industrial growth in post-unification Italy', *Economic History Review,* vol. 66, pp. 57-85.

Ciccarelli, C. and Missiaia, A. (2013), 'The industrial labour force of Italy's provinces: estimates from population censuses, 1871-1911', *Rivista di Storia Economica,* vol. 29, pp. 141-192.

Ciccarelli, C. and Pompeo, M. (2014), "Railways development in liberal Italy", *mimeo*.

Ciccarelli, C. and Proietti, T. (2013), 'Patterns of industrial specialisation in post-Unification Italy', *Scandinavian Economic History Review,* vol. 61, pp. 259-286.

Cinnirella, F. and J. Streb (2013), *The Role of Human Capital and Innovation in Prussian Economic Development,* CESifo Working Paper, no. 4391.

Cives, G. (ed.) (1990), *La scuola italiana dall'Unità ai nostri giorni*, Firenze: La nuova Italia Editrice.

Cliff, A.D. and Ord, J.K. (1981), *Spatial Processes: Models and Applications*, London, Pion.

Conti, E. (2001), *Istruzione tecnica e professionale e sviluppo economico italiano: 1859-1940*, PhD Thesis, University of Bari.

Crafts, N. and Wolf, N. (2014), 'The location of the UK cotton textiles industry in 1838: a quantitative analysis', *Journal of Economic History,* vol. 74 (4), pp. 1103-1139.

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,* vol. 30 (1), pp. 3-36.

Daniele, V. and Malanima, P. (2014b), 'Due commenti finali', *Rivista di Storia Economica*, vol. 30 (2), pp. 243-248.

Feldman, M. and Kogler, D. (2010), 'Stylized facts in the geography of innovation' in B.Hall, B. and N. Rosenberg (eds.), *Handbook of Economics of Innovation,* Amsterdam: North Holland, vol. I, pp. 381-410.

Felice, E. (2012), 'Regional convergence in Italy, 1891-2011: testing human and social capital', *Cliometrica,* vol. 6, pp. 267-306.

Felice, E. (2013), *Perchè 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*, vol. 30 (2), 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*, vol. 19 (1), pp. 44-66.

Fenoaltea, S. (2011), *The Reinterpretation of Italian Economic History: from Unification to the Great War*, New York, Cambridge University Press.

Ferrovie dello Stato, Ufficio Statistica (1911), *Ferrovie Italiane 1861-1909. Riproduzione dei lavori grafici presentati all’ Esposizione Internazionale di Torino del 1911*, Roma.

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*, vol. 15 (4), pp. 360-375.

Khan, Z. (2015), ‘Knowledge, Human Capital and Economic Development: Evidence from the British Industrial Revolution, 1750-1930’, NBER Working paper, n. 20853.

Kelly, M., J. Mokyr and O’Grada C. (2015), *Roots of the Industrial Revolution*, UCD Centre for Economic Research Working Paper series, no. WP15/24.

Lacaita, C.G. (1993), ‘Ingegneri e scuole politecniche nell’Italia liberale’, in S. Soldani and G. Turi (eds.), *Fare gli italiani. Scuola e cultura nell’Italia contemporanea. I. La nascita dello Stato nazionale*, Bologna: Il Mulino, pp. 213-253.

Lacaita, C. and P.P. Poggio (2011), *Scienza, Tecnica e Industria nei 150 anni di Unità d’Italia*, Milano: Jaca Book.

Ministero di Agricoltura, Industria e Commercio – Direzione Generale di Statistica (n.d.), *Salari. Prezzi medii 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. Popolazione,* Torino.

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. Popolazione classificata per età sesso, stato civile e istruzione elementare,* Roma.

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] (1902-1916), *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. Popolazione presente classificata per sesso, età, stato civile ed istruzione,* Roma.

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.

Mokyr, J. (1976), *Industrialization in the Low Countries, 1795-1850,* New Haven, Yale University Press.

Mokyr, J. (1995), 'Urbanization, technological progress, and economic history', in H. Giersch (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. 1113-1180.

Mokyr, J. (2005b), 'The Intellectual Origins of Modern Economic Growth', *Journal of Economic History*, Vol. 65, pp. 285–351.

Nagaoka, S., Motohashi, K. and Goto, A. (2010), 'Patent statistics as an innovation indicator', in B.H. Hall and N. Rosenberg (eds.), *Handbook of Economics of Innovation*, vol. 2, Amsterdam: North Holland, pp. 1083-1127.

Nuvolari, A. and Vasta, M. (2015a), 'Independent invention in Italy during the Liberal Age, 1861-1913', *The Economic History Review*, vol. 68 (3), pp. 858-886.

Nuvolari, A. and Vasta, M. (2015b), 'The Italian patent system(s) during the long nineteenth century (from the Congress of Vienna to World War I), *mimeo*.

Putnam, R.D., Leonardi, R. and Nanetti, R.Y. (1993), *Making democracy work: civic traditions in modern Italy*, Princeton: Princeton University Press.

Roses, J.R., 'Measuring the contribution of human capital to the development of the Catalan factory system (1830-61)', *European Review of Economic History*, vol. 2 (1), pp. 25-48.

Schmookler, J. (1966), *Invention and Economic Growth,* Cambridge (MA): Harvard University Press.

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.

Wengenroth, U. (2000), 'Science, Technology, and Industry in the 19th Century', Working Paper, Munich Center for the History of Science and Technology.

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. 137-178.

Zamagni, V. (1993), *The economic history of Italy 1860-1990*, Oxford: Oxford University Press.

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.

*Table 2. Distribution of patents according to pre-unitary states borders (1864/65-1881)**

	1864-65			1881		
	No. patents	Present population	Patents per million population	No. patents	Present 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

* 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.
Source: our own elaboration.

Table 3. Index of Relative Patenting (IRP) versus Index of Relative Industrialization (IRI), 1864/65-1911

Year	Coefficient IRI	R^2
1864-65	2.680 (0.502)	0.333
1881	2.106 (0.334)	0.368
1891	2.041 (0.248)	0.503
1902	1.727 (0.213)	0.495
1911	1.697 (0.198)	0.522
Pooled	1.957 (0.127)	0.416

Source: our own elaboration on patents database and Ciccarelli and Fenoaltea (2013).

Table 4. Spatial autocorrelation for IRI and IRP

IRI	Moran I statistic	Normalized	Randomized
1864-65	0.265	***	***
1881	0.222	***	***
1891	0.291	***	***
1902	0.341	***	***
1911	0.422	***	***
IRP	Moran I statistic	Normalized	Randomized
1864-65	-0.006		
1881	0.028		
1891	0.006		
1902	0.046		
1911	0.006		

Notes: *,**,*** indicate significance levels of 10%, 5% and 1% respectively.
Standard errors in parentheses.

Table 5. Patenting activity and human capital formation (pooled OLS regressions)

Variables	(1) 1864/65-1911	(2) 1881-1911	(3) 1881-1911	(4) 1881-1911	(5) 1881-1911	(6) 1881-1902	(7) 1911
s_literacy	2.132*** (0.167)	1.482*** (0.199)	1.473*** (0.200)	1.293*** (0.192)	1.148*** (0.181)	1.035*** (0.187)	1.858*** (0.591)
s_tech_stud		1.418*** (0.188)	1.442*** (0.193)	1.200*** (0.183)	0.743*** (0.185)	0.706*** (0.187)	0.938 (0.653)
s_university_stud			-0.0207 (0.0395)				
s_tech_university_stud				0.209*** (0.0390)	0.110*** (0.0371)	0.091** (0.0393)	0.174* (0.101)
s_population	0.494* (0.281)	-0.591* (0.318)	-0.561* (0.324)	-0.669** (0.303)	-0.612* (0.337)	-0.711** (0.358)	-0.926 (0.967)
s_urban_population					-0.217* (0.116)	-0.145 (0.124)	-0.420 (0.320)
s_rail					-0.476*** (0.166)	-0.343* (0.180)	-0.721* (0.415)
s_newspaper					0.846*** (0.0878)	0.862*** (0.0937)	0.708*** (0.233)
Constant	-2.425*** (0.288)	-1.898*** (0.274)	-1.933*** (0.282)	-1.497*** (0.271)	-0.786*** (0.248)	-0.717*** (0.274)	-0.883 (0.571)
Observations R-squared	335 0.662	276 0.738	276 0.739	276 0.764	276 0.827	207 0.836	69 0.827
Diagnostics for spatial autocorrelation							
Moran's I in the residuals (1864-65)	-0.029						
Moran's I in the residuals (1881)	-0.113	-0.123	-0.124	-0.112	-0.202**	-0.201**	
Moran's I in the residuals (1891)	-0.084	-0.055	-0.057	-0.041	-0.062	-0.075	
Moran's I in the residuals (1902)	-0.140*	0.058	0.063	0.027	-0.084	-0.125	
Moran's I in the residuals (1911)	-0.209***	-0.189**	-0.191**	-0.165**	-0.187**		-0.146*

Notes: pooled OLS regressions (dependent variable is the provincial share of patents on the national total), *,**,*** indicate significance levels of 10%, 5% and 1% respectively. Standard errors in parentheses.

Table 6. High quality patenting activities and human capital formation (pooled OLS regressions)

Variables	(1) 1864/65-1911	(2) 1881-1911	(3) 1881-1911	(4) 1881-1911	(5) 1881-1911	(6) 1881-1902	(7) 1911
s_literacy	1.920*** (0.207)	1.143*** (0.228)	1.124*** (0.229)	0.976*** (0.226)	0.890*** (0.246)	0.680*** (0.238)	2.027** (0.898)
s_tech_stud		2.209*** (0.216)	2.257*** (0.222)	2.016*** (0.215)	1.983*** (0.251)	1.892*** (0.238)	2.582** (0.992)
s_university_stud			-0.0431 (0.0453)				
s_tech_university_stud				0.185*** (0.0457)	0.166*** (0.0503)	-1.017** (0.456)	-1.879 (1.469)
s_population	0.822** (0.351)	-1.309*** (0.365)	-1.246*** (0.371)	-1.378*** (0.356)	-0.952** (0.457)	0.179*** (0.0501)	0.0876 (0.154)
s_urban_population					-0.209 (0.157)	-0.194 (0.158)	-0.267 (0.486)
s_rail					-0.434* (0.225)	-0.223 (0.230)	-0.896 (0.631)
s_newspaper					0.240** (0.119)	0.293** (0.119)	0.0466 (0.353)
Constant	-2.601*** (0.359)	-1.515*** (0.314)	-1.587*** (0.323)	-1.161*** (0.318)	-0.993*** (0.336)	-0.885** (0.350)	-1.026 (0.867)
Observations R-squared	335 0.550	276 0.704	276 0.705	276 0.721	276 0.728	207 0.759	69 0.697
Diagnostics for spatial autocorrelation							
Moran's I in the residuals (1864-65)	0.076						
Moran's I in the residuals (1881)	-0.162*	-0.098	-0.099	-0.104	-0.144*	-0.164*	
Moran's I in the residuals (1891)	-0.127	-0.064	-0.065	-0.060	-0.081	-0.123	
Moran's I in the residuals (1902)	-0.129	-0.03	-0.029	-0.020	-0.035	-0.048	
Moran's I in the residuals (1911)	-0.196**	-0.208***	-0.211***	-0.183**	-0.190**		-0.150*

Notes: pooled OLS regressions (dependent variable is the provincial share of 'high quality' patents on the national total), *,**,*** indicate significance levels of 10%, 5% and 1% respectively. Standard errors in parentheses.

Table 7. Patenting activity and human capital formation (panel data estimations)

Variables/Sample period	(1) 1864/5-1911	(2) 1881-1911	(3) 1881-1911	(4) 1881-1891	(5) 1881-1911	(6) 1881-1891	(7) 1864/5-1911	(8) 1881-1911	(9) 1881-1902
s_literacy	1.385*** (0.313)	-0.0846 (0.241)	-0.0682 (0.240)	1.337** (0.580)	-0.120 (0.225)	0.938 (0.600)	2.182*** (0.353)	-0.0447 (0.227)	0.911** (0.402)
s_tech_stud			-0.139 (0.0883)	0.0733 (0.136)	-0.131 (0.0827)	0.125 (0.135)			0.0650 (0.0992)
s_tech_university_stud					0.199*** (0.0367)	0.252** (0.125)		0.186*** (0.0377)	0.0859* (0.0461)
s_population	1.218* (0.694)	1.289*** (0.451)	1.455*** (0.461)	3.312** (1.581)	0.881** (0.445)	1.568 (1.772)	-0.240 (0.682)	0.649 (0.454)	-1.040 (0.903)
s_urban_population							0.643*** (0.184)	0.102 (0.139)	0.204 (0.191)
s_rail							0.00136 (0.0994)	-0.132 (0.0882)	-0.100 (0.105)
s_newspaper							-0.446*** (0.0753)	0.231*** (0.0781)	0.334*** (0.0877)
Constant	-2.389*** (0.866)	-0.294 (0.582)	-0.357 (0.582)	-5.397** (2.252)	0.251 (0.557)	-2.731 (2.569)	-1.700** (0.806)	0.0125 (0.594)	0.784 (1.033)
Observations	335	276	276	138	276	138	335	276	207
R-squared	0.149	0.044	0.055	0.186	0.174	0.233	0.284	0.210	0.177
Number of provincia1	69	69	69	69	69	69	69	69	69

Notes: panel fixed effect regressions (dependent variable is the provincial share of patents on the national total), *,**,*** indicate significance levels of 10%, 5% and 1% respectively. Standard errors in parentheses.

Table 8. High quality patenting activity and human capital formation (panel data estimations)

Variables/Sample period	(1) 1864/5-1911	(2) 1881-1911	(3) 1881-1911	(4) 1881-1891	(5) 1881-1911	(6) 1881-1891	(7) 1864/5-1911	(8) 1881-1911	(9) 1881-1911
s_literacy	0.8 (0.492)	-0.102 (0.615)	-0.208 (0.594)	2.032 (1.333)	-0.324 (0.564)	0.641 (1.323)	2.367*** (0.513)		-0.129 (0.564)
s_tech_stud			0.890*** (0.218)	1.180*** (0.312)	0.907*** (0.207)	1.361*** (0.298)		0.769*** (0.206)	0.771*** (0.207)
s_tech_university_stud					0.447*** (0.0918)	0.878*** (0.275)		0.412*** (0.0934)	0.413*** (0.0937)
s_population	3.420*** (1.090)	4.957*** (1.151)	3.895*** (1.139)	5.319 (3.635)	2.604** (1.113)	-0.764 (3.905)	0.332 (0.992)	1.536 (1.033)	1.648 (1.146)
s_urban_population							1.469*** (0.267)	0.638* (0.345)	0.633* (0.347)
s_rail							0.0192 (0.144)	-0.297 (0.215)	-0.287 (0.220)
s_newspaper							-0.905*** (0.109)	-0.560*** (0.197)	-0.561*** (0.198)
Constant	-4.805*** (1.361)	-5.586*** (1.487)	-5.186*** (1.437)	-10.92** (5.177)	-3.820*** (1.391)	-1.619 (5.661)	-3.405*** (1.172)	-2.172 (1.471)	-2.161 (1.476)
Observations	335	276	276	138	276	138	335	276	276
R-squared	0.092	0.100	0.168	0.320	0.255	0.412	0.347	0.298	0.298
Number of provincia1	69	69	69	69	69	69	69	69	69

Notes: panel fixed effect regressions (dependent variable is share of ‘high quality’ patents on the national total), *,**,*** indicate significance levels of 10%, 5% and 1% respectively. Standard errors in parentheses.

Table 9. Determinants of industrialization 1911

Variables	(1)	(2)	(3)
literacy 1881	0.110*** (0.0238)	0.117*** (0.0243)	0.140*** (0.0247)
water resources	-0.000950 (0.00144)	-0.000668 (0.00149)	-0.000958 (0.00159)
real wages a. 1880	-3.575 (2.789)	-3.338 (2.867)	-3.551 (3.068)
Patpop	0.127*** (0.0269)		
years patpop		0.0241*** (0.00571)	
patpop 10 years			0.446*** (0.163)
Constant	6.576*** (1.566)	6.323*** (1.606)	6.110*** (1.717)
Observations *R-squared*	69 0.573	69 0.548	69 0.483

Notes: OLS regressions (dependent variable is the provincial share of industrial employment), *,**,*** indicate significance levels of 10%, 5% and 1% respectively. Standard errors in parentheses.

Table 10. Determinants of industrialization's growth 1881-1911

Variables	(1)	(2)	(3)	(4)	(5)	(6)
literacy 1881	0.0123*** (0.00147)	0.0127*** (0.00146)	0.0135*** (0.00140)	0.00806*** (0.00188)	0.00836*** (0.00189)	0.00883*** (0.00189)
water resources	-1.78e-05 (8.58e-05)	-1.22e-05 (8.71e-05)	-2.11e-05 (8.85e-05)	-7.01e-05 (8.13e-05)	-6.71e-05 (8.27e-05)	-7.62e-05 (8.34e-05)
real wages a. 1880	-0.103 (0.167)	-0.0932 (0.169)	-0.0891 (0.172)	0.135 (0.172)	0.147 (0.173)	0.159 (0.175)
share ind empl 1881	-3.171*** (0.436)	-3.099*** (0.438)	-2.971*** (0.438)	-2.397*** (0.469)	-2.317*** (0.470)	-2.189*** (0.466)
patpop	0.00315* (0.00167)			0.00256 (0.00156)		
years patpop		0.000500 (0.000346)			0.000378 (0.000324)	
patpop 10 years			0.00441 (0.00926)			0.00282 (0.00857)
South (dummy)				-0.216*** (0.0655)	-0.219*** (0.0662)	-0.227*** (0.0665)
Constant	0.883*** (0.118)	0.864*** (0.118)	0.837*** (0.118)	0.900*** (0.109)	0.883*** (0.110)	0.862*** (0.110)
Observations R-squared	69 0.768	69 0.763	69 0.756	69 0.803	69 0.799	69 0.795

Notes: OLS regressions (dependent variable is the growth rate of the provincial share of industrial employment), *,**,*** indicate significance levels of 10%, 5% and 1% respectively. Standard errors in parentheses.

Figure 1. Geographical distribution of patents per million present population, 1864/65-1911

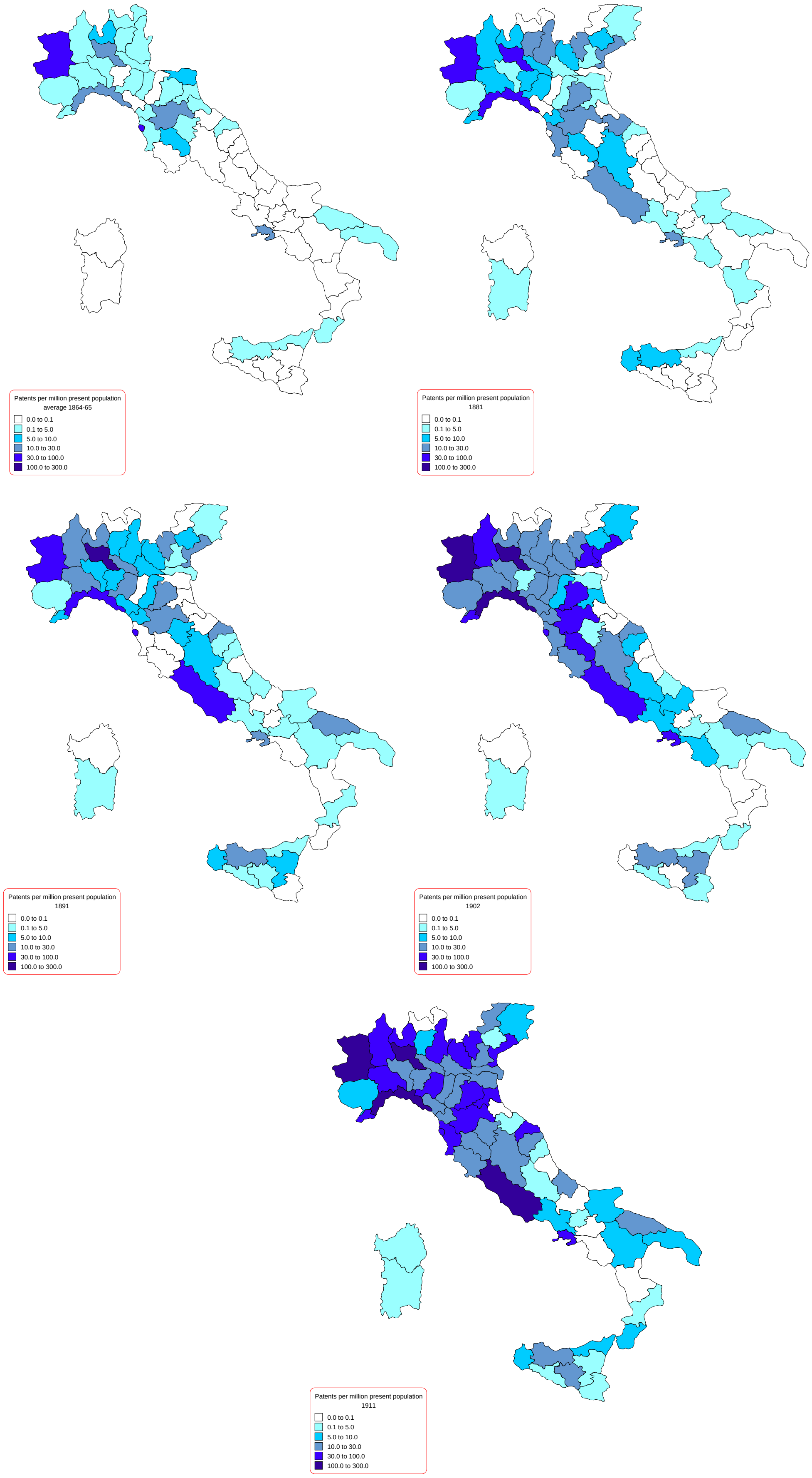


Figure 2. Geographical distribution of quality adjusted patents per million present population, 1864/65-1911

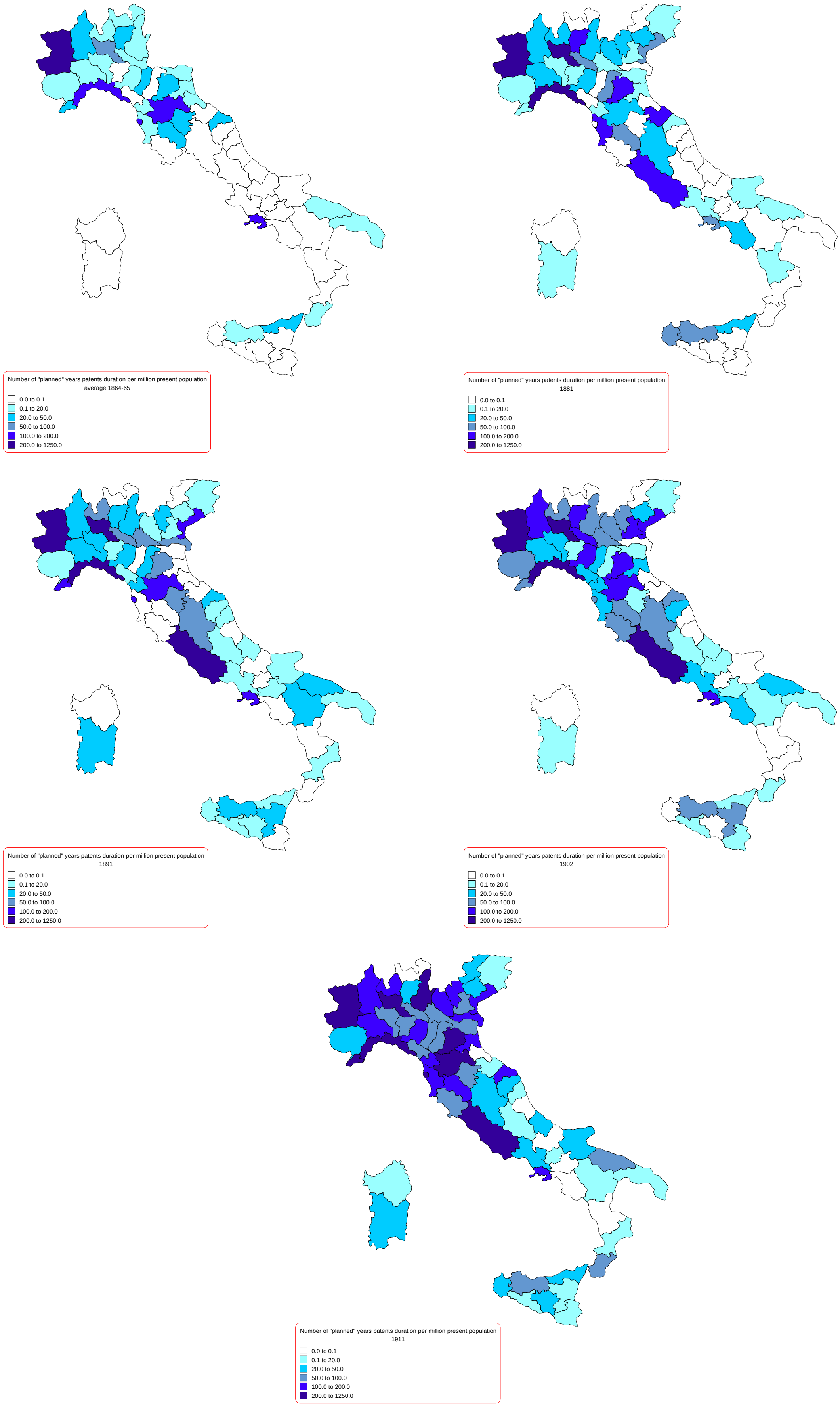


Figure 3. Patents per million present population, 1864/65-1911

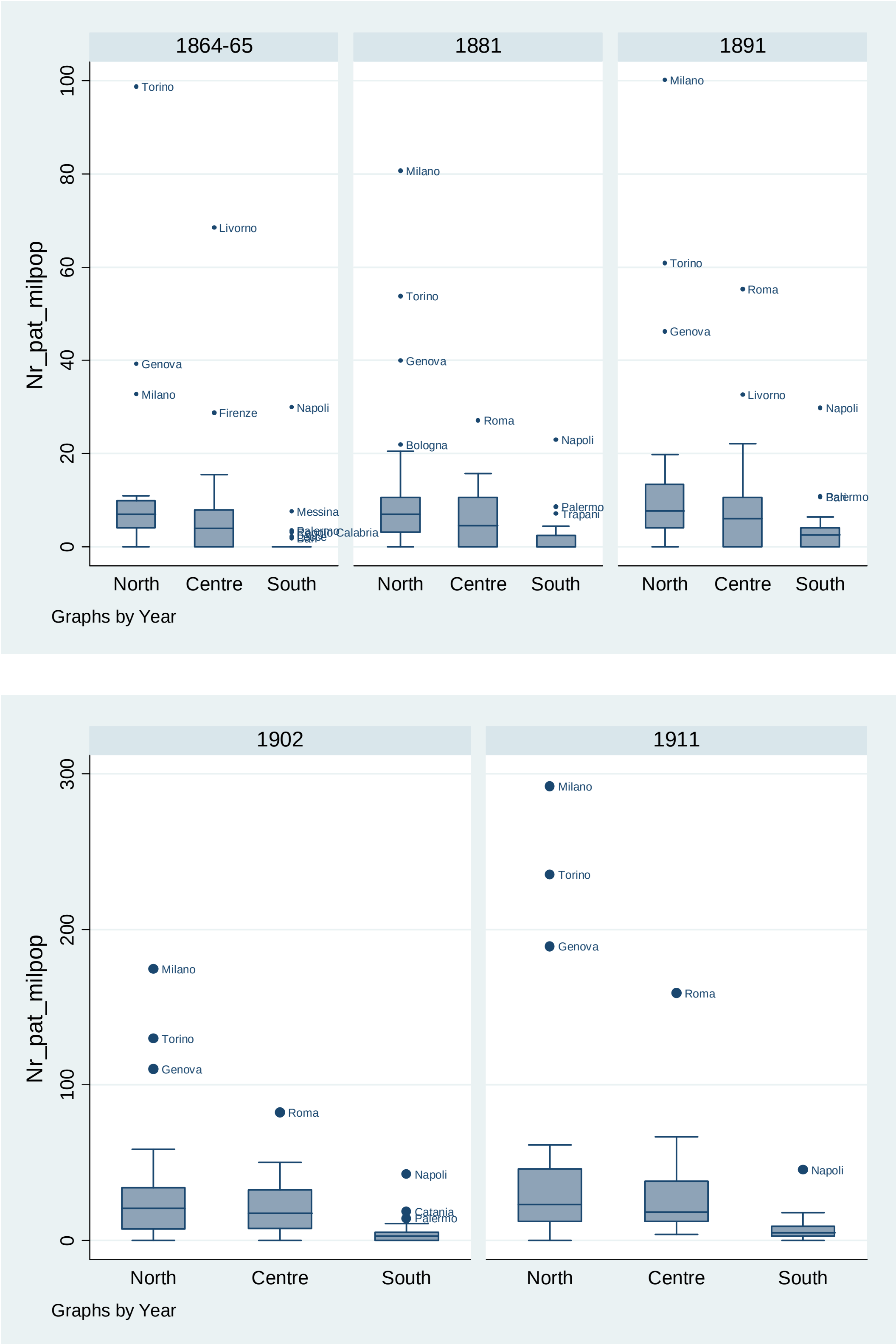


Figure 4. Quality adjusted patents per million present population, 1864/65-1911

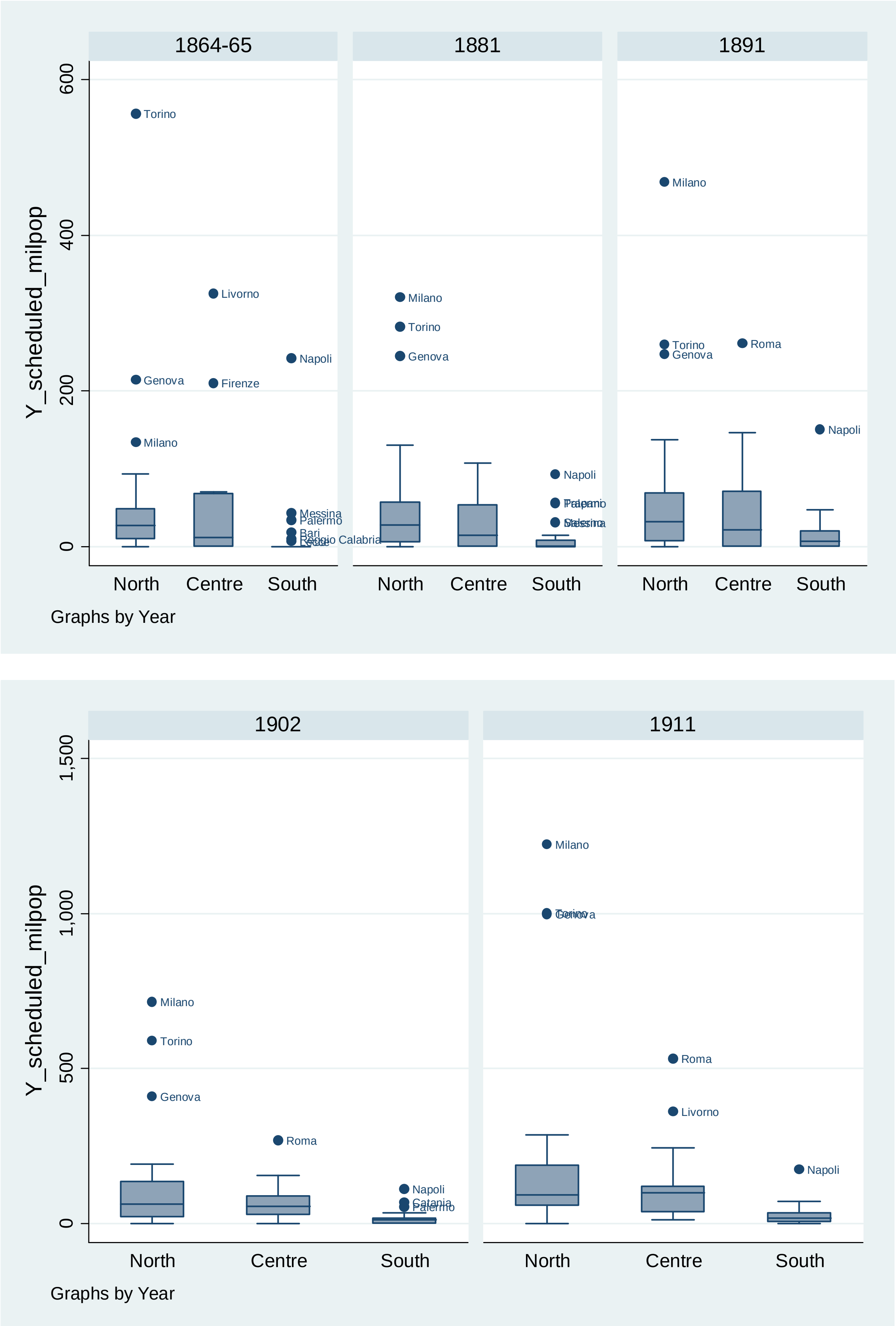
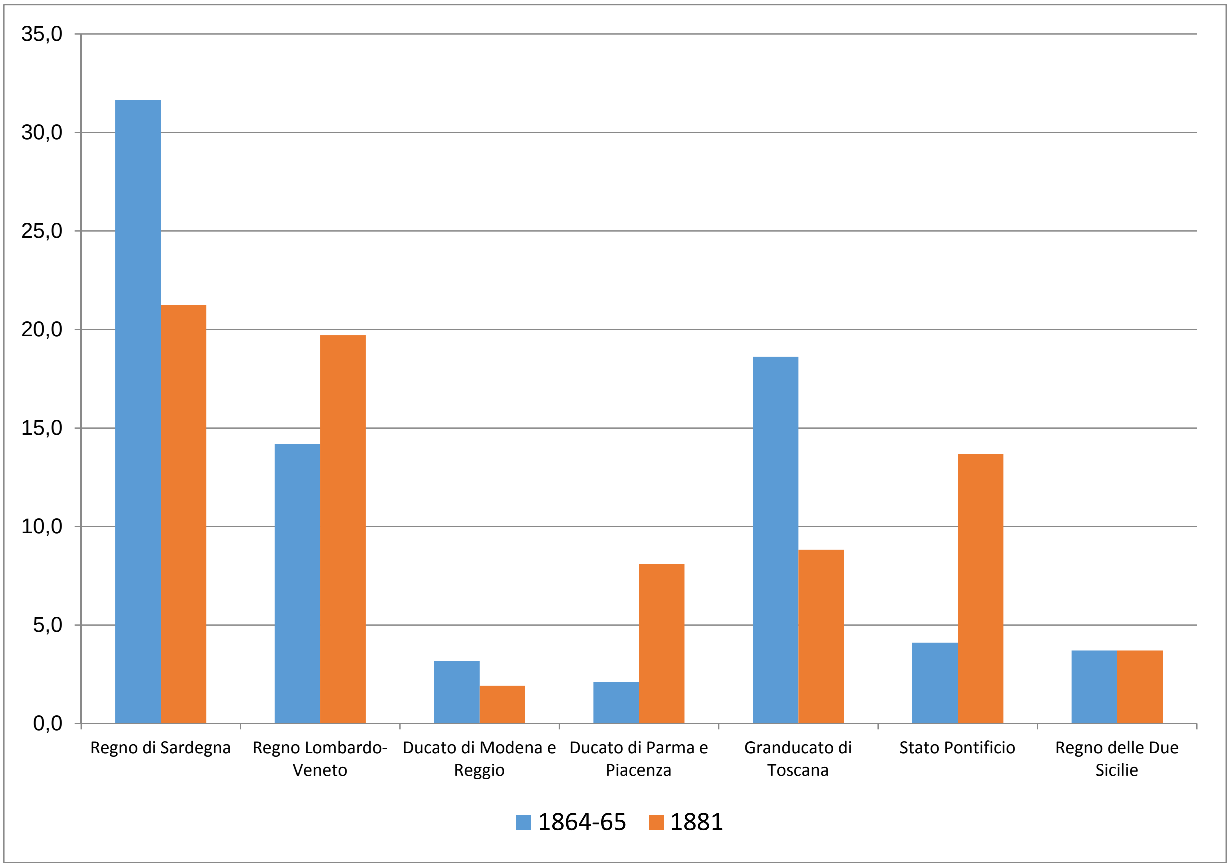


*Figure 5. Distribution of patents according to pre-unitary states borders (1864/65-1881)**



* 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.

Figure 6. Index of Relative Patenting (IRP) versus Index of Relative Industrialization (IRI), 1864/65-1911

Figure 7. Maps of provinces industrialization and their determinants