# The location of the Italian manufacturing industry, 1871-1911: a quantitative analysis

Roberto Basile*  Carlo Ciccarelli†

May 2015

**Abstract**

This paper analyzes spatial location patterns of manufacturing activity in Italy during the period 1871-1911. To this end we use a recent dataset of value added at 1911 prices covering the whole set of 69 Italy's provinces (NUTS 3 units). Total manufacturing value added is disaggregated into 10 industrial sectors. We test the relative effect of domestic market potential and factor endowment, focusing in particular on water supply. The results show that, as transportation costs decreased and barriers to domestic trade were eliminated, Italian provinces became more and more specialized, and manufacturing activity became increasingly concentrated in a few provinces, mostly belonging to the North-West part of the country. The estimation results corroborate our hypothesis that both comparative advantages (water endowment effect) and market potential (home-market effect) have been responsible of the above process of spatial concentration. The location of some traditional industries (such as clothing) was mainly driven by water endowment, while the location of fast growing new sectors (such as engineering, metalmaking, chemicals, and paper) was mainly driven by the domestic market potential. Our results also suggest that the domestic market potential was also a key determinant of the location of the textile industry.

*Keywords*: Market potential, Factor endowment, Concentration, Italy.
*Jel codes*: R12, R15, N13

PAPER PREPARED FOR THE 56th ANNUAL CONFERENCE of the SIE - NAPLES 22-24 OCTOBER 2015.

PRELIMINARY AND INCOMPLETE VERSION - PLEASE DO NOT QUOTE WITHOUT AUTHORS PERMISSION.

*Department of Economics, Second University of Naples, Corso Gran Priorato di Malta, 1 - 81043, Capua (CE), Italy. *Email*: roberto.basile@unina2.it

†Department of Economics and Finance, University of Rome "Tor Vergata", Via Columbia 2, 00133 Rome, Italy. carlo.ciccarelli@uniroma2.it

# 1 Introduction

Economic theory suggests that the key determinants of industrial location are the uneven spatial distribution of natural resources (including the endowment of water, timber, and mineral deposits), the asymmetric spatial distribution of market access which induces firms to concentrate in the regions with higher market potential to benefit from increasing returns, or both. These agglomeration (or centripetal) forces tend to contrast the market competition (centripetal) forces arising from the concentration of firms and inducing to lower local market prices and higher local factor prices. The equilibrium between these forces mainly depends on the degree of market integration and transportation costs. Under autarky each region produces essentially all the goods that it consumes, the location of industries is stable, and the level of industrial concentration is low. When trade costs decreases and product markets tend to integrate, the neoclassical trade model (Heckscher-Ohlin) predicts that regional specialization will arise as regions produce and export products that are relatively intensive in their abundant resources (comparative advantages). Moreover, when transportation costs decrease, the New Economic Geography literature predicts that manufacturing activities characterized by increasing returns to scale tend to concentrate in the regions with higher demand (home-market effect), while all other regions suffer de-industrialization. Therefore, a regional division of labor spontaneously arises through a process of uneven development.

The relative effect of comparative advantages (factor endowment) and increasing returns (market access) can be hardly quantified using empirical data for the modern economies as the two factors tend to coexist and interact in a very complex way along with the effect of (endogenous) policy interventions to determine the location of industry. For example, as discussed by Midelfart-Knarvik and Overman (2002), the European industrial policy strongly influenced the industrial location patterns across the EU regions; Basile et al. (2008) also provide evidence that EU Structural funds significantly affected the location choice of multinational firms in Europe. To the extent that these industrial policies were endogenously driven by the actual spatial distribution of economic activity, it turns to be quite hard to quantify the effect of comparative advantage and/or market potential net of the effect of industrial policies. However, the use of historical data covering the years following the political unification of a country (that is when the domestic market integration was full of obstacles, and virtually no national industrial policy took place), may provide an opportunity to better contrast the different explanations for the spatial concentration of industry and, in particular, to appreciate the increasing role of the home-market effect. Examples of studies in this direction are Venables and A'Hearn (2013), Wolf (2007), and Rosés (2003).

Following these recent contributions, in the present paper we analyze the spatial location patterns of industrial activity in Italy during the period 1871-1911. Over this period the country experimented a strong process of economic integration, with the dismantling of institutional barriers to domestic trade, the adoption of a common currency, the harmonization of local institutions and the creation of a national railway network. Our first hypothesis is that this process of economic integration has amplified the role of domestic market potential as a key determinant of manufacturing activity, especially for the fast growing modern sectors. On the other hand, the lack of a modern electricity transmission system during the sample period imposed the need to exploit local sources of power. In the Italian case, as widely pointed out (see, for instance, Cafagna 1989, Bardini 1997), abundance of water represented an essential source of power in a country without coal.[1] Our second hypothesis is therefore that the location of manufacturing activity was strongly correlated to water supply.

We test the relative importance of factors' endowment (above all water) and market potential in the early phases of Italian industrialization using recently developed value added data at 1911 prices at provincial level (for 69 provinces) for the following 12 manufacturing sectors: foodstuffs, textile, tobacco, clothing, leather, wood, Non-metallic mineral products, engineering, metal, chemicals, paper and sundry (Ciccarelli and Fenoaltea, 2013, 2014).

Our results clearly shows that as transportation costs decreased and barriers to domestic trade were eliminated, Italian provinces became more and more specialized, and manufacturing activity became increasingly concentrated in a few provinces, mostly belonging to the North-West part of the country. The estimation results corroborate our hypothesis that both comparative advantages (water endowment effect) and market potential (home-market effect) have been responsible of this process of spatial concentration.

The work Ciccarelli and Fenoaltea (2013) considers total value added growth in the manufacturing industry from 1871 to 1911. The authors compare the growth in the provincial share (of the regional total) of manufacturing value added, and that of the provincial share (of the regional total) of male labor force, used as a proxy for the growth of provincial GDP. Among the results, the authors point to the benefit of a free trade regime for Southern provinces. They show in particular that a rapid industrial and overall growth characterized many Southern provinces, and suggest that this might be related to the extension of the low import duties from the northern provinces to the southern ones in the aftermath of the country unification (1861). Ciccarelli and Proietti (2013) considers instead the internal composition of the manufacturing industry. The authors propose a graphical tool named "dynamic specialization biplots" that allows them to visualize the changing specialization of the Italian provinces among 12 manufacturing sectors during the 1871-1911. The authors find that much of the variability in the data is explained by the foodstuffs, textile, and engineering sectors, and that trade policy much contributes to explain the changing specialization trajectories of selected Northern provinces. The work by Ciccarelli and Missiaia (2013) is then particularly useful in that provides explicitly the labor force figures by gender and industrial sectors for the census years 1871, 1881, 1901 and 1911. The advantages and disadvantages of the Italian census data are also discussed.[2]

Our paper is close in spirits to Ciccarelli and Proietti (2013). that also refer to industrial value added in Italian provinces, but essentially in a descriptive way. The present paper provides for the first time an economic analysis of the main determinants of early industrial location at the provincial level. To this end we gathered entirely new data on local water endowment (both in terms of waterfalls and rivers), and market potential. Behind that, our study uses very up-to-date value added estimates obtained by incorporating those provided by Ciccarelli and Fenoaltea

[1]Coal was largely imported from the UK. It was delivered to the main Italian ports of Genoa and Naples, and distributed across the country by railways. The port of Genoa was connected by railways to Turin in 1853. Turin and Milan were connected by railways between 1855 and 1859. These connections contribute substantially to the early industrial development of the North-West area.

[2]A non-technical summary of the quantitative literature on Italy's industry at the provincial and regional level during the 19th century can be also find in Ciccarelli (2015)

(2014). The rest of the paper is organized as follows. Section 2 describes the basic historical and economic feature of the manufacturing industry. Section 3 illustrate the spatial distribution and diffusion of industrial activity. Section 4 present the result based on the estimation of a Geoadditive model of industrial location considering factor endowment and market potential.

# 2 Setting the scene

This section gives detailed historical background of the post-unification period, and reviews the evidence on economic integration across its provinces. Italy was unified in 1861, although Venetia and Latium were annexed to the country only in, respectively, 1866 and 1870. In official statistics, the Italian territory is divided into 69 administrative units (provinces, roughly NUTS 3 units). The territorial borders of province did not change during the 1871-1911 here considered. The literature often aggregates Italian provinces into four macro areas: the North-West, the North-East, the Centre, and the South. Maps 1 and 2 illustrate the Italian provinces at 1911 borders. Map 1 group them according to the seven pre-unitarian States, while map 2 also include provincial labels.

Insert Figures 1 and 2 about here

Distribution of comparative advantages (water).

Insert Figure 3 about here

Level of economic integration/autarky among provinces in 1870 and its dynamics onwards:

- tariff and non-tariff barriers to domestic trade and their removal,
- factor mobility (labor and capital), integration of capital markets, internal migration flows
- creation of a common currency area
- transport and communication facilities and the decrease of transportation and communication costs due to investments in the railway network and telegraph mileage,
- institutional differences across provinces and their harmonization

Distribution of market potential.

Insert Figure 4 about here

## 2.1 The structure of economic activity

To measure the level of economic activity in Italy and in its provinces over the period after the unification, we make use of the most disaggregated data set available. First, we have data for the four population census periods 1871, 1881, 1901 and 1911 that cover employment (approximated by the male labor force) in all sectors. Overall, there is information on three aggregates:

Agriculture, Industry (of which Manufacturing) and Services and for twelve manufacturing industries: Food, Tobacco, Textile, Clothing, Wood, Leather, Metal, Engineering, Non-mineral Metals, Chemicals, Paper, and Sundry. Second, for these 12 manufacturing sectors we also have value added estimates elaborated by Ciccarelli and Fenoaltea (2013).

At the beginning of our sample period (1871), that is just after the unification of Italy, 62% of the male labor force was occupied in agriculture, 21% in the secondary sectors (of which 15% in manufacturing) and the remaining 17% in services (see Table 1). Over the following three population census periods (1881, 1901 and 1911) the share of agriculture decreased up to 54% in favor of both secondary and tertiary activities. NB: in 1901 there is no difference between our aggregate industry (and services) and that of Vitali, while agriculture coincides.

Insert Table 1 about here

Within manufacturing, in 1871 the male labor force was concentrated in the "traditional" sectors, namely Wood (20%), Leather (18%), Food (16%), Textile (10%) and Clothing (10%) (see see Table 2). The remaining 25% was distributed between Engineering, Metalmaking and other "modern" sectors (Chemicals, Paper and Non-metallic mineral products). Over the sample period, however, the weight of the above mentioned "traditional" sectors in the total manufacturing labor force decreased (except for leather) in favor of the "modern" sectors, which in 1911 reached the 40%. A similar distribution and a similar dynamics can be observed using value added data: the output of all "modern" industries increased over time, while the value added of traditional sectors strongly decreased (in particular in the Food industry) or remained stable (Textile, Clothing and Wood).

Insert Tables 2 and 3 about here

An alternative classification of manufacturing sectors on the base of the horse-power per worker (see ...):

- High horse-power per worker
  - Metal: 2.62
  - Chemicals: 1.30
  - Food: 0.98
  - Paper: 0.73
- Medium horse-power per worker
  - Textile: 0.52
  - Engineering: 0.50
- Low horse-power per worker
  - Non-metallic mineral products: 0.36
  - Wood: 0.23
  - Leather: 0.09
  - Clothing: 0.07

# 3 Division of labor and spatial diffusion of manufacturing activity

In this section, we use the value added data described above to answer several questions about the division of labor and the spatial diffusion of industrial activity in the early Italy. How similar were the industrial structures across different provinces in Italy? Do we find a regional division of labor across Italy? Did this change during the 1871-1911 period? How concentrated was manufacturing activity as a whole and how concentrated was a given industry? Which industries tend to agglomerate, which industries were rather dispersed? And do we find an increase or decrease in concentration over time?

In order to answer these questions, we clarify some measurement issues. First, we compute the share of a certain manufacturing industry $k$ in the total manufacturing activity of province $i$ $\left(s_i^k(t)\right)$, defined as

$$s_i^k(t) = \frac{x_i^k(t)}{\sum_k x_i^k(t)} \tag{1}$$

where $x_i^k(t)$ measures the level of economic activity $k$ at location $i$ and time $t$.

Second, we compute the share of a certain location $i$ in the total manufacturing activity of industry $k$ as

$$l_i^k(t) = \frac{x_i^k(t)}{\sum_i x_i^k(t)} \tag{2}$$

Third, we define the location quotient as

$$LQ_i^k(t) = \frac{l_i^k(t)}{\sum_k x_i^k(t) / \sum_i \sum_k x_i^k(t)} = \frac{s_i^k(t)}{\sum_i x_i^k(t) / \sum_i \sum_k x_i^k(t)} \tag{3}$$

In section 4, we will use the location quotient as our dependent variable in an econometric analysis of the determinants of industrial location.

## 3.1 Regional specialization and geographical concentration of economic activities

Using the shares $s_i^k(t)$ we can address the question of how specialized were Italian provinces by using Krugman's specialization index $K_i(t)$, defined as:

$$K_i(t) = \sum_k |(s_i^k(t) - s_i^{-k}(t))| \tag{4}$$

where $s_i^{-k}(t)$ is the share of industry $k$ in the total production of all provinces except province $i$. Thus, the Krugman index summarizes a province's difference in specialization with respect to the rest of Italy over all industries. It takes the value of zero if a province's industrial structure is identical to the rest of Italy, and the value of two if the region has no industries in common with the rest of Italy. Table 4 shows means and standard deviations of Krugman index for the four sample periods, while figure 5 maps its spatial distribution. The results clearly show that on average the degree of industrial specialization of Italian provinces has monotonically increased over the sample period (from 0.26 to 0.38), while its dispersion slightly raised. The increment in the value of Krugman index were particularly strong for some provinces in the North-West partition of the Country (namely, Turin, Milan, Genova, Novara, Como, Cremona, Bergamo, Massa Carrara, Lucca, Pisa and Livorno).

Insert Table 4 about here

Insert Figure 5 about here

Next, we try to assess whether this increase in specialization corresponded to a higher spatial concentration of industries. We constructed a relative Theil index of industrial concentration as:

$$C_k(t) = \sum_i \frac{v_i}{V} LQ_i^k(t) log \left( LQ_i^k(t) \right) \quad (5)$$

where $v_i$ is the value added in province $i$ and $V$ is the total value added in Italy. The higher the value of $C_k(t)$, the higher the concentration of industry $k$.

Table 5 reports the value of the Theil index for the whole manufacturing sector (in which case the location quotient is computed by normalizing with respect to the total economic activity) and for the 12 individual industries (in which case the location quotient is computed by normalizing with respect to the total manufacturing activity) over the 4 sample periods. We observe that, along with the progressive integration of commodities and factor markets (falling transport costs), the manufacturing activity became slightly more concentrated throughout the entire period (the value of the Theil index for the manufacturing industry as a whole was 0.03 in 1871 and became 0.08 in 1911).

The aggregate coefficient hides the notable tendency towards the concentration of several industries (namely, Food, Textile, Wood, Leather, Metal and Engineering). Instead, the relative level of concentration remained stable in two sectors (Clothing and Chemicals) and only three industries became relatively more dispersed (Tobacco, Paper and Non-metallic mineral products). Nevertheless, Tobacco remained the most concentrated sector, followed by Metal and Textile. The most dispersed sectors were instead Wood, Food, and Engineering. However, the ranking of the Theil indices remained very stable over time (the Spearman rank correlation between the Theil indices in 1871 and in 1911 is 0.95 with a p-value of 0.000). The stability in concentration level observed across industries for the early Italy is a pattern also found in studies for more recent periods for Italy (Arbia et al., 2011) as well as for other countries (Dumais et al., 2002; Devereux et al., 2004; Alonso-Villar et al., 2004).

Insert Table 5 about here

## 3.2 Global spatial dependence

The relative Theil index $C_k(t)$ provides useful information about the extent to which industries in the early Italy were concentrated in a limited number of areas, but does not take into consideration whether those areas were close together or far apart. In other words, it does not take into account the spatial structure of the data. Two industries may appear equally geographically concentrated, while one was located in two neighboring regions, and the other split between the northern and the southern part of the country. Every region is treated as an isolated island, and its position in space relative to other regions is not taken into account. Thus, the relative Theil index $C_k(t)$ is an a-spatial measure of concentration: the same degree of concentration can be compatible with very different localization schemes.

As pointed out by Basile and Mantuano (2010), Arbia, Basile and Mantuano () and Arbia et al. (2011), a more accurate analysis of the spatial distribution of economic activities requires the combination of traditional measures of geographical concentration and methodologies that account for spatial dependence, in that they provide different and complementary information about the concentration of the various sectors. Spatial autocorrelation is present when the values of one variable observed at nearby locations are more similar than those observed in locations that are far apart. More precisely, positive spatial autocorrelation occurs when high or low values of a variable tend to cluster together in space and negative spatial autocorrelation when high values are surrounded by low values and vice-versa. Among the spatial dependence measures, the most widely used is the Moran's $I$ index (Moran, 1950):

$$I = \left( \frac{N}{\sum_i \sum_j w_{ij}} \right) \left( \frac{\sum_i \sum_j w_{ij} \left( LQ_i - \overline{LQ} \right) \left( LQ_j - \overline{LQ} \right)}{\sum_i \left( LQ_i - \overline{LQ} \right)^2} \right) \tag{6}$$

where $N$ is the total number of provinces, $LQ_i$ and $LQ_j$ are the observed values of the location quotient for the locations $i$ and $j$ (with mean $\overline{LQ}$), and the first term is a scaling constant. This statistic compares the value of a continuous variable at any location with the value of the same variable at surrounding locations. The spatial structure of the data is formally expressed in a spatial weight matrix $W$ (Anselin, 1988) with generic elements $w_{ij}$ (with $i \neq j$). In the rest of the paper we will employ row-standardized spatial weights matrix ($W$), whose elements $w_{ij}$ on the main diagonal are set to zero whereas $w_{ij} = 1$ if $d_{ij} < \overline{d}$ and $w_{ij} = 0$ if $d_{ij} > \overline{d}$, with $d_{ij}$ the great circle distance between the centroids of region $i$ and region $j$ and $\overline{d}$ a cut-off distance (equal to ...).

Table 6 reports the values of the Moran's $I$ statistics for the whole manufacturing sector and for the 12 sector computed for each year, as well the respective $p$-values. First of all, it is worth noticing that the degree of spatial autocorrelation of $LQ$ for the manufacturing as whole increased over time passing from a null to a statistically significant and positive value. Thus, both the a-spatial and the spatial concentration of manufacturing activity increased over time.

An increase of the degree of spatial autocorrelation can also be observed for all traditional sectors (Leather, Wood, Clothing, Textile, and Food). In 1911, Textile and Leather were the sectors with the highest levels of spatial autocorrelation.[3] For three sectors (Non-metallic mineral products, Metalmaking and Engineering) we found a decrease of spatial autocorrelation, albeit the Moran's $I$ statistic remained positive and statistically significant except for Engineering. Only for three sectors (Paper, Chemicals and Tobacco) we found no evidence of significant

[3]Interstingly, an upsurge of spatial dependence in Italian traditional sectors characterized by the use of basic technologies also emerged for more recent periods (Arbia et al., 2011). These are sectors in which operate firms of small and medium size localized in well-defined industrial clusters in the Northern and Central part of the country (Emilia Romagna, Tuscany and Marches).

spatial dependence over the whole sample period. Also the rank correlation between the Moran's *I* indices in 1871 and in 1911 is positive and quite high (0.67) with a *p*-value of 0.015.

Insert Table 6 about here

As a further step of our analysis, we can jointly consider the results produced by the Theil measure of geographical concentration and the Moran's *I* measure of spatial dependence described above. Similarly to Guillain and Le Gallo (2006), we can identify for different patterns in the distribution of economic activities where either or both geographical concentration and spatial dependence occur: (*i*) high concentration and low spatial dependence, (*ii*) high concentration and high spatial dependence, (*iii*) low concentration and high spatial dependence, (*iv*) low concentration and low spatial dependence. To this purpose, we combine the a-spatial Theil index and the Moran's *I* index in a scatterplot (Figure 6) for each period, excluding Tobacco and Sundry. The four patterns in the distribution of economic activities are identified by including a vertical and an horizontal dashed line at the median values of the each indicator.

Important changes in the composition of the four patterns of spatial distribution occurred over the sample period. Three sectors (Chemicals, Paper and Metalmaking) registered a high level of concentration and a low degree of spatial dependence in 1911 (pattern *i*)). This means that these three industries concentrated their activity in a small number of areas that were not close to each other. In these sectors, indeed, economies of scale may be reached only by increasing the plant size and concentrating the production in a small number of locations. However, in 1871, Metalmaking belonged to the pattern *ii*) (high concentration and high degree of spatial dependence), along with Textile and Non-mineral metals. Textile was still in this second group in 1911, along with Leather. Thus, Textile was the mostly "spatially" clustered sector over the whole sample period, being highly concentrated and at the same time strongly agglomerated in all census periods. This feature still characterizes the Textile sector in the modern period (see Arbia et al. 2011). Clothing was instead the sector that persistently belonged to the third spatial pattern (low concentration and high spatial dependence) over the while sample period. In 1871, Leather was in the low-concentration/high-dependence cluster, but then moved to the high-high group in 1911. In 1911, instead, Food and Wood were in the low-concentration/high-dependence group, but they belonged to the low-low club in 1871. Finally, we find that Engineering, Food and Wood were the mostly "spatially" spread industries (pattern *iv*)) in 1871. Engineering was still in this group in 1911, along with Non-mineral metals that was in pattern *ii*) (high Theil and high Moran) in 1871.

Insert Figure 6 about here

### 3.3 Local spatial patterns in the distribution of economic activities

The Moran's *I* statistic of spatial autocorrelation that we used above captures the overall spatial clustering in the data. A positive and significant measure of Moran's *I* captures the existence of both high-value clustering and low-value clustering, while a negative autocorrelation captures the presence of high-values next to low-values. In other words, only one dominant type of autocorrelation can be detected. If two structures, such as high-value clustering and low-value clustering, coexist, the global statistic of spatial autocorrelation cannot distinguish them (Zhang and Lin, 2007). In contrast, the local version of the Moran's $I$ statistic is specifically designed to find evidence of local spatial patterns in the empirical data, that is to detect local spatial associations ('hot spots'), such as high-value clusters, low-value clusters, and negative autocorrelations (Anselin, 1995). The local Moran statistic for an observation $i$ is defined as (Anselin, 1995):

$$I_i = LQ_i \sum_{j \neq i} w_{ij} LQ_j \tag{7}$$

The results of the local Moran's $I$ tests are shown in Figures ..., along with the map of the spatial distribution of the location quotients. Significant values of the local Moran are classified as: $HH$ for locations with high levels of the $LQ$ for a specific sector surrounded by regions with high levels of the $LQ$; $LL$ for locations with low levels surrounded by locations where the $LQ$ are also low; $HL$ for locations with high values surrounded by locations with low values, and $LH$ for locations with low values surrounded by locations with high values. While the first two typologies (namely $HH$ and $LL$) suggest clustering of similar values, the last two situations ($HL$ and $LH$) capture the presence of regional outliers in the spatial distribution of economic activities.

Spatial distribution of specialization in manufacturing (maps of Location Quotients - LQ - of labor force in manufacturing): 4 panels maps in the four census (see Figure 7).

Insert Figure 7 about here

Summing up, the empirical spatial data analysis provides a clear evidence of an upsurge of the interregional division of labor and of the spatial concentration of industrial activity across the Italian provinces during post-unification period 1871-1911. However, this evidence does not point to any particular set of explanations. An increased interregional division of labor might be seen as evidence in favor of HO-type mechanisms of industrial location. It might equally be seen as the other side of a process of concentration in some industries, due to NEG-type mechanisms. What forces dominate is left to an econometric analysis carried out in the next section.

# 4 Water endowments, Market potential and industry location: econometric results

In this section we examine the relevance of comparative advantages (factor endowment; HO hypothesis) and of market potential (increasing returns to scale; NEG hypothesis) in shaming the location of industry across Italian provinces during the period 1871-1911. We also intend to assess whether the effect of market potential changed over time due to the process of economic integration.

Industrial location is measured in relative terms, that is using the logarithm of the location quotients $LQ_i^k(t)$, based on value added data, as defined in section ... As for factor endowment, we first focus on water abundance, measured using the logarithm of the variable $RIVER$ described above and its interaction with the dummy variable $WATERFALLS$, indicating the existence of waterfalls in the province. In the next section, we will consider also the effect of other production factors (land, skilled labor and innovations). Market potential is measured by the logarithm of the Harris market potential ($MKTPOT$) described in section ...

In order to avoid mis-specification biases due to a wrong function form, we adopt a semi-parametric additive model specification where all relevant variables enter additively as smooth functions, without imposing any specific functional relationship (linear or nonlinear) with the response variable. In order to omitted time-related factors biases, we introduce a time fixed effect. Finally, in order to avoid mis-specification biases due to unobserved spatial heterogeneity, we introduce a smooth spatial trend (a smooth interaction between the spatial coordinates of the provinces) in the right-hand-side of the model, thus specifying what is known in the literature as a Geoadditive model (Basile et al, 2015). The most general specification used in this phase of the analysis is therefore:

$$\begin{aligned} \ln(LQ_{i,t}) &= \alpha + f_1\left[\ln(RIVER_i|WATERFALLS_i = 0)\right] \\ &\quad + f_2\left[\ln(RIVER_i|WATERFALLS_i = 1)\right] \\ &\quad + \sum_t^T f_{3,t}\left[\ln(MKTPOT_{i,t})\right] \\ &\quad + h(no_i, e_i) + \gamma_t + \varepsilon_{i,t} \\ \varepsilon_{i,t} &\sim iid\mathcal{N}\left(0, \sigma_\varepsilon^2\right) \qquad i = 1, \ldots, N \qquad t = 1, \ldots, T \end{aligned} \tag{8}$$

where $f_{[k]}s$ represent unknown smooth functions of the univariate terms. Namely, $f_1(.)$ - $f_2(.)$ - captures the smooth effect of water endowment (measured by the RIVER variables) for those provinces which do not - do have - have waterfalls; while $\sum_t^T f_{3,t}(.)$ capture the smooth effect of market potential in each period. $h(no_i, e_i)$ is a spatial trend term capturing the effect of unobserved spatial heterogeneity, $\gamma_t$ is a time fixed effect capturing common shocks, and $\varepsilon_{i,t}$ is a random error term assumed to be *iid*.[4] In more restrictive specifications, we will we impose the assumption of temporal homogeneity of the effect of market potential and we will impose linearity in the functional form. We will formally test this restrictions.

For the manufacturing sector as a whole, Table .. gives the estimation results of the two semi-parametric model with and without the assumption of time heterogeneity of the effect of market potential (Semipar. M.1 and Semipar. M.2, respectively), and of their parametric counterparts (Param. M.1 and Param. M.2). For the semiparametric models, the table shows the results of the F test for the overall significance of the smooth terms (and the corresponding p-value) as well as the estimated degree of freedom (*edf*), a broad measure of nonlinearity (an *edf* equal to 1 indicates linearity, while a value higher than 1 indicates nonlinearity). For the parametric linear model specifications, we report the estimated coefficients and the corresponding p-values. As mentioned above, all the estimated models include time dummies the spatial trend, but the results for these controls are not reported (they are available upon request).

The estimation results of semiparametric model 1, obtained by pooling over the 69 provinces and the 4 years in our sample, clearly corroborate the hypothesis that both water abundance and

[4]A part from the semiparametric form, this specification is different from the one used in the related literature, for example by Wolf (2007), Ellison and Glaeser (1999), and Midelfart-Knarvik, Overman, and Venables (2000, 2001). In those studies th authors pool all the data by regions, sectors and time and regress the location quotient on a set of interactions between the vectors of location characteristics (factors endowment and market potential) and a vector of industry characteristics (measuring industries' factor intensities and the share of intermediate inputs in GDP). In our case, we do not need to consider such an interaction, since we do not pool the data and estimate a separate model for each sector.

market potential affected the location of industry over the sample period. The F tests indicate that the three smooth terms enter significantly the model, while the *edf* suggest that their effect is nonlinear. The linearity assumption is also confirmed by a more formal F test, comparing Semipar. M1 and Param. M1 (see Table ..). The positive relationship between industrial location and market potential confirms that even at the early period of unification in Italy firms tended to settle at the regions with the highest market potential to minimize costs. The results of the Semiparam. M2 and of its parametric counterpart however show that the effect of market potential did not remain stable over time, but increased, suggesting that the process of economic integration progressively affected the location of industry through the removal of trade barriers. As discussed in Section .. it was the North West the area with the highest market potential and the integration process improved its relative position within the domestic market.

Tables .. report the results for the three main industries, namely Textile, Engineering and Food as well as for the other sectors aggregated in two groups, high K/L and low K/L sectors.

Moreover, we examine whether the respective impact of these two factors changed over time due to the ongoing process of economic integration (or other time-specific factors).

# 5 Conclusions

We test the relative effect of domestic market potential and factor endowment, focusing in particular on water supply. The results show that, as transportation costs decreased and barriers to domestic trade were eliminated, Italian provinces became more and more specialized, and manufacturing activity became increasingly concentrated in a few provinces, mostly belonging to the North-West part of the country. The estimation results corroborate our hypothesis that both comparative advantages (water endowment effect) and market potential (home-market effect) have been responsible of the above process of spatial concentration. The location of some traditional industries (such as clothing) was mainly driven by water endowment, while the location of fast growing new sectors (such as engineering, metalmaking, chemicals, and paper) was mainly driven by the domestic market potential.

# Data Appendix

This section describes the sources and methods used to build the provincial dataset used in this paper.

## Manufacturing value added

Industrial value added data are from Ciccarelli and Fenoaltea (2013), duly updated with the new regional estimates by Ciccarelli and Fenoaltea (2014). The basic idea is to allocate to provinces the regional value added estimates at 1911 prices as follows. For each of the 12 manufacturing sectors, provincial labor force shares of the regional total are computed. Sector-specific regional value added estimates are then allocated to provinces using the above province-specific labor force shares. The calculation has been done separately for each benchmark year (1871, 1881, 1901, and 1911). Using formulas:

$$va_{p,j,t} = VA_{r,j,t} * lfsh_{p,j,t}$$

were $p = 1, 2, \ldots, 69$ denotes provinces; $j = 1, 2, \ldots, 12$ denotes manufacturing sectors, $t = 1871, 1881, 1901$, and 1911 denotes the time index, $va$ denote provincial value added, $VA$ denotes regional value added, and $lfsh$ denote the labor force share (of the regional total).

Compare the spatial distribution of industry labor force in 1911 as resulting from the population census and the spatial distribution of industry employment as resulting from the first industrial census in 1911 (see figure 25).

Insert figure 25 about here

## Water endowment

The water endowment variables entering the regression models were obtained as follows.

**Rivers**:

Data on Italian rivers are from the Italian Geographic Military Institute: Istituto Geografico Militare (2007), *Classificazione dei corsi d'acqua d'Italia*, Florence. The source include a detailed list of the most important 100 rivers classified by province. To each river the experts of the IGM assign a score (ranging from 1 to 5) depending on "its length in km, and its socio-economic relevance". The data were inspected again historical sources such as the *Annuario statistico italiano*, various years. Many other sources also provided us useful information. The main consulted are: i) Baccarini A. (1877), *Appunti di statistica idrografica italiana: i fiumi*, Rome; iii) Ministero Agricoltura Industria e Commercio (1884), "Forza idraulica utilizzata nelle diverse provinci d'Italia", Bollettini di notizie agrarie, luglio, pp. 893-896; iii) Nitti, F.S. (1905), *La conquista della forza : l'elettricitÃă a buon mercato : la nazionalizzazione delle forze idrauliche*, Roux e Viarengo; iv) Ministero dei lavori pubblici (1926), *Statistica delle grandi utilizzazioni idrauliche per forza motrice, Rome*;

**Waterfalls**:

Quantitative data (height and main jumps, in meters) on Italian waterfalls are from the book: Pavolini, M. (1995), *Cascate d'Italia*, Udine. The data were compared, whenever possible, against the world waterfalls database (WWD) available at www.worldwaterfalldatabase.com, last accessed March 2015.

## Market potential

The market potential variable is measured by the logarithm of the Harris market potential.

## References


A'Hearn, B., Venables, A.J., (2013), "Regional disparities: Internal geography and external trade", in G. Tniolo (ed.), *The Oxford Handbook of the Italian Economy Since Unification*, Oxford University Press.

Alonso-Villar, O., J.-M. Chamorro-Rivas, X. Gonzalez-Cerdeira (2004), "Agglomeration Economies in Manufacturing Industries: The Case of Spain", *Applied Economics*, 36, pp. 2103-2116.

Arbia, G., Basile, R. and Mantuano M. (2005), "Does spatial concentration foster economic growth ? Empirical evidence from EU regions", in Trivez F. J., Mur J., Angulo A. Kaabia M., and Catalan B. (eds), *Contributions in Spatial Econometrics*, pp. 301-325.

Basile R., Kayam S., Mínguez R. Montero J.M, Mur, J. et al. (2015), "Semiparametric Spatial Autoregressive Geoadditive Models" in Commendatore P., Kayam S. and Kubin I. (eds), *Complexity and Geographical Economics: Topics and Tools*,

Basile R., Mantuano M., (2008), "La concentrazione geografica dell'industria in Italia: 1971-2001", Scienze Regionali, 7, 3, pp. 5-28.

Bardini C. (1997), *Senza carbone nell'etá del vapore*, Milan, Mondadori.

Cafagna, L. (1989), *Dualismo e sviluppo nella storia d'Italia*, Venice, Marsilio.

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.

Ellison G., Glaeser, E. L. (1997), "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach", *Journal of Political Economy*, 105, 5, pp. 889-927.

Midelfart-Knarvik, K., Overman, H. (2002). "Delocation and European integration", *Economic Policy*, pp. 323-59.

Midelfart-Knarvik, K. H., Overman. H. G. and Venables A. J. (2001), "Comparative advantage and economic geography: estimating the determinants of industrial location in the EU," LSE Research Online Documents on Economics 677, London School of Economics and Political Science, LSE Library.

Missiaia, A. (2013), "Market vs. endowment: Explaining early industrial location in Italy (1871-1911)", London School of Economics , mimeo.

Moran, P. A. P. (1950). “Notes on Continuous Stochastic Phenomena”, *Biometrika*, 37, 1, pp. 17-23.

Rosés, J. R. (2003), “Why Isn’t the Whole of Spain Industrialized ? New Economic Geography and Early Industrialization, 1797-1910”, *Journal of Economic History*, pp. 995-1022.

Wolf, N. (2007), “Endowments vs. market potential: What explains the relocation of industry after the Polish reuniïňĄcation in 1918 ?”, *Explorations in Economic History*, 44, pp. 22-42.

| Sectors | 1871 | 1881 | 1901 | 1911 |
|---|---|---|---|---|
| Agriculture | 62.0 | 58.2 | 58.8 | 54.2 |
| Industry | 21.0 | 24.8 | 23.8 | 25.8 |
| ... of which Manufacturing | 15.3 | 16.9 | 16.9 | 18.8 |
| Services | 17.1 | 17.1 | 17.3 | 20.0 |

TABLE 1
Agriculture, Industry and Services: employment shares

| Sectors | 1871 | 1881 | 1901 | 1911 |
|---|---|---|---|---|
| Food | 16.0 | 15.7 | 14.5 | 11.6 |
| Tobacco | 0.2 | 0.1 | 0.1 | 0.1 |
| Textile | 10.2 | 8.5 | 6.5 | 5.6 |
| Clothing | 10.6 | 9.8 | 8.3 | 7.9 |
| Wood | 20.5 | 20.8 | 21.3 | 16.6 |
| Leather | 18.2 | 19.4 | 18.9 | 19.5 |
| Metal | 0.8 | 0.8 | 1.2 | 2.2 |
| Engin | 14.7 | 15.4 | 18.3 | 21.1 |
| NonMet | 5.8 | 6.0 | 7.0 | 9.8 |
| Chem | 0.6 | 0.7 | 0.8 | 1.8 |
| Paper | 1.9 | 2.1 | 2.5 | 3.2 |
| Sundry | 0.7 | 0.7 | 0.6 | 0.6 |

TABLE 2
Manufacturing sectors: employment shares

| Sectors | 1871 | 1881 | 1901 | 1911 |
|---|---|---|---|---|
| Food | 33.5 | 30.4 | 25.4 | 21.5 |
| Tobacco | 1.6 | 1.3 | 0.9 | 0.7 |
| Textile | 10.3 | 10.3 | 12.8 | 11.1 |
| Clothing | 6.9 | 7.4 | 6.8 | 6.3 |
| Wood | 10.0 | 9.3 | 9.7 | 10.0 |
| Leather | 10.5 | 11.5 | 11.4 | 7.8 |
| Metal | 0.6 | 1.0 | 1.7 | 3.1 |
| Engin | 17.5 | 17.8 | 18.6 | 21.5 |
| NonMet | 3.6 | 4.2 | 4.2 | 6.6 |
| Chem | 2.1 | 2.4 | 3.0 | 4.3 |
| Paper | 2.7 | 3.5 | 4.8 | 6.3 |
| Sundry | 0.7 | 0.7 | 0.6 | 0.7 |

TABLE 3
Manufacturing sectors: value added shares

| Year | 1871 | 1881 | 1901 | 1911 |
|---|---|---|---|---|
| Mean | 0.26 | 0.29 | 0.35 | 0.38 |
| St.dev | 0.10 | 0.11 | 0.14 | 0.13 |
| CV | 0.38 | 0.36 | 0.39 | 0.35 |

TABLE 4
Industry specialization of Italian provinces. Krugman index (value added data)

| Sectors | 1871 | 1881 | 1901 | 1911 |
|---|---|---|---|---|
| Food | 0.02 | 0.02 | 0.04 | 0.06 |
| Tobacco | 1.23 | 1.05 | 1.04 | 0.81 |
| Textile | 0.26 | 0.3 | 0.45 | 0.45 |
| Clothing | 0.11 | 0.13 | 0.1 | 0.11 |
| Wood | 0.02 | 0.02 | 0.03 | 0.05 |
| Leather | 0.05 | 0.06 | 0.11 | 0.16 |
| Metal | 0.38 | 0.57 | 0.86 | 0.74 |
| Engin | 0.05 | 0.03 | 0.06 | 0.07 |
| NonMet | 0.17 | 0.17 | 0.19 | 0.09 |
| Chem | 0.18 | 0.16 | 0.23 | 0.17 |
| Paper | 0.21 | 0.21 | 0.17 | 0.16 |
| Sundry | 0.33 | 0.89 | 0.62 | 0.46 |
| Manufacturing | 0.03 | 0.04 | 0.07 | 0.08 |

TABLE 5
Industry concentration: Theil index (value added data)

| Sectors | 1871 | 1881 | 1901 | 1911 |
|---|---|---|---|---|
| Food | 0.15 | 0.31 | 0.25 | 0.29 |
| | 0.02 | 0.00 | 0.00 | 0.00 |
| Tobacco | -0.04 | 0.00 | 0.02 | 0.01 |
| | 0.63 | 0.41 | 0.32 | 0.38 |
| Textile | 0.33 | 0.38 | 0.47 | 0.49 |
| | 0.00 | 0.00 | 0.00 | 0.00 |
| Clothing | 0.29 | 0.30 | 0.40 | 0.35 |
| | 0.00 | 0.00 | 0.00 | 0.00 |
| Wood | 0.11 | 0.10 | 0.22 | 0.37 |
| | 0.05 | 0.06 | 0.00 | 0.00 |
| Leather | 0.60 | 0.62 | 0.69 | 0.73 |
| | 0.00 | 0.00 | 0.00 | 0.00 |
| Metal | 0.28 | 0.26 | 0.17 | 0.17 |
| | 0.00 | 0.00 | 0.00 | 0.00 |
| Engin | 0.17 | 0.19 | 0.12 | 0.02 |
| | 0.01 | 0.00 | 0.03 | 0.34 |
| NonMet | 0.20 | 0.08 | 0.12 | 0.11 |
| | 0.00 | 0.03 | 0.00 | 0.02 |
| Chem | -0.04 | 0.15 | 0.00 | 0.05 |
| | 0.61 | 0.02 | 0.40 | 0.19 |
| Paper | 0.11 | 0.07 | 0.06 | 0.07 |
| | 0.05 | 0.12 | 0.17 | 0.13 |
| Sundry | 0.05 | -0.00 | 0.01 | 0.15 |
| | 0.19 | 0.37 | 0.36 | 0.01 |
| Manufacturing | 0.05 | -0.00 | 0.01 | 0.15 |
| | 0.47 | 0.41 | 0.09 | 0.02 |

TABLE 6
Spatial autocorrelation: global Moran's I index (value added data)

| Variable | Year | Semipar. M.1 F test | Semipar. M.1 edf | Semipar. M.2 F test | Semipar. M.2 edf | Param. M.1 Coeff | Param. M.2 Coeff |
|---|---|---|---|---|---|---|---|
| ln(River)\|Waterfalls=0 | Pooled | 4.247 | 1.905 | 4.080 | 1.953 | 0.042 | 0.043 |
| | | (0.014) | | (0.016) | | (0.135) | (0.128) |
| ln(River)\|Waterfalls=1 | Pooled | 15.940 | 3.474 | 15.060 | 3.544 | 0.017 | 0.018 |
| | | (0.000) | | (0.000) | | (0.279) | (0.264) |
| ln(Mktpot) | Pooled | 62.228 | 2.677 | | | 0.441 | |
| | | (0.000) | | | | (0.000) | |
| | 1871 | | | 19.143 | 2.461 | | 0.391 |
| | | | | (0.000) | | | (0.000) |
| | 1881 | | | 20.062 | 2.939 | | 0.384 |
| | | | | (0.000) | | | (0.000) |
| | 1901 | | | 26.364 | 3.074 | | 0.445 |
| | | | | (0.000) | | | (0.000) |
| | 1911 | | | 116.770 | 1.000 | | 0.524 |
| | | | | (0.000) | | | (0.000) |
| R-sq.(adj.) | | 0.699 | | 0.713 | | 0.614 | 0.619 |
| Dev.expl.(%) | | 0.732 | | 0.752 | | 0.647 | 0.656 |

TABLE 7
Estimation results. Whole manufacturing

| Variable | Year | Semipar. M.1 F test | Semipar. M.1 edf | Semipar. M.2 F test | Semipar. M.2 edf | Param. M.1 Coeff | Param. M.2 Coeff |
|---|---|---|---|---|---|---|---|
| ln(River)\|Waterfalls=0 | Pooled | 5.610 | 1.000 | 5.304 | 1.000 | -0.140 | -0.140 |
| | | (0.019) | | (0.022) | | (0.014) | (0.015) |
| ln(River)\|Waterfalls=1 | Pooled | 1.682 | 1.000 | 1.346 | 1.000 | 0.020 | 0.020 |
| | | (0.196) | | (0.247) | | (0.523) | (0.521) |
| ln(Mktpot) | Pooled | 55.586 | 2.086 | | | 0.722 | |
| | | (0.000) | | | | (0.000) | |
| | 1871 | | | 34.836 | 1.326 | | 0.791 |
| | | | | (0.000) | | | (0.000) |
| | 1881 | | | 27.864 | 1.558 | | 0.758 |
| | | | | (0.000) | | | (0.000) |
| | 1901 | | | 22.230 | 1.647 | | 0.691 |
| | | | | (0.000) | | | (0.000) |
| | 1911 | | | 18.200 | 1.573 | | 0.668 |
| | | | | (0.000) | | | (0.000) |
| R-sq.(adj.) | | 0.501 | | 0.491 | | 0.484 | 0.480 |
| Dev.expl.(%) | | 0.538 | | 0.535 | | 0.519 | 0.521 |

TABLE 8
Estimation results. High K/L ratio sectors

| Variable | Year | Semipar. M.1 F test | Semipar. M.1 edf | Semipar. M.2 F test | Semipar. M.2 edf | Param. M.1 Coeff | Param. M.2 Coeff |
|---|---|---|---|---|---|---|---|
| ln(River)\|Waterfalls=0 | Pooled | 1.694 | 2.702 | 2.480 | 2.903 | 0.028 | 0.027 |
| | | (0.166) | | (0.056) | | (0.152) | (0.163) |
| ln(River)\|Waterfalls=1 | Pooled | 4.276 | 2.295 | 3.586 | 2.245 | 0.010 | 0.010 |
| | | (0.009) | | (0.021) | | (0.326) | (0.315) |
| ln(Mktpot) | Pooled | 34.705 | 2.029 | | | -0.197 | |
| | | (0.000) | | | | (0.000) | |
| | 1871 | | | 7.759 | 1.000 | | -0.119 |
| | | | | (0.006) | | | (0.003) |
| | 1881 | | | 12.497 | 1.000 | | -0.146 |
| | | | | (0.000) | | | (0.000) |
| | 1901 | | | 36.413 | 1.000 | | -0.234 |
| | | | | (0.000) | | | (0.000) |
| | 1911 | | | 51.216 | 1.000 | | -0.270 |
| | | | | (0.000) | | | (0.000) |
| R-sq.(adj.) | | 0.601 | | 0.604 | | 0.572 | 0.585 |
| Dev.expl.(%) | | 0.629 | | 0.632 | | 0.593 | 0.610 |

TABLE 9
Estimation results. Low K/L ratio sectors

| Variable | Year | Semipar. M.1 F test | Semipar. M.1 edf | Semipar. M.2 F test | Semipar. M.2 edf | Param. M.1 Coeff | Param. M.2 Coeff |
|---|---|---|---|---|---|---|---|
| ln(River)\|Waterfalls=0 | Pooled | 0.837 | 1.685 | 0.771 | 1.663 | -0.036 | -0.036 |
| | | (0.429) | | (0.456) | | (0.698) | (0.698) |
| ln(River)\|Waterfalls=1 | Pooled | 5.368 | 3.059 | 4.616 | 3.054 | 0.091 | 0.091 |
| | | (0.001) | | (0.003) | | (0.070) | (0.069) |
| ln(Mktpot) | Pooled | 11.636 | 1.000 | | | 0.414 | |
| | | (0.001) | | | | (0.000) | |
| | 1871 | | | 4.008 | 1.000 | | 0.438 |
| | | | | (0.046) | | | (0.027) |
| | 1881 | | | 3.209 | 1.428 | | 0.420 |
| | | | | (0.051) | | | (0.031) |
| | 1901 | | | 4.608 | 1.000 | | 0.428 |
| | | | | (0.033) | | | (0.022) |
| | 1911 | | | 3.675 | 1.000 | | 0.374 |
| | | | | (0.056) | | | (0.040) |
| R-sq.(adj.) | | 0.642 | | 0.639 | | 0.615 | 0.611 |
| Dev.expl.(%) | | 0.666 | | 0.668 | | 0.636 | 0.636 |

TABLE 10
Estimation results. Textile

| Variable | Year | Semipar. M.1 F test | Semipar. M.1 edf | Semipar. M.2 F test | Semipar. M.2 edf | Param. M.1 Coeff | Param. M.2 Coeff |
|---|---|---|---|---|---|---|---|
| ln(River)\|Waterfalls=0 | Pooled | 1.098 | 1.771 | 0.764 | 1.651 | -0.049 | -0.049 |
| | | (0.335) | | (0.458) | | (0.060) | (0.051) |
| ln(River)\|Waterfalls=1 | Pooled | 6.580 | 2.896 | 7.655 | 2.928 | -0.054 | -0.055 |
| | | (0.000) | | (0.000) | | (0.000) | (0.000) |
| ln(Mktpot) | Pooled | 21.622 | 2.688 | | | 0.102 | |
| | | (0.000) | | | | (0.001) | |
| | 1871 | | | 0.798 | 1.623 | | -0.046 |
| | | | | (0.442) | | | (0.385) |
| | 1881 | | | 1.452 | 1.725 | | 0.024 |
| | | | | (0.236) | | | (0.641) |
| | 1901 | | | 6.682 | 1.974 | | 0.115 |
| | | | | (0.001) | | | (0.022) |
| | 1911 | | | 18.140 | 2.169 | | 0.281 |
| | | | | (0.000) | | | (0.000) |
| R-sq.(adj.) | | 0.445 | | 0.437 | | 0.322 | 0.375 |
| Dev.expl.(%) | | 0.485 | | 0.486 | | 0.353 | 0.411 |

TABLE 11
Estimation results. Engineering

| Variable | Year | Semipar. M.1 F test | Semipar. M.1 edf | Semipar. M.2 F test | Semipar. M.2 edf | Param. M.1 Coeff | Param. M.2 Coeff |
|---|---|---|---|---|---|---|---|
| ln(River)\|Waterfalls=0 | Pooled | 5.311 | 3.520 | 3.726 | 3.437 | 0.083 | 0.083 |
| | | (0.001) | | (0.007) | | (0.001) | (0.001) |
| ln(River)\|Waterfalls=1 | Pooled | 4.015 | 2.806 | 3.811 | 2.796 | -0.022 | -0.022 |
| | | (0.008) | | (0.010) | | (0.110) | (0.105) |
| ln(Mktpot) | Pooled | 24.353 | 1.718 | | | -0.202 | |
| | | (0.000) | | | | (0.000) | |
| | 1871 | | | 10.271 | 1.000 | | -0.162 |
| | | | | (0.002) | | | (0.001) |
| | 1881 | | | 6.296 | 1.448 | | -0.146 |
| | | | | (0.004) | | | (0.003) |
| | 1901 | | | 16.360 | 1.000 | | -0.189 |
| | | | | (0.000) | | | (0.000) |
| | 1911 | | | 41.893 | 1.000 | | -0.291 |
| | | | | (0.000) | | | (0.000) |
| R-sq.(adj.) | | 0.526 | | 0.530 | | 0.493 | 0.501 |
| Dev.expl.(%) | | 0.569 | | 0.577 | | 0.530 | 0.543 |

TABLE 12
Estimation results. Food

| Models | Whole Man. | High K/L | Low K/L | Textile | Engineering | Food |
|---|---|---|---|---|---|---|
| Semipar. M.1 vs. Semipar. M.2 | 2.829 | NA | 5.121 | 0.404 | 0.181 | 1.853 |
| | (0.008) | (NA) | (0.050) | (0.777) | (0.960) | (0.147) |
| Semipar. M.1 vs. Param. M.1 | 12.087 | 6.857 | 4.898 | 6.209 | 9.448 | 4.756 |
| | (0.000) | (0.004) | (0.000) | (0.000) | (0.000) | (0.000) |
| Param. M.1 vs. Param. M.2 | 2.119 | 0.302 | 3.857 | NA | 8.246 | 2.327 |
| | (0.097) | (0.823) | (0.011) | (NA) | (0.000) | (0.075) |
| Semipar. M.2 vs. Param. M.2 | 9.086 | 3.199 | 5.979 | 5.777 | 4.426 | 4.676 |
| | (0.000) | (0.033) | (0.001) | (0.000) | (0.000) | (0.001) |

TABLE 13
Model comparison

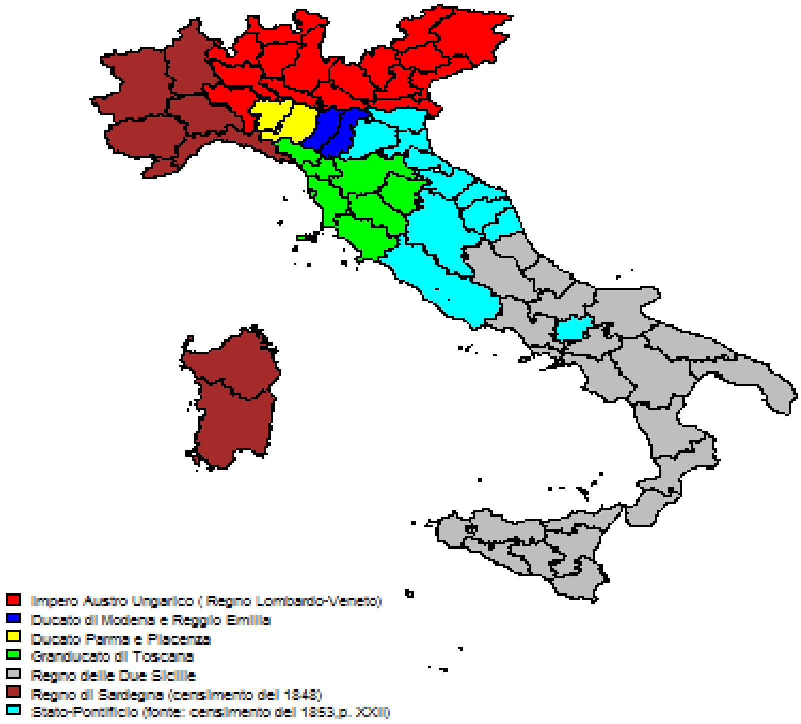



FIGURE 1
Italian provinces grouped into pre-unitarian States (1850 ca.)

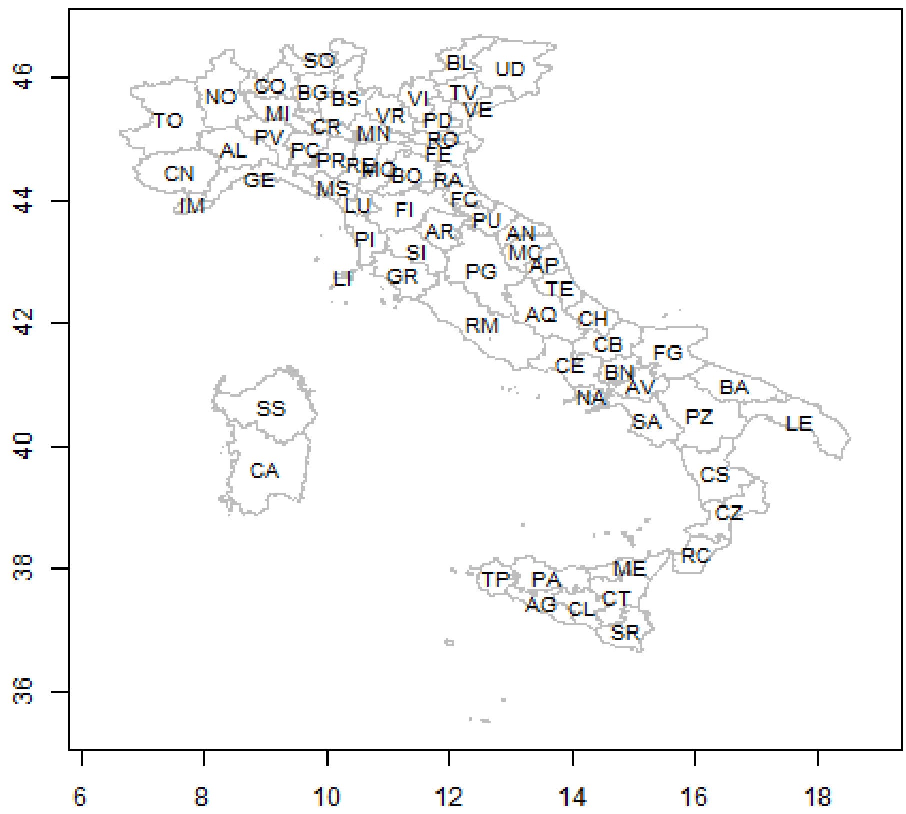
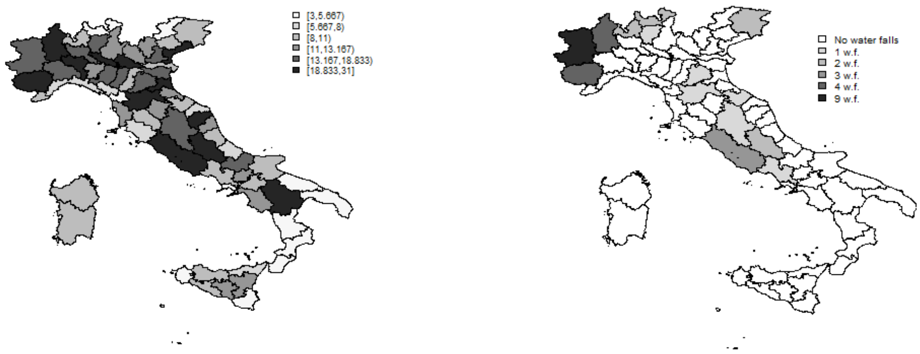


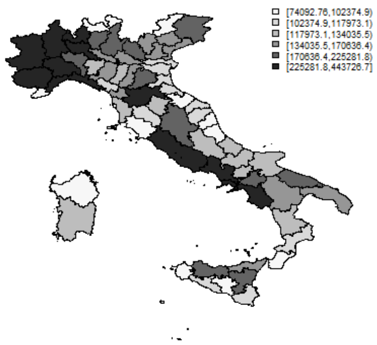

FIGURE 2
Italian provinces



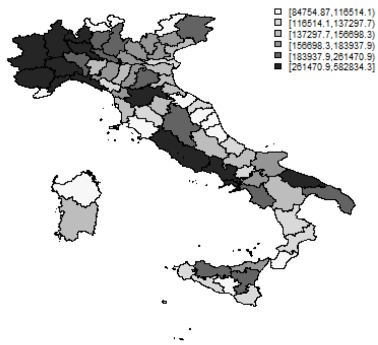

(a) Rivers

(b) Water falls

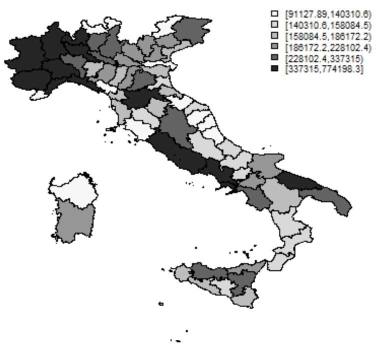
FIGURE 3
Water endowment



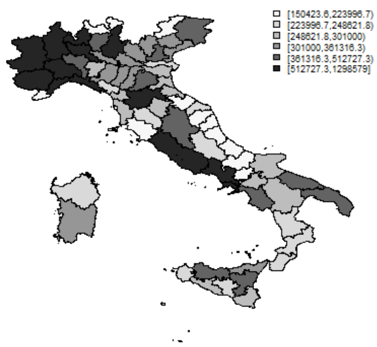

(a) Year: 1871




(b) Year: 1881

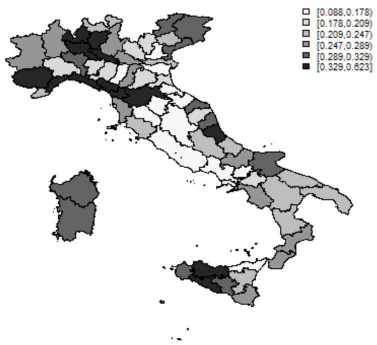



(c) Year: 1901

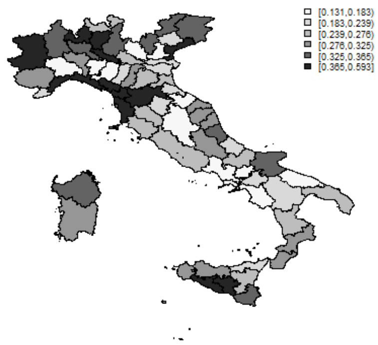



(d) Year: 1911

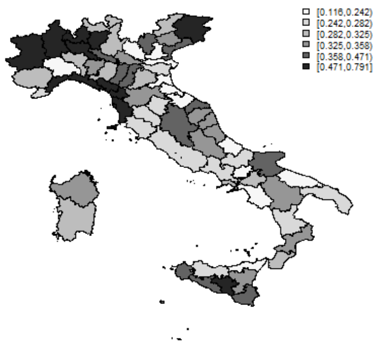
FIGURE 4
Harris market potential



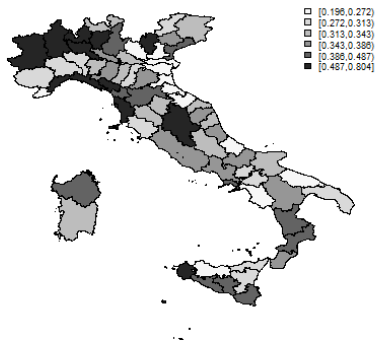

(a) Year: 1871




(b) Year: 1881

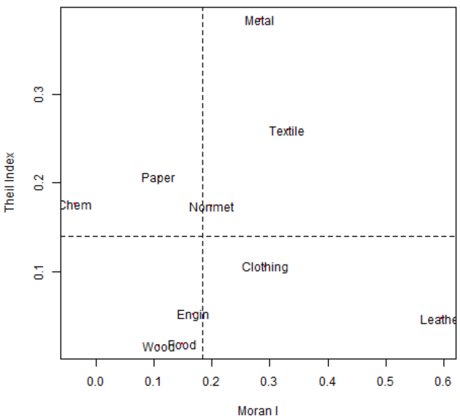



(c) Year: 1901

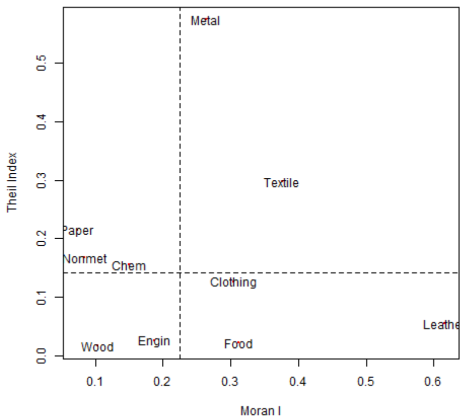



(d) Year: 1911

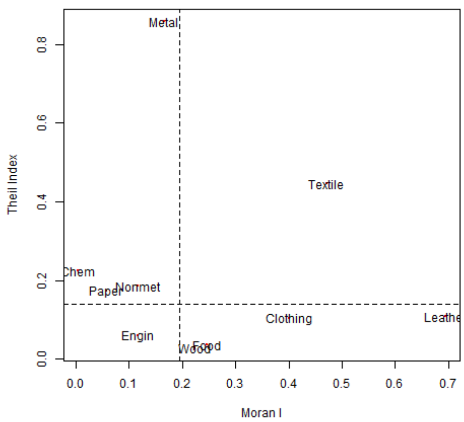
FIGURE 5
Krugman specialization index



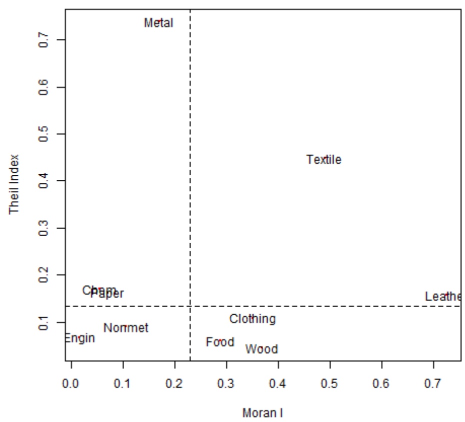

(a) Year: 1871




(b) Year: 1881

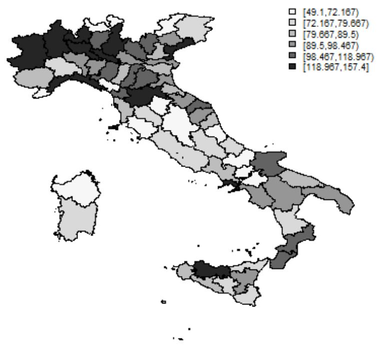



(c) Year: 1901

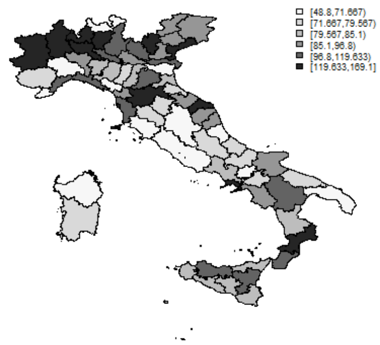



(d) Year: 1911

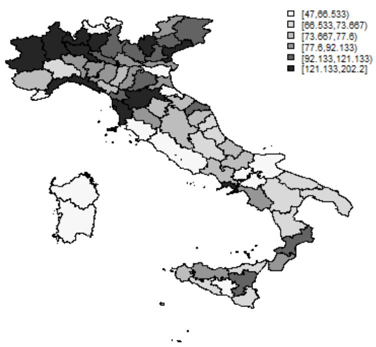
FIGURE 6
Scatterplot between the a-spatial Theil index of concentration and the global I Moran index of spatial autocorrelation. Value added data



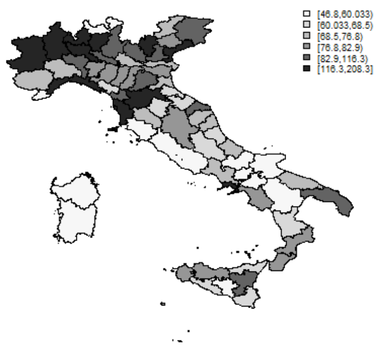

(a) Year: 1871




(b) Year: 1881

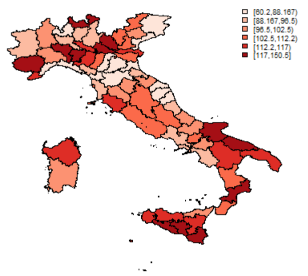



(c) Year: 1901

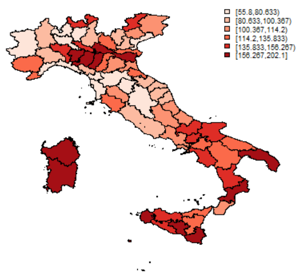



(d) Year: 1911

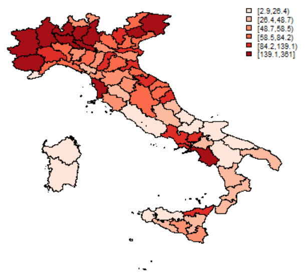
FIGURE 7
Specialization in manufacturing



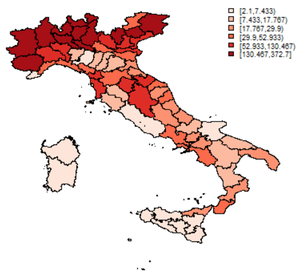

(a) FOOD: 1871




(b) FOOD: 1911

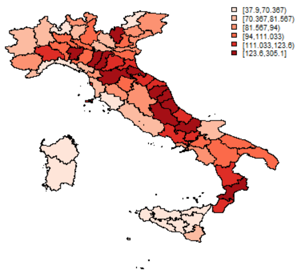



(c) TEXTILE: 1871

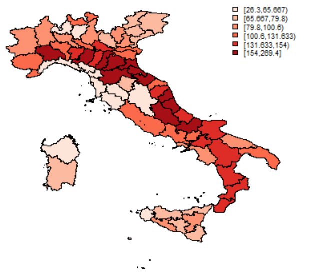



(d) TEXTILE: 1911

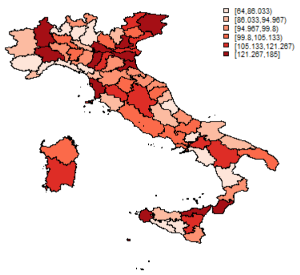
FIGURE 8
Choroplet maps of LQ values. Value added data



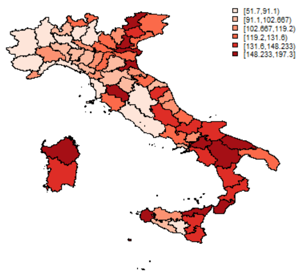

(a) CLOTHING: 1871




(b) CLOTHING: 1911

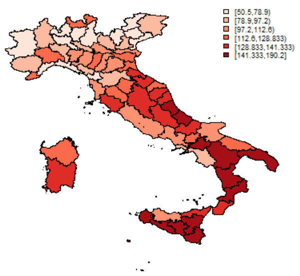



(c) WOOD: 1871

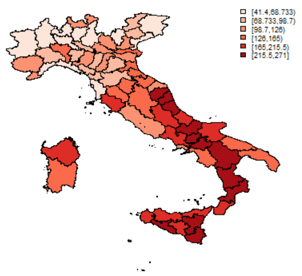



(d) WOOD: 1911

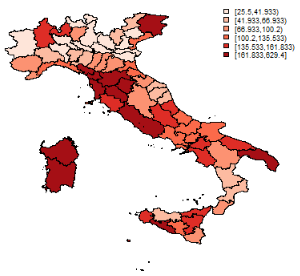
FIGURE 9
Choroplet maps of LQ values. Value added data



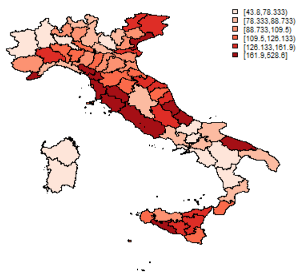

(a) LEATHER: 1871




(b) LEATHER: 1911

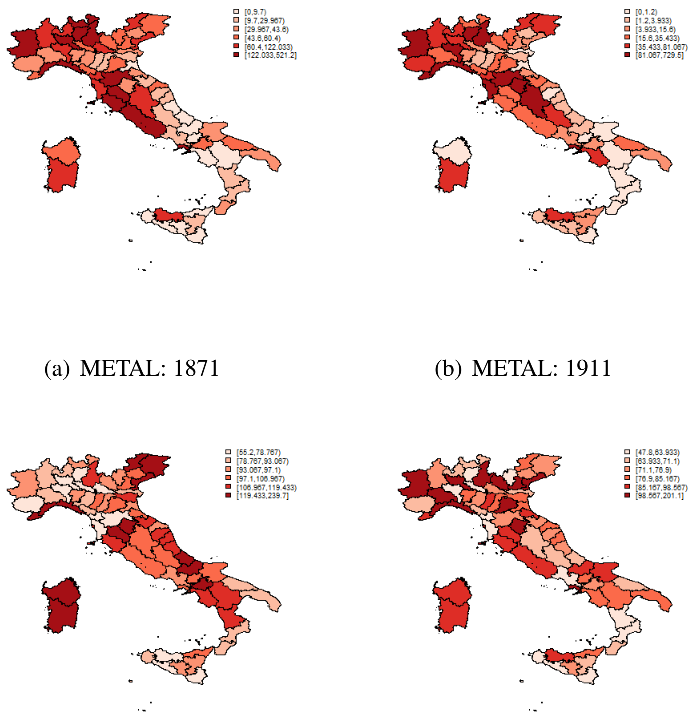



(c) NONMET: 1871




(d) NONMET: 1911

FIGURE 10
Choroplet maps of LQ values. Value added data



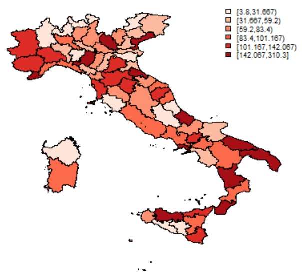

(a) METAL: 1871

(b) METAL: 1911

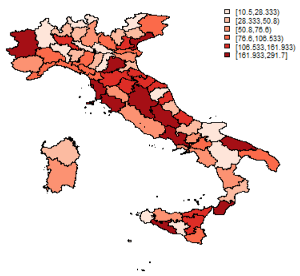
(c) ENGIN: 1871

(d) ENGIN: 1911

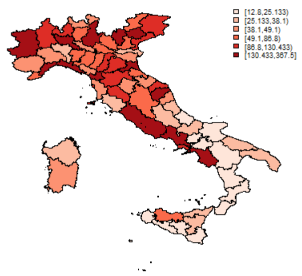
FIGURE 11
Choroplet maps of LQ values. Value added data



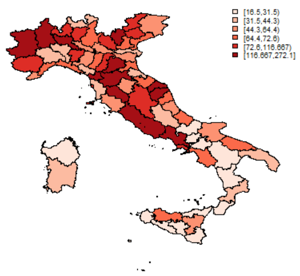

(a) CHEM: 1871




(b) CHEM: 1911

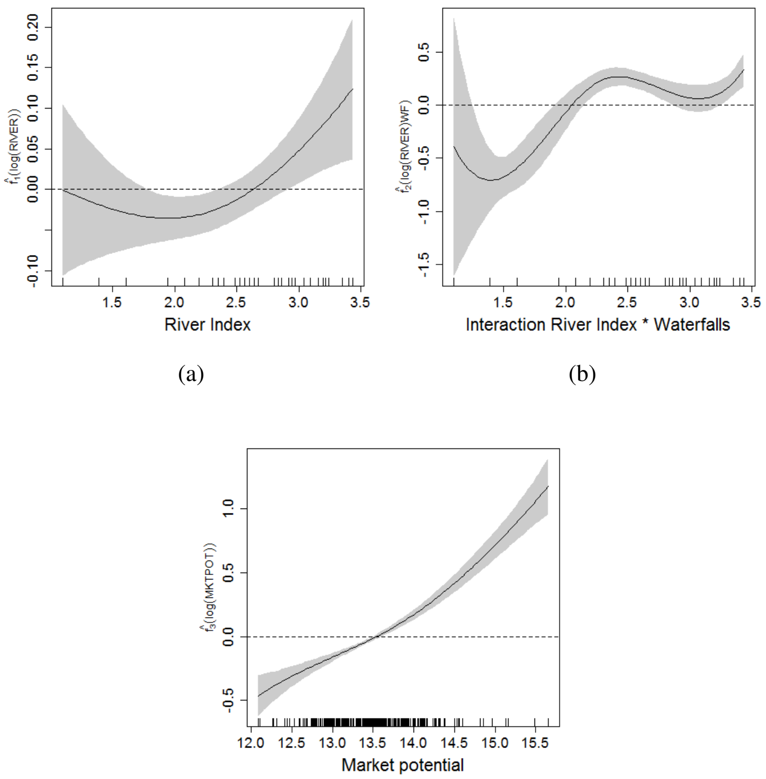



(c) PAPER: 1871




(d) PAPER: 1911

FIGURE 12
Choroplet maps of LQ values. Value added data




(a) (b) (c)

FIGURE 13

**Whole manufacturing**. Smooth effects from semipar. M.1

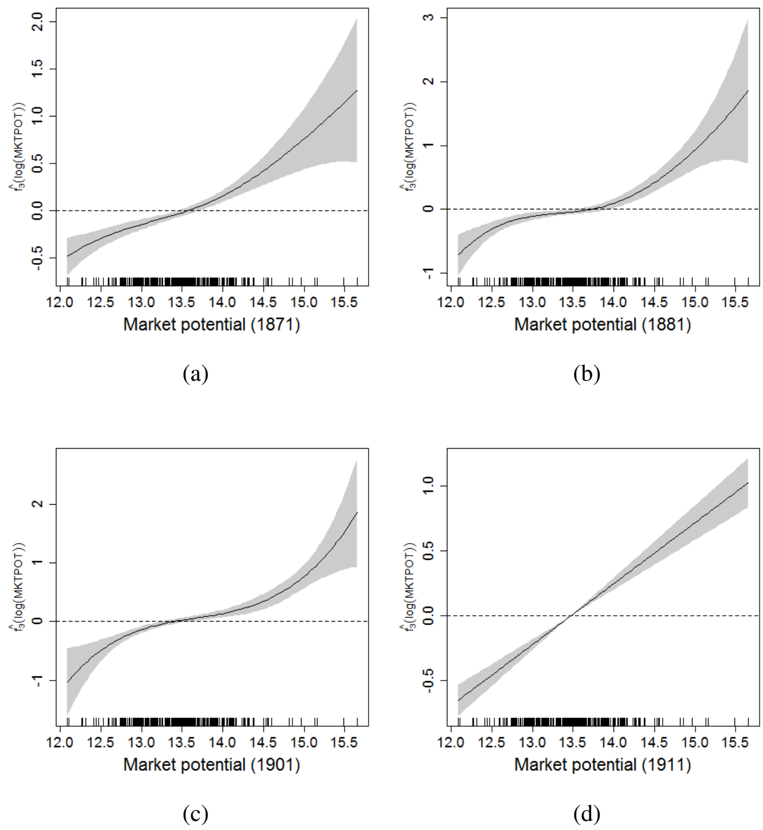
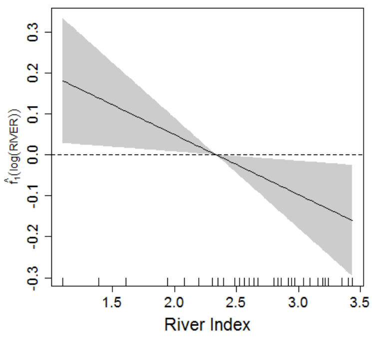
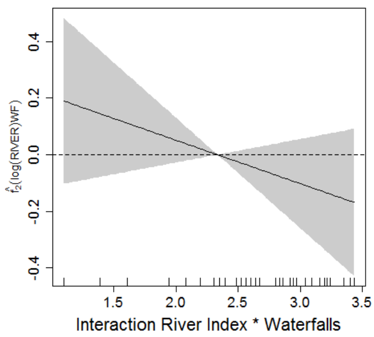



FIGURE 14

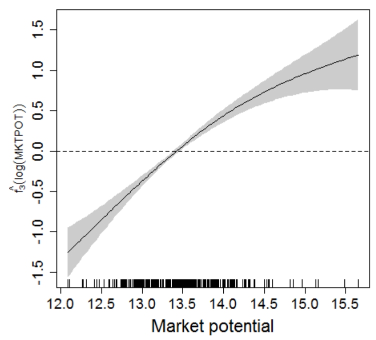
**Whole manufacturing**. Smooth effects from semipar. M.2




(a)




(b)




(c)

FIGURE 15

**High K/L sectors**. Smooth effects from semipar. M.1

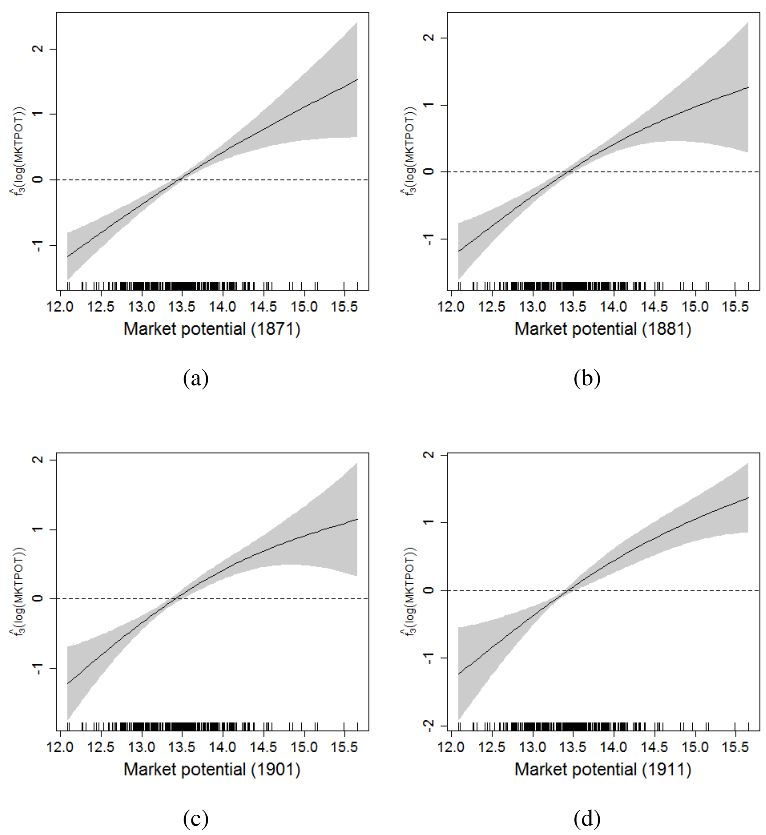
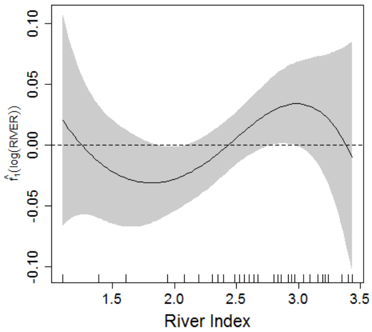
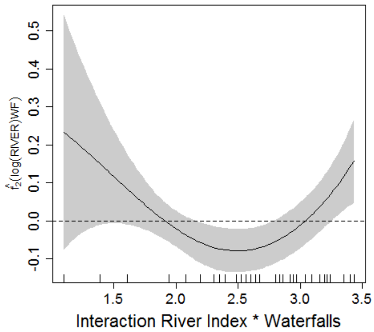



FIGURE 16

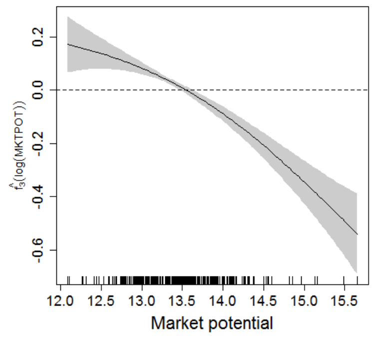
**High K/L sectors**. Smooth effects from semipar. M.2




(a)




(b)




(c)

FIGURE 17

**Low K/L sectors**. Smooth effects from semipar. M.1

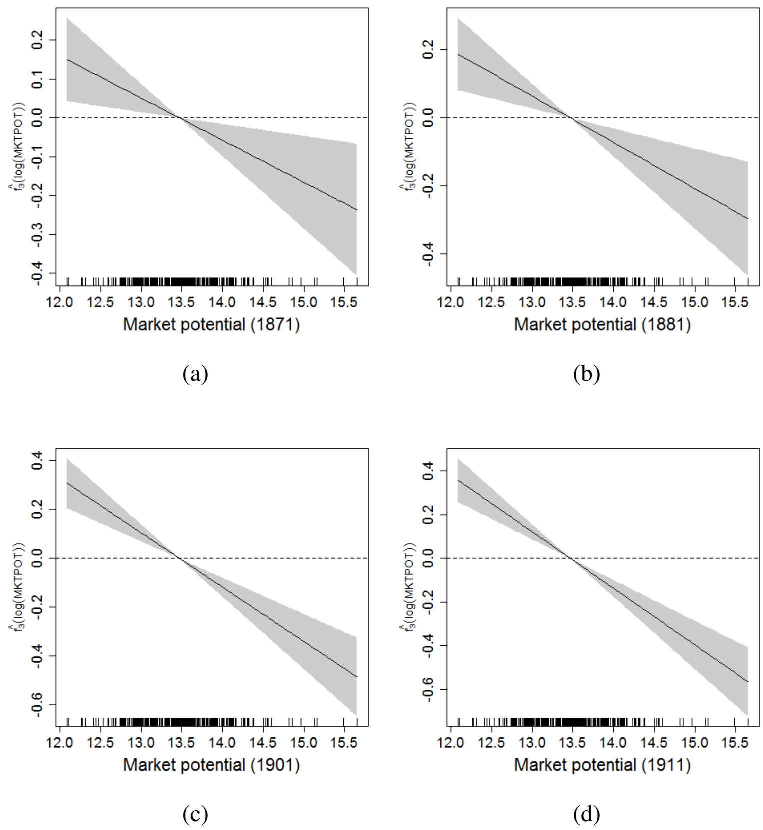
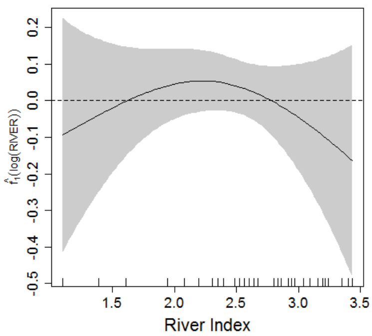
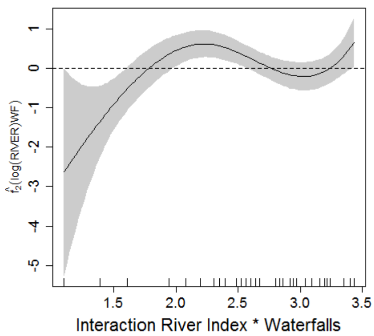


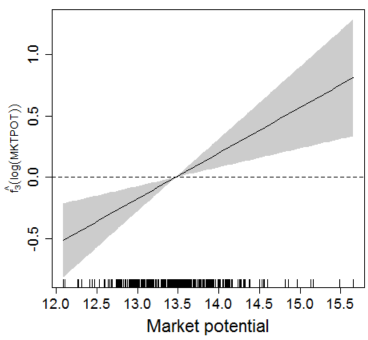

FIGURE 18
**Low K/L sectors**. Smooth effects from semipar. M.2




(a)




(b)




(c)

FIGURE 19

**Textile**. Smooth effects from semipar. M.1

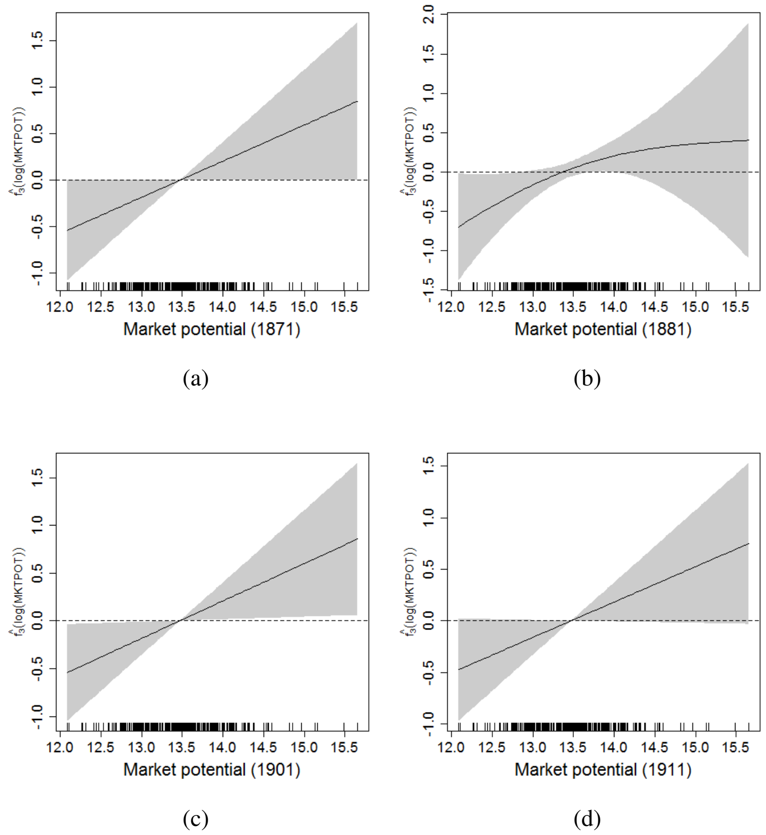
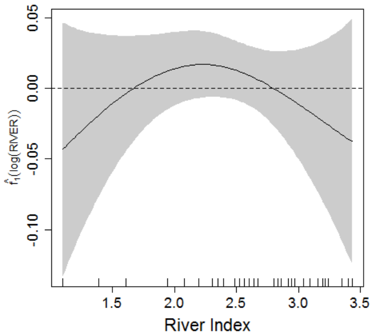
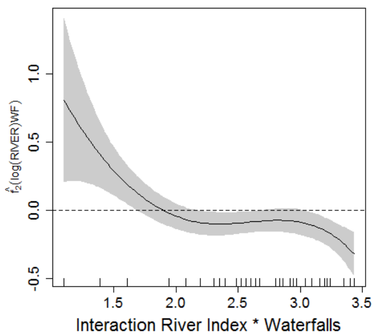



FIGURE 20

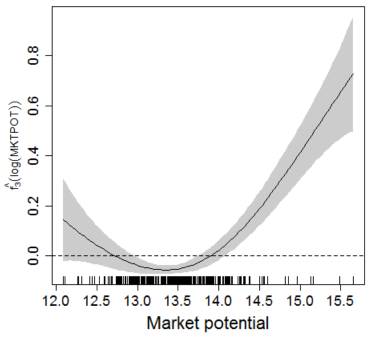
**Textile**. Smooth effects from semipar. M.2




(a)




(b)




(c)

FIGURE 21

**Engineering**. Smooth effects from semipar. M.1

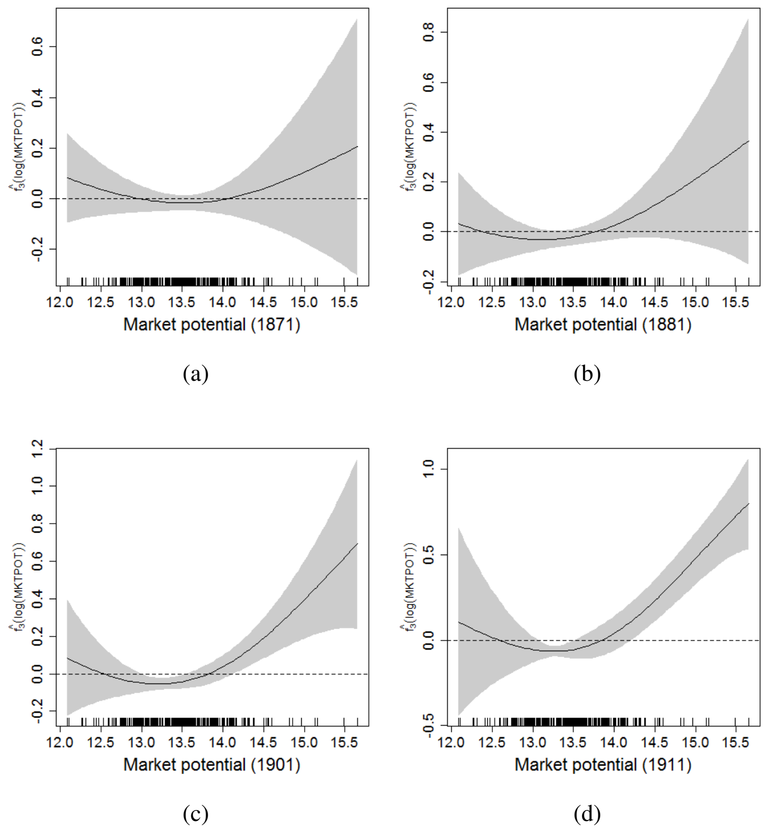
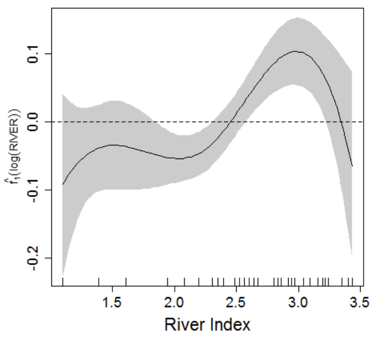
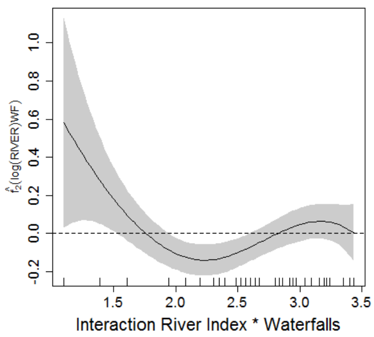


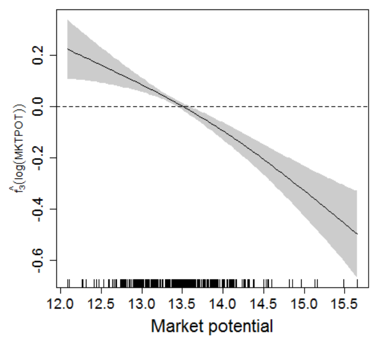

FIGURE 22
**Engineering**. Smooth effects from semipar. M.2




(a)




(b)




(c)

FIGURE 23

**Food**. Smooth effects from semipar. M.1

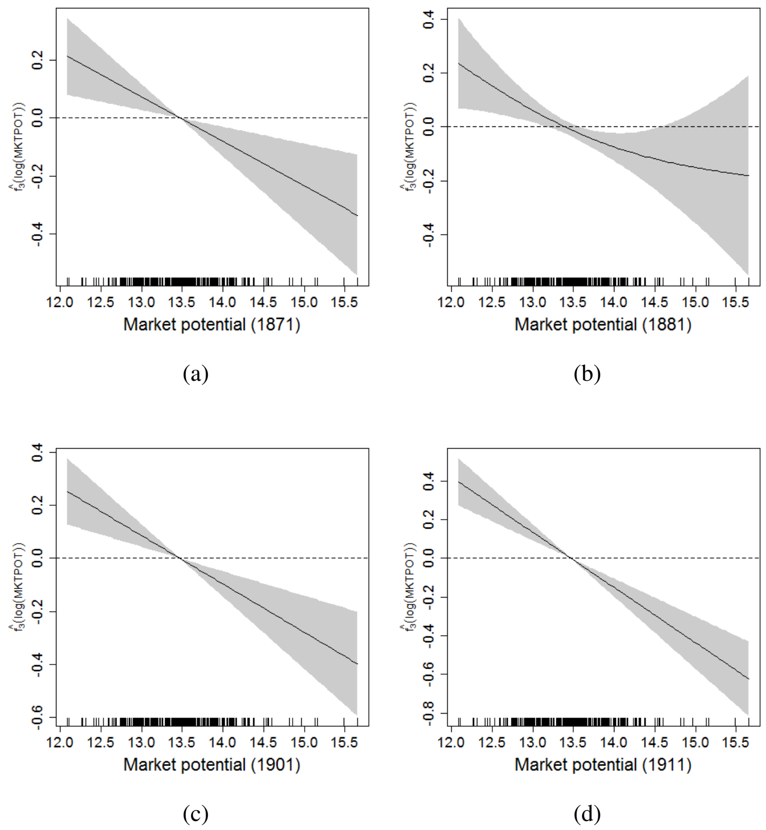
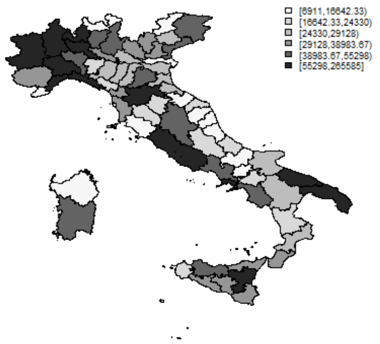
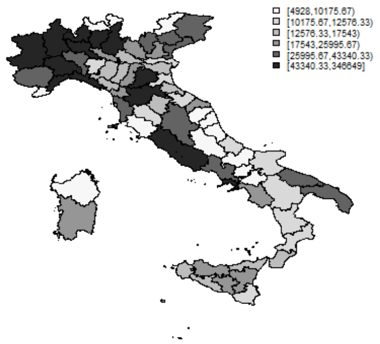


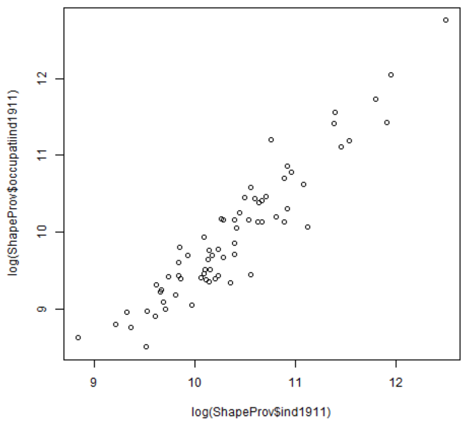

FIGURE 24
**Food**. Smooth effects from semipar. M.2



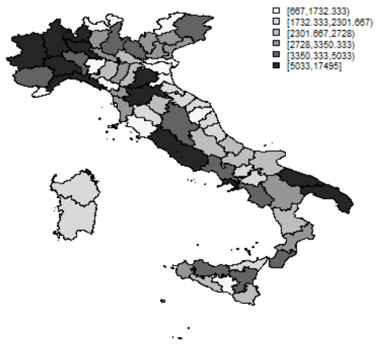

(a) Industry labor force 1911 (Pop. census)




(b) Industry employment 1911 (Ind. census)

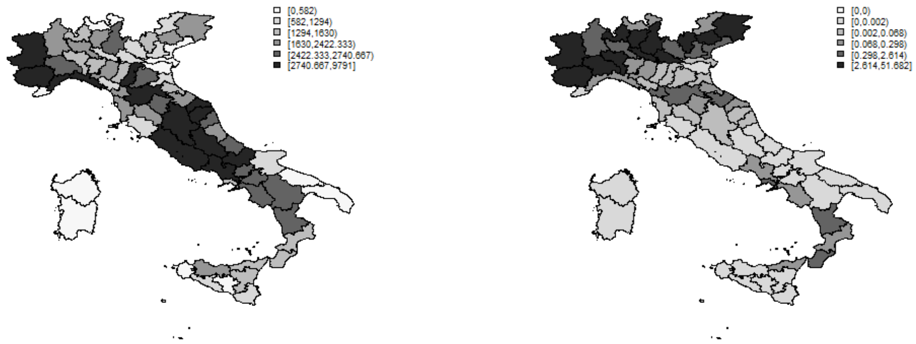



(c) Relation pop.census and ind. census




(d) Industry firms 1911

FIGURE 25
Spatial distribution of industry according to the two censuses




(a) Potenza mulini

(b) Filande

FIGURE 26
Mills and spinning mills