

# Industrial growth and spatial spillovers in XIX century Italy\*

Carlo Ciccarelli<sup>†</sup>

Università di Roma “Tor Vergata”

Stefano Fachin<sup>‡</sup>

Università di Roma “La Sapienza”

June 30, 2014

## Abstract

This paper represents a first attempt to investigate the early steps of Italian industrialization at the local level accounting for spillovers. The analytical framework is the conditional convergence model augmented to account for spatial effects. The geographic unit of analysis is constituted by provinces (NUTS 3 units) and the time period considered goes from 1871 to 1911. Estimation of the conditional convergence model augmented with human capital, social capital, and social overhead capital suggests that education, a cooperative culture, and spatial spillovers are able to explain, other things being equal, much of the variability of value added growth in the manufacturing industry in 19th century Italy. Since northern provinces generally had a more educated population and a more cooperative culture the polarization between dynamic, high-growth areas and backward areas seems to go back to the very beginning of Italy's economic history.

---

\*Paper prepared for the AISRe's 35th Annual Meeting - Padua, 11-13 September, 2014. We thank Brian A'Hearn for kindly providing data on age-heaping. Research supported by MIUR PRIN (Programmi di Ricerca Scientifica di Rilevante Interesse Nazionale) grant 2010J3LZEN. **PRELIMINARY VERSION. PLEASE DO NOT QUOTE. COMMENTS WELCOME.**

<sup>†</sup>Università di Roma “Tor Vergata”, Dipartimento di Economia e Finanza, Via Columbia 2, I-00133 Rome, Italy; email: [carlo.cicarelli@uniroma2.it](mailto:carlo.cicarelli@uniroma2.it).

<sup>‡</sup>Università di Roma “La Sapienza”; email: [stefano.fachin@uniroma1.it](mailto:stefano.fachin@uniroma1.it).

*The making of Italy – the merging of its elementary components, the harmonization of the South and the North – is a complicated task. It is more difficult than dealing with Austria and the Catholic Church.*

(C. Cavour, Prime Minister of Italy, 1860)

## 1 Introduction

This paper represents a first attempt to investigate the early steps of industrialization in Italy at the local level within a well defined quantitative setting. It does so by estimating a conditional convergence model of value added in the manufacturing industry of Italy's provinces (NUTS 3 units) allowing for spatial effects.<sup>1</sup> The time period covered goes from 1871 to 1911. Exploiting brand new data gathered from primary historical sources on human capital, social capital, and infrastructures endowment at the local level, we provide new empirical evidence on the main determinants of local industrial growth in 19th century Italy.

The literature on the historical roots of the regional divide, briefly summarized in the next section, is surely sizeable, to say the least. However, a key point here is that the geographical disaggregation of the data examined is typically low: at most *regioni*, NUTS 2 level. As a consequence, the geographical dynamics have not been so far adequately examined. The task to fill this gap in the literature may now be tackled using the recent 1911 prices manufacturing value added estimates at the level of *province*, Italy's NUTS 3 units<sup>2</sup>, for pre-WWI census years presented in Ciccarelli and Fenoaltea (2013). These data offer a wealth of new insights on the early stages of 19th century Italian industrial development.

Anticipating things a bit, but the point is crucial to our analysis, Figure 1 illustrates a Moran scatterplot of the growth of per capita value added in the manufacturing industry in Italy's provinces from 1871 to 1911. The figure includes four panels, using the same scale. Panel A refers to the whole set of 69 Italian provinces. Panels B1, B2, and B3 refer, respectively, to northern, center, and southern provinces.<sup>3</sup> Panel A also reports a weighted smoother (lowess) fitted through the value added data; for the sake of comparison the same smoother is also added in the macro-regional panels. The y-axis refers to the "neighbors" growth (provinces within a distance of 100 km, as detailed later) while the x-axis refers to each province own growth. A few features appear immediately. First, "high" provincial growth rates (measured along the x-axis) are only present among northern and center provinces, with all southern province below the 0.3 growth threshold. Second, and at this point inevitably, high "neighbors" growth rates (measured along the y-axis) are again only present among northern and center provinces. Third, a certain amount of within macro-area heterogeneity characterizes northern and central provinces, while the bulk of southern provinces appear well concentrated around a limited growth area surrounded by stagnating neighbors.

The bring-home conclusion of the above introductory analysis is that a reasonable model of Italian industrialization during the second half of the 19th century has to necessarily account for some form of spatial dependence.

The main goal of the present piece is, as anticipated, to shed light on the main determinants of the early steps of Italy's industrial growth. The economic growth literature,

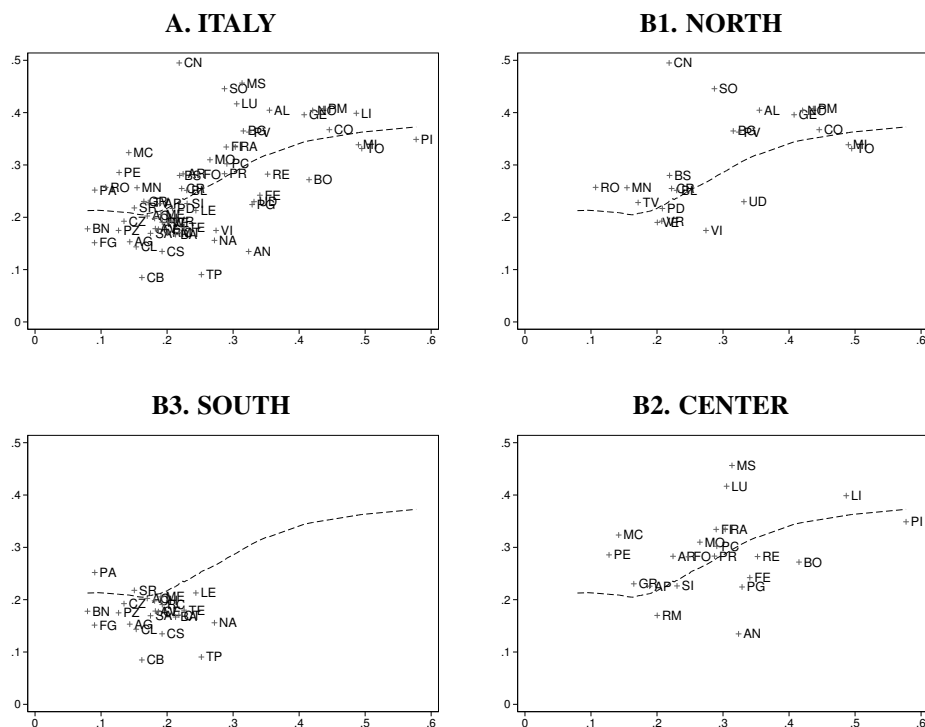
---

<sup>1</sup>Reasons to concentrate our analysis on manufacturing instead of total industry are given in the Appendix.

<sup>2</sup>For the sake of comparison, this is the level of *départments* in France, *kreise* in Germany, and *districts* in England.

<sup>3</sup>A detailed map of 19th century Italy's provinces is reported in the Appendix. For the sake of completeness, the 69 provinces are there also grouped into the corresponding 16 regions of the time.

**Figure 1.** Moran scatterplot. Per capita value added in manufacturing: neighbors growth (y-axis) vs growth (x-axis) by macro-areas, 1871-1911.



Source: see text.

and in particular the branch on endogenous growth models, points to human capital as a one of the key engine of growth. (For a review see, *e.g.*, Kruger and Lindahl 2001). According to Aghion and Howitt (1998) the relation between human capital and growth can be essentially evaluated from two different perspectives. The first one (see, *e.g.*, Lucas 1988) considers the accumulation of human capital over time as a determinant of sustained growth; the second one (see, *e.g.*, Romer 1990) considers instead human capital as a stock generating innovations, contributing thus to a country's ability (especially so in case of latecomers) to imitate and adapt new technology. Essentially the first approach extend the concept of capital to embrace human capital, while in the second the link between human capital and sustained growth pass through technological progress.

A related branch of the literature, relatively less developed, relates economic growth to social capital, perhaps a more slippery theme. According to Temple (2001, p. 58), much in line with Putnam (1993, p. 167), "social capital can be thought of as capturing such things as the extent of trustworthiness, social norms, and participation in networks and associations." Of much interest here, Temple (2001, p. 89) notices that it is generally difficult to discriminate among alternative approaches using macroeconomic data at the national level and that studies at the local (regional or provincial) level can be more informative. The study by Guiso *et al.* (2005) provides a celebrated example for the case of Italy.

The third branch of the literature that is relevant to our research relates economic growth to infrastructures. Already about half a century ago Eckaus noticed that “the stock of social capital available to region is regarded in the current development literature as on of the most important determinants of the growth potential of a region (Eckaus 1961, p. 288). Decades later the seminal work of Aschauer (1989) provided quantitative evidence on the positive effect of public infrastructures on U.S. total factor productivity. A recent and detailed survey on the topic is given in Romp and de Hann (2005). Turning to the Italian case, study at the sub-national level include, among others, Cannari *et al.* (2002), Picci (1999, 2002), and the more recent Bronzini e Piselli (2009), which consider the relation between R&D, human capital and public infrastructure accounting for regional spillovers in Italy’s regions between 1980 and 2001 is particularly close in spirits with the aim of our research. The study by Ciccarelli and Fenoaltea (2008), of more historical nature, provides time series evidence on social overhead capital in the regions of 19th century Italy, and is also related to our research. The relevance of infrastructures on growth can be also highlighted from a policy point of view. In recent years the World Bank allocated about 20 percent of its loans to finance infrastructures of various countries, an amount greater than loans to education, and health (World Bank 2007).

From a practical point of view, a relevant part of our research effort consisted in finding (and understanding) appropriate historical quantitative sources to reconstruct plausible indicators able to represent adequately theoretical concepts. We reconstructed in particular proxy variables at the provincial level for human capital, social participation, and infrastructure endowment in 1871. In this way we complemented conveniently the Ciccarelli and Fenoaltea (2013) value added dataset. Our empirical strategy can be easily summarized. We started by estimating the long established conditional convergence model, one the cornerstones of the growth and regional economics literature of the last decades (see, *e.g.*, Magrini, 2004). In its standard version this model relates income growth of a group of regions to both initial level of income and a set of conditioning variables. Albeit this model has proved to be powerful, it ignores altogether the spatial relations, a fundamental dimension of regional dynamics. Spatial spillovers are instead explicitly<sup>4</sup> considered in the generalized version introduced by the more recent literature. As a subsequent step, following in particular Rey and Montuori (1999), we estimated accordingly a generalized version of the conditional growth model allowing explicitly for spatial effects.

To summarize, this paper estimates a generalized conditional convergence model enriched to allow for spatial spillovers to study manufacturing valued added growth in Italy’s provinces (NUTS 3 units) from 1871 to 1911. The dataset used has been largely compiled from primary historical sources, and is thus mostly original. The conditioning variables include both material (roads and railways) and non-material capital (measures of educational levels, and of social and political participation). We find that the much of the growth variability in manufacturing at the local level can be attributed to the variability of these conditioning variables. In other terms, consistently with the results by and Rey and Montuori (1999) and Bauer *et al.* (2012), in their celebrated studies on US states during the second half of the 20th century, we find that both knowledge and spatial spillovers matter. Northern provinces, above all selected provinces of the North-West, grew more than those of the South because, other things being equal, were endowed with labor forces with better educational levels, had a stronger social

---

<sup>4</sup>Interestingly, Capello (2009) points out that the concept of spatial spillovers is implicit in many growth theories (*e.g.*, the celebrated potential development theory of Isard, 1954, and Giersch, 1949) but it was largely ignored until the early ’90s, when it was pushed by the developments of spatial econometrics.

participation, and were surrounded by more dynamic neighbors.

The remaining of the paper is organized as follows. Section 2 provides a brief review of the quantitative literature on the early step of economic development in 19th century Italy at the local level. Sections 3.1 and 3.2 describes the main quantitative features of the new dataset, and illustrate in particular the spatial distribution of both manufacturing value added and socio-economic variables used as controls in the empirical exercise. Section 3.3 presents results from the estimation of conditional convergence models accounting for spatial dependence. Our conclusions are summarized in Section 4.

## 2 A review of the literature on early Italian industrial development

The contributions on the origin of the Italian regional imbalances are numerous. We provide here a very brief account, by no means exhaustive, and focuses on two ideal strands of the literature depending on which ideal partition of the country's territory is considered able to better reflect the extent of a regional divide.<sup>5</sup>

The first strand, more consolidated, includes the writing of contemporaries. According to a young Luigi Einaudi, among the founding fathers of the Italian Republic “one cannot deny that about thirty years ago Northern Italy represented the more advanced part of the country” (Einaudi 1959, p. 197).<sup>6</sup> Clough and Livi (1956) and Eckaus (1961) are generally credited to have given the first quantitative accounts of the North-South differential in the aftermath of Italy's unification (1861). Both studies consider a wide set of indicators (including construction of railroads, tax receipts, state expenditures, share of labor force by sector of activity, and number of joint stock companies) and conclude that “a difference in per capita income between North and South of between 15 and 25 percent seems plausible” (Eckaus, 1961, p. 300). Moreover, leading economic historians claimed repeatedly that the origin of the North-South divide, particularly evident in the pre-WWI year, dated well before the country's unification of 1861. The secular tradition of textile (above all silk) exports was taken as an example of the ability Northern regions (above all Lombardy) to compete in international markets (Cafagna, 1998 pp. 307-308). A recent monograph summarizing the state of play in terms of statistical reconstruction on historical economic aggregates for the case of Italy and reports that after half a century of quantitative research on regional economic indicators Eckaus findings about the extent of the regional divide in the aftermath of the unification are essentially confirmed. (Felice, 2014, p. 40)

The second strand of the literature, more recent, annoverate a limited number of contributions. Fenoaltea (2003) on the basis of new regional estimates of industrial value added show that the regional divide that most characterized Italy in the early years after its unification was that between the Western and Eastern part of the country. The author notices that industry was then largely artisanal and largely concentrated

---

<sup>5</sup>For a more detailed survey of the quantitative literature on industrial development of 19th century Italy at the provincial level we refer the reader to Ciccarelli and Proietti (2013), and Ciccarelli and Missiaia (2013). Daniele and Malanima (2011), and the more recent Felice (2014) are instead useful references on the long-term economic development of Italy's regions.

<sup>6</sup>The quote in the main text refers to a piece written by Einaudi during the 1890s. The quoted volume, collecting Einaudi's writings of the period 1893-1902, was only published in 1959. The reader should conveniently keep this warning in mind when, next in this paper, we will refer again to 19th century writings by Einaudi as to Einaudi (1959).

near the capitals (Turin, Florence, Rome, Naples, and Palermo) of pre-unitarian states, which were all distributed along the West of the country. The so-called “industrial triangle” formed by the regions of Piedmont, Liguria, and Lombardy took shape only at the beginning of the 20th century thanks to the development of proper industrial plants tied to the mechanization of the production processes and the related reduction in transport costs. The recent Brunetti *et al.* (2011) providing the very first estimate of regional GDP in 1871 confirm the relevance of the West-East divide highlighted by Fenoaltea (2003) for industry alone.

Turning to the determinants of the industrial growth at the local level, economic historians noticed that the North had environmental advantages, including more fertile lands for agriculture and hydraulic resources for motive power (Cafagna, 1989, pp. 194-202). Furthermore, the presence of better governments and institutions characterizing northern Italy has also been stressed repeatedly.<sup>7</sup> Enough to recall here Einaudi (1959, pp. 197-198), who noticed that for a set of historical reasons northern regions (when compared to southern regions) were characterized by “better Governments, reduced distance from prosperous nations, higher self-confidence, and a geographical location prone to rapid and fruitful economic trades.” A recent contribution (Ciocca, 2007) summarizes the above literature and concludes that in the early 1860s Northern and Southern regions were different ever since along many dimensions, including educational levels, social participation, and respect for the rule of law, all affecting the growth potential<sup>8</sup> of the two macro regions, thus supporting both Einaudi’s and Cafagna’s view. Classic references on this line of research are Banfield (1958) and Putnam (1993) and, for the case of 19th century Italy, A’Hearn (1998), and de Blasio and Nuzzo (2009). Leonardi (1995) stress that in the case of Italian South in recent decades, social norms emphasizing collective action (*i.e.* social capital) as a viable means of achieving societal goods are absent. Rather, the South is permeated by a culture emphasizing individual norms oriented towards short-term individual gains. In a similar fashion, Felice (2012, and 2014) points to social participation and institutions as relevant variables to look at when investigating the long-term development of the North-South divide. Felice (2014, Appendice statistica, pp. 14-15) presents estimates of long-term regional human capital and social capital indices. According to these estimates for the year 1871 the human capital index (Italy = 100) amounted to 53 for southern regions, and to 96 for northern and center regions; similarly, the social capital index (Italy = 100) amounted to 67 for southern regions, and to 117 for northern and center regions. An early quantitative account on the issue is given in Zamagni (1990), p. 40, showing that northern regions scored better than southern ones in terms of a battery of socio-economic indicators (ranging from railways and streets extension, to illiteracy rates, and to enrollment rates in primary and secondary school.) The point that social variables matter is moreover not new. Contemporaries (see *e.g.* Franchetti, 1875, p. 40) noticed that besides economic factors the North and the South were different in terms of “intellectual and moral” terms. Einaudi (1959, p. 197), after noticing that “only a barker pseudo-sociology can distinguish between two races one prone to progress and another destined to savagery”, reports that admittedly Northern regions, representing the core of Italy’s industry, “felt higher incentive to reach an appropriate cultural level”.

<sup>7</sup>Felice 2014 provides a detailed analysis on the matter.

<sup>8</sup>But for a critical view see Fenoaltea (2011), p. 230 who argue that “Italy’s dualism developed after Unification because the evolution of the technology of production, and of organization, increasingly favored location in the North.”

### 3 Modeling early Italian industrial development

From the summary of the literature on regional imbalances of the previous section one learns that while the North-South divide that characterizes Italy in present day is out of dispute, the debate on its origin and causes of regional imbalances is instead still essentially open. This paper contributes to the above debate by considering for the first time the early steps towards Italy's industrialization within a spatial analysis framework. Following Rey and Montouri (1999), our benchmark model is the basic convergence model. The latter will represent a convenient starting point, to be generalized in successive steps to include both control variables (thus allowing for steady states different across provinces) and spatial effects.<sup>9</sup> In order to estimate such a model we thus need manufacturing value added data, available from Ciccarelli and Fenoaltea (2013), and additional control variables as well. The variables used in the conditional convergence model to allow for the possible existence of multiple steady states typically include measures of material and human capital, and in some cases variables meant to capture political and business climate (for a review see Islam, 2003). In a study of Italian regions over 1950-1990, Helliwell and Putnam (1995) emphasized the role of social capital, measured on the basis of a variety of indicators. Obtaining data with disaggregation at NUTS 3 level for 1871 is clearly a challenging task. However, the dataset of controls we managed to construct can be considered reasonably comprehensive when compared to those commonly found in the literature concerning more recent times. The control variables used in this paper include education, social and political participation, and, to account for infrastructure endowment of the various provinces, postal offices, roads and railways extension.<sup>10</sup> While details on sources and methods are given in the Appendix we proceed now to a description of the data, and then to the findings from model estimates.

#### 3.1 The geography of the Italian manufacturing industry in the late XIX century

The wide changes that occurred in the provincial distribution of per capita value added in the manufacturing industry between 1871 and 1911 are clearly summarized by the non-parametric estimates of the density functions reported in Figure 2.<sup>11</sup> The 1871 density resembles clearly that of a mixture of two different distributions, as if in the early step of Italy's industrialization the various provinces belonged to two different populations.<sup>12</sup> In 1911, after the generalized mechanization of production processes and the diffusion of modern industrial plants, the distribution shifted to right and its support considerably widened, so that both location and scale parameters are considerably higher than forty years before.<sup>13</sup>

---

<sup>9</sup> As noticed, we concentrate our analysis on the manufacturing sector alone; the reasons for this choice are discussed in some length in the Appendix.

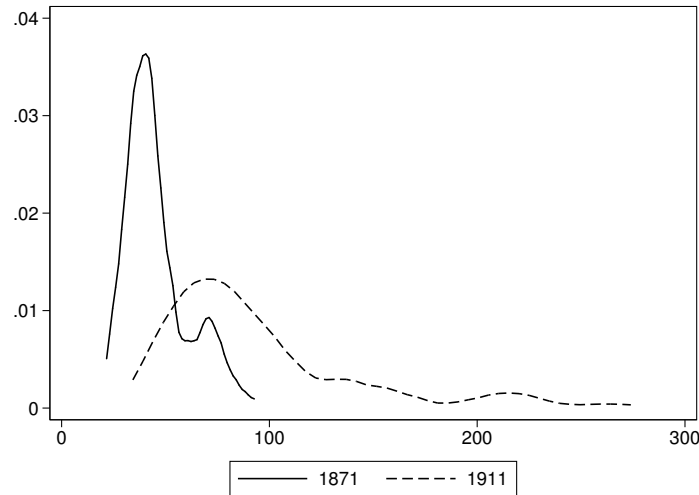
<sup>10</sup> We instead failed in gathering useful quantitative information on important variables, such as the local availability of hydraulic resources for motive power.

<sup>11</sup> Despite the predominant share of agriculture on GDP in 19th century Italy, development patterns are clearly better visible in the more dynamic sectors of the economy. Additional reasons for focusing on manufacturing instead of GDP in the context of convergence analysis are given in Rodrik (2013).

<sup>12</sup> Note that this provides strong support to Cafagna's view of the North-South divide as pre-dating the unification of the country in 1861

<sup>13</sup> As it is rather obvious, many features of Italian industry changed over the 1871-1911 period. To mention a few: i) the internal mix of industrial sectors changed: with traditional sectors tied to the production of consumption goods (*e.g.* foodstuffs) reducing their share in favor of the fast-growing "high-tech" sector tied to the production of durable goods (*e.g.* engineering); ii) coal power was gradually replaced by hydro-

**Figure 2.** Per capita value added in manufacturing industry, 1871 and 1911 (lire at 1911 prices) <sup>a</sup>



<sup>a</sup>Non parametric kernel density estimate with Silverman (1986) plug-in smoothing parameter.  
Source: see text.

In fact, from Table 1 we can see that the ratio between maximum values of per capita manufacturing value added in 1911 and in 1871 is about 1.8 (Table 1, col. 1), approximately two for corresponding average values (Table 1, col. 2), and nearly three for corresponding minimum values (Table 1, col. 4). Both dispersion and concentration, as measured respectively by the standard deviation (normalized by mean values), and by the Gini coefficient, also increased considerably (about 1.5 times, Table 1, cols. 5 and 6).

**Table 1.** Manufacturing industry: per capita value added, descriptive statistics.<sup>a</sup>

	(1)	(2)	(3)	(4)	(5)	(6)
	<i>min</i>	<i>median</i>	<i>mean</i>	<i>max</i>	$\sigma/\text{mean}$	<i>Gini</i>
1871	25.78	41.15	45.67	88.59	0.32	0.17
1911	46.03	81.12	94.75	262.20	0.50	0.25
1911/1871	1.8	2.0	2.1	3.0	1.6	1.4

<sup>a</sup> Lire at 1911 prices. The 1911/1871 row at the bottom of the table reports the (rounded) ratio of figures appearing in the preceding rows (so that for instance  $1.8 = 46.03/25.78$ ).  
Source: see text.

Figure 3 presents things from a spatial perspective.<sup>14</sup> The left hand side refers to 1871 and illustrates the geographical distribution of  $y_j^{71}$ , representing here the “initial value” of per capita value added in manufacturing. A North-South divide is clearly

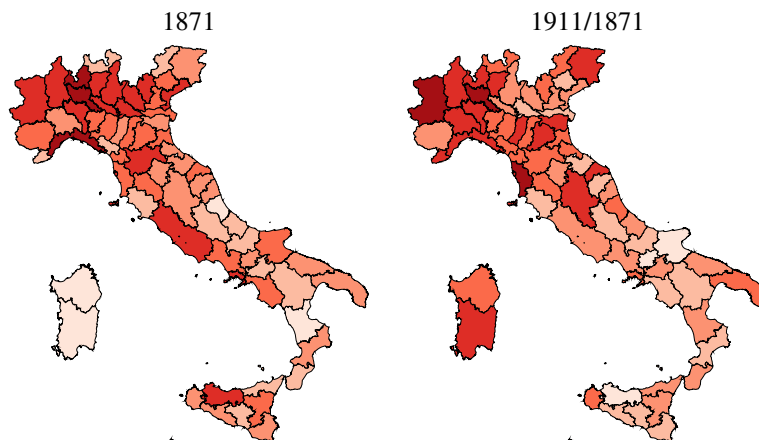
electric power; iii) Italy’s international trade policy shifted from free-trade to protectionism, with uneven net protection guaranteed to the various industrial sectors, possibly affecting total industrial output.

<sup>14</sup>A detailed map of 19th century Italy’s provinces is reported, as noticed, in the Appendix.



evident, with northern provinces typically belonging to the upper part of the distribution (see again Figure 2 from this perspective). With the exception of Naples and

**Figure 3.** Per capita value added in manufacturing (lire at 1911 prices)<sup>a</sup>



<sup>a</sup> The maps use as a class breaks the percentiles 5, 25, 50, 75, 95 of the distributions of manufacturing value added (map on the left hand side) and of end-to-end growth rate (map in the right hand side).

Source: see text.

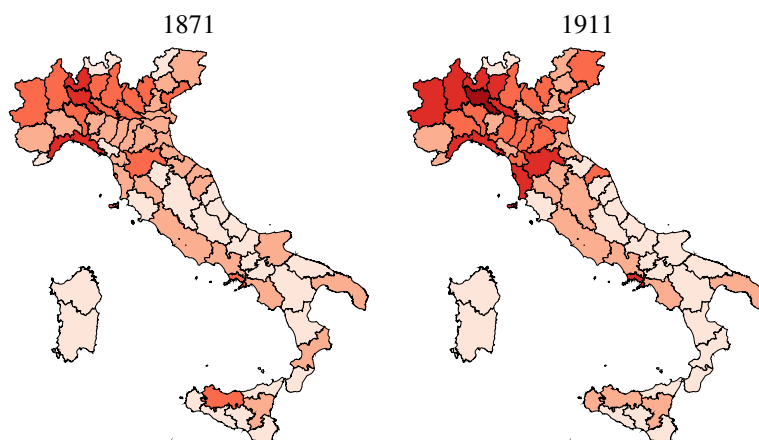
Palermo (the capitals of the former Kingdom of two Sicilies) the only provinces above the third quartile of the distribution (about 50 lire per capita in 1871, corresponding to the province of Rome) belong to the North. Within the advanced North only Genoa (in Liguria) and Milan (in Lombardy) are above the (arbitrary) threshold of 80 lire and represent clear outliers. But there is more than that. The considerable amount of heterogeneity within both North and South reveals that the simplistic partition of the country between North and South, so often adopted in the historical literature on the so called *Mezzogiorno* problem, seems inappropriate. To exemplify, within Piedmont (in the very North-West part of the map) northern provinces (Turin and Novara) score better of southern provinces (Alessandria and Cuneo). Within Liguria (immediately below Piedmont in the map), the provinces of Genova – with important industrial plants such as the Ansaldo factory – and Porto Maurizio, at the border with France on the seaside, differ considerably. Similarly, in the North-East the Alpine provinces of Venetia appear relatively backward when compared to Vicenza, center of important textile industries, and Venice, with its celebrated Arsenal representing a large-scale industrial enterprise all along. The case of Lombardy (located in the map between the regions of Piedmont and Venetia) presents instead a more homogeneous geographical distribution of industrial value added, with six out of eight provinces above median values. Differences within Southern regions are also present. Here suffice to consider the case of Sicily with its seven provinces grouped in the map in five different intervals of per capita manufacturing value added. The 1871 map also confirms the existence of an early West-East gradient related to the artisanal nature of early industry, closely tied to marked proximity and thus located near the capitals of preunitarian states (Turin,

Florence, Naples, Rome, and Palermo), all in the Western part of the country.<sup>15</sup> To summarize, a North-South industrial divide in the aftermath of the country's unification is clearly present but, a point worth repeating, with a considerable amount of within regions heterogeneity.

The map on the right hand side of Figure 3 gives a geographical representation of the dependent variable of our convergence regression models  $\Delta y_j^{71-11} = (y_j^{11} - y_j^{71})/4$  (the average growth for a decade), where  $y$  is log value added per capita. It shows that in the four decades between 1871 and 1911 the growth of per capita manufacturing value added has been considerably unequal across provinces, with provinces of the North-West rising their shares of the total. The considerable industrial growth occurred in Italy in the four decades here considered is, with a few exceptions, driven by north/western provinces. Within the "industrial triangle" formed by the regions of Piedmont, Lombardy, and Liguria, the provinces of Turin and Milan score particularly well.

Finally, Figure 4 reports the indices of value added obtained by using the minimum values in 1871 and 1911 as reference level. The normalization of each provincial value added by the minimum value of the distribution (as a matter of fact, given by the province of Sassari, in northern Sardinia, both in 1871 and 1911) highlights more clearly that in 1871, and even more in 1911, the North, and above all the provinces of North-West, stands out from the mass. In this metric, industry in 1871 appears largely

**Figure 4.** Per capita value added in manufacturing (lire at 1911 prices): indices (min = 100)<sup>b</sup>



<sup>b</sup> The maps report indices of manufacturing value added (minimum provincial value in 1871 and 1911 equal to 100); the six intervals are defined by the values 100, 150, 200, 300, 500, and 600.

Source: *see text*.

concentrated in subalpine provinces and Genova, less so in northern provinces of the Po valley.<sup>16</sup> With the exception of the province of Naples, an end-to-end darkening of

<sup>15</sup>This interpretation of the early location of industry in Italy's regions is proposed in Fenoaltea (2003).

<sup>16</sup>Essentially the part of northern Italy going from the province of Cuneo, at the western border with France, to the provinces of Ferrara, Ravenna and Rovigo on the east side of the country, along the Adriatic coast

the map is registered only by (the majority of) provinces above the 43rd parallel north. With the marked diffusion of industry in Emilia (in the Po valley) within North heterogeneity reduces over time. Below 43rd parallel north the map present mainly light colors; a cluster of provinces along the adriatic coast down to Calabria, the toe of the Italian boot, form the less industrialized part of the country.

Overall, the value added maps suggest rather clearly that considerable industrial divergence took place between the unification of the country and the eve of WWI. Clearly, the maps suggest a variety of additional research questions that go well beyond the analysis of economic disparities between macro-areas, as non negligible growth differentials are visible also within the group of the Northern and Southern provinces.

### 3.2 Control variables

With the aim of capturing dimensions such as social participation, human and social capital, and infrastructure endowment we collected data from historical sources ranging from population censuses, to official railways and electoral publications, to individual contributions of various scholars of the time. As a result, the present piece represents, to the best of our knowledge, the first attempt ever to investigate the main determinants of local (NUTS 3) industrial growth in 19th century Italy in a well defined quantitative setting. In order to increase the efficiency of our estimation procedure in some cases we resort to principal component analysis (PCA), a common way to reduce the dimensionality of the problem (for instance, it is used in a very similar set-up by Tabellini, 2010).

The rest of this section illustrates the geographical distribution of the socio-economic variables as reported in Figure 5. We consider entirely new synthetic indices of human capital, social capital, and social overhead capital. In the lack of similar evidence at the provincial level the reference to an existing literature appear inevitably problematic. Details on sources and methods are reported in the Appendix.

**HUMAN CAPITAL.** The human capital control used in the regression models is the first principal component emerging from PCA on the logs of (standardised) illiteracy rates, age heaping, number of pupils and number of teachers in primary school. “Age heaping” (A’Hearn *et al.*, 2013) is defined as an estimate of the excess frequencies of population reporting age in round numbers. This phenomenon is obviously in an inverse relationship with the numeracy of the population. Hence, age heaping and illiteracy rates measure the education level of the population, while number of pupils and teachers measure the size of the education sector. Figure 5, panel 1, show that a North-South divide in term of human capital is clearly evident. Within Northern Italy, regions in the West (Piedmont, Lombardy and Liguria) score particularly well, especially so in the case of selected Alpine provinces of Lombardy and Venetia (both part of the Habsburgh empire till 1866)- At the opposite side of the distribution, provinces of southern Sicily performs instead poorly.<sup>17</sup> One cannot avoid to notice, with inevitably exceptions, the close similarity with the geographical pattern presented in Figure 3 concerning the distribution of manufacturing value added across Italy’s provinces. The unequal geographical distribution of human capital in 19th century Italy was moreover well known to contemporaries at the point that Einaudi (1959), p. 198, in his positive evaluation of the State education policy carried over by Italian policy makers of the time noticed that “it was more useful to spend money in the making of a secondary

---

<sup>17</sup>For reasons of space, provincial maps on elementary variables used to build the human capital index (illiteracy rates, age heaping, number of students, and number of teachers in primary school in 1871) are not reported. They are available upon request.

education system in Northern Italy [than in Southern Italy] to avoid wasting the fruits from primary education there well established” and that “in Southern Italy [...] given the lack of industry [...] the educated young would most likely languish either in bureaucracy or find a way out in some liberal profession [had the State pursue there the making of a secondary education system].”<sup>18</sup>

**SOCIAL CAPITAL.** The social capital controls can be divided in two groups: (i) social participation; (ii) political participation. The first group includes the number of published newspapers and magazines (used also by Helliwell and Putnam, 1995), and membership of mutual societies (both suitably standardized by the population size); the second group includes the number of voters (per 100 registered voters) in the local and national elections of mid 1860s and 1870. PCA suggests to keep these two groups separated, so that we constructed a social participation variable as the simple average of the number of newspapers and magazines titles and membership of mutual societies, and a political participation variable as the simple average of the number of voters (per 100 registered voters) in the local and national elections of mid 1860s and 1870. To better frame the role of social capital in 19th century Italy it seems important to recall some basic facts on the matter. First, the age of mass-press was only moving its first steps and the shift from elite to mass culture was far to come. Our sample includes nevertheless more than a thousand titles including newspaper, magazines, and other publications. Second, in Italy as elsewhere<sup>19</sup> only adult males were enfranchised to vote. Of these adult males, mostly were landowners or belonging to some other well-off category. As a result, only some two percent of the population (about 26 millions in 1870) was involved in elections. One could further argue, legitimately, that in the lack of universal suffrage electoral participation correlates weakly with social participation. As a matter of fact, the participation rate in our dataset is about equal to .4 in the local elections and of about .5 in national elections, both held in 1865 ca (see Appendix for extra details). Third, welfare state was then very limited; and in the lack of public programs, mutual societies constituted effective tools to provide to its members insurance against various individual risks (such as those related to unemployment, health conditions, and retirement). Our sample records about 238,000 members of mutual societies distributed in the various provinces, and ranging from the no-member-at-all case of Campobasso (Abruzzi) to the about 17,000 members in the case of Milan.

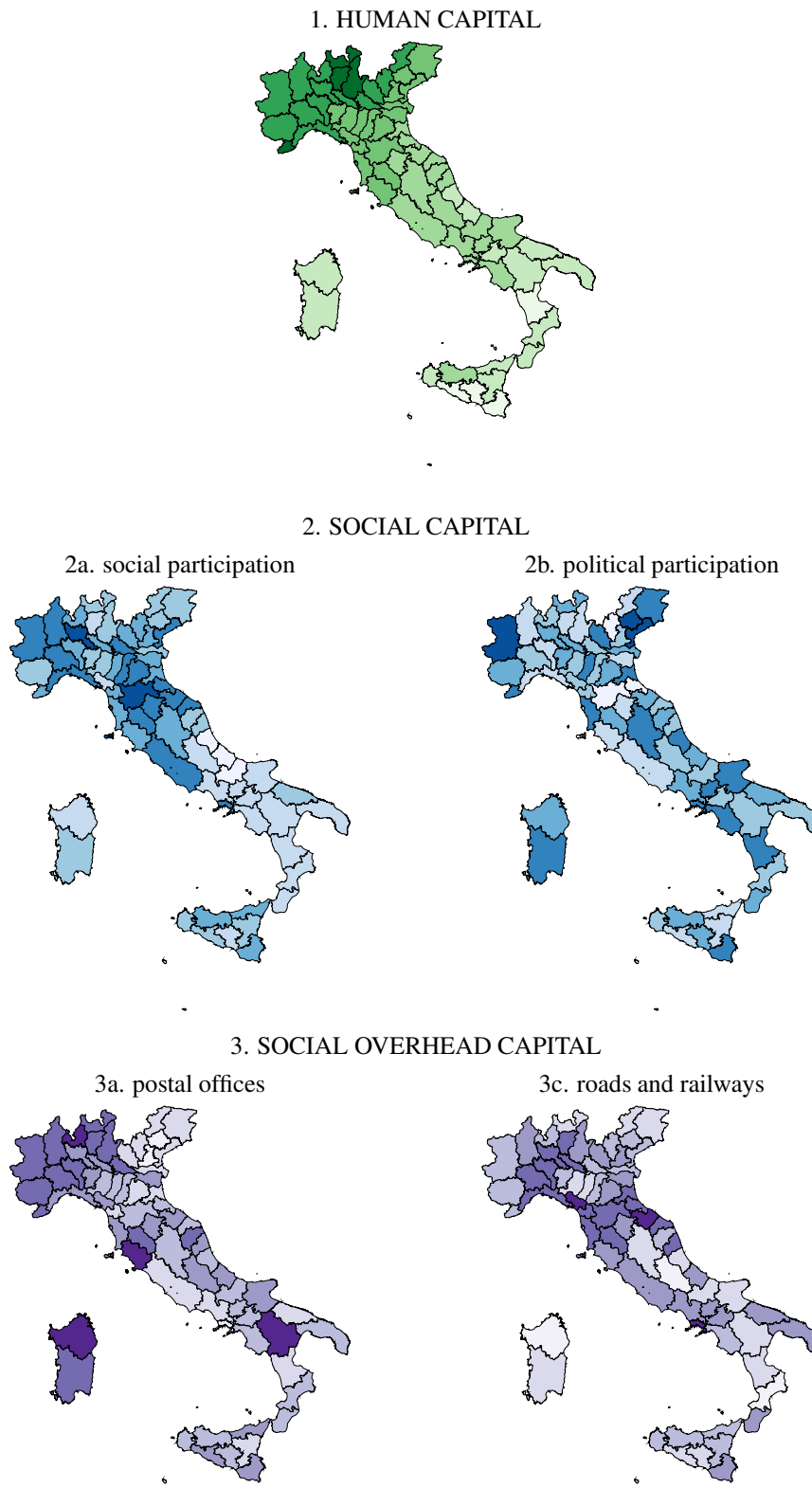
**SOCIAL OVERHEAD CAPITAL.** Infrastructure capital includes local and national roads, railroad extension (all measured in linear km's per square km of provinces surface, as in Capello, 2009), and the number of post offices (standardized by the population size). Following PCA results, the proposed regressions include the simple average of (standardized) roads and railways, and, as a separated variable, the number of post offices per 100.000 habitants. A detailed analysis of social overhead capital in Italy's regions from 1861 to 1913, including reference to historical sources, is given in Ciccarelli and Fenoaltea (2009).

---

<sup>18</sup>The finding that the Sondrio and Porto Maurizio, far from representing the prototype of modern economies, performed so well according to the human capital index here proposed was admittedly a surprise for the present writers. In the lack of better explanation, after judiciously double checking the sources and methods behind the result, the above finding appears classifiable within the “data oddities” category.

<sup>19</sup>See, for instance, Goldstein (1983), p. 4, reporting the percentages of population enfranchised for lower legislative chambers in various European countries between 1815 and 1915.

**Figure 5.** Control variables used in regression models (year = 1871 ca.) <sup>a</sup>



<sup>a</sup> The maps use as a class breaks the percentiles 5, 25, 50, 75, 95 of the relevant distribution.  
*Source:* see text.

It is here important to stress that a consistent part of the Italian railroad network was built after 1871 (the extension of the Italian network was of roughly of 1,200 in 1855, 2,400 in 1860, 6,400 in 1870, 9,100 in 1880, 12,200 km in 1890, 14,400 in 1900, and 15,300 in 1910, about at the end point of our 1871-1911 reference period (Istat, 1958, p. 137).<sup>20</sup>

The previous maps were hopefully useful to illustrate patterns of spatial clustering in the data. Table 2, more rigorously, reports Moran's I spatial autocorrelation coefficient applied to the same set of variables. It emerges that the rules is constituted by

**Table 2.** Spatial autocorrelation: Moran's I.<sup>a</sup>

	$\Delta \mathbf{y}^{71-11}$	$\mathbf{y}^{71}$	Education	Social partic.	Roads & rails	Post offices	Political particip.
Moran's I	0.49	0.21	0.84	0.28	0.41	0.57	-0.26
p-value	0.00	0.02	0.00	0.00	0.00	0.00	0.98

<sup>a</sup>p-values per Ho: no autocorrelation (1-tail test). The spatial weight matrix used in the computation is defined in section 3.3.

Source: see text.

(positive) spatial autocorrelation, especially so for industrial growth, post offices, and education. The exception is instead represented by the political participation variable. For the latter we failed to reject the null of no spatial autocorrelation.

### 3.3 Model estimation

The starting step of our study is the simplest growth model<sup>21</sup> with the initial level of per capita industrial value added at 1911 prices as the only explanatory variable. Taking as dependent variable the average growth for a decade,  $\Delta y_j^{71-11} = (y_j^{11} - y_j^{71})/4$ , where  $y$  is log value added per capita<sup>22</sup>, denoting in boldcase the  $N \times 1$  vectors of the various variables ( $N = 69$ ), and letting  $\varepsilon = [\varepsilon_1 \dots \varepsilon_N]'$ :

$$\Delta \mathbf{y}^{71-11} = \phi \mathbf{y}^{71} + \varepsilon \quad (1)$$

Estimates are reported in Table 3, panel A, col. 1. In view of the descriptive analysis of the previous section the positive sign of the coefficient is of course not

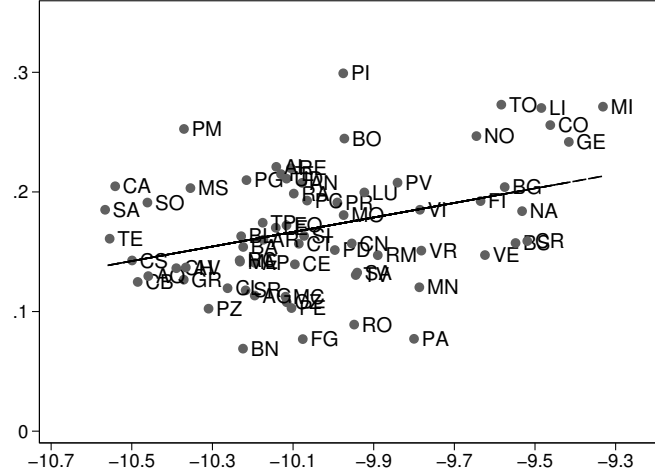
<sup>20</sup>Beside roads and railroads, ports were also a relevant component of the national endowment of infrastructures given the importance of costal shipping, and more generally water transports. Quantitative information on the matter do exist (as for instance those in the annual publication *Statistica del Regno d'Italia. Navigazione nei porti del regno, 1872*. We found in particular data on the numbers of vessels (about 17,000 units) and their gross tonnage. The source is very detailed; we know for instance that about 1 percent of vessels in 1871 was constituted by steamships, most of which metal-hulled. However, the data are reported in the source at the level of *compartimento marittimo* (sea district) and not at the provincial level, and were thus not used here. Data on the number and relevance of major ports (Genoa, Messina, Leghorn, Palermo, Naples, Venice, Catania, Ancona, Brindisi, and Cagliari) in 1871 are also available, but were not used to avoid using in the empirical application a heavily skewed to right variable, given that major ports were simply absent in the vast majority of the 69 provinces here considered.

<sup>21</sup>Since we know that divergence forces have in action we prefer to refer to model (1) as "growth model," rather than as " $\beta$ -convergence model," as standard in the literature.

<sup>22</sup>Two remarks are in order here: first, population acts purely as a scaling term capturing the size of the provinces; second, in the rest of the paper we will often refer to this variable simply as value added.

surprising: we have seen that provinces with higher starting income grew more than those with a lower one. Figure 6 illustrates this, and further highlights the advantages of considering industrial growth at the provincial level. Within Lombardy only Milan and Como are well above the national average in both the growth and initial values dimensions. Similarly, only selected provinces of Piedmont – Turin (TO) and Novara (NO) – and of Liguria (Genoa, GE) score particularly well. A provincial “industrial triangle” thus (Genoa, Turin, and Milan) more than a regional one (Liguria, Piedmont, and Lombardy) often considered in the literature.

**Figure 6.** Per capita value added in manufacturing industry:  $\Delta y_j^{71-11}$  (y-axis) vs  $y_j^{71}$  (x-axis).



Source: see text.

However, we should not forget that the above simple model ignores spatial spillovers altogether. Before running any inference on the estimates we thus need to check carefully for the presence of spatial autocorrelation in the residuals. To this end we briefly recall the key points of spatial modelling, so to establish notation. First of all, let  $\mathbf{W}$  be a  $N \times N$  spatial weights matrix which “expresses the strength of potential interaction between each observation and its neighbors” (Anselin and Rey, 1991, p. 117); the simplest example is the binary contiguity matrix with elements  $w_{ij} = 1$  if  $i$  and  $j$  have a common border, and 0 else. Then a first way to introduce the spatial dimension is allowing for spatial dependence in the errors of (1), obtaining the spatial error model (SEM, see, e.g., Rey and Montuori, 1999, eq. 6-8):

$$\Delta \mathbf{y}^{71-11} = \phi \mathbf{y}^{71} + \varepsilon \quad (2a)$$

$$\varepsilon = \rho \mathbf{W} \varepsilon + \eta \quad (2b)$$

where  $\rho$  the spatial autoregressive coefficient and  $\eta$  a random noise  $N(\mathbf{0}, \sigma^2 \mathbf{I})$ . Substituting from (2b) into (2a) we obtain

$$\Delta \mathbf{y}^{71-11} = \phi \mathbf{y}^{71} + (\mathbf{I} - \rho \mathbf{W})^{-1} \eta \quad (3)$$

From equation (3) it is clear that the spatial error model emphasizes shock propagation, as a shock to a given province is assumed to spill over to the province's neighbors according to the weights given by the elements of the matrix  $(I - \rho\mathbf{W})^{-1}$ .

Alternatively, we can augment model (1) with a spatially lagged dependent variable, obtaining the spatial autoregressive model (SAR, see, *e.g.*, Rey and Montuori, 1999, eq. 9):

$$\Delta\mathbf{y}^{71-11} = \rho\mathbf{W}\Delta\mathbf{y}^{71-11} + \phi\mathbf{y}^{71} + \varepsilon \quad (4)$$

in which the question of if, and how, growth in one provinces has been influenced by growth of its neighbors can be directly tackled.

We can test for the presence of spatial autocorrelation in the regression residuals with three different statistics: Moran's  $I$ , which does not assume a specific form of spatial dependence,  $LM^{SEM}$  against the alternative of spatial error dependence, and  $LM^{SAR}$  against the alternative of spatial lag model (see, *e.g.*, Anselin, 2008, Anselin and Rey, 1991). The spatial weights matrix used has been constructed assuming all provinces with a distance smaller than a threshold  $c$  to be neighbors (hence,  $w_{ij} = 1$  if  $d(i, j) < c$ , 0 else). The threshold  $c$  has been fixed *ex-post* at the value maximizing spatial interaction, i.e. the spatial autocorrelation statistics. This value turned out to be 100 km, a plausible distance. The two provinces of Sardinia, which obviously have as a unique neighbor the other province of the island, have been dropped from the sample. Finally, the distances are measured from the administrative centers, an acceptable simplification in view of the small dimension of the provinces. As to be expected, two of these statistics suggest strong residuals spatial autocorrelation (Table 3, panel B, col. 1), so that we definitely need to account in some way for spatial effects. The simplest possible way is to include some dummies for the constant, *i.e.* to assume the existence of groups of provinces, or macro-areas, with different steady states. From Table 2, panel A, col. 2 we can appreciate that allowing for three groups, North, Center, and South, has the effect of halving the divergence effect of 1871 value added. The Center dummy is not significant while that of the South is negative and significant, so that Central and Northern provinces appear to share the same steady state value added per capita while in the South this seems to be definitely lower. The spatial autocorrelation diagnostics are now marginally not significant (Table 3, panel B, col. 2), so that this rather naive model might be considered a reasonably good description of regional industrial growth in 19th century Italy.

However, both the clustering in different steady states groups and spatial spillovers deserve to be explicitly modeled. The starting level of value added clearly acts as a proxy for the many dimensions of productive potential, while spatial interaction occurs at a much finer level than those three broad areas. Before turning to conditional convergence models we consider the effect of allowing for spatial interaction through SEM and SAR models. The estimates are reported in Table 3, panel A, cols. 3-4. The coefficients of the initial level of manufacturing value added ( $y_j^{71}$ ) are still positive and significant, pointing to divergence. Turning to spillover effects, the estimated coefficient (Table 3, panel A, col. 4), is about 0.4, suggesting that a ten percent increase in the industrial growth of the neighboring provinces implies approximately a 4 percent increase of provincial growth.

In the next step we temporarily omit the spatial dimension, estimating conditional growth equations which relate per capita value added growth over 1871-1911 to initial level of per capita log value added ( $y_j^{71}$ ) and the controls discussed in the previous section: Education ( $Edu_j^{71}$ ), Social participation ( $Soc_j^{71}$ ), Political participation ( $Pol_j^{71}$ ), Post offices ( $Post_j^{71}$ ), Roads and railways ( $RR_j^{71}$ ):



$$\Delta y_j^{71-11} = \alpha + \phi y_j^{71} + \beta_1 Edu_j^{71} + \beta_2 Soc_j^{71} + \beta_3 Pol_j^{71} + \beta_4 Post_j^{71} + \beta_5 RR_j^{71} + \varepsilon_j \quad (5)$$

where  $j = 1, \dots, 69$ . This model will then be generalized to allow for spatially autocorrelated errors or a spatially lagged dependent variable.

Model selection on the basis of significance of the individual coefficients and minimization of AIC suggests for model (5) the restricted specification reported in Table 4, Panel A, col 1. The retained variables are *Edu* (educational system) and *Soc* (social participation), with *RR* (roads and railways) only very marginally significant. Thus, non-material (human and social) capital seems to matter more than communication infrastructures. This may appear puzzling, but we should not forget that the vast majority of the Italian railways were built after 1870 and are thus not captured by our *RR* variable. Remarkably, all diagnostics are largely non significant (Table 4, Panel B, col. 1). Since the spatial autocorrelation tests rely on asymptotic distributions it is nevertheless instructive to estimate the two specifications explicitly allowing for spatial effects. Maximum Likelihood estimation of the error and spatial models with starting values given by either OLS estimates or random numbers generated from a uniform distribution, yields very similar results (Table 4, Panel A, cols. 2-3). Essentially, the coefficients *Edu* and *Soc* are very similar to those obtained earlier, and the marginally significant spatial effects replace the *RR* variable (which in fact had a  $p$ -value of 0.16 in the conditional growth model). AIC is minimized by the simpler conditional convergence model with no spatial effects, but the explanatory power of both models with spatial effects is obviously much higher than that of the base model.

We can now draw some reasonably robust conclusions. First, quoting Bauer *et al.* (2012), “knowledge matters”: the variable *Edu*, which synthesizes size and quality of the educational system is always strongly significant and has sizeable effects on valued added growth. The gap in value added growth explained by the *Edu* differential between the two provinces with the largest and smallest values (respectively, Bergamo, in Lombardy, and Siracusa, in Sicily) can be estimated as approximately 30 per cent over the entire 1871-1991 period<sup>23</sup>. Of course, the finding that literacy has significant effects on growth is consistent with many other studies. Di Liberto (2008) shows that improvements in literacy rates had a strong impact on the growth of Southern Italian regions in the 1960s, when the income differentials with the rest of the country temporarily shrank (on this, see Paci and Saba, 2008). On a much larger scale, Gennaioli *et al.* (2013) show that education is the single most determinant of the current world regional development differentials, and Tabellini (2010) that the literacy differentials can have highly persistent effects on income differentials, up to the point regional per capita output in the European regions at the end of the 20th century is significantly related to literacy rates of over a century before, which shape culture and institutions in the long-run.

Second, active citizenship also matters: press diffusion and social awareness, as measured by the diffusion of mutual societies, have a significant effect on growth. The size of the effect is smaller than that of the *Edu* variable, but still sizeable: the growth differential between top (Leghorn in Tuscany) and bottom (Campobasso in Abruzzi) provinces explained by *Soc* is close to 20 per cent over 1871-1911<sup>24</sup>. The relevance of this variable is consistent with results in Helliwell and Putnam’s (1995) study of income

<sup>23</sup>More precisely, it is 7.5 per cent per decade on the basis of the estimate of the coefficient  $\beta_1$  from the SEM model and 6.8 per cent per decade on the basis of the estimate from the SAR model

<sup>24</sup>About 5 and 3.5 per cent per decade on the basis respectively of the SEM and SAR coefficient estimates.

convergence in the Italian regions between 1950 and 1990, and again with Tabellini (2010). Both these papers emphasize the positive effects on growth of the cultural aspects often labeled “social capital” which our two variables attempt to measure. Third, the initial endowment of roads and railways did not seem to have a significant influence on development. As anticipated above, most of the Italian railways were built after 1871, so this is actually not surprising. Fourth, consistently with what found by Rey and Montuori (1999) for the US states in the 20th century, spatial spillovers matter. Here the explained 1871-1911 growth differential between top (Cuneo in Piedmont) and bottom (Trapani in Sicily) is over 15 per cent<sup>25</sup>.

Summing up, the polarization of the distribution, with inequality growing strongly between 1871 and 1911, seems to have been a consequence of the fact that most of the provinces of the North had not only higher starting values of value added, but also superior human capital endowments, a more cooperative culture, and were on average more exposed to positive spatial spillovers. This last point raises an interesting question: are these spillover effects possibly non linear? A first answer is provided by the so-called Moran scatterplot, a plot of the average growth of the neighbors,  $\mathbf{W}\Delta\mathbf{y}^{71-11}$ , versus  $\Delta\mathbf{y}^{71-11}$  which is reported in Figure 1 along with a non parametric loess estimate of the regression function. From the scatterplot and the estimate of the regression function it is evident that for low values of income growth (approximately lower than 15 percent for decade) there is no essentially no relationship between growth in one province and that of its neighbors, while for higher values there is almost a one-to-one relationship. Not surprisingly, most of the provinces of the Center-North (more precisely, 37 out of 44) are in the higher growth group; it is thus interesting to estimate the spatial models restricting the sample to these provinces only. From columns 4 and 5 of Table 4, Panel A, we can see that, as expected, the estimates of the spatial effects are larger and, above all, much more significant. To conclude, we note that the spatial error model seems to suffer from heteroskedasticity and, surprisingly, spatial autocorrelation of the residuals; however, it delivers estimates very close to those of the spatial lag model, whose diagnostics do not signal any problems.

---

<sup>25</sup> About 4 per cent per decade on the basis of the coefficient estimate from the SAR model.

**Table 3.** Basic growth models, 1871-1911. Dependent variable:  $\Delta y_j^{71-11} = (y_j^{11} - y_j^{71})/4$ .

A: regression results <sup>a</sup>

	(1)	(2)	(3)	(4)
	base model	base model + regio dummies	SEM	SAR
$y_j^{71}$	0.071	0.035	0.048	0.051
	0.002	0.020	0.019	0.018
	[<0.01]	[0.09]	[0.01]	[0.01]
center = 1	n.a.	-0.003	n.a.	n.a.
		0.014		
		[0.85]		
south = 1	n.a.	-0.050	n.a.	n.a.
		0.014		
		[<0.01]		
constant	0.882	0.543	0.611	0.615
$\mathbf{W}\varepsilon$	n.a.	n.a.	0.207	n.a.
			0.060	
			[0.01]	
$\mathbf{W}\Delta \mathbf{y}^{71-11}$	n.a.	n.a.	n.a.	0.381
				0.098
				[<0.01]

B: goodness of fit and diagnostic tests <sup>b</sup>

	(1)	(2)	(3)	(4)
	base model	base model + regio dummies	SEM	SAR
<i>goodness of fit:</i>				
$R^2$	0.16	0.33	0.94	0.95
$LogL$	109.74	117.12	115.36	115.99
$AIC$	-215.48	-226.24	-222.72	-223.98
<i>diagnostic tests:</i>				
<i>White</i>	1.42	6.48	4.27	4.28
	[0.49]	[0.37]	[0.64]	[0.64]
$I$	2.78	1.04	1.12	-0.52
	[<0.01]	[0.15]	[0.13]	[0.70]
$LM^{SAR}$	0.70	.62	9.81	2.06
	[0.40]	[0.43]	[<0.01]	[0.15]
$LM^{SEM}$	7.84	2.39	1.49	0.81
	[<.01]	[0.12]	[0.22]	[0.37]

<sup>a</sup> The variable  $y$  represents per capita value added in the manufacturing industry at 1911 prices. All variables in logs. Figures underneath the estimates represent asymptotic standard errors. Figure in brackets represent p-values of t-tests of significance.

<sup>b</sup> Figure in brackets represent p-values of t-tests of significance. *White* denotes White's heteroskedasticity test.  $I$ ,  $LM^{SAR}$ ,  $LM^{SEM}$  are test statistics of spatial autocorrelation tests, and  $H_0$  : no residual spatial autocorrelation;  $I$  :  $H_1$  : residual spatial autocorrelation;  $LM^{SAR}$  :  $H_1$  : spatial lag model;  $LM^{SEM}$  :  $H_1$  : spatial error model.

Source: see text.

**Table 4.** Conditional growth models, 1871-1911. Dependent variable:  $\Delta y_j^{71-11} = (y_j^{11} - y_j^{71})/4$ .

A: regression results <sup>a</sup>

	(1)	(2)	(3)	(4)	(5)
	Italy			Centre-North	
	<i>base</i>	<i>SEM</i>	<i>SAR</i>	<i>SEM</i>	<i>SAR</i>
$y_j^{71}$	-	-	-	-	-
$Edu_j^{71}$	0.013	.011	0.010	0.013	0.012
	0.003	0.004	0.003	0.006	0.006
	[<0.01]	[<0.01]	[<0.01]	[0.03]	[0.03]
$Soc_j^{71}$	0.012	0.015	0.010	0.023	0.020
	0.007	0.007	0.003	0.011	0.011
	[0.11]	[0.03]	[<0.01]	[0.03]	[0.05]
$Pol_j^{71}$	-	-	-		
$Post_j^{71}$	-	-	-		
$RR_j^{71}$	0.016				
	0.011	-	-		
	[0.16]	-	-		
<i>constant</i>	0.167	0.155	0.139	0.140	0.132
$\mathbf{W}\varepsilon$	<i>n.a.</i>	0.010	<i>n.a.</i>	0.170	
		0.066		0.077	
		[0.13]		[0.03]	
$\mathbf{W}\Delta\mathbf{y}^{71-11}$	<i>n.a.</i>	<i>n.a.</i>	0.193		0.221
			0.116)		0.116
			[0.10]		[0.06]

B: goodness of fit and diagnostic tests <sup>b</sup>

	(1)	(2)	(3)	(4)	(5)
	Italy			Centre-North	
	<i>base</i>	<i>SEM</i>	<i>SAR</i>	<i>SEM</i>	<i>SAR</i>
<i>goodness of fit:</i>					
$R^2$	0.42	0.95	0.95	0.95	0.96
$LogL$	122.05	121.00	121.19	78.67	77.45
$AIC$	-236.10	-232.00	-232.38	-139.35	-144.88
<i>diagnostic tests:</i>					
<i>White</i>	5.16	5.23	4.75	17.37	3.33
	[0.52]	[0.51]	[0.58]	[0.01]	[0.77]
<i>I</i>	0.63	0.45	-0.04	1.35	0.94
	[0.26]	[0.33]	[0.52]	[0.09]	[0.17]
$LM^{SAR}$	0.24	1.89	0.43	22.31	0.23
	[0.63]	[0.17]	[0.51]	[0.01]	[0.63]
$LM^{SEM}$	0.51	0.18	0.17	1.59	0.60
	[0.48]	[0.67]	[0.68]	[0.21]	[0.44]

<sup>a, b</sup> All abbreviations and symbols: see legend at the bottom of Table 2.  
Source: see text.

## 4 Conclusions

This paper represents a first attempt to investigate the early steps of Italian industrialization at the local level accounting for spillovers. The analytical framework is the conditional convergence model augmented to account for spatial effects. The geographic unit of analysis is constituted by provinces (NUTS 3 units) and the time period considered goes from 1871 to 1911. Estimation of the conditional convergence model augmented with human capital, social capital, and social overhead capital suggests that education, a cooperative culture, and spatial spillovers are able to explain, other things being equal, much of the variability of value added growth in the manufacturing industry in 19th century Italy. The growth differential between the provinces best and worst endowed in terms of education and social capital can be estimated to be respectively about 20 and 15 percent over the entire 1871-1911 period, while the influence of spatial spillovers over 15 percent. Since the northern provinces generally had a more educated population and a more cooperative culture the polarization between dynamic, high-growth areas and backward areas thus seems to go back to the very beginning of Italy's economic history.

## Appendix: Sources and methods.

This appendix documents the sources and the methods behind the provincial dataset used in this paper.

### Manufacturing industry: provincial value added, 1871-1911

Industrial value added estimates at 1911 prices for the 69 Italian provinces, for the years 1871 and 1911, are those provided by Ciccarelli and Fenoaltea (2013). The estimates are available at <http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0289.2011.00643.x/supinfo>, last accessed June 2014. The authors report a breakdown of industry into its four major components (extractive, manufacturing, constructions, and utilities).

We here focus on the manufacturing industry (net of tobacco) for three reasons. First, the extractive industries (of coal-less Italy) were of traditional nature, locally highly concentrated (particularly so in Sardinia and Sicily), and most of the production (sulphur is a leading example) was exported, rather than processed by the national industry.

Second, the utilities industry in 1871 consisted in the distribution of water through a supply network often dating back to the ancient Rome, while in 1911 it included electricity, a new product. This large change in the industry mix suggested to exclude utilities (accounting only for about 4 percent of industrial value added in 1911) from our analysis.

Third, the construction sector is simply too cyclical: in 1871 this sector was at the early stage of a rising cycle, while in 1911 it was at the top of a sizeable boom. The relevant volatility of the construction sector, accounting for some 15 percent of industrial value added in 1911, suggested us to exclude it from the analysis.

Finally, the tobacco sector, which was of extremely limited size (representing in 1911 less than one percent of value added of total manufacturing), it was managed by the State under monopoly conditions, with the production concentrated in a dozen of industrial plants (*manifatture dei tabacchi*).

Summing up, we examine the aggregate formed by 12 industrial sectors: foodstuffs, textiles, clothing, leather, wood, metalmaking, engineering, non-metallic mineral products, chemicals and rubber, paper, printing (including sundry manufacturing).

The sectoral shares of value added mentioned above are those reported in Ciccarelli and Proietti (2013), Table 1, providing descriptive statistics on Italy's industry for the years 1871, 1881, 1901, and 1911.

Ciccarelli and Fenoaltea (2013) obtained their provincial estimates by allocating existing regional value added estimates at 1911 prices to the various provinces using their labor-force share of the regional total, separately by sector of activity. For this reason we accounted for the different size of Italian provinces by simply dividing value added by population, separately by province.

We next turn to the description of sources and methods related to the control variables. As illustrated in the main text we collected data from historical sources on a number of variables. To increase the efficiency of the estimation procedure, rather than including them directly in the regressions, we constructed synthetic indicators using Principal Components Analyses (PCA).

## Human capital

The data on illiterates for 1871 are from Ministero di Agricoltura, Industria e Commercio (1874–76), vol. 2, Introduzione, pp. B-I. Data on primary education in 1871 (number of students, and number of teachers) are from Antonielli (1872), pp. 282-283. The data on age heaping in 1871 were kindly provided by Brian A’Hearn (see A’Hearn *et al.* 2013).

The elementary variables used as inputs to the principal component analysis providing as output the human capital indicator used in the convergence models are:

1. Illiteracy rate (*Illit*)
2. Age heaping (*Age*)
3. Primary school pupils divided by population in the 5-12 age bracket (*Pupils*)
4. Primary school teachers divided by population in the 5-12 age bracket (*Teach*)

The last two variables measure the size of the education sector, while the first two the quality of its output (with a negative sign). The results of PCA applied to the logs of these elementary variables are reported in Table A1. The first principal component turns out to be able to explain over 80% of the variance; all variables have approximately the same weights in absolute value, with opposite signs as to be expected for the first and the second pair. We thus decided to include the first principal component, labeled as *Edu* (Education), in the convergence regressions. The provincial distribution of this variable is illustrated in Figure 5, panel 1.

**Table A1.** Human capital: Principal component analysis

	elementary variables			
	<i>Illit</i>	<i>Age</i>	<i>Pupils</i>	<i>Teach</i>
first princ. component:				
weights	-0.451	-0.521	0.524	0.500
variance explained (%):	83.7			

Source: see text.

## Social capital

The data on the sales of journals and magazines are from Ottino (1875), Allegato 4, “Prospetto statistico della stampa periodica, della tipografia e della libreria in Italia.”, on the membership of mutual societies from Ministero di Agricoltura, Industria e Commercio (1875), pp. 200-203, while on the number of voters in local and national elections of 1865 from Antonielli (1872), p. 146 Election data for the province of Rome, annexed to the country in 1870, refer to the national elections of 1870 and the local elections of 1872, as reported in Correnti (1873), pp. 87b and 86.

The elementary variables used as inputs to the principal component analysis providing as output the social capital indicators used in the convergence models are:

1. Number of journals and magazines published per 100.000 residents (*Journals*)
2. Members of mutual societies per 100.000 residents (*Mutual*)

3. number of voters at local elections per 100 registered voters (*Local\_vot*)
4. number of voters at national elections per 100 registered voters (*Nat\_vot*)

The results of applying PCA on the above elementary variables are summarized in Table A2. The first principal component turned out to be able to explain only half of the variance of this set of variables. As it can be seen, press diffusion and membership of mutual societies have positive weights, while the rate of participation to local and national elections negative ones: hence, there is no obvious interpretation. We thus opted for splitting this set of variables in two pairs, with the simple means of the *Journals* and *Mutual* (suitably standardized to have unit standard deviation) capturing Social participation (*Soc*) and that of voters to the two elections Political participation (*Pol*).

$$Soc = 0.5 \left( \frac{Journals}{\sigma_{Journals}} + \frac{Mutual}{\sigma_{Mutual}} \right)$$

$$Pol = 0.5 \left( \frac{Local\_vot}{\sigma_{local\_vot}} + \frac{Nat\_vot}{\sigma_{nat\_vot}} \right)$$

The provincial distribution of social participation (*Soc*) and political participation (*Pol*) is illustrated in Figure 5, panel 2.

**Table A2.** Social capital: Principal component analysis

	elementary variables			
	<i>Journals</i>	<i>Mutual</i>	<i>Local_vot</i>	<i>Nat_vot</i>
first princ. component:				
weights	-0.375	-0.550	0.543	0.511
variance explained (%):	46.1			

Source: see text.

## Social overhead capital

The data on the local and national roads, measured in kilometers, are from Correnti (1873), pp. 122-125. The data on the number of postal offices are from Direzione generale delle Poste (1873), p. XLIV. The extension of the railway network by province for the year 1871 is obtained adding to the figures for 1861, taken from Ferrovie dello Stato (1911) the lines opened in each province between 1861 and 1871, as reported in Ministero delle comunicazioni (1927).

The elementary variables used are thus the following:

1. Local and national roads, km's normalized by surface of the province in km<sup>2</sup> (*Roads*)
2. Railways, km's normalized by surface of the province in km<sup>2</sup> (*Rail*)
3. Number of post offices per 100.000 residents (*Post*)

The results of applying PCA on the above elementary variables are summarized in Table A3. The first principal component explains about 42% of the variance of the entire set of variables, and it is positively correlated with *Roads* and *Rail* (which have almost equal weights) and negatively with *Post* (which has a smaller weight).



In fact, the second PC (not shown here) is very close to the latter variable. We thus decided to split this set of indicators as well, including the post offices variable directly, and the simple mean of the *Roads* and *Rail* variables (*RR*), duly scaled, capturing communication infrastructures:

$$RR = 0.5 \left( \frac{Roads}{\sigma_{Roads}} + \frac{Rail}{\sigma_{Rail}} \right)$$

The provincial distribution of both social overhead capital indicators, *Post* and *RR*, is illustrated in Figure 5, panel 3.

**Table A3.** Social overhead capital: Principal component analysis

	elementary variables		
	<i>Roads</i>	<i>Rail</i>	<i>Post</i>
first princ. component:			
weights	0.658	0.673	-0.339
variance explained (%):	42.2		

Source: see text.

## Historical sources

Antonielli, E. (1872), *Annuario statistico delle provincie italiane per l'anno 1872*, Florence.

Correnti, C. (1873), *L'Italia economica nel 1873*, Rome.

Ferrovie dello Stato (1911), *Ferrovie Italiane 1861-1909*, Riproduzione dei lavori grafici presentati all'Esposizione internazionale di Torino del 1911.

Franchetti, L. (1875), *Condizioni economiche ed amministrative delle province napoletane*, Tipografia della Gazzetta d'Italia, Florence.

Ministero delle comunicazioni (1927), *Sviluppo delle ferrovie italiane dal 1839 al 31 dicembre 1926*, Rome.

Istat (1958), *Sommario di statistiche storiche italiane, 1861-1955*, Rome.

Ministero di Agricoltura Industria e Commercio (1867), *Statistica del Regno d'Italia. Elezioni politiche e amministrative. Anni 1865-66*. Florence.

Ministero di Agricoltura, Industria e Commercio (1874–76) *Popolazione. Censimento 31 dicembre 1871. 3 voll.* Direzione Generale della Statistica, Rome, Italy.

Direzione generale delle Poste (1873), *Nona relazione sul servizio postale in Italia*, Rome, .

Ministero di Agricoltura, Industria e Commercio (1875), *Statistica delle società di mutuo soccorso*, Rome.

Ottino, G. (1875), *La stampa periodica, il commercio dei libri e la tipografia in Italia*, Milan.

The appendix ends by providing a map of Italy's provinces in 1911. The name of each region, in bold, is followed by the name (and tag) of its provinces. Different colors simply highlight regional borders.

**PIEDMONT:** Alessandria (AL), Cuneo (CN), Novara (NO), Turin (TO)

**LIGURIA:** Genoa (GE), Porto Maurizio (PM)

**LOMBARDY:** Bergamo (BG), Brescia (BS), Como (CO), Cremona (CR), Mantua (MN), Milan (MI), Pavia (PV), Sondrio (SO)

**VENETIA:** Belluno (BL), Padua (PD), Rovigo (RO), Treviso (TV), Udine (UD), Venice (VE), Verona (VR), Vicenza (VI)

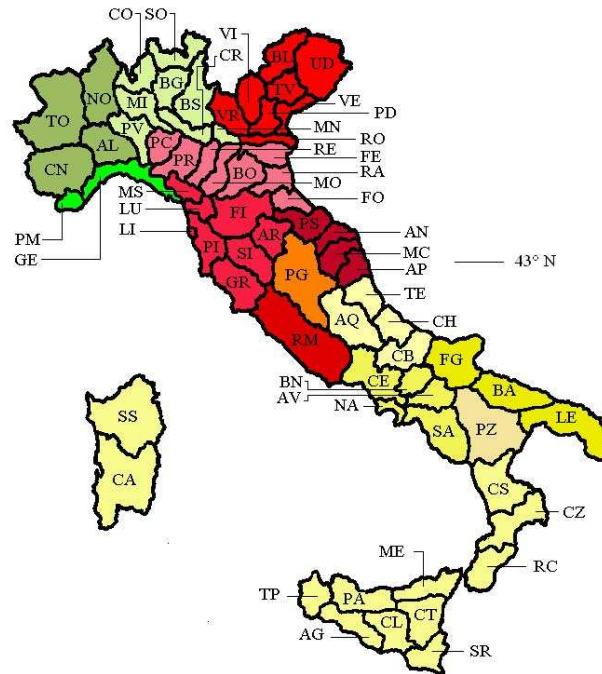
**EMILIA:** Bologna (BO), Ferrara (FE), Forlì (FO), Modena (MO), Parma (PR), Piacenza (PC), Ravenna (RA), Reggio Emilia (RE)

**TUSCANY:** Arezzo (AR), Florence (FI), Grosseto (GR), Leghorn (LI), Lucca (LU), Massa Carrara (MS), Pisa (PI), Siena (SI)

**MARCHES:** Ancona (AN), Ascoli Piceno (AP), Macerata (MC), Pesaro (PE)

**UMBRIA:** Perugia (PG)

**LATIUM:** Roma (RM)



**ABRUZZI:** Aquila (AQ), Campobasso (CB), Chieti (CH), Teramo (TE)

**CAMPANIA:** Avellino (AV), Benevento (BN), Caserta (CE), Naples (NA), Salerno (SA)

**APULIA:** Bari (BA), Foggia (FG), Lecce (LE)

**BASILICATA:** Potenza (PZ)

**CALABRIA:** Catanzaro (CZ), Cosenza (CS), Reggio Calabria (RC)

**SICILY:** Caltanissetta (CL), Catania (CT), Girgenti (AG), Messina (ME), Palermo (PA), Syracuse (SR), Trapani (TP)

**SARDINIA:** Cagliari (CA), Sassari (SS)

## References

- Aghion, P. and P. Howitt (1998), *Endogenous growth theory*. Cambridge, MA:MIT Press.
- A'Hearn B. (1998), "Institutions, externalities and economic growth in southern Italy. Evidence from the cotton textile industry, 1861/1914," *Economic History Review*, , pp. 734-762.
- A'Hearn, B., Delfino, A. and Nuvolari, A. (2013), " 'Uniting Souls' and counting skills: age-heaping in the first Italian national censuses, 1861-1881', paper presented at the Valencia FRESH meeting, 13-14 June, 2013.
- Anselin, L. (2008) "Spatial regression," in Fotheringham, A.S, and P.A. Rogerson (eds.) *The SAGE Handbook of Spatial Analysis*, SAGE, Los Angeles and London, pp. 255-276
- Anselin, L. and Rey, S. (1991) "Properties of tests for spatial dependence in linear regression models," *Geographical Analysis*, 23, pp. 112-131
- Aschauer, D. A. (1989), "Is Public Expenditure Productive?," *Journal of Monetary Economics*, 23, pp. 177-200.
- Banfield, E. C. (1958), *The moral basis of a backward society*, Free Press, Glencoe.
- Bauer, Paul B., Mark E. Schweitzer, and Scott A. Shane. (2012), "Knowledge matters: the long-run determinants of state income growth," *Journal of Regional Science*, 52, 240-255.
- Bronzini R, and Piselli, P. (2009), "Determinants of long-run regional productivity with geographical spillovers: the role of R&D, human capital, and public infrastructure," *Regional Studies and Urban Economics*, 39, pp. 187-199.
- Brunetti A., Felice E. and G. Vecchi (2011), "Reddito," in G. Vecchi (ed.), *In ricchezza e in povertà. Il benessere degli italiani dall'Unità a oggi*, Bologna: Mulino.
- Cafagna, L. (1989), *Dualismo e sviluppo nella storia d'Italia*, Marsilio, Padova.
- Cafagna, L. (1998), "Contro tre pregiudizi sulla storia dello sviluppo economico italiano," in P. Ciocca and G. Toniolo (eds.), *Storia economica d'Italia, 1: Interpretazioni*, Laterza, Rome-Bari, pp. 297-325.
- Cannari, L., S. Chiri, (2002), "Le infrastrutture economiche dall-Unità," in P. Ciocca and G. Toniolo (eds.), *Storia economica d'Italia. 3. Industrie, mercati, istituzioni. 1. Le strutture dell'economia*, pp. 225-297. Milan-Rome-Bari:IntesaBci and Laterza.
- Capello, R. (2009) "Spatial spillovers and regional growth: a cognitive approach" *European Planning Studies*, 17, pp. 639-658.
- Ciccarelli, C. Fenoaltea, S. (2008), "Social-overhead construction in Italy's regions, 1861-1913," *Research in Economic History*, 26, pp. 1-80.
- Ciccarelli C., S. Fenoaltea (2013) "Through the magnifying glass: Provincial aspects of industrial growth in post-Unification Italy," *Economic History Review*, 66, pp. 57-85.

- Ciccarelli C., Missiaia A. (2013), "The industrial labor force of Italy's provinces: Estimates from the population censuses, 1871-1911," *Rivista di Storia Economica*, 29, 2, pp. 141-192.
- Ciccarelli C., Proietti T. (2013), "Patterns of industrial specialisation in post-unification Italy," *Scandinavian Economic History Review*, 61, 3, pp. 259-286.
- Ciocca P. (2007), *Ricchi per sempre ?*, Bollati Boringhieri, Turin.
- Clough, S. B, and Livi, C. (1956), "Economic growth in Italy: An analysis of the uneven development of North and South," *Journal of Economic History*, 16, 3, pp. 334-349.
- Daniele V., Malanima P. (2007), "Il prodotto delle regioni e il divario Nord-Sud in Italia (1861-2004)," *Rivista di politica economica*, 97, no. 3-4, pp. 267-340.
- Daniele V., Malanima P. (2011), *Il divario nord-sud in Italia, 1861-2011*, Rubbettino, Soveria Mannelli.
- de Blasio G., Nuzzo. G. (2009), "Historical traditions of civicness and local economic development," *Journal of Regional Science*, 50, 4, pp. 883-857.
- Di Liberto, A. (2008), "Education and Italian regional development," *Economics of Education Review*, 27, pp. 94-107.
- Eckaus, R. S. (1961), "The North-South differential In Italian economic development," *Journal of Economic History*, 21, 3, pp. 285-317.
- Einaudi, L. (1959), *Cronache economiche e politiche di un trentennio (1893-1925). Vol. I (1893-1902)*, Einaudi, Turin.
- Farneti, P. (1971), *Sistema politico e società civile*, Giappichelli, Turin.
- Felice E. (2012), "Regional convergence in Italy, 1891-2001: testing human and social capital," *Cliometrica*, 6, pp. 267-306.
- Felice E. (2014), *Perché il Sud è rimasto indietro*, Il Mulino, Bologna.
- Fenoaltea, S. (2003), "Peeking backward: Regional aspects of industrial growth in post-unification Italy," *Journal of Economic History*, 63, 4, pp. 1059-1102.
- Fenoaltea, S. (2011), *The reinterpretation of Italian economic history: from unification to the great war*, New York: Cambridge University Press.
- Gennaioli, N., R. La Porta, F. Lopez-De-Silanes (2013), "Human capital and regional development," *Quarterly Journal of Economics*, 128, pp. 105-164.
- Giersch, H. (1949), "Economic union between nations and the location of industries," *Review of Economic Studies*, 17, pp. 87-97.
- Goldstein, R. L. (1983), *Political repression in 19th century Europe*, Beckenham: Croom Helm.
- Guiso, L. Sapienza, P., and Zingales, L. (2005), "The role of social capital in financial development," *American Economic Review*, 94, pp. 526-556.

- Helliwell, J.H, R.D. Putnam (1995), "Economic growth and social capital in Italy," *Eastern Economic Journal*, 21, pp. 295-307.
- Krueger, A. B. and Lindahl, M. (2001), "Education for growth ? Why and for whom ?," *Journal of Economic Literature*, 39, pp. 1101-1136.
- Isard, W. (1954), "Location theory and trade theory: A short run analysis," *Quarterly Journal of Economics*, 68, pp. 305-320.
- Islam, N. (2003), "What have we learnt from the convergence debate?," *Journal of Economic Surveys*, 17, 309-362.
- Leonardi R. (1995), "Regional development in Italy: Social capital and the Mezzogiorno," *Oxford Review of Economic Policy*, 11, 2, pp. 165-79.
- Lucas, R. (1988), "On the mechanics of economic development," *Journal of Monetary Economics*, 22, pp. 3-42.
- Magrini, S. (2004), "Regional (di)convergence," in J. V. Henderson and J.-F. Thisse, (eds.) *Handbook of Regional and Urban Economics*, Elsevier, Vol. 4, pp. 2741-2796.
- Paci, R. and A. Saba (1998), "The empirics of regional economic growth in Italy: 1951-1993," *Rivista Internazionale di Scienze Economiche e Commerciali*, 5, pp. 513-542.
- Picci, L. (1999), "Productivity and infrastructure in the Italian regions," *Giornale degli Economisti e Annali d'Economia*, 58, 3-4, pp. 329-353.
- Picci, L. (2002), "Le opere pubbliche dall'Unità d'Italia: l'informazione statistica," *Rivista di Politica Economica*, 92, pp. 29-80.
- Putnam, R. D. (1993), *Making democracy work: civic traditions in modern Italy*, Princeton University Press, Princeton (with Leonardi R. and Nanetti R.Y.)
- Rey, J. S. and Montouri, B. D. (1999), "US regional income convergence: a spatial econometric perspective," *Regional Studies*, 33, 2, pp. 143-156.
- Rodrik, D. (2013), "Unconditional convergence in manufacturing," *Quarterly Journal of Economics*, 128, 1, pp. 165-204.
- Romer, P. (1990), "Endogenous technological change," *Journal of Political Economy*, 89, pp. S71-S102.
- Romp, W. and J de Hann (2005), "Public capital and economic growth: a critical survey," *EIB papers*, 10 (1), pp. 40-70.
- Silverman, B. W. (1986), *Density estimation for statistics and data analysis*, Chapman & Hall, London.
- Tabellini, G. (2010), "Culture and institutions: Economic development in the regions of Europe," *Journal of the European Economic Association*, 8, pp. 677-716.
- Temple, J. (2001), "Growth effects of education and social capital in the OECD countries," *OECD economic studies*, no. 33 (2001/II), pp. 57-101.

World Bank (2007), *Evaluation of World Bank support to transportation infrastructure*, Washington DC: World Bank Publications.

Zamagni, V. (1990), *Dalla periferia al centro. La seconda rinascita economica dell'Italia, (1861-1990)*, Il Mulino, Bologna.