ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Missiaia, Anna

Working Paper

Where do we go from here? Market access and regional development in Italy (1871 - 1911)

LEM Working Paper Series, No. 2015/18

Provided in Cooperation with:
Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies

Suggested Citation: Missiaia, Anna (2015) : Where do we go from here? Market access and regional development in Italy (1871 - 1911), LEM Working Paper Series, No. 2015/18, Scuola Superiore Sant'Anna, Laboratory of Economics and Management (LEM), Pisa

This Version is available at:
https://hdl.handle.net/10419/119861

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

INSTITUTE OF ECONOMICS

Scuola Superiore Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM
WORKING PAPER SERIES

Where Do We Go From Here? Market Access and Regional Development in Italy (1871-1911)

Anna Missiaia °

° Lund University, Sweden

2015/18 **June 2015**

ISSN(ONLINE) 2284-0400

Where Do We Go From Here? Market Access and Regional Development in Italy (1871–1911)

Anna Missiaia*

Lund University

Abstract

Italy has been characterized, throughout its history as a unified country, by large regional differentials in the levels of income, industrialization and socio-economic development. This paper aims at testing the New Economic Geography hypothesis on the role of market access in explaining these regional differentials. We first quantify market access of the Italian regions for benchmark years from 1871 to 1911 following Harris (1954). We then use these estimates to study the causal link between GDP per capita and market potential following Head and Mayer (2011). The main result of this paper is that only domestic market potential, which represents the home market, shows a "traditional" North-South divide. When international markets are introduced, the South does not appear to lag behind. Regression analysis confirms that market potential is a strong determinant of GDP per capita only in its domestic formulation. This suggests that the home market in this period mattered far more for growth than the international markets, casting new light on one of the classical explanations to the North-South divide.

JEL Classification Numbers: N33, N73, N93, O18, R11.

Keywords: Economic Geography; Economic History of Italy; Market Potential.

*Lund University, Economic History Department, Room 2110, Alfa 1, Scheelevägen 15 B, Lund, Skåne 22363, Sweden, e-mail:*anna.missiaia@ekh.lu.se*. The author is grateful to Brian A'Hearn, Carlo Ciccarelli, Emanuele Felice, Giovanni Federico, Stefano Fenoaltea, Steve Gibbons and Max Schulze for advice and suggestions, and to participants of the seminars at the London School of Economics, Bocconi University, EHS Annual Conference, Carlos III Summer Schools and European Business History Association Summer School, FRESH Winter School for helpful comments.

1 Introduction

Italy's political Unification dates back to 1861. It was the result of two independence wars against the Habsburg Empire and the Expedition of the Thousand through which Southern regions were annexed to the newly unified North. At the time of Unification, the regions of Italy were not at all homogeneous in terms of their economic structure and performance. Although they did not show as large a North-South divide in terms of GDP per capita as today's, it is in this period that the gap starts widening (see Figure 1).

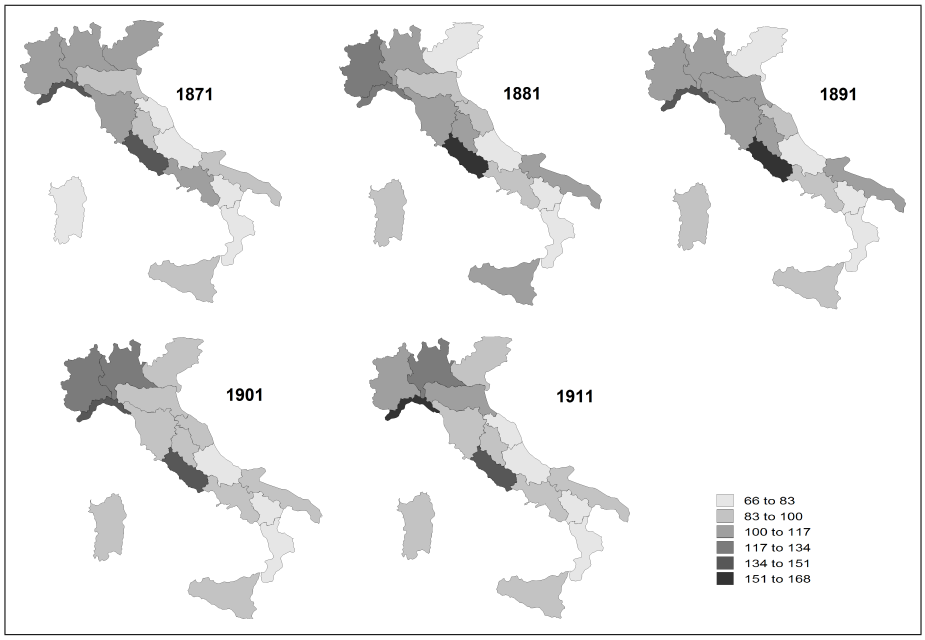
It is in the 1880s that the Italian regions start to polarize in terms of GDP per capita and even more sharply in the level of industrial production. During the first decades after Unification, most of the economic gap was between the Northwestern regions (Piedmont, Lombardy and Liguria) versus the rest of the country. If we look at the indicators of economic development, such as the literacy rate, the level of regional inequality appears evident at the time of Unification. Figure 2 shows the literacy rates for each region in the benchmark years 1871, 1881, 1901 and 1911. Finally, looking at the industrial value added per capita (which corresponds specifically to the part of GDP produced by industry) in Figure 3, the formation of the gap during this period is much more evident, due to the process of industrialization which was largely concentrated in the Northwestern regions of Piedmont, Lombardy and Liguria.

The historiography has brought forward several explanations for what is often referred to as the "Questione Meridionale", the Southern Question. The hypotheses proposed by scholars range from colonial exploitation of the South by the North (Villari (1979) and Sonnino (1877)), to differences in agricultural structure (Cafagna (1989) and Zamagni (1990)), to differences in the institutional framework (Felice and Vasta (2012)) all the way to cultural differences (Galassi (2000) and A'Hearn (2000)).[1]

Another line of research has been focusing on the physical geography of Italy as the primary cause of the backwardness of the South. For instance, Fenoaltea (2006, p. 261) gives a possible explanation for the regional patterns of industrialization in Italy before the First World War based on the comparative advantages of the North, in particular in terms of energy endowment from water, but he does not translate it into a formal model. Daniele and Malanima (2007, p. 181) discuss the role of the physical distance of the southern regions to the centre of Europe and they claim that the position of the South constituted a natural disadvantage for its industrialization. This explanation is strongly rejected by Felice (2013, p. 201) who uses several other historical examples, i.e.

[1] For a more detailed overview of these views see Missiaia (2014).

Figure 1: GDP per capita of the Italian regions, 1871–1911 (constant 1911 prices, Italy=100).

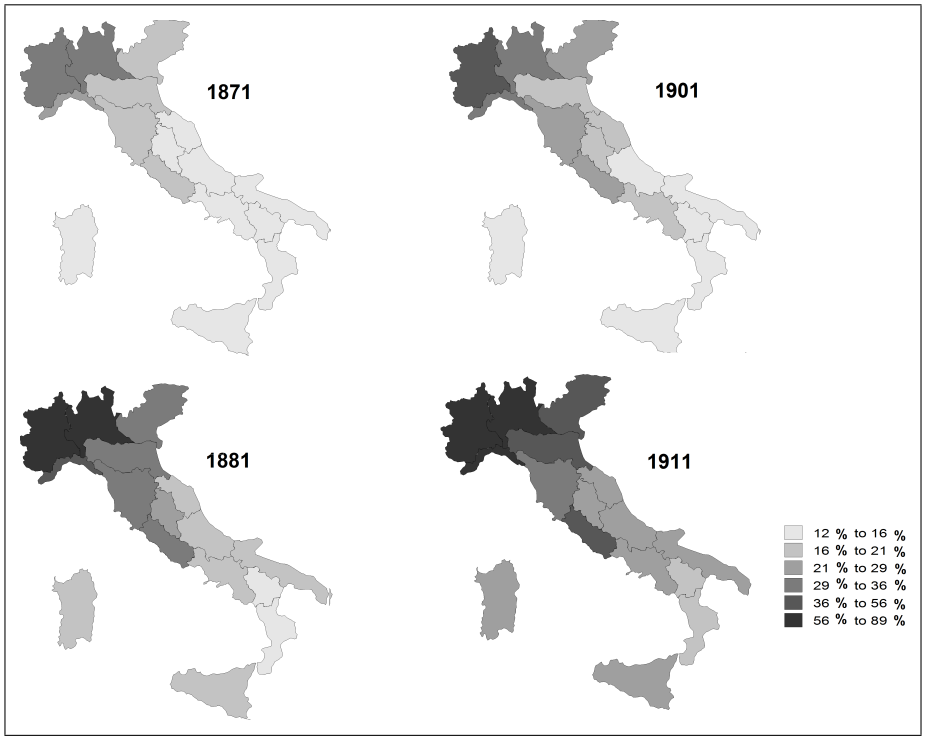


Source: Felice (2009) for 1881 and 1901 and Brunetti et al. (2011) for 1871, 1891, 1901 and 2009.

California, Japan and Southeast Asia, to demonstrate that mere physical distance from the core does not necessary imply a disadvantage. One other existing piece of research that links Economic Geography to the regional disparities in Italy is a working paper by A'Hearn and Venables (2011). The paper explores the relationship between economic disparities, internal geography and external trade for the 150 years of the unitary history of Italy. The authors propose different explanations for what drove economic activity across Italian regions in different periods: in the period 1861–1890 the main driver was natural advantage; in the period 1890–1950 it was access to the domestic market; and finally in the post-war period it was the access to foreign markets. The purpose of this paper is to contribute to the existing debate on the role of geography in the Italian North–South divide.

In Economic Geography, we know that there are two competing views on why certain regions attract economic activity more than others. The traditional Heckscher-Ohlin (H-O) view focuses on factor endowment to explain the location of economic activity, meaning that regions with a higher endowment of natural, financial or human resources attract economic activity. Opposite to this, the New Economic Geography (NEG) view considers market access as the main force. While measuring endowments is a fairly straightforward procedure, if not in terms of data collection at least in terms of methodology, measuring

Figure 2: Literacy rates in the Italian regions, 1871–1911.

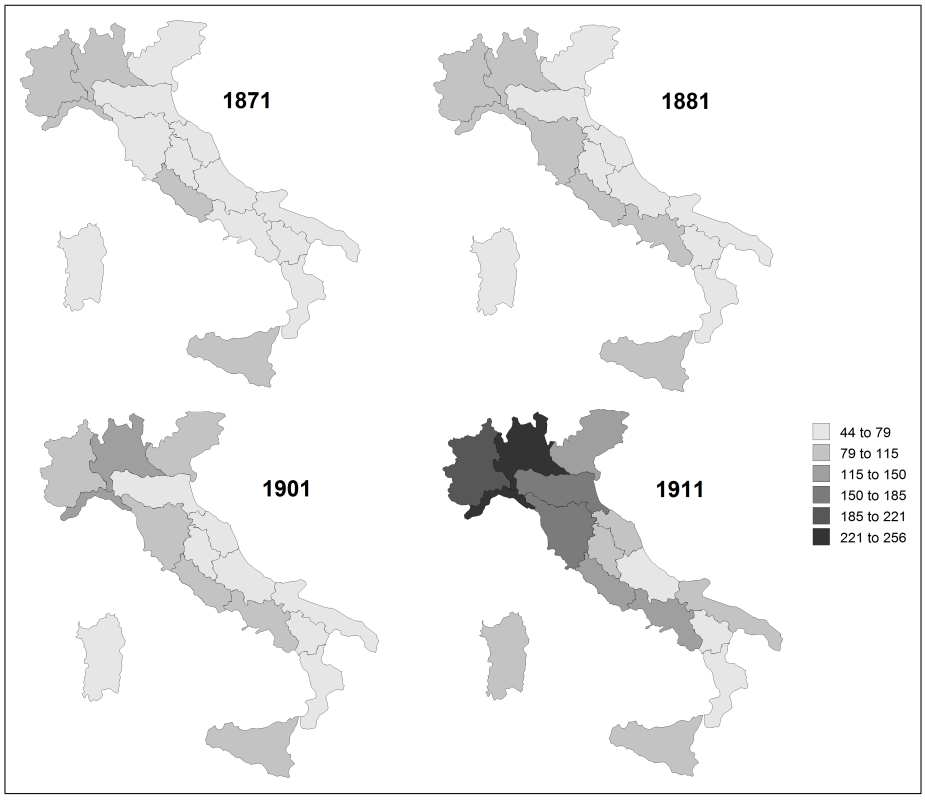


Source: A'Hearn et al. (2011).

market access is far more complicated. Reliable measures of market access are an essential starting point in order to evaluate NEG forces. This paper looks at the role of market access in the regional divergence of Italian regions, using the concept of market potential. This is a measure of the centrality of a region in terms of its access to markets. Given the lack of trade volumes data for the Italian regions, this measure will be based on the GDP of each region and the GDP of the adjacent regions, weighted by their distance. The formulation used here dates back to the seminal work by Harris (1954), adjusted by the several developments and extensions since then. This paper will estimate the market potentials of all Italian regions for a series of ten year benchmarks, from 1871 to 1911, following the methodology by Crafts (2005) and Schulze (2007). These estimates will allow us to look at the market access of different regions both before the process of industrialization gained ground and during its evolution.

The estimation of market access through market potentials has fruitful applications beyond the mere quantification of the relative position of the regions. In this paper we use both domestic and total market potential in order to explain regional GDP per capita and regional industrial value added per capita. The main result is that regional GDP per capita in this period was affected by domestic market potential more strongly than by total market potential, which includes trading partners. In particular, it is the case

Figure 3: Industrial value added per capita in the Italian regions, 1871–1911 (constant 1911 prices, Italy=100).



Source: Fenoaltea (2003) and MAIC (1874, 1883, 1902, 1914).

when the model is run in first differences, which corresponds to looking at growth rates. The second economic indicator that we try to explain is industrial value added per capita, which in terms of growth rates seems to be much less affected by any type of market potential. This measure is included in the analysis to help us focus on the part of GDP generated by the secondary sector, which appears to be more geographically polarized.

This paper is organized as follows. Section 2 discusses the methodology used to calculate market potentials for the Italian regions; Section 3 illustrates the sources used; Section 4 shows the estimates and provides a commentary on the results; Section 5 applies these estimates to explain GDP per capita and industrial value added per capita; Section 6 concludes.

2 Market access in regional analysis: empirical framework

In the literature, market potential has been calculated following several methodologies. The most basic one goes back to Harris (1954) and uses retail sales weighted by distances to evaluate the market access of US regions. This type of formulation has been used

in more recent works such as Midelfart et al. (2000) on the European Union, Crafts (2005) on Britain (1870–1910), Schulze (2007) on Austria-Hungary (1870–1910), Martinez-Galarraga (2012) on Spain (1856–1929) and Klein and Crafts (2012) on the US (1880–1920). All these works use the total GDP of the regions weighted by distance.[2] The alternative approach to calculating market potentials is to use a gravity model to estimate the functional form of market potential. A gravity model explains the volumes of trade among regions using mainly the size of the regions and the distance between each region, jointly with some dummy variables and controls such as adjacency or the presence of a border. Market potential in this case is calculated through the parameters estimated by means of the gravity model. Examples of works using this methodology are Redding and Venables (2004) at world level, Head and Mayer (2004) on Japanese firms in the European Union, Hanson (2005) on the United States, and in historical perspective, Wolf (2007) on Poland (1925–1937) . This exercise is not possible in the case of Italian regions because data on volumes of trade within the Italian borders are not available for this period. The next section illustrates the methodology adopted here, which largely follows Crafts (2005) and Schulze (2007).

2.1 Modelling market potential

In its original formulation, the market potential of region A was defined as the sum of the GDP of all the adjacent regions, each weighted by their distance from region C, plus the GDP of region C adjusted by a coefficient that takes account of its size. The calculation of market potential, following Harris (1954), is shown in Equation (1):

$$\mathrm{MP}_c = \sum_w \mathrm{GDP}_w \times D_{c,w}^{\gamma} \tag{1}$$

with $D_{c,w}^{\gamma}$ the distance between region c and w. The parameter γ is set as -1 and is defined as in Equation (2):

$$D_{c,c} = 0.333 \times \sqrt{\frac{\mathrm{Area}_c}{\pi}}\,. \tag{2}$$

The idea behind market potential is quite straightforward. First, for each region C, we take the main node. The second step is to calculate the distances between the node of region C and the nodes of each other adjacent region.[3] For a given region C, the larger the GDP of the other regions, the better the access of the regions to markets; the larger the distance between region C and the other regions and the lower the weight of the GDP

[2] All the works actually use transport cost adjusted distances; the only exception is Klein and Crafts (2012), which uses straight line distances. The appropriateness and difference between these approaches is discussed later in the paper.

[3] In this case all the Italian regions plus the main trading partners of Italy.

of each of these regions in the market access of the region concerned. Finally, Equation (1) shows how to deal with the own GDP of the region, which represents the contribution of the home GDP to the overall market access: Harris (1954) proposes a formula for own distance that takes into account the size of the region, so that the larger the region, the lower the weight of its own GDP. The rationale here is that for a given level of GDP, the larger the region the more spread out, and therefore harder to access, is its own GDP.

Although a gravity model cannot be used because of the lack of internal volumes of trade data, several refinements are still possible. First of all, distance in our case is weighted by transport costs as normally done in the literature on market integration and market potential when volumes of trade are not available. In this case, we decided to take into account both ground distances, which are assumed to be covered by railway and sea distances which are assumed to be covered by ship. For each pair of nodes, we calculate the cheapest combination of railway and shipping. To do so, we apply to all distances both a variable component (cost per km) and a terminal component (a lump sum cost when using each given mean of transportation). Whenever a part of the distance is assumed to be covered with a different mean of transportation, the corresponding terminal cost is applied. Finally, the last cost to take into account is the existence of trade barriers between nodes. It is not the case for Italian nodes, but whenever one of the two nodes is a foreign city, a correction is needed. Following Crafts (2005) and Schulze (2007), tariffs are converted into distance equivalents. This procedure is based on the coefficients of the gravity model by Estevadeordal et al. (2002). The elasticities of the model are used to convert ad valorem tariffs into a distance equivalent measure to be added to the regular terminal component of the transport cost. The tariff between Italy and each trading partner is computed as the ratio of the total custom revenues of the trading partner over its total imports. This gives an average tariff level for each country.

2.2 Testing the effect of market potential on economic development

The goal of this paper is to study the relationship between economic development and market access in the Italian regions in the period 1871–1911. The empirical framework we use is taken from the work by Head and Mayer (2011) on market potential and economic development in the period 1965–2003. This work focuses on the calculation of market potentials for all countries in the world, relating them to GDP per capita. The main methodological difference between our work and that by Head and Mayer (2011) is the calculation of market potentials: here the calculation of market potentials follows Harris

(1954) using GDP figures and transport costs; Head and Mayer (2011) use a gravity model based on trade data.

After obtaining market potential estimates, we implement the model; our goal is to cast light on the relationship between economic development and market access. We rely on the following base line specification:

$$\begin{aligned} \ln(GDPpc_i)_t = \beta_t \text{ Market Potential}_\text{i} \\ + \alpha_t \text{ Region Controls}_\text{i} \\ + \sum_t \theta_t \text{ Year} + \epsilon_t \,. \end{aligned} \tag{3}$$

The model described in Equation 3 aims at explaining the GDP per capita of region i through market potential as the main explanatory variable. The region controls are as follows: South, which is equal to 1 if the region is in the South (we also show a version of this model with region fixed effects and with latitude as a control); literacy, which is the literacy rate in a given region and the share of arable land. This latter is used in 1871 level to explain all years as a method to avoid endogeneity.

Equation 3 can thus be expanded in the following estimating equation:

$$\begin{aligned} \ln(\text{GDPpc}_i)_t = \beta_t \text{ Market Potential}_\text{i} \\ + \gamma_{1t} \text{ Latitude}_\text{i} + \alpha_{1t} \text{ Literacy}_\text{i} \\ + \alpha_{2t} \text{ Share Arable Land}_\text{i} \\ + \sum_t \theta_t \text{ Year} + \epsilon_t \,. \end{aligned} \tag{4}$$

In Section 5 various specifications of this model are shown. We also split the sample in North and South and we show the same regression with industrial value added per capita as a dependent variable. All the models are also run in first differences to address collinearity concerns. Before we show the results, the next section illustrates in detail the sources used.

3 Sources

In this section we describe the sources used for testing the model of Equation 4. All monetary measures are taken in constant 1911 lire. The regions considered in this work are the sixteen regions created in 1870 after the annexation of Rome. Figure 4 shows the boundaries of the Italian regions which did not change over the period.[4]

Starting the analysis on market potentials in 1871 and ending it in 1911 is both historically and practically useful. From a pragmatic point of view, the 1871–1911 period is convenient because borders did not have any variation and because 1871 was the year of the first census after the main annexations. In fact, Italy was formally unified in 1861 but its borders changed twice, in 1866 and in 1870, when Veneto and Latium were annexed. All the main regions of Italy (except Trentino Alto Adige and Venezia Giulia) were part of Italy in 1871. From an historical point of view, ending the analysis in 1911 allows us to isolate this period of the early industrialization from the effects of the First World War and Fascism.

3.1 Market potentials

The variable at the core of this model is market potential. Market potential is calculated using regional the GDP, the GDP of trading partners, transport cost adjusted distances and tariffs. The next sections illustrates the sources and how they are used for each of the components of market potential.

Regional and Foreign GDP Estimates. The first and main ingredient of market potential is GDP. The estimation of regional disparities for Italy in the period 1871–1911 has been a matter for discussion among scholars for a long time.[5] For the GDP estimates for the Italian regions the latest available series are those from the work of Emanuele Felice. The estimates of regional GDP used here come from Felice (2009) for 1881 and 1901 and Brunetti et al. (2011), of which Felice is a coauthor, for 1871, 1891 and 1911. The data provided here on the regional disparities of GDP per capita are the starting point for deriving GDP estimates in levels for this period. The next step in the procedure is to apply these per capita disparities to the national GDP per capita. We decided to use for this step the GDP estimates published by Baffigi (2011) within the broader project of the Bank of Italy for the 150th anniversary of the Unification. These are the latest estimates

[4] These regions are quite similar to present regions, with the exception of Venezia Giulia and Trentino Alto Adige which were not yet Italian and Valle Aosta and Molise which were at the time parts of Piedmont and Abruzzi, respectively.

[5] See Fenoaltea (2003), Zamagni (1978), Felice (2007), Felice (2009) and Brunetti et al. (2011).

Figure 4: Italian regions, 1870–1918.

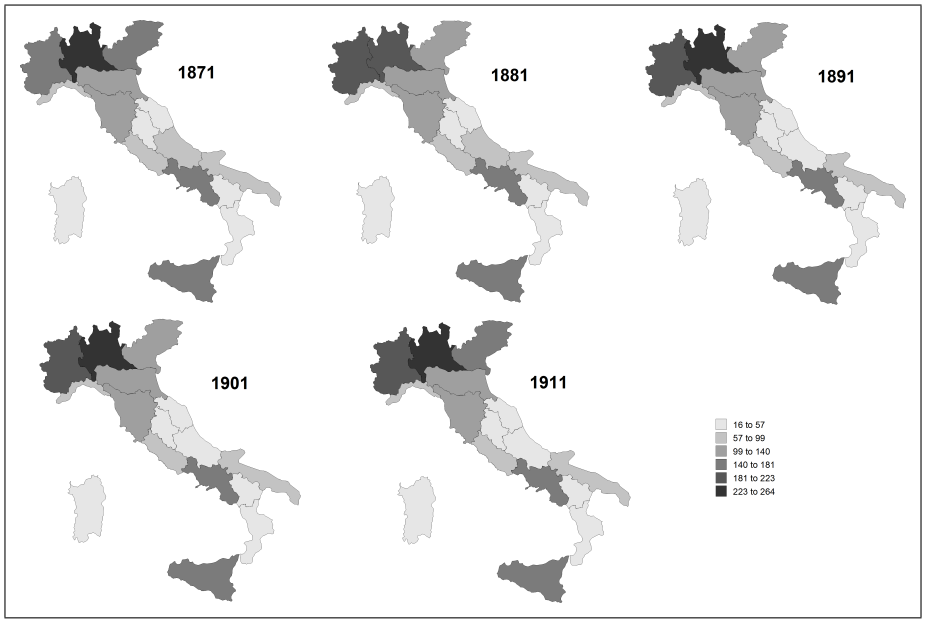
for the Italian national income and are published both in constant and current prices. In this paper we use the GDP figures in 1911 constant lire. Starting from the national GDP estimates by Baffigi (2011), we calculate the national GDP per capita estimates by dividing them by the present Italian population. This figure is then multiplied by the coefficients of regional disparities provided by Felice (2009) and Brunetti et al. (2011) to work out all the level of the regional GDP per capita. Finally, the per capita figures are multiplied by the regional population figures to obtain the total GDP of each region in levels at constant 1911 prices.[6] The regional GDP disparities from Felice (2009) and Brunetti et al. (2011) are shown in Figure 5.

With regard to the GDP of foreign trading partners, the procedure is the following. Using data on exports from the Annuario Statistico Italiano, for each benchmark year, we take all the countries that cover 80% of Italian foreign trade. The trading partners included are: Austria-Hungary, France, Germany, the United Kingdom, Switzerland, Argentina and the United States.[7] Figure 6 shows the Italian exports to these countries as the share of total exports for each benchmark year.

[6]To illustrate, assume that the national GDP per capita in a given year is 100 lire. If the GDP per capita coefficient for a region A is 1.20, the GDP per capita is 120. If the present population of the region in 1871 is 1,000,000 than the total GDP of region A in 1871 is derived by multiplying these three figures, returning 1,200,000,000 lire.

[7]The only exception to the 80% criterion is Turkey, for which GDP estimates for this period in current prices are not easily available from either source.

Figure 5: Total GDP of the Italian regions, 1871–1911 (constant 1911 prices, Italy=100).



Source: Felice (2009) for 1881 and 1901 and Brunetti et al. (2011) for 1871, 1891 and 1911.

The main source for the GDP of foreign partners is Crafts (2005), who relies primarily on the work of Prados de la Escosura (2000). Argentina and Switzerland come directly from Prados de la Escosura (2000) for the years 1881–1911.[8] We also consider Austria and Hungary separately. In order to do so, we split the estimate by Crafts (2005).[9]

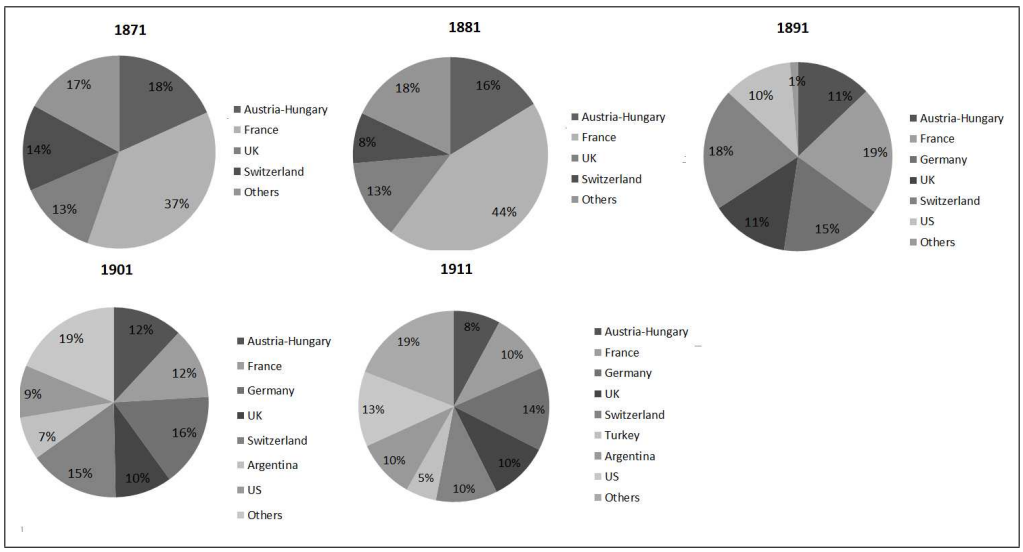
Figure 7 shows the magnitude of the Italian GDP compared to foreign GDP. Looking at Figure 7, the different magnitude of the US, for 1911 in particular, stands out. Including such a large GDP compared to that of the Italian regions could be problematic in the sense that most of the market access could be driven by the US. To underplay the role of the US, we split it in four macro-regions, Northeast, Midwest, West and South. The implication of this choice is discussed in more detail in the next section. The regional disparities for the US are worked out from Klein (2009). These are the same estimates as those on which the market potentials of Klein and Crafts (2012) are based for 1881–1911. To work out 1871, we use the same disparities as for 1881.

Distances. In order to weight the sum of GDP, a distance measure is required. As in previous works by Crafts (2005) and Schulze (2007), geographical distance is replaced by transport costs. This is suitable because straight line distances per se do not take into account differences of costs for alternative means of transportation nor of the existence

[8]For 1871, estimates were not available; therefore we used the relative disparity among countries of 1881 and applied it to 1871.

[9]The two GDP are worked out by looking at the relative size in each year from Schulze (2007).

Figure 6: Italian export shares, 1871–1911.



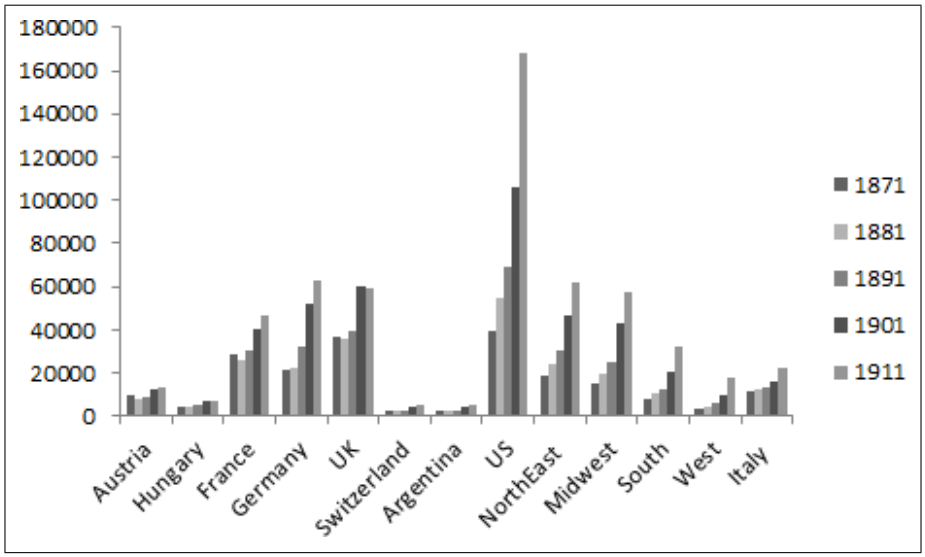
Source: Annuario Statistico Italiano 1877, 1881, 1892, 1911.

of railway lines and ports. The first step is to choose a node for each region and foreign trading partner. Then the distance in terms of railway line or sea route (or the two together) between each pair of nodes is computed. The administrative regions considered are shown in Figure 4.

The general rule in selecting the nodes is to take the most populous centre which often corresponds to the administrative centre. Exceptions are Terni and Pescara for Umbria and Marches. The reason for this choice is that the actual administrative centres, Perugia and L'Aquila, were not the economic centres of the regions and were not very well integrated in its transport network. Using them would have created too high a penalty for Umbria and Abruzzi in terms of market access. For the US, none of the nodes is the administrative centre since in the US the largest cities almost never correspond to the capital. The two means of transportation considered in this period are railways and shipping. For railways, the length of lines has been worked out from the following sources: Bradshaw's Continental Guide of 1914 and the publication on the all lines opened by Ferrovie dello Stato (1927). Relying on sources that cover the whole period is very important, because this takes into account the construction of new lines. For shipping, distances were easily computed from the website www.dataloy.com, which provides the length of maritime routes between all the main ports worldwide.

Transport Costs. Once a matrix of distances is computed, the next step is to quantify the rate per tonne per kilometre. The rate taken into account is the average rate between

Figure 7: GDP of Italy and its main trading partners, 1871–1911 (constant 1911 million lire).



Source: Prados de la Escosura (2000) and Crafts (2005).

coal and wheat, which are considered here the two representative goods.[10] For Italian railways, the source is a publication on railway rates by Ferrovie dello Stato (1912). This publication is quite detailed, providing terminal and variable components of the rate of transportation for a variety of goods. For the rates of foreign countries, we rely on the work by Schulze (2007) which uses on the information from the US Bureau for Railway Economics (1915) and Noyes (1905) to compute terminal and variable components. The first source provides an overview of 1914 rates for different countries and for both coal and wheat, separating the cost in terminal and variable component for different city pairs in each country. From this, the terminal and variable components for various countries are worked out. The next step is to project these estimates back in time. Noyes (1905) provides average rates starting from 1870. This information is converted into an index and used jointly with the 1914 baseline to extrapolate terminal and variable components for the whole period. For the US, which is not covered by Schulze (2007), we rely on the information from Noyes (1905) and assume that the US rate is 50% of the general European rate. For shipping, there are no sources specific to Italy. The estimates for international ocean shipping from Kaukiainen (2003) are used for all routes. This is a

[10] We would of course like to take into account the transportation cost of all industrial goods. However, collecting information on transport rate for all goods would be extremely data and time consuming. Moreover, it would not be clear how to use this information in a synthetic measure. Therefore, we follow the existing literature such as Crafts (2005) and Schulze (2007) and adopt the standard solution.

Table 1: Regions, trading partners and nodes.

	Region/Country	**Node**
Nodes with sea access	Liguria	Genoa
	Venetia	Venice
	Marches	Ancona
	Campania	Naples
	Apulia	Bari
	Calabria	Reggio Calabria
	Sicily	Palermo
	Sardinia	Cagliari
	United Kingdom	London
	Turkey	Istanbul
	Argentina	Buenos Aires
	Unites States	New York
	Northeast	New York
	South	New Orleans
	West	San Francisco
Nodes without access to the sea	Piedmont	Turin
	Lombardy	Milan
	Emilia	Bologna
	Tuscany	Florence
	Umbria	Terni
	Latium	Rome
	Abruzzi	Pescara
	Basilicata	Potenza
	Austria-Hungary	Vienna
	France	Paris
	Germany	Berlin
	Switzerland	Zurich

widely used source for this type of research, including works on Italy.[11] The transport costs used are set out in Tables 2, 3 and 4.

Table 2: Shipping rate per tonne per km, in constant 1911 lire, 1871-1911.

	Terminal Component	Cost per km
1871	19.67	0.00307
1881	14.71	0.00180
1891	11.14	0.00116
1901	9.06	0.00111
1911	7.29	0.0011

Source: Kaukiainen (2003)

[11]Federico (2007) uses the same source to study the market integration of Italy in the 19th century.

Table 3: Italian railway rates per tonne per km, in constant 1911 lire, 1871-1911.

	Terminal Component	Cost per km
1871	1.93	0.05688
1881	1.83	0.05386
1891	1.51	0.04456
1901	1.56	0.04588
1911	1.26	0.03731

Source: Ferrovie dello Stato (1912), Schulze (2007), Noyes (1905)

Table 4: Railway and Shipping rates per tonne per km, in constant 1911 lire 1871-1911.

	Austria Railway Rates		France Railway Rates		Germany Railway Rates	
	Terminal	Variable	Terminal	Variable	Terminal	Variable
1871	5.44	0.08342	7.67	0.01937	6.56	0.03344
1881	4.35	0.06633	8.69	0.02195	6.69	0.03379
1891	3.11	0.04860	7.61	0.01924	6.21	0.03149
1901	2.98	0.04618	7.00	0.01753	5.96	0.03057
1911	2.60	0.04013	5.73	0.01450	4.90	0.02505
	Europe Railway Rates		UK Railway Rates		US Railway Rates	
	Terminal	Variable	Terminal	Variable	Terminal	Variable
1871	7.55	0.034289	2.35	0.052741	7.99	0.03614
1881	7.58	0.034409	2.34	0.05054	4.88	0.022046
1891	6.32	0.028692	2.04	0.045535	3.63	0.016437
1901	6.19	0.028006	2.29	0.0491	3.10	0.014003
1911	5.10	0.023171	1.87	0.041229	2.55	0.011585

Source: Ferrovie dello Stato (1912), Schulze (2007), Noyes (1905) and US Bureau for Railway Economics (1915)

Tariffs. The last cost to be taken into account is the one originated by trade barriers between nodes. It is not the case for Italian nodes, but whenever one of the two nodes is a foreign city, a correction is needed. Following Crafts (2005) and Schulze (2007), tariffs are converted into distance equivalents. The source used is Mitchell (2003) except for the United Kingdom, where Mitchell (1988) is used. Capie (1994) gives data for 1870 Germany and Ferreres (2005) for Argentina. Whenever a year is missing, either the closest available year is used or the gap is filled by interpolation. Table 5 shows the level of tariffs used.

3.2 Region controls

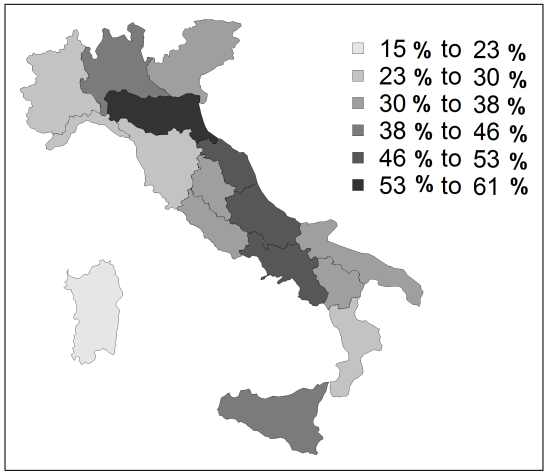
In Section 5, GDP per capita and industrial value added per capita are explained using market potentials, geographical controls (such as dummies for macro areas, region fixed effects and latitude) and two other controls; literacy rates and share of arable land. The share of arable land is shown in Figure 8 while literacy rates in Figure 2 are derived from (A'Hearn et al., 2011).

Table 5: Ad valorem tariffs, 1871–1911.

	Austria-Hungary	France	Germany	UK	Switzerland	Argentina	US
1871	2%	3%	8%	6%	2%	22%	29%
1881	0%	6%	6%	5%	2%	27%	24%
1891	3%	8%	9%	5%	3%	20%	20%
1901	6%	9%	9%	9%	4%	29%	20%
1911	6%	10%	8%	12%	4%	2%	15%

Source: Mitchell (2003) for Austria-Hungary, France and Germany, Mitchell (1988) for the United Kingdom, Capie (1994) for 1870 Germany and Ferreres (2005) for Argentina.

Figure 8: Share of arable land by province, 1870.



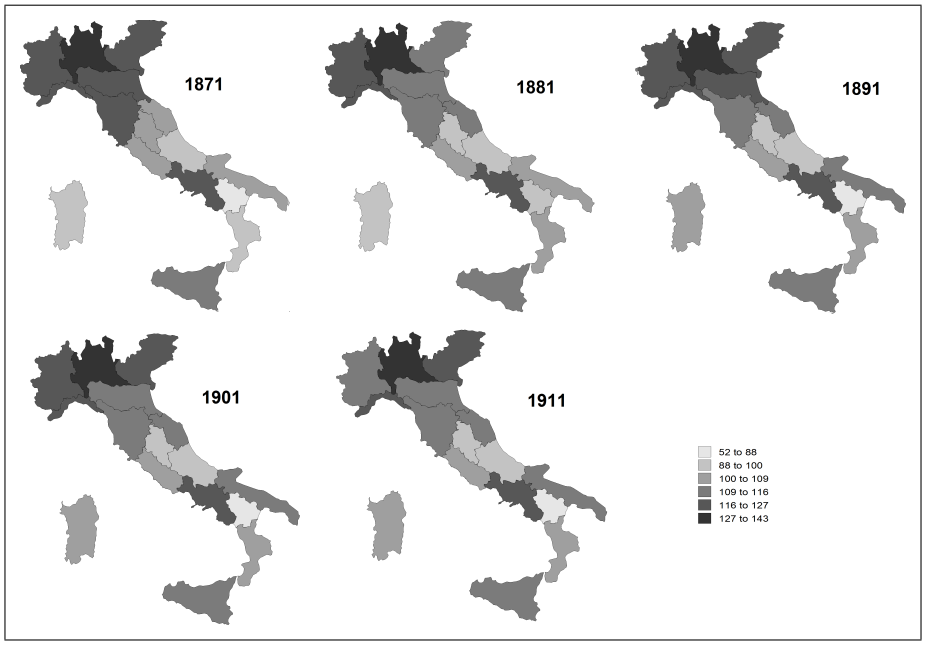
Source: MAIC (1976).

A last remark is on the year 1891. The 1891 population census has not been carried out owing to budget restrictions. For this reason, explanatory variables such as literacy rates are not available. In order to avoid relying heavily on interpolations, we decided to leave 1891 out of the regression analysis. The next section shows the results of the market potential calculations.

4 Market potential of Italian regions, 1871–1911: empirical results

This section shows and compares different versions of market potential for the Italian regions that will be used in the next section as explanatory variable for GDP per capita and industrial value added per capita. The first one proposed is a domestic market potential, exclusively taking into account the Italian regions. The second one is a repetition of the domestic market potential using straight distances and shows how different results can

Figure 9: Domestic market potential in Italian regions, 1871–1911 (constant 1911 prices, Italy=100).



Source: our own calculations.

emerge without controlling for transport costs. The third is total market potential, which comprises all the main trading partners of Italy along with the Italian regions.

Let us start with the first version. Figure 9 and Table 6 provide the estimates of domestic market potential for the Italian regions between 1871 and 1911.

The domestic market potential shows two main results. The first is that at the beginning of the period the picture is quite in line with the classic North-South divide, with the North showing a higher level of market access and the South lagging behind. The exceptions in the South are Campania and Sicily. The reason why these two regions have levels comparable to the Northern regions is that their total GDP at the beginning of the period were comparable to those of the regions in the North and that these two regions have sea-ports as their economic centre. This makes them able to exploit shipping, which is cheaper than railways in this period. The second result is that there is a tendency over the period for the market access of the North to worsen with respect to the South. This tendency can be explained if we look at the transport costs in Tables 2, 3 and 4. Shipping costs drop relatively faster in this period than railway costs, giving an advantage to regions that have their node on the coast. This is the case of the main Southern regions. A similar result, connected to the access to shipping, is discussed in Schulze (2007) on the Habsburg Empire. In its case, the only two regions with access to the sea (Littoral and Dalmatia) have a persistent advantage over the others in terms of market potential in spite

Table 6: Domestic market potential in the Italian regions, 1871–1911.

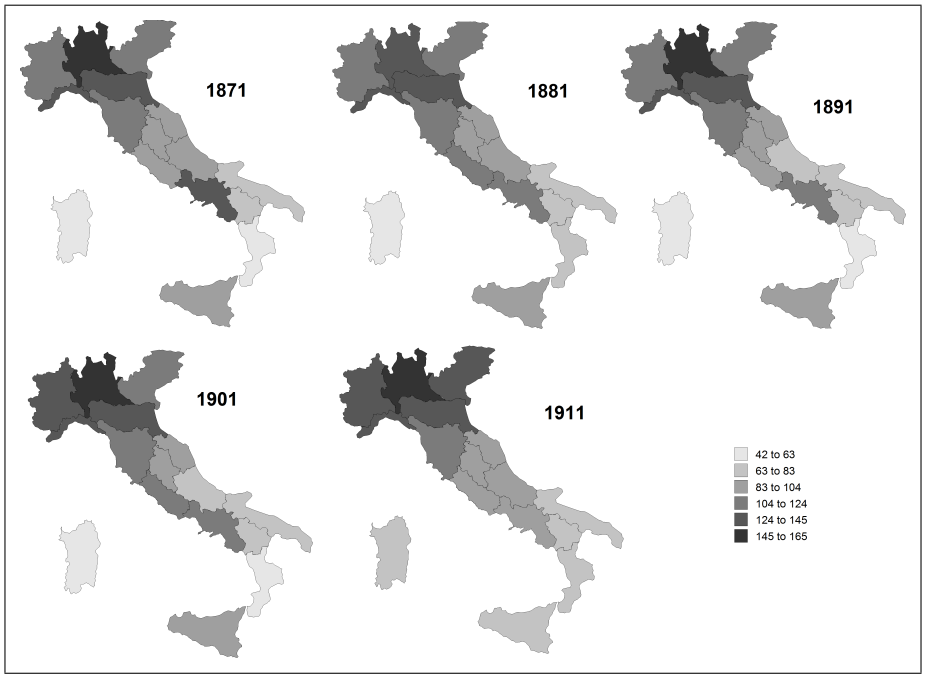
	1911 million lire					Italy=100				
	1871	1881	1891	1901	1911	1871	1881	1891	1901	1911
Piedmont	766	932	1264	1528	2552	120	118	114	113	109
Liguria	768	946	1385	1702	2973	121	119	123	126	127
Lombardy	908	1033	1496	1779	3094	142	130	133	131	132
Veneto	777	893	1274	1570	2716	122	113	114	116	116
Emilia	783	889	1270	1414	2550	123	112	114	104	109
Tuscany	733	863	1197	1340	2365	115	109	106	99	101
Marches	612	774	1115	1374	2343	96	98	99	101	100
Umbria	541	593	811	916	1588	85	75	72	68	68
Latium	560	728	1013	1182	2063	88	92	89	87	88
Abruzzi	509	608	826	951	1634	80	77	74	70	70
Campania	747	915	1298	1574	2712	117	115	116	116	116
Apulia	547	765	1105	1401	2352	86	96	98	103	101
Basilicata	336	536	746	874	1489	53	68	67	64	64
Calabria	512	701	998	1307	2234	80	88	89	96	96
Sicily	641	869	1228	1527	2593	101	110	108	113	111
Sardinia	457	642	952	1247	2139	72	81	85	92	91
Italy	637	793	1124	1355	2337	100	100	100	100	100

Source: our own calculations.

of their lower GDP. This relationship between sea access and market potential is clear when trading partners are taken out of the sample. This is because Littoral and Dalmatia are the only regions in the sample directly connected by sea to the foreign trading partners. Taking trading partners out, the two regions no longer have such advantage. In Italy this advantage is present even without the inclusion of trading partners because shipping is an available option in internal trade as well.

The second set of estimates uses straight line distances between nodes. Figure 10 and Table 7 show the domestic market potentials when we take straight line distances instead of transport costs. Here we see that the process of worsening market access worsening in the North and improving in the South is not taking place. This version of market access also shows a much wider gap between North and South, with most Southern regions doing sensibly worse than is implied by the calculation adjusted for transport costs. This version of market potential is the most similar to the one proposed by A'Hearn and Venables (2011). The authors calculate market potentials using the same GDP estimates and weight them by straight line distances. Their results show a large gap between North and South: the authors claim that domestic market access was the driving force for the location of industries in the period 1890–1950. Looking at the

Figure 10: Domestic market potential in Italian regions, 1871–1911 (constant 1911 prices, Italy=100, straight line distances).



Source: our own calculations.

difference between the domestic market potentials calculated with transport costs adjusted distances and with straight line distances, the claim that market potentials were moving in the same direction as GDP or industrial value added appears to be supported much less by the former calculation.[12]

The third version proposed here, which we call total market potential, shows the market potentials calculated including the main trading partners of Italy.[13] Figure 11 and Table 8 show the results.

This version of market potential shows a quite different picture from the domestic one. The South here appears not only to perform increasingly better in the period but also as starting from a higher level than the North, which experiences the opposite evolution. These estimates, which pool the GDP of Italian regions and that of very large trading partners, are very much influenced by the GDP of the latter, in particular the US, which has a very high GDP compared to Italy (see 12). Given that most trading partners are reached at least partially by sea, the sea effect is even higher in this formulation of market potential. Taking the US as one unit would force us to pick only one node and weight the entire GDP of the US according to this node, which in our case is New York. This

[12]If we compared total market potential calculated with transport costs and with straight line distances the difference would look even more marked.

[13]The choice of trading partners to include is discussed in Section 3 above.

Table 7: Domestic market potential in the Italian regions, 1871–1911 (straight line distances).

	1911 million lire					Italy=100				
	1871	1881	1891	1901	1911	1871	1881	1891	1901	1911
Piedmont	61	74	81	95	143	116	123	120	128	121
Liguria	70	78	95	106	177	134	130	138	142	150
Lombardy	78	86	105	121	194	149	144	154	163	164
Veneto	61	62	71	82	131	116	104	105	110	111
Emilia	73	81	97	103	172	140	135	142	138	146
Tuscany	63	71	81	85	139	120	119	118	114	117
Marches	47	53	61	65	102	90	89	90	88	86
Umbria	48	53	59	63	98	92	89	86	84	83
Latium	51	63	71	77	124	99	105	101	104	105
Abruzzi	45	52	56	59	93	87	86	82	80	78
Campania	67	73	82	85	134	129	122	121	114	113
Apulia	38	49	55	59	89	74	82	80	80	75
Basilicata	35	42	46	49	75	68	70	69	66	63
Calabria	31	38	40	43	67	60	64	59	58	57
Sicily	44	55	61	64	98	84	93	87	87	83
Sardinia	22	27	31	34	53	43	45	45	45	45
Italy	52	60	68	74	118	100	100	100	100	100

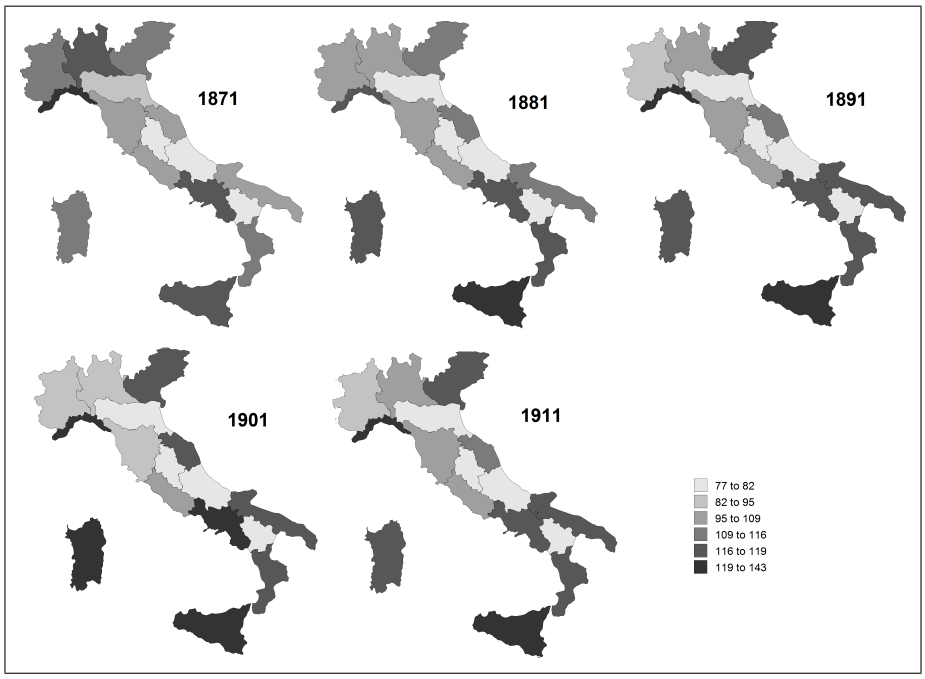
Source: our own calculations.

practically corresponds to assuming that the whole GDP of the US is as easily accessible as if it was all located on the East Coast. This of course is far from the fact. This is why we decided to consider the US as four separate macro regions (Northeast, Midwest, South and West). Doing so allows us to reduce the role of the US in the calculation without ignoring the presence of the US in the foreign markets.[14]

The next section move our focus to the relationship between these estimates and regional economic development.

[14]The same problem could be posed by Austria and Hungary: we decided to keep them separate in all calculations as in this period, it must be remembered that the two regions of the Habsburg Empire became a dual monarchy in 1867 and therefore were quite economically independent.

Figure 11: Total market potential in Italian regions, 1871–1911 (constant 1911 prices, Italy=100).



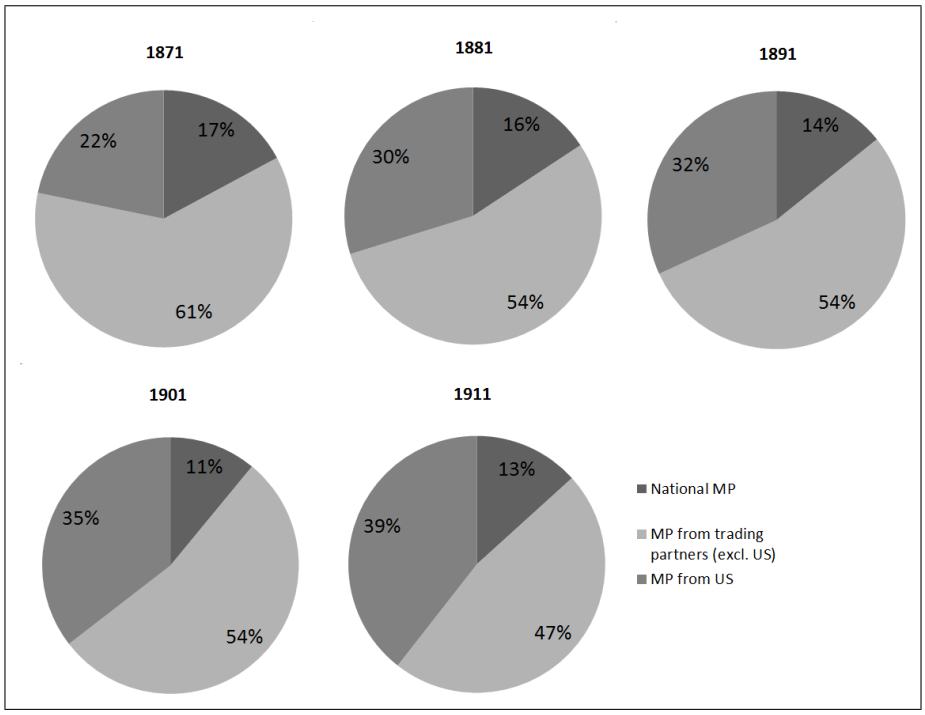
Source: our own calculations.

Table 8: Total market potential in the Italian regions, 1871–1911.

	1911 million lire					Italy=100				
	1871	1881	1891	1901	1911	1871	1881	1891	1901	1911
Piedmont	3,976	4,739	7,020	10,609	15,071	108	93	88	85	86
Liguria	4,322	5,912	9,361	14,902	20,886	118	116	117	119	119
Lombardy	4,010	5,032	7,697	11,533	16,665	109	99	96	92	95
Venetia	3,900	5,506	8,775	13,763	19,136	106	108	110	110	109
Emilia	3,402	4,312	6,589	9,744	14,216	93	84	82	78	81
Tuscany	3,628	4,874	7,523	11,470	16,421	99	95	94	92	93
Marches	3,673	5,363	8,582	13,573	18,790	100	105	107	108	107
Umbria	3,028	4,136	6,408	9,744	13,895	82	81	80	78	79
Latium	3,452	4,877	7,601	11,812	16,775	94	96	95	94	95
Abruzzi	3,048	4,235	6,577	10,086	14,293	83	83	82	81	81
Campania	4,025	5,791	9,188	14,600	20,372	109	113	115	117	116
Apulia	3,673	5,445	8,709	13,869	19,193	100	107	109	111	109
Basilicata	2,934	4,230	6,588	10,177	14,418	80	83	82	81	82
Calabria	3,767	5,548	8,846	14,251	19,767	102	109	111	114	112
Sicily	4,159	6,036	9,515	15,404	21,391	113	118	119	123	122
Sardinia	3,837	5,646	9,027	14,640	20,340	104	111	113	117	116
Italy	3,677	5,105	8,000	12,511	17,602	100	100	100	100	100

Source: our own calculations.

Figure 12: Contribution of trading partners to total market potential, 1871–1911.



Source: our own calculations.

5 Market access and economic development: empirical results

Section 4 presented market potential estimates for the Italian regions for five benchmark years, between 1871 and 1911. In this section we explore the relationship between market potential and regional GDP per capita, which represents a fundamental measure of economic development. We then show the same model using industrial value added per capita as the dependent variable, which represents the part of the GDP of a region that comes from the secondary sector. To do so, we use both basic pooled OLS and first difference OLS. The main explanatory variable in all specifications is market potential, along with other geographic and economic controls. In the next section we show the results on GDP per capita.

5.1 Market potential and GDP per capita

In this section we show the empirical results based on the model of section 2.2. Table 9 shows the OLS estimates of Equation 4. All years are pooled and standard errors are heteroskedasticity-robust.[15] Columns 1 and 2 show the basic pooled OLS regression with GDP per capita explained by domestic and total market potential. Here there are no controls. The relationship appears positive and significant at the 1% level for domestic market potential as well as for total market potential. The coefficients are expressed in logs and therefore they can be interpreted as elasticities. Domestic market potential has an elasticity of 0.351 and total market potential has an elasticity of 0.235. The R^2 is about 10% higher for the specification with domestic market potential, which is near 0.5, suggesting that the former has more explanatory power. The comparison with previous literature appears challenging, because the existing works are at country level rather than sub-national level. Redding and Venables (2004, p. 65) provide a coefficient of 0.146 for domestic market potential and 0.395 for total market potential. Head and Mayer (2011, p. 289) find a coefficient of 0.80 for the specification without specific country controls. As admitted by the authors, the difference in the magnitude of the coefficients is more than double that of Redding and Venables (2004) and "this difference in the coefficients stems mainly from the different construction of the market potential variable". Therefore, differences in the type of sample, in the historical periods and in the techniques used to calculate market potentials allow for very different magnitudes in the coefficients without affecting the validity of the different works. In columns 3 and 4 of Table 9 we add year

[15] The pooling of the four available years (1871, 1881, 1901 and 1911) is necessary in order to increase the number of observations. This procedure is the one followed by Head and Mayer (2011).

dummies to take into account the fact that the data are pooled across years. In this version, the results change: domestic market potential increases its elasticity to 0.544 significant at 1% while total market potential is not significant in this case. The next step is to introduce further controls to capture differences across regions. The first candidate is physical geography. In columns 5 and 6 we introduce a dummy for belonging to different macro areas.[16] In the regression, we decided to use the North-West as our baseline and include the dummies corresponding to the other two macro areas. The result is that the domestic market potential decreases its coefficient size and level of significance while the total market potential moves in the opposite direction. Both columns have market potentials significant at the 5% level while the dummy South is highly significant with a negative sign. The North-East-Centre is also negative but with a considerably smaller coefficient. The negative signs on both macro areas are not surprising, since we chose the richest macro area as the baseline. The fact that being in the South has the strongest effect is also expected because of the generally poorer conditions of Southern regions in this period. The use of these macro areas should be considered in light of the types of difference we want to capture. If the aim is to describe geographic differences across Italy, avoiding other elements that could be endogenous to GDP, the imposition of these macro borders is not necessarily the best option in spite of the strong results. This is because macro areas are often designed ex-post to collect regions with similar economic conditions.[17] In order to avoid this endogeneity problem, columns 7 and 8 use latitude as the geographic control. Latitude is a control that is completely exogenous, being the distance from the Equator, and has a positive effect on GDP per capita. The result is that both coefficients are around 0.4, with domestic market potential significant at the 1% and total market potential at 5%. The R^2 decreases slightly compared to that in columns 6 and 7, but are very much higher than the specification without geographic controls. From now on, whenever we decide to include a control for the geographic position of the regions we will use latitude.

The previous table aimed at exploring the relationship between market potential and GDP per capita using controls that are possibly not endogenous to GDP per capita. In Table 10 we try to introduce further controls and discuss whether or not these are appropriate in our case. The first control that we can think of including, to capture

[16] The classification of regions in the three different macro areas is taken from Felice (2007) and was later used in his other works. The Northwest includes Piedmont, Lombardy and Liguria; the Northeast-Centre includes Venetia, Emilia, Marches, Tuscany, Umbria and Latium; the South comprises Campania, Apulia, Basilicata, Calabria, Sicily and Sardinia.

[17] See for example the inclusion of the North-East in the same macro area of the Centre and the decision to keep the regions of the Industrial Triangle separate from the rest.

Table 9: GDP per capita and market potential, 1871–1911 (pooled OLS regression with geographic controls).

GDPpc	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Domestic MP	0.351***		0.544***		0.217**		0.431***	
	(0.0452)		(0.0853)		(0.0939)		(0.0929)	
Log Total MP		0.235***		0.191		0.306**		0.411**
		(0.0450)		(0.194)		(0.115)		(0.172)
Northeast-Centre					-0.142**	-0.163***		
					(0.0566)	(0.0597)		
South					-0.315***	-0.389***		
					(0.0535)	(0.0406)		
Latitude							0.0207**	0.0452***
							(0.00801)	(0.00961)
Constant	-1.189	0.768	-5.049***	1.773	1.760	-0.529	-3.634**	-4.983
	(0.941)	(1.022)	(1.732)	(4.275)	(1.920)	(2.545)	(1.792)	(4.070)
Year Fixed Effects	no	no	yes	yes	yes	yes	yes	yes
Region Fixed Effects	no	no	no	no	no	no	no	no
Observations	64	64	64	64	64	64	64	64
R^2	0.491	0.309	0.537	0.376	0.667	0.671	0.563	0.519

Notes: Heteroskedastic robust standard errors in parentheses. * , ** and *** correspond to a coefficient significantly different from zero with a 10%, 5% and 1% confidence level respectively. The dependent variable is the log of GDP per capita in 1911 prices. Market potentials are in 1911 prices.

other fundamental variables that can affect GDP per capita is some measure of human capital. In modern studies, such as Head and Mayer (2011), average years of schooling is the best variable to use. For the case of pre-First World War Italy we decided to use a more basic output measure that is often used to capture human capital in the Italian case: literacy rates. When literacy rates are introduced in the regression, they are highly significant with a coefficient of around 0.6 in both specifications. Market potentials loose their significance and latitude is incorrectly signed. These results clearly suffer from a bias. The reason why the plain inclusion of this variable is problematic is two-fold: literacy is endogenous to GDP per capita and it is also collinear to other variables that present a North-South gradient.[18] This is, for instance, the case of domestic market potential, which has a geographic distribution that appears very similar to that of literacy rates. In this case, it is very hard to disentangle the effect of literacy from that of market potential; moreover, it is very hard to establish the direction of the causality between literacy and GDP per capita. Columns 3 and 4 show the inclusion of the share of arable land in 1871 as a control.[19] Including arable land at the beginning of the period is a way to avoid

[18]We also tried to included similar controls such as the share of labour force in agriculture; the problems encountered in the estimation were very similar and we decided to deal with literacy rates only in the estimation to avoid making the issue even more problematic by including further controls.

[19]The share of arable land is also used as a control by Redding and Venables (2004). The authors here use a number of other controls, such as fraction of land in tropical areas, prevalence of malaria and

Table 10: GDP per capita and market potential, 1871–1911 (pooled OLS regression with further controls).

Log GDPpc	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Domestic MP	-0.0531		0.431***		0.430***		0.430**	
	(0.125)		(0.0843)		(0.127)		(0.181)	
Log Total MP		-0.155		0.335**		-0.206		-0.206
		(0.158)		(0.162)		(0.212)		(0.286)
Log Literacy	0.623***	0.640***						
	(0.118)	(0.0998)						
Latitude	-0.0522***	-0.0595***	0.0241***	0.0468***				
	(0.0148)	(0.0178)	(0.00825)	(0.00961)				
Log Share of Arable Land 1871			-0.192***	-0.153**				
			(0.0513)	(0.0627)				
Liguria					0.165**	0.245***	0.165***	0.245***
					(0.0665)	(0.0816)	(0.0127)	(0.0695)
Lombardy					-0.0766	0.00284	-0.0766**	0.00284
					(0.0773)	(0.0588)	(0.0280)	(0.0181)
Venetia					-0.262***	-0.223***	-0.262***	-0.223***
					(0.0690)	(0.0766)	(0.00276)	(0.0451)
Emilia					-0.153**	-0.184***	-0.153***	-0.184***
					(0.0680)	(0.0488)	(0.00469)	(0.0282)
Tuscany					-0.0989*	-0.129***	-0.0989***	-0.129***
					(0.0517)	(0.0459)	(0.0149)	(0.00718)
Marches					-0.269***	-0.308***	-0.269***	-0.308***
					(0.0508)	(0.0454)	(0.0273)	(0.0366)
Umbria					0.00337	-0.188***	0.00337	-0.188***
					(0.0659)	(0.0489)	(0.0682)	(0.0411)
Latium					0.388***	0.283***	0.388***	0.283***
					(0.0605)	(0.0433)	(0.0467)	(0.00728)
Abruzzi					-0.269***	-0.483***	-0.269***	-0.483***
					(0.0763)	(0.0638)	(0.0796)	(0.0345)
Campania					-0.175***	-0.127*	-0.175***	-0.127**
					(0.0610)	(0.0704)	(0.00215)	(0.0597)
Apulia					-0.128*	-0.174**	-0.128***	-0.174***
					(0.0705)	(0.0659)	(0.0318)	(0.0407)
Basilicata					-0.201**	-0.494***	-0.201*	-0.494***
					(0.0922)	(0.0483)	(0.112)	(0.0360)
Calabria					-0.366***	-0.437***	-0.366***	-0.437***
					(0.0649)	(0.0599)	(0.0443)	(0.0480)
Sicily					-0.204***	-0.177**	-0.204***	-0.177**
					(0.0631)	(0.0720)	(0.0106)	(0.0723)
Sardinia					-0.153**	-0.251***	-0.153**	-0.251***
					(0.0596)	(0.0752)	(0.0575)	(0.0545)
Constant	7.257***	9.838**	-3.100*	-2.845	-2.631	10.66**	-2.631	10.66
	(2.631)	(3.856)	(1.564)	(3.856)	(2.570)	(4.661)	(3.689)	(6.287)
Year Fixed Effects	yes	yes	yes	yes	yes	yes	yes	yes
Region Fixed Effects	no	no	no	no	yes	yes	yes	yes
Region Clustering	no	no	no	no	no	no	yes	yes
Observations	64	64	64	64	64	64	64	64
R^2	0.760	0.763	0.599	0.541	0.955	0.948	0.955	0.948

Notes: Heteroskedastic robust standard errors in parentheses. * , ** and *** correspond to a coefficient significantly different from zero with a 10%, 5% and 1% confidence level respectively. The dependent variable is the log of GDP per capita in 1911 prices. Market potentials are in 1911 prices. Literacy is expressed as rate.

endogeneity as the share of arable land in subsequent years may have been affected by changes in the economic or technological conditions of the regions. The share of arable land is significant and negatively signed in both specifications. The negative sign can be explained by interpreting the share of arable land as the presence of agricultural activity in a region. Since agriculture has lower value added levels than industries, leading to lower levels of GDP per capita, it has a larger weight in the regional economy. Here latitude is correctly signed and significant again, suggesting that endogeneity and multicollinearity are less of an issue in this specification. Columns 4–8 take a different approach to the specification. In the last four columns we introduce region fixed effects. This strategy is also used by Head and Mayer (2011). However, in their case the inclusion of these dummies leads to much less of an increase in the R^2 than our case shows. Here, region fixed effects lead to an increase from about 0.5 to values well over 0.90. This increase is explained by the number of regional dummies that are highly significant. The baseline is Piedmont, a fairly rich Northern region. This is why most other significant regional dummies have a negative sign (with the exceptions of Liguria and Latium which do show quite high levels of GDP per capita even compared to Piedmont in this period). Southern regions all have negative signs, which is expected given the lower levels of GDP per capita in this part of the country. The high explanatory power of these fixed effects suggests that the inclusion of any further control other than market potential should be dropped in order to avoid the risk of over-identifying the model. For this reason, given the high share of variation explained, we assume that the region fixed effects capture a sufficient part of the regional differences (although they are not able to disentangle them) and we do not include further controls. Columns 6 and 7 show the inclusion of fixed effects without regional clustering, while Columns 8 and 9 show the results with clustering. Market potential here shows similar results to the previous specification: domestic market potential is positive and significant with an elasticity of 0.430; total market potential is not significant.

Summing up the results of the first two tables, we notice that domestic market potential is significant and correctly signed in all specifications except when literacy rates are introduced while total market potential is not significant in a number of specifications, most notably when fixed effects are used, appearing less robust as the explanatory variable for GDP.

We now move to the issues of endogeneity and multicollinearity that affect the model when literacy rates are introduced. On the first point, Head and Mayer (2011) attempt

risk of expropriation that appear sensible when doing a cross-country analysis but would lead to difficult quantification or have little meaning in the case of Italian regions in this period.

two different approaches. First, they substitute total market potential with foreign market potential, which is market potential calculated with own GDP removed.[20] This solution has been attempted for the case of the Italian regions but the results do not hold. According to Head and Mayer (2011, p. 291),"[foreign market potential] has nice features [in dealing with endogeneity], but is clearly not ideal as a replacement or instrument for RMP". The solution that Head and Mayer (2011) adopt in their work is to use an instrument for market potential. In particular, they take geographic centrality proxied by two instruments: the sum of the inverse of straight line distances and the inverse of transport costs. In their case the former is ruled out and the solution adopted is to use the latter. Unfortunately, this instrument does not appear to work for the Italian regions in this period. Looking at the levels of our variables, in the absence of reliable instruments, we cannot establish the direction of causality between GDP per capita, domestic market potential and literacy. This is evident when looking at the geographical distribution of the levels in Figures 1, 2 and 10, where the North-South gradient appears clear. One solution here is to move our focus from levels to changes in the time of the variables of interest. By transforming all the logs of the variables in first differences we can effectively interpret the model in terms of growth rates. Running the model in first differences allows us to address the issue of multicollinearity, since the common long-term trend in the variables is not present in the growth rates.[21]

Table 11 shows the basic formulation without region or geographic controls in Columns 1 and 2. The result is a positive and significant coefficient for the domestic market potential and very low and insignificant coefficients on the total market potential.[22] Columns 3–8 replicate the main specifications proposed so far: latitude as control with literacy rates first and then with the share of arable land and, in the last two columns, region fixed effects as only controls. Unlike the results with levels, first difference regressions show little effect from the controls, while domestic market potentials are always correctly signed and significant with a coefficient around 0.8 in all the specifications. Total market potential does not appear significant in any specification.

[20]This leaves us effectively with a market access measured only accounting for the GDP of foreign countries in the sample, ignoring the internal market.

[21]In particular, if we look at Figures 1, 2 and 10, we notice that literacy rates and domestic market potential in the South tend to converge with those of the North while GDP per capita tends to diverge.

[22]The coefficients here can be interpreted as percentage point increases in growth rate of GDP per capita when an explanatory variable increases by 1%. Therefore, in column 1, if the growth rate of domestic market access increases by 1%, the growth rate of GDP per capita increases by 0.649%.

Table 11: GDP per capita, market potential and literacy rates, 1871–1911 (pooled OLS regression in first differences).

Log GDPpc	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DLog Domestic MP	0.649***		0.849***		0.816***		0.883**	
	(0.206)		(0.235)		(0.249)		(0.345)	
DLog Total MP		0.0557		0.0731		0.111		0.370
		(0.332)		(0.399)		(0.373)		(0.489)
DLog Literacy			0.0645	-0.245				
			(0.272)	(0.304)				
Latitude			0.0124	-0.00195	0.0112*	0.00296		
			(0.00812)	(0.0102)	(0.00612)	(0.00761)		
Share of Arable Land 1871					-0.000347	-0.00146		
					(0.00109)	(0.00127)		
Constant	-0.0764	0.258**	-0.724	0.392	-0.629*	0.167	-0.133	0.216
	(0.113)	(0.114)	(0.449)	(0.575)	(0.335)	(0.395)	(0.190)	(0.129)
Year Fixed Effects	yes	yes	yes	yes	yes	yes	yes	yes
Region Fixed Effects	no	no	no	no	no	no	yes	yes
Observations	48	48	48	48	48	48	48	48
R^2	0.619	0.520	0.651	0.531	0.651	0.532	0.696	0.592

Notes: Heteroskedastic robust standard errors in parentheses. * , ** and *** correspond to a coefficient significantly different from zero with a 10%, 5% and 1% confidence level respectively. The dependent variable is the log of GDP per capita in 1911 prices. Market potentials are in 1911 prices. Literacy is expressed as rate. All variables are in first differences.

5.2 Market potential in the North and South

Tables 12 and 13 show our model splitting the sample in the North (which merges the North-West and North-East-Centre) and the South.[23] As a result, the number of observations goes down to 21 for the South and 27 for the North. Let us start with the South. In Table 12 we notice that the level of GDP per capita in Southern regions is much more clearly affected by market potential when the regression is in levels (Columns 1–4) than in first differences (Columns 5–8). This is the case both without geographic controls and when we include latitude and share of arable land (Columns 3–4). Moreover, the coefficients are slightly higher for total market potential than for the domestic one. In first differences, the results do not hold, suggesting that after Unification market potentials do nothing to explain the growth rates of GDP per capita.

The picture looks quite different in Table 13, where the sample is restricted to the northern regions. In this case, total market potential is never significant while domestic market potential is significant in both levels and first differences, with fairly

[23] Venetia, Lombardy, Piedmont, Liguria, Emilia, Tuscany, Latium, Umbria and Marches are considered North; the rest of the regions are considered South as in footnote 16.

Table 12: GDP per capita and market potential, southern regions, 1871–1911 (pooled OLS regression and first differences).

Log GDPpc	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Domestic MP	0.514***		0.642***					
	(0.0647)		(0.0996)					
Log Total MP		0.624***		1.030***				
		(0.104)		(0.176)				
Latitude			0.0361**	0.0589***			0.00476	0.00591
			(0.0146)	(0.0164)			(0.0111)	(0.0106)
Log Share of Arable Land 1871			-0.132**	0.0298				
			(0.0539)	(0.0508)				
DLog Domestic MP					0.617**		0.489	
					(0.260)		(0.302)	
DLog Total MP						1.316		0.976
						(0.914)		(1.033)
DLog Literacy					0.539	0.427	0.523	0.446
					(0.437)	(0.419)	(0.455)	(0.435)
Share of Arable Land 1871							-0.00205	-0.00291
							(0.00205)	(0.00221)
Constant	-4.486***	-7.906***	-8.046***	-19.31***	-0.227	-0.303	-0.269	-0.324
	(1.295)	(2.266)	(2.203)	(4.336)	(0.212)	(0.324)	(0.539)	(0.608)
Year Fixed Effects	yes	yes	yes	yes	yes	yes	yes	yes
Observations	28	28	28	28	21	21	21	21
R^2	0.824	0.750	0.866	0.870	0.700	0.652	0.719	0.697

Notes: Heteroskedastic robust standard errors in parentheses. * , ** and *** correspond to a coefficient significantly different from zero with a 10%, 5% and 1% confidence level respectively. The dependent variable is the log of GDP per capita in 1911 prices. Market potentials are in 1911 prices. Literacy is expressed as rate. Columns 5–8 show the variables in first differences.

high coefficients. Market potential is therefore a stronger predictor of GDP per capita within northern regions compared to the South.

Considering the North and South of Italy separately brings some interesting insights. The effect of market access on the GDP per capita of these two different parts of the country is not the same: in the South, both domestic and total market potential have a strong role in determining the levels of GDP per capita but they do not seem to impact on the growth rates. For the North, only the domestic market potential is significant in both in levels and in first differences. This difference suggests that the results at national level are driven by the Northern part of the country, where the access to markets outside Italy was often more expensive because of the presence of rich but landlocked regions.

The next section moves from GDP to industrial production, repeating the exercise using industrial value added per capita as the dependent variable.

5.3 Market potential and industrial value added per capita

The growth of the industrial output is considered a strong driver of economic growth for Italy in this period. One further question we can address is how far the factors that explain

Table 13: GDP per capita and market potential, northern regions, 1871–1911 (pooled OLS regression and first differences).

Log GDPpc	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Domestic MP	0.120		1.458***					
	(0.210)		(0.212)					
Log Total MP		0.194		0.0767				
		(0.275)		(0.259)				
Latitude			-0.221***	-0.0517			0.00255	0.0000621
			(0.0366)	(0.0325)			(0.0181)	(0.0283)
Log Share of Arable Land 1871			-0.318***	-0.400***				
			(0.0889)	(0.107)				
DLog Domestic MP					1.552***		1.586***	
					(0.338)		(0.343)	
DLog Literacy					-0.241	-0.109	-0.356	-0.176
					(0.291)	(0.354)	(0.451)	(0.584)
DLog Total MP						-0.146		-0.128
						(0.449)		(0.515)
Share of Arable Land 1871							0.00152	0.000694
							(0.00206)	(0.00238)
Constant	3.632	1.811	-12.75***	8.108	-0.530***	0.355**	-0.698	0.332
	(4.298)	(6.063)	(3.062)	(5.572)	(0.179)	(0.167)	(0.813)	(1.352)
Observations	36	36	36	36	27	27	27	27
R^2	0.450	0.453	0.804	0.633	0.762	0.569	0.770	0.570

Notes: Heteroskedastic robust standard errors in parentheses. * , ** and *** correspond to a coefficient significantly different from zero with a 10%, 5% and 1% confidence level respectively. The dependent variable is the log of GDP per capita in 1911 prices. Market potentials are in 1911 prices. Literacy is expressed as rate.

GDP per capita can also explain industrial value added per capita. Let us start with Table 14, in which the main OLS regressions in levels are reported. Columns 1 and 2 show the baseline model with year dummies as the only controls. Here domestic market potential is positive and significant at 1% while total market potential is significant at the 5% level. The elasticities are considerably higher than the ones for GDP per capita. Columns 3 and 4 introducing the dummy for latitude show the same result; Columns 5 and 6 report the specification with the share of arable land and latitude together. It appears that in all specifications, both domestic and total market potentials are positive and significant with high coefficients (between 0.6 and 1) and a 1% level of significance. However, the results do not hold when fixed effects are introduced. This last two Columns provide a counter intuitive result for total market potential, which is negative and significant. To investigate further, we turn our attention to Table 15, where the same exercise is repeated in first difference. Here again the results on market potentials do not hold. As in the case of the GDP per capita of Southern regions, market potential seems to possibly explain levels of GDP per capita but not the changes from period to period. The next section sums up the results.

Table 14: Industrial value added per capita and market potential, 1871–1911 (pooled OLS regression).

Log Ind VA pc	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Domestic MP	1.136***		0.997***		0.997***		-0.369	
	(0.142)		(0.153)		(0.151)		(0.240)	
Log Total MP		0.651**		1.059***		1.022***		-0.835***
		(0.272)		(0.207)		(0.214)		(0.214)
Latitude			0.0255**	0.0840***	0.0289**	0.0847***		
			(0.0110)	(0.0132)	(0.0116)	(0.0135)		
Log Share of Arable Land 1871					-0.193***	-0.0741		
					(0.0656)	(0.0885)		
Constant	-18.96***	-10.30*	-17.22***	-22.84***	-16.68***	-21.80***	11.84**	22.62***
	(2.897)	(5.989)	(2.950)	(4.779)	(2.842)	(4.963)	(4.895)	(4.707)
Year Fixed Effects	yes	yes	yes	yes	yes	yes	yes	yes
Region Fixed Effects	no	no	no	no	no	no	yes	yes
Observations	64	64	64	64	64	64	64	64
R^2	0.801	0.538	0.817	0.742	0.832	0.744	0.953	0.957

Notes: Heteroskedastic robust standard errors in parentheses. * , ** and *** correspond to a coefficient significantly different from zero with a 10%, 5% and 1% confidence level respectively. The dependent variable is the log of the industrial value added per capita in 1911 prices. Market potentials are in 1911 prices. Literacy is expressed as rate.

Table 15: Industrial value added per capita, market potential and literacy rates, 1871–1911 (pooled OLS regression in first differences.

Log Ind VA pc	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DLog Domestic MP	-0.0138		0.378		0.281		0.406*	
	(0.195)		(0.251)		(0.249)		(0.216)	
DLog Total MP		-0.291		0.268		0.275		0.565*
		(0.253)		(0.304)		(0.281)		(0.323)
Latitude			0.0244***	0.0223***	0.0248***	0.0241***		
			(0.00615)	(0.00658)	(0.00605)	(0.00614)		
Share of Arable Land 1871					-0.00268*	-0.00307**		
					(0.00144)	(0.00136)		
Constant	0.389***	0.481***	-0.980***	-0.888**	-0.722*	-0.621*	0.141**	0.311***
	(0.107)	(0.0878)	(0.299)	(0.343)	(0.362)	(0.329)	(0.0519)	(0.0849)
Year Fixed Effects	no	no	yes	yes	yes	yes	yes	yes
Region Fixed Effects	no	no	no	no	no	no	yes	yes
Observations	48	48	48	48	48	48	48	48
R^2	0.483	0.492	0.622	0.603	0.655	0.648	0.796	0.792

Notes: Heteroskedastic robust standard errors in parentheses. * , ** and ** correspond to a coefficient significantly different from zero with a 10%, 5% and 1% confidence level respectively. The dependent variable is the log of industrial value added per capita in 1911 prices. Market potentials are in 1911 prices. Literacy is expressed as rate. All variables are in first differences.

6 Conclusions

This paper presented the estimates of market potentials of the Italian regions for 10-year benchmarks in the period 1871–1911 and used market potential to explain GDP per capita and industrial value added per capita. Market potentials are based on constant 1911 price estimates of regional GDP for Italy from Felice (2009) and Brunetti et al. (2011) and the GDP of the main trading partners of Italy in the period from 1871 to 1911 are from Prados de la Escosura (2000) and Crafts (2005). In the calculation of market potentials, the GDP are weighted by distance-adjusted transport costs to take into account the actual distance among regions in terms of the costs of transporting one unit of a representative good. In the case of foreign partners a distance-equivalent tariff is calculated and added to the transport cost. In this paper we proposed different specifications ranging from domestic market potential, which comprises Italian regions only to total market potential that includes all Italian regions and all the main trading partners. The bottom line of these calculations is the following: domestic market potentials at the beginning of the period show a more traditional picture of Italy, with the North presenting higher values than the South, in particular at the beginning of the period. What contradicts the North-South traditional image is that North and South converge. This is explained by the relative cost of shipping, which goes down, unlike that of railways, over the period. This is an advantage for the main Southern regions such as Campania and Sicily, which have good access to the sea along with a fairly high total GDP, while Piedmont and Lombardy for instance do not. The role of transport costs is easily appreciated when we compare the same calculation with straight line distances: the picture is reversed, showing the strong and consistent position of the Northern regions. These comparisons show how useful it is, at least for the case of Italy and other countries with more than one viable mean of transportation, to correctly account for transport costs and existing lines. Introducing foreign markets in the total market potential, the results change for the first year, but the trend in time goes in the same direction. The North has no advantage in terms of market access at the beginning of the period and, as for the domestic one, it worsens its position compared to the South.

Once the estimates of market potentials have been produced, we used them as explanatory variables for the main measure of economic performance: GDP per capita. We first proposed a baseline specification in levels and we then added further controls. In the specifications in levels, we found that domestic market potential has a stronger and more significant positive effect on GDP per capita than total market potential. These

findings are in line with our expectations from the descriptive analysis of the data. However, the inclusion of economic controls such as literacy rates cause concerns over the multicollinearity and endogeneity that could be affecting the results. To deal with these issues, we ran the model in first differences. Taking the first differences of the logs corresponds to considering the variables in growth rates. This procedure leaves out the common trends that the variables might present in the long run and restricts the focus to the change from one benchmark year to another. The model in first differences confirms that the domestic market potential is positive and significant when trying to explain the growth rate of GDP per capita of the Italian regions.

The next step was to divide the sample into North and South, running the model separately. It is quite useful to look at different parts of the country separately whenever the number of observations allows us to do so. In this case it is especially useful to separate levels and growth rates as the main result here is that total market potential determined the level of GDP per capita in the South but it did not in the North. Moreover, growth rates in market potential of both types do not affect the growth rates of GDP per capita in the South but they do in the North when we consider domestic market potential. This insight tells us that in the South, the total market potential was positively correlated to the levels of GDP per capita but its growth rate did not drive the growth rate of GDP per capita. In the rest of the country, both the level and the growth rates of total market potential are not significant in explaining GDP per capita. This suggests that the access to international markets did matter more for the South compared to the rest of Italy. This result is in line with the view that the Southern economy was, until the invasion of agricultural products from the US in the late 19th, quite open to international trade. A well known study by Morilla et al. (1999, pp. 333–337) takes into account the case of southern Italy as exporter of high value added agricultural products. The authors give the example of citrus production in Sicily, which supplied 95% of lemons and 16% of oranges consumed in the US in 1890. However, the trend over the period before the First World War was negative for the southern Italian exporters because of the competition from California. We believe that this decline in exports is informative on the first difference regressions results on total market potential. Finally, we ran the model in levels and first differences, using industrial value added per capita as the dependent variable. Here the main result was that both domestic and total market potentials were positive and significant in explaining the levels of industrial value added per capita, but the results were not robust to the use of fixed effect and first differences. This suggests that market potential, in its domestic

formulation, is a more clear predictor of GDP per capita than industrial value added per capita.

The findings of this paper contribute to the existing literature in two ways. First of all, the paper suggests that the quantification of market access can be quite sensitive to the inclusion of trading partners as well as to the proper quantification of transport costs rather than simple geographical distances. Furthermore, the different role of domestic and international market potentials challenges the idea that the South of Italy has been penalized by its geographical position in terms of access to the international markets as suggested for instance by Daniele and Malanima (2007).

Primary Sources

[1] Bureau of Railway Economics, Comparison of railway freight rates in the United States, the principal countries of Europe, South Australia, and South Africa. Washington, 1915.

[2] Ferrovie delo Stato (FS), Tariffe e condizioni pei trasporti sulle Ferrovie dello Stato. Roma, 1912.

[3] Ferrovie delo Stato (FS), Sviluppo delle ferrovie italiane dal 1839 al 31 dicembre 1926. Roma, 1927.

[4] Istituto Nazionale di Statistica (ISTAT), Annuario statistico italiano, various years.

[5] Ministero di agricoltura, industria e commercio (MAIC), Relazione sulla condizione dell'agricoltura nel quinquiennio 1870–1874. Roma, 1876

[6] Ministero di agricoltura, industria e commercio (MAIC), Censimento generale al 31 dicembre 1871. Roma, 1874.

[7] Ministero di agricoltura, industria e commercio (MAIC), Censimento della popolazione del Regno d'Italia al 31 dicembre 1881. Roma, 1883.

[8] Ministero di agricoltura, industria e commercio (MAIC), Censimento della popolazione del Regno d'Italia al 10 febbraio 1901. Roma, 1902.

[9] Ministero di agricoltura, industria e commercio (MAIC), Censimento della popolazione del Regno d'Italia al 10 giugno 1911. Roma, 1914.

References

A'Hearn, B. (2000): "Could Southern Italians Cooperate? Banche Popolari in the Mezzogiorno," *Journal of Economic History*, 60, 67–93.

A'Hearn, B., C. Auria, and G. Vecchi (2011): *In ricchezza e in povertà*, Il Mulino, Bologna, chap. Istruzione.

A'Hearn, B. and A. Venables (2011): "Internal Geography and External Trade: regional disparities in Italy, 1861-2011," Economic History Working Papers 12, Bank of Italy.

Baffigi, A. (2011): "Italian National Accounts, 1861-2011," Economic History Working Papers 18, Bank of Italy.

Brunetti, A., E. Felice, and G. Vecchi (2011): *In ricchezza e in povertà*, Il Mulino, Bologna, chap. Reddito.

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

Capie, F. (1994): *Tariffs and Growth: Some Illustrations from the World Economy, 1850-1940*, Manchester University Press, Manchester.

Crafts, N. (2005): "Market potential in British regions, 1871-1931," *Regional Studies*, 39, 1159–1166.

Daniele, V. and P. Malanima (2007): "Il prodotto delle regioni italiane e il divario nord-sud in Italia, 1861-2004," *Rivista di Politica Economica*, 47, 267–315.

Estevadeordal, A., B. Frantz, and A. M. Taylor (2002): "The Rise and Fall of World Trade, 1870-1939," Working Papers 9318, NBER.

Felice, E. (2007): *Divari regionali e intervento pubblico: per una rilettura dello sviluppo in Italia*, Il Mulino, Bologna.

——— (2009): "Estimating regional GDP in Italy (1871-2001): sources, methodology and results," Working Papers in Economic History 0907, Universidad Carlos III.

——— (2013): "Regional income inequality in Italy in the long run (1871-2001). Patterns and determinants," UHE Working Papers 08, Universitat Autònoma de Barcelona.

Felice, E. and M. Vasta (2012): "Passive Modernization? The new Human Development Index and its Components in Italy's Regions (1871-2007)," UHE Working Papers 10, Universitat Autònoma de Barcelona.

Fenoaltea, S. (2003): "Peeking Backward: Regional Aspects of Industrial Growth in Post-Unification Italy," *Journal of Economic History*, 63, 1059–1102.

——— (2006): *L'economia italiana dall'Unità alla Grande Guerra*, Collezione storica, Laterza, Bari.

Ferreres, O. (2005): *Dos siglos de economía argentina, 1810-2004: : historia argentina en cifras*, Fundaciøn Norte y Sur, Buenos Aires.

Galassi, F. (2000): "Measuring Social Capital: Culture as an Explanation of Italy's Economic Dualism," Department of economic working papers, University of Warwick.

Hanson, G. (2005): "Market potential, increasing returns and geographic concentration," *Journal of International Economics*, 67, 1–24.

Harris, C. D. (1954): "The Market as a Factor in the Localization of Industry in the United States," *Annals of the Association of American Geographers*, 44, 315–348.

Head, K. and T. Mayer (2004): "Market Potential and the Location of Japanese Investment in the European Union," *Review of Economics and Statistics*, 86, 959–972.

——— (2011): "Gravity, market potential and economic development," *Journal of Economic Geography*, 11, 281–294.

Kaukiainen, M. (2003): "How the price of distance declined: ocean freights for grain and coal from the 1870s to 2000," Department of history, University of Helsinki, MIMEO.

Klein, A. (2009): "Personal Income of U.S. States: Estimates for the Period 1880–1910," Economics Research Paper Series (TWERPS) 916, University of Warwick.

Klein, A. and N. Crafts (2012): "Making sense of the manufacturing belt: determinants of U.S. industrial location, 1880–1920," *Journal of Economic Geography*, 12, 775–807.

Martinez-Galarraga, J. (2012): "The determinants of industrial location in Spain, 1856–1929," *Explorations in Economic History*, 49, 255–275.

Midelfart, K., H. Overman, S. Redding, and A. Venables (2000): "The Location of European Industry," European Economy Economic Papers 142, European Commission.

Missiaia, A. (2014): "Industrial location, market access and economic development: regional patterns in post-unification Italy," PhD Thesis, London School of Economics and Political Science.

Mitchell, B. (1988): *British Historical Statistics*, Cambridge University Press, Cambridge.

——— (2003): *International historical statistics: Europe, 1750-2000*, Palgrave Macmillan, New York.

Morilla, C., A. Olmstead, and P. Rhode (1999): "Horn of Plenty: The Globalization of Mediterranean Horticulture and the Economic Development of Southern Europe, 1880-1930," *Journal of Economic History*, 59, 316–352.

Noyes, W. (1905): *American railroad rates*, Little, Brown and Company, New York.

Prados de la Escosura, L. (2000): "International Comparisons of Real Product, 1820-1990: an Alternative Data Set," *Explorations in Economic History*, 37, 1–41.

Redding, S. and A. J. Venables (2004): "Economic geography and international inequality," *Journal of International Economics*, 62, 53–82.

Schulze, M. (2007): "Regional income dispersion and market potential in the late nineteenth century Hapsburg Empire," Working Paper 106-07, London School of Economics.

Sonnino, S. (1877): *La Sicilia nel 1876*, Barbera, Firenze, chap. I contadini in Sicilia.

Villari, P. (1979): *Le Lettere Meridionali ed Altri Scritti sulla Questione Sociale in Italia*, Guida, Napoli.

Wolf, N. (2007): "Endowments vs. market potential: What explains the relocation of industry after the Polish reunification in 1918?" *Explorations in Economic History*, 44, 22 – 42.

Zamagni, V. (1978): *Industrializzazione e squilibri regionali in Italia: bilancio dell'età giolittiana*, Il Mulino, Bologna.

——— (1990): *Dalla periferia al centro: la seconda rinascita economica dell'Italia (1861-1990)*, Il Mulino, Bologna.