European
Historical
Economics
Society

EHES Working Paper | No. 229 | October 2022

Market access, the skill premium and human capital in Spain (1860-1930)

Rafael González-Val,
Universidad de Zaragoza and Institut d'Economia de Barcelona (IEB)

Pau Insa-Sánchez,
Universitat de València

Julio Martinez-Galarraga,
Universitat de Barcelona

Daniel A. Tirado-Fabregat,
Universitat de València

EHES Working Paper | No. 229 | October 2022

Market access, the skill premium and human capital in Spain (1860-1930)*

Rafael González-Val[1],
Universidad de Zaragoza and Institut d'Economia de Barcelona (IEB)

Pau Insa-Sánchez,
Universitat de València

Julio Martinez-Galarraga[2],
Universitat de Barcelona

Daniel A. Tirado-Fabregat,
Universitat de València

Abstract

This paper explores the relationship between market access and education levels in the context of an industrializing economy, in this case Spain between the late nineteenth and early twentieth centuries. Specifically, we examine whether differences in regional accumulations of human capital could be related to market access, which would explain the divergent trajectories of regional economic growth in Spain. To do this, we empirically test the relationship between education variables and market potential for Spanish provinces between 1860 and 1930. We then focus on the mechanism that may be mediating this relationship, i.e. the skill premium. The results suggest that there were sizeable provincial differences in the return on investment in education, the explanation for which would be that those provinces with the highest market potential specialized more in skill-intensive sectors in which higher wages were paid.

JEL Codes: I25, N90, O15, R12

Keywords: Economic history, market access, human capital

[1]Corresponding Author: Rafael González-Val (rafaelg@unizar.es)

[2]Corresponding Author: Julio Martinez-Galarraga (julio.martinez@ub.edu)

* The authors would like to thank the Bank of Spain and the FBBVA for their financial support through the "Programa de Ayudas a la Investigación en Macroeconomía, Economía Monetaria, Financiera y Bancaria, e Historia Económica" and the "Programa de Ayudas a equipos de investigación científica en Humanidades Digitales" respectively. Financial support from the Spanish Ministry of Science and Innovation, projects ECO2015-65049-C12-1-P, ECO2017-90848-REDT, PGC2018-095821-B-I00, PID2020-114354RA-I00 (MCIU/AEI/FEDER,UE), FPU18/01210; the Gobierno de Aragón (ADETRE research group); and the Generalitat Valenciana (AICO/2018/1930 and PROMETEO/2020/083) is also gratefully acknowledged.

Notice
The material presented in the EHES Working Paper Series is property of the author(s) and should be quoted as such. The views expressed in this Paper are those of the author(s) and do not necessarily represent the views of the EHES or its members

1. Introduction

This paper explores the possible existence of a relationship between a region's market access and its population's level of education. Basing our study on Spain over the period 1860-1930, we examine whether those provinces (NUTS3) with the greatest market access or market potential (the two terms are used interchangeably in the context of this paper) also had the greatest accumulations of human capital, as measured by literacy rates and numeracy skills. We then focus on the economic mechanism that may explain this relationship, looking in particular at the connection between market potential and the skill premium of skilled industrial workers versus unskilled workers employed in both agriculture and industry. In short, we test whether or not one of the hypotheses deriving from the new economic geography (NEG) literature can be confirmed, i.e. that there is a link between a region's market potential and its accumulation of human capital (Redding and Schott, 2003).

The theoretical and empirical literature of the NEG has given us an economic interpretation of the spatial inequality that characterizes the world today. The basis for this is the belief that interaction between the economies of scale typical of certain activities and lower trade costs makes it more likely that companies and workers will become concentrated in those locations with the best access to the markets of goods and inputs, this being a cumulative process (Krugman, 1991; Fujita et al., 1999; Puga, 1999; Combes et al., 2008). The existence of this relationship would help us to understand not only the marked inequality that characterizes the distribution of activity across territories, but also the appearance of divergent growth paths within them (Krugman and Venables, 1995; Redding and Venables, 2004; Ottaviano and Pinelli, 2006; Combes and Gobillon, 2015).

Using this economic foundation as a starting point, the empirical literature has made it possible to establish two relevant aspects. First, there is evidence pointing to the existence of a relationship between market potential and regional production specialization (Midelfart-Knarvik et al., 2002), the direction of capital flows (Crozet et al., 2004) and workers (Crozet, 2004), and regional differences in wage levels (Hanson, 1998, 2005; Roos, 2001; Brakman et al., 2004) and income (Redding and Venables, 2004; Breinlich, 2006; Niehbur, 2006; Head and Mayer, 2011). And second, these processes appear to be cumulative, which implies that the inequality that characterizes most developed economies today may have its roots in the early stages of the economic development processes (Crafts and Venables, 2003; Wolf, 2007; Liu and Meissner, 2015). In this regard, for many of today's developed countries the nineteenth century was the first time the two fundamental elements of the NEG models

coincided: technological change, typical of the first and second industrial revolutions, and market integration (reduction in trade costs) for goods and inputs.

The NEG literature has thus sufficiently proven that today's regional inequality is linked to differences in market access resulting from a process that has its roots in the past. However, a great deal less interest has been shown in identifying and studying the mechanisms that would become the basis of any such sustained relationship over time. In the search for these mechanisms, one body of work has connected two apparently disparate lines of literature. Firstly, the literature on economic growth. The endogenous growth theory has repeatedly pointed out that the availability of human capital is a determinant of growth over the long term (Romer, 1986; Lucas, 1988; Galor, 2011; Barro, 2013; Gennaioli et al., 2013). Thus the existence of marked differences in the regional availability of human capital could be one of the main reasons for the differences in regional economic trajectories and therefore a key element for understanding current regional inequality.

And secondly, the new economic geography literature. Since this first made its appearance, progress has been made in the theoretical formulation and empirical testing of new hypotheses. As pointed out by Redding and Schott (2003), given that better market access may encourage the accumulation of human capital, the incentive for individuals to invest in education should be related to their geographical location. The economic mechanism behind this would be as follows. Greater market access would favour regional specialization in those sectors characterized by the presence of increasing returns. In a two-sector theoretical framework in which manufacturing uses both skilled and unskilled labour and is the sector with increasing returns, and in which agriculture has constant returns to scale and uses unskilled labour for production purposes, it is to be expected that the higher the market potential, the higher the specialization in industrial production and the higher the wage premium for skilled workers. In these conditions, market potential would increase the incentives to invest in human capital in a particular location. As a result, the rate of human capital accumulation would be endogenous to the model and be related to each region's market potential.

Empirical analysis of this possible relationship has mainly involved recent periods and large samples of countries (Redding and Schott, 2003) or European regions (Faíña and López-Rodríguez, 2006; López-Rodríguez et al., 2007; Karahasan and López-Bazo, 2013; Matas et al., 2015). The results generally indicate that levels of human capital (calculated using different measures of educational attainment) are highest in those regions with the greatest

market access and decrease in line with lower market potential. Thus it has been shown that market access is a factor that influences human capital accumulation in the present time. However, the study of whether or not this relationship existed during the early stages of economic development and what form it took is still in its infancy (Diebolt and Hippe, 2018).

As its contribution to this area of study, this paper aims to analyse whether or not there was a relationship between market potential and human capital accumulation in Spanish regions during the first major stage of the industrialization process, which in the case of Spain extends from the mid-nineteenth century to the outbreak of the Civil War in 1936. There are several reasons why this case study is a useful contribution to the literature. Firstly, in this particular case the period analysed saw the beginning of the transition from an agrarian to an industrial economy that took place in parallel with an increasingly rapid integration of regional markets. It has already been verified that industrialization did not occur homogeneously across Spain and that there was in fact a marked spatial concentration of industrial production. It could therefore be said that the evolution of the Spanish economy over the period reflects a reality that is in line with that described in the NEG theoretical literature (Martinez-Galarraga et al., 2021).[1]

Secondly, approaching the analysis of the relationship between market potential and human capital accumulation from a historical and regional perspective framed in a national case offers various advantages. The first is that it ensures that the institutional framework that regulates the educational supply by the State is common to all the regions analysed, given that the presence of diverse education policies is an inherent limitation to those investigations that have addressed this type of analysis in cross-country studies. In addition, since this is a historical study, the time frame for the analysis begins before the Spanish government started its direct funding of education. Regional variations in education would therefore be fundamentally related to the conditions that existed in each region and would not be the result of government policy aimed at achieving greater territorial balance.

Finally, exploring the case of Spain over the period 1860-1930 makes it possible to avoid the effects of another element that makes empirical analysis of the model proposed in Redding and Schott (2003) difficult. As López-Rodríguez et al. (2006) point out, that model assumes an absence of mobility of qualified labour, which means that testing it involves considering evidence relating to countries or regions in which few migratory flows are recorded. There

[1] In particular, it fits the dynamics behind the model proposed by Fujita et al. (1999), on which the work by Redding and Schott (2003) is based.

was little interprovincial migration in Spain during the second half of the nineteenth and the early twentieth centuries, so we are working in a historical context in which it is reasonable to maintain the assumption of no worker mobility.[2]

Bearing these objectives and expected contributions in mind, we have assembled a provincial (NUTS3) dataset that includes a measure of market access, a series of basic human capital indicators (literacy, numeracy) and wages for different levels of skill for several benchmark years between 1860 and 1930. The results we obtain prove the existence of a relationship between provincial market potential and education levels. A mechanism that would mediate this relationship is also identified, this being regional differences in the expected return on investment in education, approximated via the wage levels of qualified industrial workers compared to those of untrained industrial or agricultural workers (Fallah et al., 2011). The analysis therefore establishes the existence of a mechanism that links market access to the unequal growth trajectory of Spanish provinces over the long term (Martinez-Galarraga et al., 2015; Díez-Minguela et al., 2018).

The work is structured as follows. Section 2 introduces the theoretical foundation of our investigation, while Section 3 provides a brief historical background describing how the location of industry and regional growth in Spain evolved over the period analysed. Section 4 describes the construction of the dataset and the characteristics of the main study variables and presents preliminary non-parametric evidence of the relationship between them. Section 5 develops the empirical test to prove the existence of a relationship between market potential and education levels, while Section 6 studies the mechanism that underlies this relationship by analysing the link between market access and the skill premium. Finally, Section 7 concludes.

2. Theoretical background

Redding and Schott (2003) were the first to consider human capital accumulation in an NEG theoretical framework. They proposed a model in which, in its initial stage, all individuals are unskilled. In its second stage, however, individuals must choose endogenously whether or not they want to invest in becoming skilled workers through the following decision rule:

[2] In the Spanish case, interprovincial migrations only reached noteworthy size in the 1920s and were mainly associated with movement towards Madrid and Barcelona (Pons et al., 2007).

$$w_i^S - w_i^U \geq \frac{h_i}{a(z)} w_i^U \quad (1)$$

The left-hand side of Equation (1) reflects the skill premium, i.e. the wage difference between skilled (w_i^S) and unskilled workers (w_i^U). This is compared to the right-hand side, which reflects the cost of investing in human capital; h_i is an inverse measure of the institutional quality of the education system, which means that higher values for h_i imply a greater cost to acquire an education (i.e. the education system is not widely established across the territory); and $a(z)$ is a measure of individual ability, so a greater ability to study will reduce the cost of acquiring human capital. According to the rule, therefore, an individual will choose to accumulate human capital as long as the skill premium is greater than the cost of accumulating that capital when still an unskilled worker.

Once the behaviour of the labour factor has been established, the authors characterize the production sector of country i using the following wage equation:

$$\left(w_i^S\right)^{\alpha}\left(w_i^U\right)^{\beta} = \zeta \frac{1}{c_i} (MA_i)^{\frac{1}{\sigma}} (SA_i)^{\frac{1-\alpha-\beta}{\sigma-1}} \quad (2)$$

where MA_i is market access and SA_i is supply access; α and β are the respective proportions of skilled and unskilled workers in the manufacturing sector; c_i is an index of technological efficiency, which may vary across territories; and σ is the elasticity of substitution between manufacturing goods. The market access and supply access variables in country i (Eq. (2)) are defined as:

$$MA_i = \sum_{j=1}^{N} \frac{E_j G_j^{\sigma-1}}{(T_{ij}^M)^{\sigma-1}} \quad (3)$$

$$(SA_i)^{\frac{1}{1-\sigma}} = G_i \quad (4)$$

where E_j is total consumer expenditure on manufactured goods in importing country j; G_j is a price index for manufactured goods in importing country j; and T_{ij}^M is the bilateral trade costs between i and j.

Following Equations (3) and (4), Redding and Schott (2003) concluded that areas with good supply access (SA_i) will have a lower manufacturing price index, thereby reducing unit production costs and increasing the maximum wage that manufacturing companies in these areas would be willing to pay. If these areas also have good market access scores (MA_i), this would further increase the maximum wage that companies would be willing to pay. Considering this characterization of the sector and combining the zero-profit condition in both agriculture and manufacturing, the authors define the equilibrium relationship between geographical location and endogenous investment in human capital as:

$$0 = \phi \frac{dw_i^S}{w_i^S} + (1 - \phi) \frac{dw_i^U}{w_i^U} \tag{5}$$

$$\alpha \frac{dw_i^S}{w_i^S} + \beta \frac{dw_i^U}{w_i^U} = \frac{1}{\sigma} \frac{dMA_i}{MA_i} + \frac{(1 - \alpha - \beta)}{\sigma - 1} \frac{dSA_i}{SA_i} \tag{6}$$

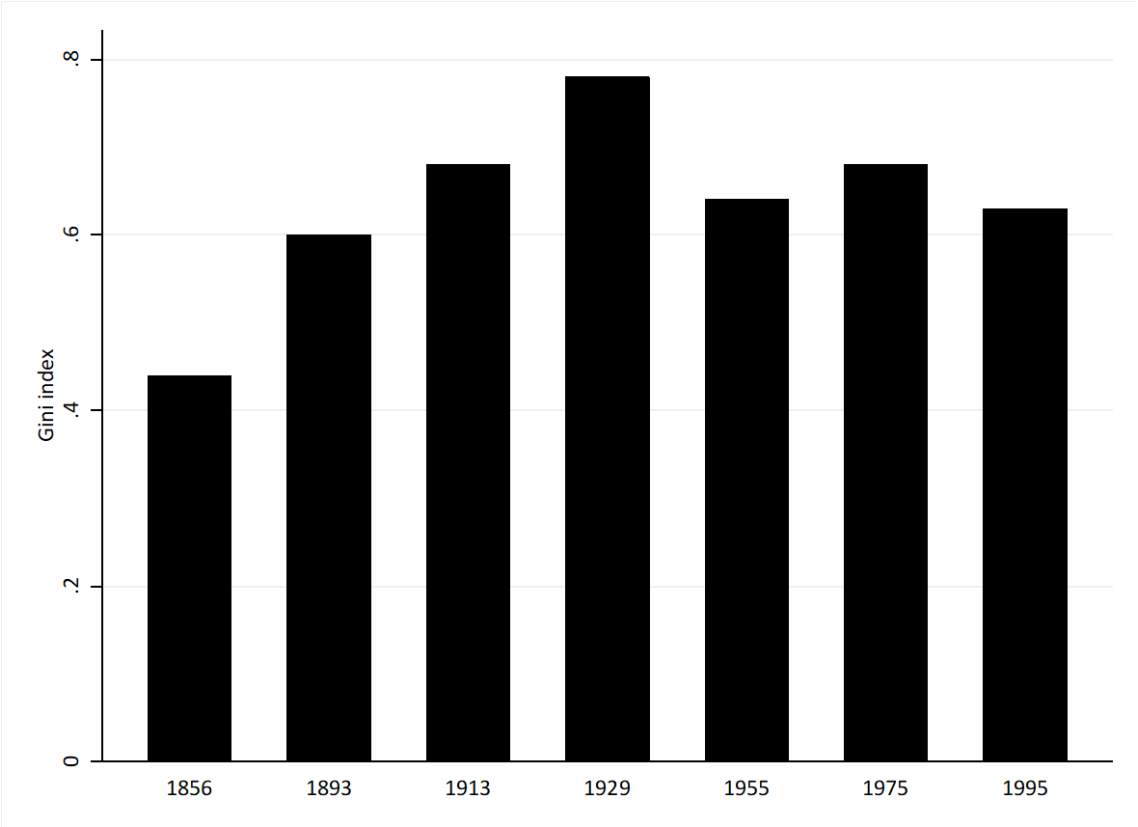
where ϕ is the proportion of skilled labour in agriculture and α and β are again the respective proportions of skilled and unskilled workers in the manufacturing sector. Thus if a peripheral region starts to increase its market access and supply access, this will increase the wages of its skilled workers. Returning to Equation (1), an increase in skilled workers' wages will increase the skill premium, the incentives to acquire human capital and eventually the numbers of skilled workers in that region.

3. Market access, industry location and regional inequality in Spain

It is a well-established fact in Spanish economic history that from the mid-nineteenth century to the outbreak of the Civil War (1936-1939), industrial production gradually agglomerated in a small number of provinces (Nadal, 1987; Paluzie et al., 2004). Economic historians have thoroughly investigated the roots and causes of the specific location of industrial activity and of the notable increase in the spatial concentration of manufacturing over this period (Figure 1). Rosés (2003), following Davis and Weinstein (1999, 2003), finds evidence that the *home*

market effect was behind early Catalan industrialization (around the 1860s). Tirado et al. (2002), in line with Kim (1995), identify scale economies and market size as determining Spain's industrial geography in the mid-nineteenth century. By the end of the century, the explanatory power of these NEG effects had increased in parallel with advances in the economic integration process.

Figure 1. Concentration of manufacturing in Spain, 1856-1995 (NUTS3)



Source: Paluzie et al. (2004).

Adopting the approach developed by Midelfart-Knarvik et al. (2002), Martinez-Galarraga (2012) confirms and extends the previous findings of Tirado et al. (2002).[3] In short, it has been concluded that, as the domestic market gradually integrated and industrialization progressed in Spain during the second half of the nineteenth century, NEG forces grew to be the main determinant of Spain's industrial landscape. In particular, it has been proven that the interaction between economies of scale and market potential favoured the concentration

[3] This methodology has repeatedly been used in economic history: Wolf (2007) analysed Poland after WWI, Crafts and Mulatu (2005, 2006) studied Britain, and Klein and Crafts (2012) applied the approach to examine the manufacturing belt in the US between 1880 and 1920. Betrán (2011) also analysed the Spanish case in the long term at regional NUTS2 level.

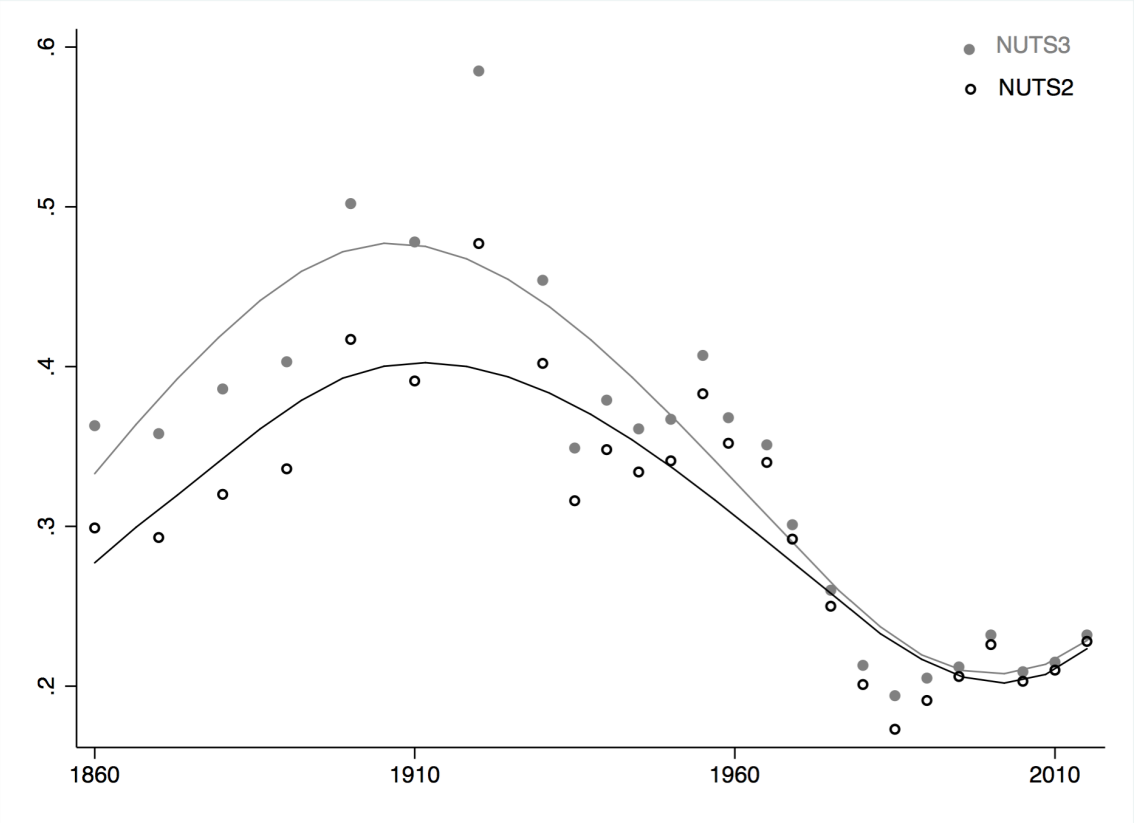
of industries with increasing returns in those provinces with the best access to demand up to the 1930s. In addition, Martinez-Galarraga et al. (2008), in line with Ciccone and Hall (1996) and Ciccone (2002), suggest that agglomeration economies were present in Spain's manufacturing during the early stages of industrialization, as those regions with a higher spatial density of economic activity also enjoyed higher industrial labour productivity.

Researchers have additionally tested not only the wage equation, i.e. the existence of higher wages in regions with greater market potential resulting from the agglomeration of manufacturing (backward linkages), but also the ability of these wages to generate migratory flows of workers (forward linkages). On this subject and following Hanson's (1998, 2005) influential research based on the Krugman wage equation, Tirado et al. (2013) examine the existence of a spatial structure in nominal wages in industry in 1920s Spain. The results confirm that wages were higher in regions with greater market potential. In addition, extending the work done by Crozet (2004), Pons et al. (2007) establish a direct relationship between migration decisions and the market potential of the host regions during the 1920s, thus verifying the presence of forward linkages in the internal migrations between Spain's provinces in the interwar years.

Beyond the manufacturing sector, Rosés et al. (2010) focus on regional inequality in terms of GDP per capita, constructing a population-weighted coefficient of variation to measure the long-term evolution of disparities in regional income per capita at provincial level. They show that, clearly echoing the growing spatial concentration of manufacturing, the second half of the nineteenth century and the first decades of the twentieth also saw a remarkable increase in regional income inequality (Figure 2). Going one step further, Diez-Minguela et al. (2018) claim that the growth in regional inequality generated a geographical pattern that appears to be well-established in 1930 and has indeed persisted right up to the present.

Finally, Martinez-Galarraga et al. (2015), in line with Ottaviano and Pinelli (2006), show how the regional inequality that characterized the Spanish economy in 1930 was due to the existence of sizeable differences in regional growth trajectories during the period 1860-1930 and the fact that these differences were related to heterogeneous access to markets. Building on Brülhart and Sbergami (2009), Díez-Minguela et al. (2016) reinforce this result, finding that agglomeration economies are important when it comes to explaining regional inequality in Spain before the Civil War (1936-39) and that industrial agglomeration during the second half of the nineteenth century and the first decades of the twentieth was therefore crucial to this process.

Figure 2. Regional per-capita income inequality in Spain, 1860-2015



Note: Population-weighted coefficient of variation.
Source: Díez-Minguela et al. (2018).

Given these conditions, the question we aim to investigate revolves around the economic mechanism that connects the concentration of manufacturing with unequal regional economic growth. Using Redding and Schott's (2003) hypothesis as a starting point, we suggest that it was differential access to markets that determined which regions would see greater advances in industrial production. In turn, the greater or lesser industrial specialization of the regions brought about the appearance of regional differences in the wages for skilled work. Under these circumstances, the incentives to invest in education were greater in those regions with the greatest market potential, an element that ultimately gave rise to differences in regional rates of human capital accumulation. In order to further investigate this hypothesis, the following section presents the database we have constructed and some descriptive evidence regarding regional differences in market access and education levels.

4. Human capital and market access: background, data and preliminary evidence

In order to understand the differences in education between regions we first need to examine the regulatory framework for primary education in Spain in 1860, which was the result of a long gestation process that culminated in 1857 with the passing of the Public Instruction Act (PIA), also known as the Moyano Law. This regulated the Spanish education system from 1857 to 1970, when it was replaced by the General Education Act (GEA).[4] This framework organized the education system, setting out its stages and establishing the content of each subject. It also listed the obligations that fell to families (the schooling of their children) and municipalities (the opening of schools and their conservation).

Certainly the provision of education infrastructures in 1860 was the responsibility of the municipalities and its funding was expected to come from these entities and families and, to a lesser extent, from secular or religious foundations. From the mid-nineteenth century and during the early twentieth, the Spanish economy and society underwent a number of important changes, including rapid population growth and increasing structural transformation (Pérez Moreda et al., 2015; Prados de la Escosura, 2017). It was in this context of far-reaching socioeconomic change that the Royal Decree of 18 April 1900 created the Ministry of Public Instruction and Fine Arts, and this was followed two years later by the central government taking over the funding of primary education.

Bearing all this in mind and with the aim of studying regional differences in human capital endowment and its determinants in greater detail, our dataset includes different indicators for different education levels and a measurement for the market potential of Spanish provinces for the period 1860-1930. Other variables such as province size, measured by total population, and population density are also included. The dataset is in the form of a panel spanning 5 benchmark years (1860, 1900, 1910, 1920 and 1930) and covering a total of 47 Spanish provinces.[5]

Human capital is a broad, complex concept (Goldin, 2015) and there are therefore different ways of approaching it (Barro and Lee, 2013; Hanushek and Woessmann, 2012). Here we focus on basic human capital and begin by calculating provincial literacy rates. Spain's population censuses provide information on self-reported literacy rates back to the mid-

[4] Successive changes were nevertheless introduced during the 113 years that the PIA remained in force. Compulsory education, for example, initially from ages 6 to 9, was extended to age 12 in 1909 and age 14 in 1964.

[5] In constructing the market potential variable we excluded the three non-mainland provinces from the analysis, and thus we work with a total of only 47 provinces.

nineteenth century, so we have been able to gather data for our benchmark years, broken down by sex. These data refer to the percentage of the population over age 10 that could read and write. It should be noted that these rates vary greatly according to sex and therefore the aggregate or total literacy rates actually capture two rather different stories, one for males and one for females. Thus we focus on male literacy rates and then use the female and total literacy rates for the purposes of robustness.

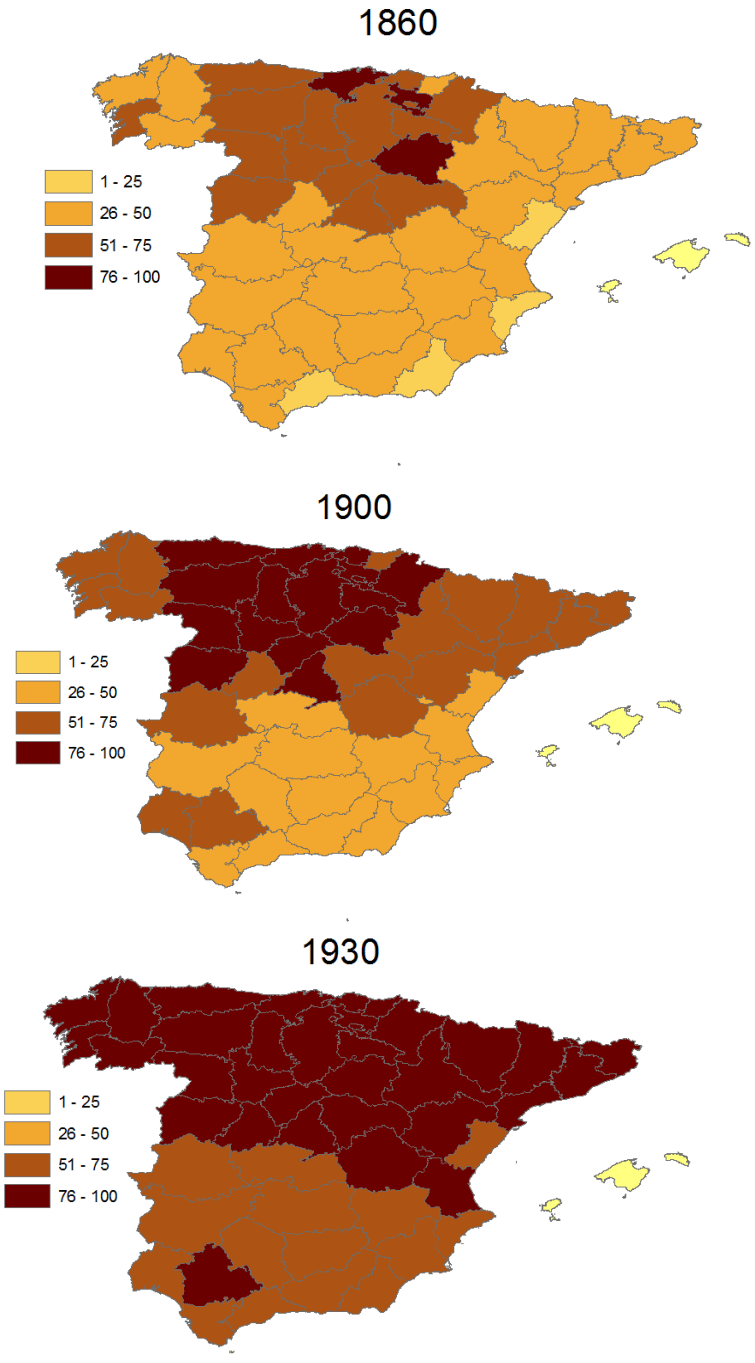
Given the virtual absence of public funding for education and the limited development of the Spanish economy, it is understandable that education levels were very low at the beginning of the period under study and that, in addition, there was much regional variation in the provision and demand for schooling (Figure 3). The male literacy rate at aggregate level was 41.7% in 1860.[6] Regions in the center-north of the country, such as Castile-Leon, Cantabria, Alava and Navarre, had the highest levels of male literacy in the mid-nineteenth century. At the other extreme were the southern regions, i.e. Andalusia, Extremadura, Castile-la Mancha, Murcia and Valencia, along with those in the north-west of the peninsula, in Galicia, for example.

Male literacy rates improved over the period studied. By 1900 they had increased to 57.1% for the country as a whole and would continue growing to reach 83.0% in 1930.[7] However, this improvement was not equally distributed across the country. To begin with, the spatial patterns identified for male literacy rates in 1860 simply became more pronounced during the final decades of the century. It appears that the Moyano Act of 1857 was unable to iron out the pre-existing imbalances in education infrastructure endowments. Indeed, it could be argued that the opposite was true. Regional differences were stronger or more obvious at the turn of the century and regional disparities were still sizeable in 1930. The image of literacy rates in 1900 had evolved towards a clear north-south divide, with the northern provinces moving towards universal literacy and the southern provinces continuing at levels which were sometimes below 60%.

[6] Female literacy, on the other hand, was only 11.9%. Taken as a whole, of a total population of almost 16 million in Spain it is estimated that only 26.5% could read and write in 1860. The provincial distribution of female and total literacy rates between 1860 and 1930 can be seen in Figures A.1 and A.2 respectively, in the Appendix.

[7] In 1930 the female literacy rate was 66.3% and total literacy was 74.4%.

Figure 3. Male literacy rates, 1860-1930



Source: Population censuses.

Our analysis also includes provincial-level numeracy indexes as a complementary measure of human capital. Numeracy skills can be defined as the ability to understand and work with numbers, so ultimately numerical knowledge and number discipline may also be crucial for industrial activities and economic growth. In order to obtain a better picture of these basic numerical skills, a growing number of works have focused on age heaping by making use of

historical data on ages (Tollnek and Baten, 2016). It has been suggested that an irregular pattern of age distribution or a preference for certain digits may indicate the presence of misreported age, which can become an indicator of a society's aggregate numeracy skills. Our data come from Beltrán Tapia et al. (2021), who calculate age heaping on the basis of the year-by-year age of the population included in the censuses published since the late nineteenth century for each of the 47 peninsular Spanish provinces, broken down by sex. In line with the extensive literature, we first calculate the Whipple index and then convert it into an ABCC index (for males), which ranges from 0 to 100.[8]

Along with measures of human capital, our analysis also relies on an indicator of provincial market access or market potential (Martinez-Galarraga, 2014). The role of market accessibility has regained significance as a factor explaining the spatial distribution of economic activity thanks to the crucial role it plays within economic geography models (Fujita et al., 1999; Combes et al., 2008). Here we measure market access using the traditional Harris market potential equation (Harris, 1954), also known as the nominal market potential. This captures the idea that the accessibility of a location depends on the size of other markets, usually measured in terms of GDP, once distances (or bilateral transport/trade costs) have been deducted.

This methodology has traditionally been used by geographers and economists (Clark et al., 1969; Keeble et al., 1982), although recently it has also grown more popular among economic historians, following in the steps of Crafts (2005) and his analysis of Victorian Britain. Harris's market potential function has also appeared in empirical works within the field of economic geography. The validity of this indicator in such a context is supported by two facts: first, the results obtained from empirical exercises in economic geography do not provide conclusive evidence that alternative measures might be more effective in terms of robustness (Head and Mayer, 2004),[9] and second, a direct relationship can be established between Harris's equation and the market access measures deriving directly from NEG models, thus providing a theoretical foundation for Harris's market potential equation (Combes et al., 2008).

[8] Here we use male ABCC indexes, given that the lack of gender gaps in age-heaping in Spain's provinces throws doubt on whether the information was actually self-reported by females (Beltrán Tapia et al., 2021). On the issue on female age heaping, see also Földvári et al. (2012) and Blum et al. (2017). The spatial patterns and evolution of the male ABCC indexes are shown in Figure A.3 in the Appendix. The first available year in this case is 1877.

[9] "Despite the fact that we bring theory to empirical implementation in a structural way, the 'correct' measure of market potential actually underperforms the atheoretical Harris (1954) measure" (Head and Mayer, 2004).

One of the noteworthy characteristics of Harris's indicator is that market potential can be broken down into domestic and foreign (referred to here as internal and external market potential). This means we can find out how much of a province's total market potential is made up of demand from other provinces (internal) and how much is linked to demand from other countries, i.e. its main trade partners (external). In addition, internal market potential not only includes all the other provinces of Spain but also the size of the market within the province in question, i.e. its self-potential. This indicator therefore provides additional information enabling us to further break down the domestic component in the analysis we carry out. In order to calculate internal market potential we need data for provincial GDP (Rosés et al., 2010) and an estimate of interprovincial transport costs. In this regard, rail access (using a network that was constantly expanding over the relevant period), sea access and cabotage transport played a crucial role.[10] When considering external market potential, on the other hand, the size of the external markets is estimated based on international maritime transport along with the tariffs applied by the main import/export markets considered.[11]

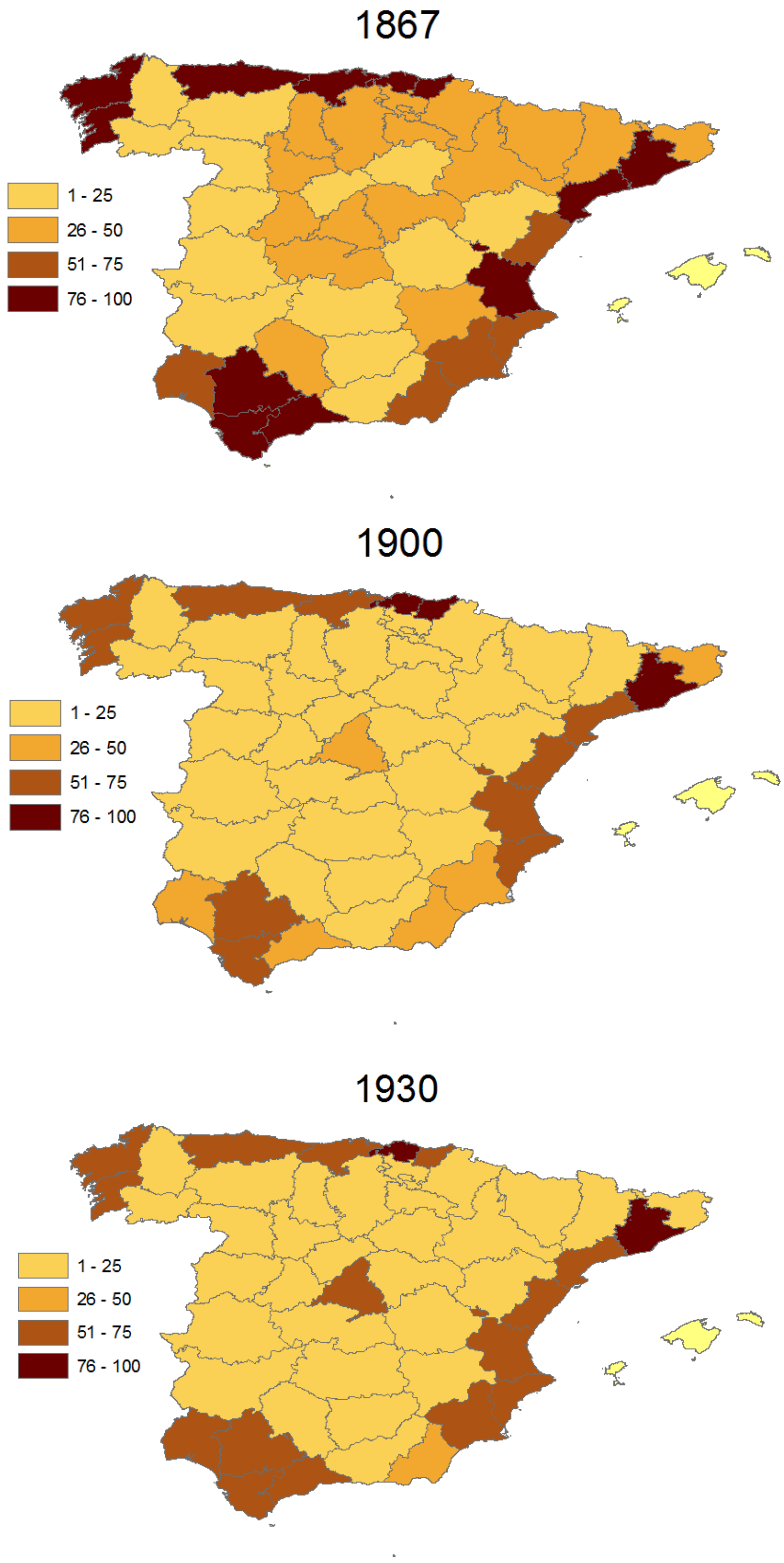
The maps in Figure 4 show the geographical pattern of regional accessibility and how it evolved from 1860 to 1930.[12] Throughout this period Barcelona had the greatest market potential, and therefore the maps are expressed in relation to this province. The evidence shows that the most significant changes in the relative accessibility of the provinces occurred in the second half of the nineteenth century in parallel with domestic market integration. A centrifugal tendency can be seen as the geographical structure evolved towards a clear division between inland and coastal provinces, with the latter showing a higher market potential than their inland counterparts, with the sole exception of Madrid. It could be hypothesized that the expansion of the railway network – all provincial capitals were connected to the network by 1901 – was responsible for much of the change in the pattern of market potential. Once this dual structure had taken shape at the end of the nineteenth century, the division between inland and coastal provinces persisted over the first few decades of the twentieth.

[10] This is why the first year for which we have information for this variable is 1867. This has a historical logic due to the fact that the first basic stage in the construction of the country's rail network had been completed by 1866, joining up the major cities and centres of economic activity.

[11] Here we follow Crafts (2005), who suggested using the elasticities associated with distance and tariffs to calculate the external market potential. The values for the elasticities come from the estimation of international trade equations by Estevadeordal et al. (2003). The details involved in the construction of the market potential variable for Spain's provinces can be found in Martinez-Galarraga (2014).

[12] In Figures A.4 and A.5 in the Appendix we include information on the geographical distribution of internal and external market potential respectively.

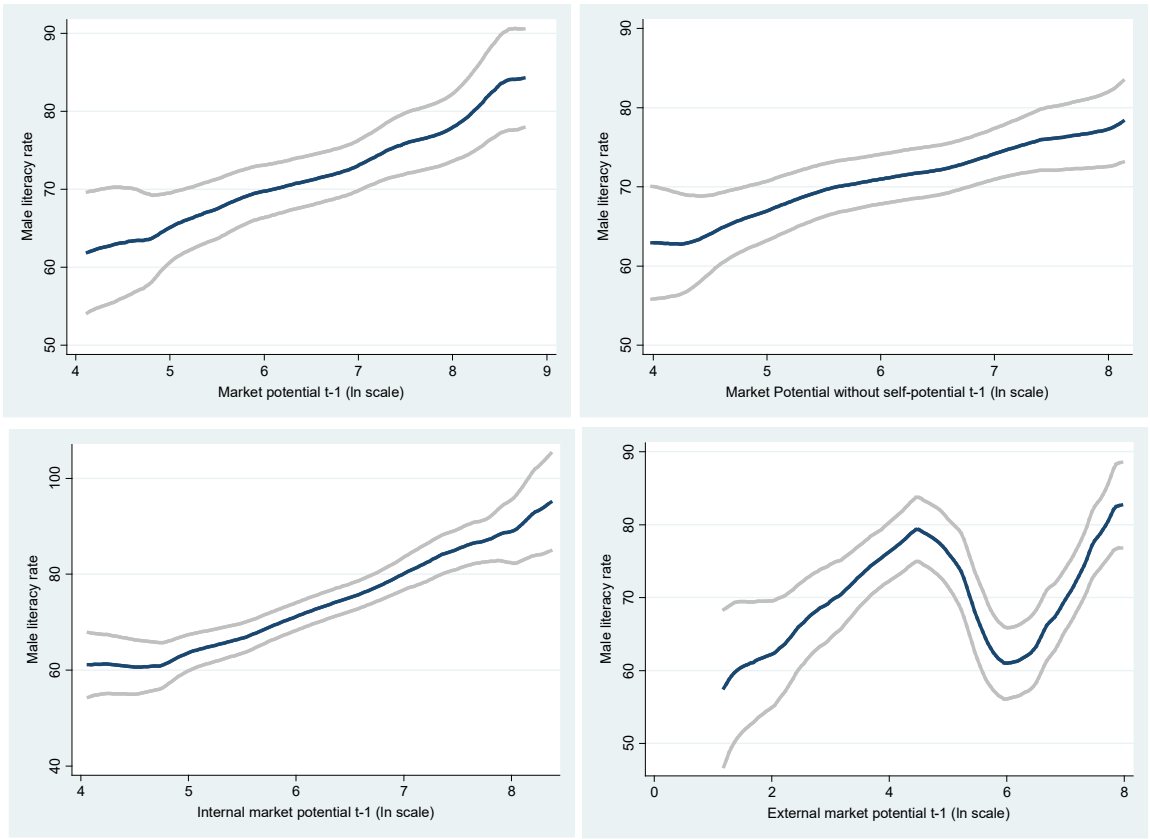
Figure 4. Market potential, 1867-1930 (Barcelona=100)



Source: Martinez-Galarraga (2014).

Now that all the main variables have been presented, the next step is to perform a preliminary exploratory approximation to discover whether there is a relationship between provincial differences in market access and the pronounced regional variability in literacy rates. First of all we present non-parametric evidence as to the existence of a relationship between the relative market potential and provincial male literacy rates (Figure 5). These graphs show how the male literacy rate varies across the market potential distribution for four different definitions of market access: total market potential, total market potential excluding self-potential, internal market potential and external market potential.[13]

Figure 5. Non-parametric relationship between male literacy rates and different measures of lagged market potential (1860-1930)



Source: see text.

In Figure 5, all the male literacy/market potential pairs from 1860 to 1930 are considered in a pool (taking into account the lagged market potentials in $t-1$). The four graphs show a similar picture: the higher the market potential, the higher the male literacy rate. This

[13] Graphs using total and female literacy rates are similar to those shown for male literacy rates. These are available from the authors on request.

increasing relationship is similar for all the market potential measures, although in the case of external market potential the relationship shows a inverted U-shaped pattern for some values. Overall, we can conclude that there was a positive relationship between market potential and male literacy rates in the period considered, and there are no signs of nonlinear effects except in the case of external market potential.[14] Nevertheless, this is only a preliminary descriptive analysis for illustrative purposes, as these graphs represent the unconditional relationship between literacy rates and market potential. The rest of the paper presents the results of an analysis that explores this relationship following a sounder theoretical and empirical method.

5. Market access and educational attainment

In the parametric analysis we make use of the panel structure of our data to test for the existence of a relationship between market potential and male literacy rates at provincial level. Following Redding and Schott (2003) and Fallah et al. (2011), given that the proportion of the literate population is bounded between 0 and 1, we use a logistic transformation of the literacy rates. We then estimate a model in which the male literacy rate is regressed on the market potential for the previous period:[15]

$$\ln\left(\frac{literacy\ rate_{it}}{1-literacy\ rate_{it}}\right)=\alpha+\beta\ln(MP_{i,t-1})+\Gamma' X_{it-1}+\Pi' T_{it}+\varphi\eta_i+\varepsilon_{it}, \qquad (7)$$

where our main explanatory variable is MP_{it-1}, i.e. the lagged market potential. As stated earlier, we consider four alternative measures of market potential corresponding to its different components: total market potential, internal market potential, external market potential and a version of total market potential that excludes each province's self-potential in order to reduce some endogeneity concerns.

[14] If we consider the ABCC index as the human capital indicator, the results are similar to those obtained with literacy rates (see Figure A.6 in the Appendix).

[15] If instead of using lagged values we take contemporaneous values of market potential – as Redding and Schott (2003) did – we obtain similar results. These are available from the authors on request.

X_{it-1} is a vector of additional (lagged) control variables at provincial level that represent factors that could attract skilled individuals or impose restrictions on the provision of primary education (Fallah et al., 2011). They form a measure of labour abundance that considers both the province's total population and its population density. On this subject, the historical literature on primary education in Spain has shown that there is a positive link between population size and education levels, pointing to a relationship between the legal framework of school supply, which favoured its spread across more populated areas, and the presence of economies of scale in education provision (Viñao, 1990; Pérez Moreda, 1997; Reher, 1997; Núñez, 2003). It has also been suggested that the costs associated with primary education are higher in rural contexts with sparse populations, where literacy rates are consistently lower, than in urban contexts, where there would be more demand for skilled work and the provision costs would be lower. Finally, η_i denotes provincial fixed effects (FE) to control for other provincial characteristics not accounted for in the specification (e.g. first-nature causes and geography) and T_{it} is a set of province-specific time trends ($\sum_{i=1}^{n-1} \text{Province}_i \cdot Time_t$) that capture the particular behaviour of the provinces in our panel over time.

Some endogeneity concerns may arise because an important component of Harris' market potential function is the contribution of its own GDP to the potential of region *i* (self-potential) and, as mentioned above, the provision of education infrastructures in 1860 was the responsibility of the local authorities. It could therefore be argued that those regions with higher GDP (and higher market potential) would spend more on education and thus have higher literacy rates. All explanatory variables are lagged one period to mitigate any direct endogeneity between the dependent and independent variables. To tackle any potential endogeneity problem we estimate Eq. (7) using panel FE and instrumental variables (IV). In addition, one of the market potential measures we consider excludes self-potential, which should avoid any possible simultaneity problems.

Columns 1 to 4 in Table 1 report the results of the panel FE estimates. The coefficient of market potential is positive and significant at a 1% level for all four different measures of market potential, indicating that greater market potential is associated with higher literacy rates. Moreover, the estimated coefficients across the different definitions of market potential are quite similar, which suggests a consistent relationship between market potential and education. Columns 5–8 of Table 1 present the IV (2SLS) results. The instruments considered for the market potential variable are the distance to the nearest main industrial

centre (Redding and Venables, 2004), a centrality index measured as $\sum_{j\neq i}^{n-1} d_{ij}^{-1}$ (Head and Mayer, 2006) and knowledge capital, measured as the number of patents per capita (Sáiz, 2005), all in logs. These three variables are used to instrument the four definitions of market potential. Table 1 also shows some statistics from the first- and second-stage regressions that indicate that our instruments perform well, since the F-test takes high values and all models pass the overidentification test (Hansen J statistic) for any level of significance. Furthermore, the main result holds in the IV regressions, given that the estimated coefficients of the market potential variables remain positive and significant, although they are slightly higher than those obtained in the panel FE regressions.

We carry out two robustness checks. First, to consider possible gender differences in education levels, we estimate Eq. (7) using the total (including both males and females) instead of just the male literacy rate. Table 2 reports the results. Both the panel FE and the IV estimations yield very similar results to those obtained using the male literacy rate. Again, the effect of market potential on education is positive and similar across the different definitions. The estimated coefficients are almost identical to those reported in Table 1 (with differences arising from the second decimal place in most cases). Furthermore, results unreported here show that, when considering the female literacy rate as the dependent variable, the coefficients are also very similar to those reported in Table 1.

Second, we consider an alternative measure of basic human capital that potentially captures numeracy levels based on age-heaping methodology, as mentioned earlier (Beltrán Tapia et al., 2021). We then rerun Eq. (7) using numeracy (ABCC index for males) instead of literacy rates as the dependent variable. Table 3 shows the results. Again, we find a positive and significant effect of market potential on education that is consistent across the different definitions. The main differences are that, first, the magnitude of the estimated effect is smaller given that the estimated coefficients for the market potential variables are lower than those obtained in Table 1, and second, in this case there are significant differences between the panel FE and IV estimates, since the IV estimated coefficients (columns 5 to 8) are approximately half the panel FE estimates (columns 1 to 4). Nevertheless, although the magnitude of the effect is different, the main result holds, given that we obtain a positive and significant effect of market potential in all cases.

Table 1. Effects of market potential on male literacy rates

	Panel FE				IV (2SLS)			
Variables:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Market potential	1.180***				1.280***			
	(0.107)				(0.121)			
Market potential without self-potential		1.160***				1.240***		
		(0.116)				(0.121)		
Internal market potential			1.259***				1.400***	
			(0.112)				(0.132)	
External market potential				1.052***				1.011***
				(0.105)				(0.092)
Labour abundance	-1.176***	-1.009**	-1.142***	-0.433	-1.209***	-1.097***	-1.231***	-0.425**
	(0.401)	(0.447)	(0.378)	(0.449)	(0.318)	(0.302)	(0.345)	(0.213)
Population density	1.536***	1.584***	1.695***	0.695	1.013	1.240*	1.076*	0.847
	(0.464)	(0.532)	(0.437)	(0.629)	(0.640)	(0.637)	(0.652)	(0.588)
Provincial fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province x Time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.958	0.956	0.959	0.956				
Observations	188	188	188	188	188	188	188	188
First stage, Shea partial R^2					0.452	0.480	0.416	0.602
First stage, F-test (p-value)					46.89 (0.000)	55.33 (0.000)	39.48 (0.000)	100.93 (0.000)
Hansen J statistic, p-value					0.270	0.234	0.285	0.202
Uncentred R^2					0.975	0.974	0.976	0.975

Notes: Dependent variable: logistic transformation of the male literacy rate. Coefficient (robust standard errors). Robust standard errors clustered by province in panel FE regressions. All variables in logarithmic scale and all models include a constant. Significant at *10%, **5% and ***1% levels. Instruments in IV regressions: distance to the nearest main industrial centre, centrality index and knowledge capital, all in logs.

Table 2. Effects of market potential on total literacy rates

	Panel FE				IV (2SLS)			
Variables:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Market potential	1.134***				1.245***			
	(0.104)				(0.086)			
Market potential without self-potential		1.137***				1.204***		
		(0.110)				(0.085)		
Internal market potential			1.192***				1.366***	
			(0.103)				(0.090)	
External market potential				1.018***				0.975***
				(0.097)				(0.072)
Labour abundance	-1.329***	-1.200***	-1.271***	-0.622**	-1.198***	-1.088***	-1.221***	-0.438*
	(0.261)	(0.266)	(0.263)	(0.280)	(0.301)	(0.283)	(0.314)	(0.226)
Population density	0.862**	0.846**	1.059***	0.027	0.195	0.425	0.244	0.076
	(0.345)	(0.362)	(0.334)	(0.463)	(0.385)	(0.366)	(0.417)	(0.324)
Provincial fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province x Time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.973	0.973	0.973	0.971				
Observations	188	188	188	188	188	188	188	188
First stage, Shea partial R^2					0.452	0.480	0.416	0.602
First stage, F-test (p-value)					46.89 (0.000)	55.33 (0.000)	39.48 (0.000)	100.93 (0.000)
Hansen J statistic, p-value					0.218	0.186	0.239	0.194
Uncentred R^2					0.977	0.978	0.976	0.977

Notes: Dependent variable: logistic transformation of the total literacy rate. Coefficient (robust standard errors). Robust standard errors clustered by province in panel FE regressions. All variables in logarithmic scale and all models include a constant. Significant at *10%, **5% and ***1% levels. Instruments in IV regressions: distance to the nearest main industrial centre, centrality index and knowledge capital, all in logs.

Table 3. Effects of market potential on numeracy levels (ABCC index)

	Panel FE				IV (2SLS)			
Variables:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Market potential	0.445***				0.248***			
	(0.065)				(0.056)			
Market potential without self-potential		0.443***				0.239***		
		(0.064)				(0.054)		
Internal market potential			0.466***				0.272***	
			(0.070)				(0.062)	
External market potential				0.358***				0.195***
				(0.059)				(0.045)
Labour abundance	-0.550***	-0.495***	-0.524***	-0.230*	-0.169**	-0.147*	-0.173**	-0.017
	(0.143)	(0.127)	(0.158)	(0.117)	(0.079)	(0.076)	(0.083)	(0.068)
Population density	0.020	0.023	0.103	-0.144	0.499**	0.546**	0.510**	0.468**
	(0.203)	(0.192)	(0.212)	(0.241)	(0.236)	(0.224)	(0.242)	(0.236)
Provincial fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province x Time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.965	0.965	0.965	0.958				
Observations	188	188	188	188	188	188	188	188
First stage, Shea partial R^2					0.452	0.480	0.416	0.602
First stage, F-test (p-value)					46.89 (0.000)	55.33 (0.000)	39.48 (0.000)	100.93 (0.000)
Hansen J statistic, p-value					0.296	0.296	0.286	0.476
Uncentred R^2					0.997	0.997	0.997	0.997

Notes: Dependent variable: logistic transformation of the numeracy levels (ABCC index). Coefficient (robust standard errors). Robust standard errors clustered by province in panel FE regressions. All variables in logarithmic scale and all models include a constant. Significant at *10%, **5% and ***1% levels. Instruments in IV regressions: distance to the nearest main industrial centre, centrality index and knowledge capital, all in logs.

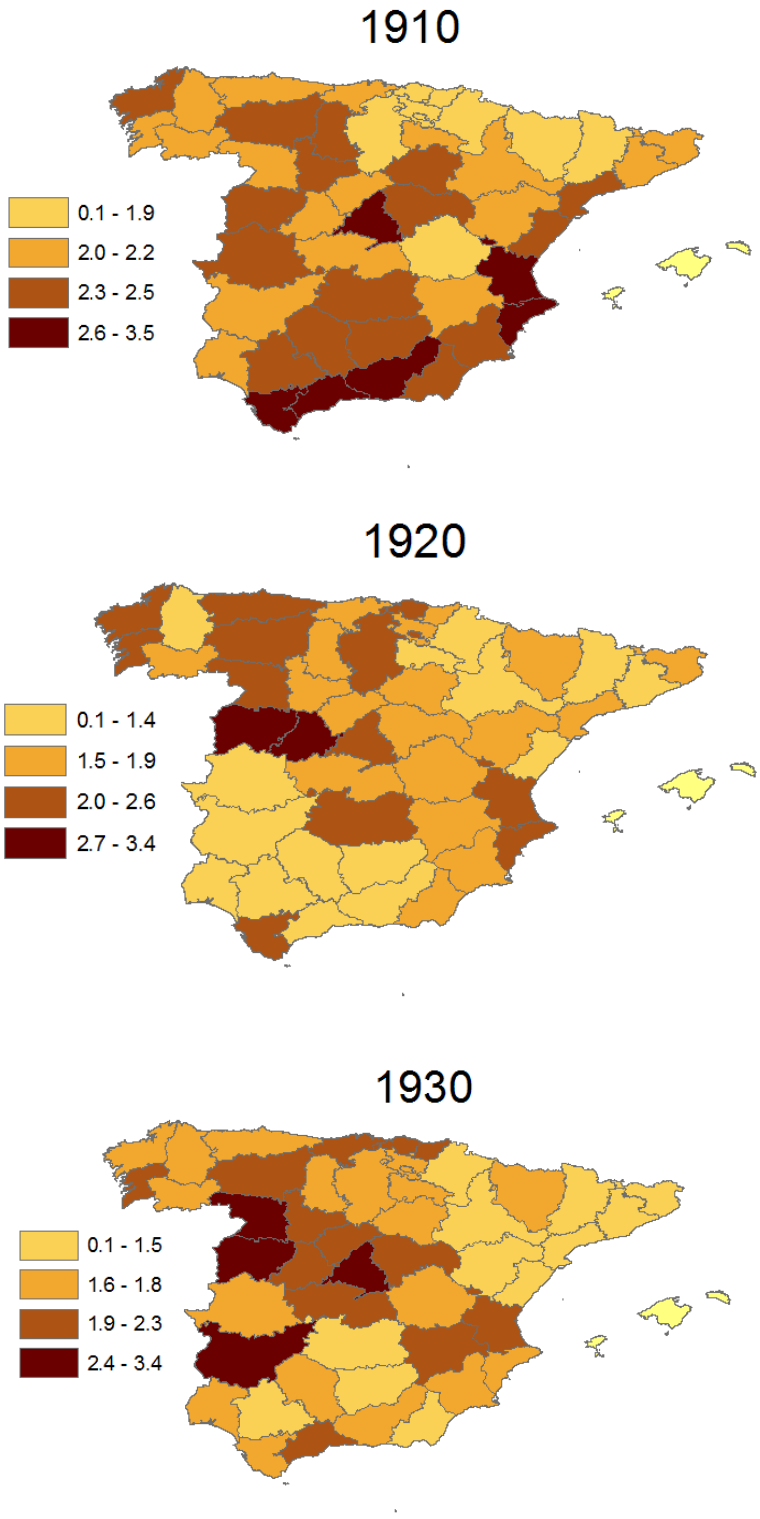
6. Exploring the economic mechanism: market access and the skill premium

Fallah et al. (2011) provide an NEG model to explain wage differentials using market access. Their theoretical framework establishes a structural relationship between nominal skill premiums and access to markets given the location of economic activities. Basically, an increase in the level of market potential generates an expansion of the manufacturing sector, and this expansion implies a higher demand for skilled labour than the agricultural sector can release at the initial equilibrium factor prices. Therefore, in the new equilibrium the nominal wage of skilled workers rises and the nominal wage of unskilled workers declines after an increase in market potential.

While Fallah et al. (2011) tested this theoretical mechanism using a cross-section of US cities in 2001, here we take advantage of our historical dataset to estimate a panel data model, although the sample period (1910 to 1930) is shorter than in previous sections because of limitations in wage data availability. We use wage per hour data (in *pesetas*) for three different types of workers: skilled (*obreros cualificados*) and unskilled industrial workers (*peones*) and agricultural workers. The data come from the Ministerio de Trabajo (1927, 1931) for industrial workers and Bringas (2000) in the case of agricultural wages. Figure 6 shows the skill premium between skilled industrial workers and agricultural workers, which is our preferred measure given that it is the most closely related to Redding and Schott's (2003) theoretical model.[16]

[16] The maps for the skill premium between skilled and unskilled industrial workers can be found in Figure A.7 in the Appendix.

Figure 6. Skill premium: skilled industrial workers v. agricultural workers



Source: Ministerio de Trabajo (1927, 1931)

At first glance the maps in Figure 6 present a complex picture. On the one hand, it is difficult to detect a clear geographical pattern in the spatial distribution of the skill premium, while on the other, extracting any conclusions from the changes observed between 1910 and 1930 does not appear straightforward. Therefore, in order to further examine this regional variation and assess the relationship between market potential and the skill premium, we estimate the following panel data model:

$$\ln\left(wage\ skilled_{it}\right) - \ln\left(wage\ unskilled_{it}\right) = \alpha + \beta \ln(MP_{i,t-1}) + \Gamma' X_{it-1} + \Pi' T_{it} + \varphi\eta_i + \varepsilon_{it}, \tag{8}$$

where the term on the left-hand side is the wage differential across two different skill groups (skilled and unskilled) again measured at provincial level, and the variables on the right-hand side are those used in Eq. (7). Again, the main explanatory variable is the lagged market potential (MP_{it-1}), for which we use the four different definitions.

Table 4 shows the results of the panel FE estimations.[17] We consider two different skill premiums: first, the premium of skilled industrial workers over agricultural workers (columns 1 to 4), and second, a measure that compares the wages of skilled versus unskilled industrial workers (columns 5 to 8). For the first we obtain a positive and significant effect of market potential (for any definition) on the wage differential between skilled industrial workers and agricultural workers. In other words, greater market access meant an increase in wages for skilled workers compared to agricultural workers. Furthermore, the magnitude of the effect is robust across all different versions of market potential. Since the variables are in logarithmic scale the estimated coefficients are the elasticities. The interpretation of this would be that, for every 1% increase in market potential, the skill premium would increase by around 0.5%. In the case of the second premium considered (wages of skilled versus unskilled industrial workers), the results are less conclusive. We still obtain a positive effect of market potential on wage differentials, but the magnitude of the effect is lower (the estimated coefficients of the market potential variables are below 0.2) and it is only significant at a 10% confidence level (columns 5 to 7).

[17] The IV estimates yield similar conclusions. These results are available from the authors on request.

Table 4. Effects of market potential on the skill premium

Dependent variable:	Wage skilled industrial workers v. wage agricultural workers				Wage skilled v. wage unskilled industrial workers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Market potential	0.503**				0.183*			
	(0.192)				(0.103)			
Market potential without self-potential		0.543**				0.198*		
		(0.216)				(0.113)		
Internal market potential			0.422**				0.158*	
			(0.182)				(0.092)	
External market potential				0.648*				0.232
				(0.327)				(0.177)
Labour abundance	1.653	1.802	1.516	2.095	-0.165	-0.112	-0.208	-0.013
	(1.178)	(1.209)	(1.221)	(1.363)	(0.842)	(0.849)	(0.852)	(0.887)
Population density	-0.163	-0.071	-0.088	-0.971	-0.337	-0.304	-0.309	-0.626
	(1.099)	(1.108)	(1.132)	(1.124)	(0.787)	(0.785)	(0.799)	(0.808)
Provincial fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province x Time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.806	0.806	0.794	0.792	0.728	0.728	0.723	0.717
Observations	141	141	141	141	141	141	141	141

Notes: Panel FE regressions. Dependent variable: ln(*wage skilled industrial* workers) – ln(*wage agricultural* workers) in columns 1 to 4, ln(*wage skilled industrial* workers) – ln(*wage unskilled industrial* workers) in columns 5 to 8. Coefficient (robust standard errors). Robust standard errors clustered by province. All variables in logarithmic scale and all models include a constant. Significant at *10%, **5% and ***1% levels.

7. Concluding remarks

The existence of notable differences between regions is one of the most striking elements of the economic growth processes in developed and developing economies. The NEG is a conceptual framework that has made it possible to obtain a well-founded explanation for the generation of regional inequalities in parallel with the advance of technological change and market integration, which are the drivers of economic development in societies. However, making the connection between the generation of regional inequalities in the early stages of the economic growth processes and their persistence over time requires us to establish links between NEG literature and economic growth theory. Endogenous growth theory indicates that differences in the rates of human capital accumulation could be at the root of the different growth paths followed by countries or regions.

In order to advance in studying the hypotheses deriving from the connection between these two strands of the literature, in this paper we use the case of Spain during the early stages of the economic development process to analyse whether or not a relationship exists between regional variability in market access and education levels. This, in turn, would act as an explanatory element for the consolidation of diverse growth paths between regions. The results obtained allow us to confirm not only the hypothesis put forward by Redding and Schott (2003) but also the presence of a mechanism that – along the lines proposed by Fallah et al. (2010) – relates the provinces' market size to their production specialization and human capital accumulation, this mechanism being regional differences in the returns on investment in education.

Thus our findings show that in the case of Spain between 1860 and 1930, those provinces with the greatest market potential would have recorded greater specialization in those production sectors that call for more intensive use of human capital. In a context of reduced worker mobility between regions, they would also have offered greater remuneration to qualified workers, thus enabling a better return on investment in education. Given these conditions, the presence of increasing returns on economic activities requiring intensive use of human capital in their production processes would make it possible to connect the unequal distribution of activity in territories (as suggested by the NEG) to the appearance of significant gaps in capital accumulation rates and in the growth trajectories of the Spanish regions (as suggested by endogenous growth theory).

References

Barro, R.J. 2013. 'Education and economic growth', *Annals of Economics and Finance* 14(2), 301-328.

Barro, R.J. and Lee, J.W. 2013. 'A new data set of educational attainment in the world, 1950-2010', *Journal of Development Economics* 104, 184-18.

Beltrán Tapia, F.J, Díez-Minguela, A., Martinez-Galarraga, J. and Tirado, D.A. 2021. 'Two stories, one fate: age-heaping and literacy in Spain, 1877-1930', *Revista de Historia Económica/Journal of Iberian and Latin American Economic History*, online first.

Betrán, C. 2011. 'Regional specialization and industry location in the long term: Spain in the US mirror (1856–2002)', *Cliometrica* 5(3), 259-290.

Blum, M., Colvin, C.L., McAtackney, L. and McLaughlin, E. 2017. 'Women of an uncertain age: quantifying human capital accumulation in rural Ireland in the nineteenth century', *The Economic History Review* 70(1), 187-223.

Brakman, S., Garretsen, H. and Schramm, M. 2004. 'The spatial distribution of wages: estimating the Helpman-Hanson model for Germany', *Journal of Regional Science* 44(3), 437-466.

Breinlich, H. 2006. 'The spatial income structure in the European Union: what role for economic geography?', *Journal of Economic Geography* 6, 593-617.

Bringas, M.A. 2000. *La Productividad de los Factores en la Agricultura Española (1752-1935)* (Madrid: Banco de España).

Brülhart, M. and Sbergami, F. 2009. 'Agglomeration and growth: cross-country evidence', *Journal of Urban Economics* 65(1), 48-63.

Can Karahasan, B. and López-Bazo, E. 2013. 'The spatial distribution of human capital: can it really be explained by regional differences in market access?', *International Regional Science Review* 36(4), 451-480.

Ciccone, A. 2002. 'Agglomeration effects in Europe', *European Economic Review* 46(2), 213-227.

Ciccone, A. and Hall, R.E. 1996. 'Productivity and the density of economic activity', *American Economic Review* 86(1), 54-70.

Clark, C., Wilson, F. and Bradley, J. 1969. 'Industrial location and economic potential in Western Europe', *Regional Studies* 3(2), 197-212.

Combes, P.P. and Gobillon, L. 2015. 'The empirics of agglomeration economies'. In G. Duranton, J.V. Henderson and W. Strange (eds.), *Handbook of Urban and Regional Economics*, volume 5 (North-Holland: Elsevier), 247-348.

Combes, P.P., Mayer, T. and Thisse, J.F. 2008. *Economic Geography. The Integration of Regions and Nations* (Princeton: Princeton University Press).

Crafts, N. 2005. 'Market potential in British regions, 1871-1931', *Regional Studies* 39(9), 1159-1166

Crafts, N. and Mulatu, A. 2005. 'What explains the location of industry in Britain, 1871-1931?', *Journal of Economic Geography* 5(4), 499-518.

Crafts, N. and Mulatu, A. 2006. 'How did the location of industry respond to falling transport costs in Britain before World War I?', *Journal of Economic History* 66(3), 575-607.

Crafts, N. and Venables, A.J. 2003. 'Globalization in history: a geographical perspective'. In M.D. Bordo, A.M. Taylor and J.G. Williamson (eds.), *Globalization in Historical Perspective*, chapter 7 (Chicago: The University of Chicago Press).

Crozet, M. 2004. 'Do migrants follow market potentials? An estimation of a new economic geography model', *Journal of Economic Geography* 4(4), 439-458.

Crozet, M., Mayer, T. and Mucchielli, J.L. 2004. 'How do firms agglomerate? A study of FDI in France', *Regional Science and Urban Economics* 34(1), 27-54.

Davis, D.R. and Weinstein, D.E. 1999. 'Economic geography and regional production structure: an empirical investigation', *European Economic Review* 43(2), 379-407.

Davis, D.R. and Weinstein, D.E. 2003. 'Market access, economic geography and comparative advantage: an empirical test', *Journal of International Economics* 59(1), 1-23.

Diebolt, C. and Hippe, R. 2018. 'Remoteness equals backwardness? Human capital and market access in the European regions: insights from the long run', *Education Economics* 26(3), 285-304.

Díez-Minguela, A., Martinez-Galarraga, J., Tirado, D.A. 2016. 'Why did Spanish regions not converge before the Civil War? Agglomeration economies and (regional) growth revisited', *Revista de Historia Económica/Journal of Iberian and Latin American Economic History* 34(3), 417-448.

Díez-Minguela, A., Martinez-Galarraga, J., Tirado, D.A. 2018. *Regional Inequality in Spain, 1860-2015* (Cham: Palgrave-Macmillan).

Estevadeordal, A., Frantz, B. and Taylor, A.M. 2003. 'The rise and fall of world trade, 1870-1939', *Quarterly Journal of Economics* 118, 359-407.

Faíña, J. and López-Rodríguez, J. 2006. 'Market access and human capital accumulation: the European Union case', *Applied Economics Letters* 13(9), 563-567.

Fallah, B.N, Partridge, M.D. and Olfert, M.R. 2011. 'New economic geography and US metropolitan wage inequality', *Journal of Economic Geography* 11, 865-895.

Földvári, P., Van Leeuwen, B. and Van Leeuwen-Li, J. 2012. 'How did women count? A note on gender-specific age heaping differences in the sixteenth to nineteenth centuries', *The Economic History Review* 65(1), 304-313.

Fujita, M., Krugman, P. and Venables, A.J. 1999. *The Spatial Economy: Cities, Regions, and International Trade* (Cambridge: The MIT press).

Galor, O. 2011. 'Inequality, human capital formation and the process of development'. In E.A. Hanushek, S. Machin and L. Woessmann (eds.), *Handbook of the Economics of Education*, vol. 4 (Amsterdam: Elsevier), 441-493.

Gennaioli, N., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. 2013. 'Human capital and regional development', *The Quarterly Journal of Economics* 128(1), 105-164.

Goldin, C. 2015. 'Human capital'. In C. Diebolt and M. Haupert (eds.), *Handbook of Cliometrics*, chapter 3 (Berlin and Heidelberg: Springer), 55-86.

Hanson, G.H. 1998. 'Market potential, increasing returns, and geographic concentration', *NBER Working Papers*, 6429.

Hanson, G.H. 2005. 'Market potential, increasing returns, and geographic concentration', *Journal of International Economics* 67(1), 1-24.

Hanushek, E. and Woessmann, L. 2012. 'Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation', *Journal of Economic Growth* 17, 267-321.

Harris, C.D. 1954. 'The market as a factor in the localization of industry in theUnited States', *Annals of the Association of American Geographers* 64, 315-48.

Head, K. and Mayer, T. 2004. 'The empirics of agglomeration and trade'. In J.V. Henderson and J.F. Thisse (eds.), *Handbook of Regional and Urban Economics 4: Cities and Geography* (North Holland: Elsevier), 2609-2669.

Head, K. and Mayer, T. 2006. 'Regional wage and employment responses to market potential in the EU', *Regional Science and Urban Economics* 36(5), 573-594.

Head, K. and Mayer, T. 2011. 'Gravity, market potential and economic development', *Journal of Economic Geography* 11, 281-294.

Karahasan, B.C. and López-Bazo, E. 2013. 'The spatial distribution of human capital: can it really be explained by regional differences in market access?', *International Regional Science Review* 36(4), 451-480.

Keeble D., Owens P.L. and Thompson D.C. 1982. 'Regional accessibility and economic potential in the European Community', *Regional Studies* 16, 419-432.

Kim, S. 1995. 'Expansion of markets and the geographic distribution of economic activities: the trends in U.S. regional manufacturing structure, 1860-1987', *Quarterly Journal of Economics* 110(4), 881-908.

Klein, A. and Crafts, N. 2012. 'Making sense of the manufacturing belt: determinants of U.S. industrial location, 1880-1920', *Journal of Economic Geography* 12(4), 775-807.

Krugman, P. 1991. 'Increasing returns and economic geography', *Journal of Political Economy* 99(3), 483-499.

Krugman, P. and Venables, A.J. 1995. 'Globalization and the inequality of nations', *Quarterly Journal of Economics* 110(4), 857-880.

Liu, D. and Meissner, C.M. 2015. 'Market potential and the rise of US productivity leadership', *Journal of International Economics* 96, 72-87.

López-Rodríguez, J., Faíña, J.A. and López-Rodríguez, J. 2007. 'Human capital accumulation and geography: empirical evidence from the European Union', *Regional Studies* 41(2), 217-234.

Lucas, R. 1988. 'On the mechanics of economic development', *Journal of Monetary Economics* 22(1), 3-42.

Martinez-Galarraga, J. 2012. 'The determinants of industrial location in Spain, 1856-1929', *Explorations in Economic History* 49(2), 255-275.

Martinez-Galarraga, J. 2014. 'Market potential estimates in history: a survey of methods and an application to Spain, 1867-1930', *European Historical Economics Society (EHES) Working Papers in Economic History*, 51.

Martinez-Galarraga, J., Tirado, D.A. and González-Val, R. 2015. 'Market potential and regional economic growth in Spain (1860-1930)', *European Review of Economic History* 19(4), 335-358.

Martinez-Galarraga, J., Paluzie, E., Pons, J., Silvestre, J. and Tirado, D.A. 2021. 'New economic geography and economic history: a survey of recent contributions through the lens of the Spanish industrialization process', *Cliometrica* 15(3), 719-751.

Martinez-Galarraga, J., Paluzie, E., Pons, J. and Tirado, D.A. 2008. 'Agglomeration and labour productivity in Spain over the long term', *Cliometrica* 2(3), 195-212.

Matas, A., Raymond, J.L. and Roig, J.L. 2015. 'How market access shapes human capital accumulation in a peripheral country: the case of Spain', *Applied Economics* 47(11), 1118-1132.

Midelfart-Knarvik, K.H., Overman, H., Redding, S. and Venables, A.J. 2002. 'The location of European industry'. In European Commission (ed.), *European Economy. European*

Integration and the Functioning of Product Markets, 2/2002 (Brussels: Directorate-General for economic and financial affairs), 213-269.

Ministerio de Trabajo, Comercio e Industria. 1927. *Estadística de Salarios y Jornadas de Trabajo Referida al Período 1914-1925* (Madrid).

Ministerio de Trabajo y Previsión. 1931. *Estadística de Salarios y Jornadas de Trabajo Referida al Período 1914-1930* (Madrid).

Nadal, J. 1987. 'La industria fabril española en 1900: una aproximación'. In C. Sudrià, J. Nadal and A. Carreras (eds.), *La Economía Española en el Siglo XX: Una Perspectiva Histórica* (Barcelona: Ariel), 23-61.

Niebuhr, A. 2006. 'Market access and regional disparities', *Annals of Regional Science* 40(2), 313-334.

Núñez, C.E. 2003. 'Schooling, literacy and modernization: a historian's approach', *Paedagogica Historica* 39(5), 535-558.

Ottaviano, G. and Pinelli, D. 2006. 'Market potential and productivity: evidence from Finnish regions', *Regional Science and Urban Economics* 36(5), 636-657.

Paluzie, E., Pons, J. and Tirado, D.A. 2004. 'The geographical concentration of industry across Spanish regions, 1856-1995', *Jahrbuch Für Regionalwissenschaft* 24(2), 143-160.

Pérez Moreda, V. 1997. 'El proceso de alfabetización y la formación de capital humano en España', *Papeles de Economía Española* 73, 243-253.

Pérez Moreda, V., Reher, D. S. and Sanz Gimeno, A. 2015. *La Conquista de la Salud. Mortalidad y Modernización en la España Contemporánea* (Madrid: Marcial Pons).

Pons, J., Paluzie, E., Silvestre, J. and Tirado, D.A. 2007. 'Testing the new economic geography: migrations and industrial agglomerations in Spain', *Journal of Regional Science* 47(2), 289-313.

Prados de la Escosura, L. 2017. *Spanish Economic Growth, 1850-2015* (Cham: Palgrave-Macmillan).

Puga, D. 1999. 'The rise and fall of regional inequalities', *European Economic Review* 43(2), 303-334.

Redding, S. and Schott, P.K. 2003. 'Distance, skill deepening and development: will peripheral countries ever get rich?', *Journal of Development Economics* 72, 515-41.

Redding, S. and Venables, A.J. 2004. 'Economic geography and international inequality', *Journal of International Economics* 62(1), 53-82.

Reher, D.S. 1997. 'La teoría del capital humano y las realidades de la historia', *Papeles de Economía Española* 73, 254-261.

Romer, P.M. 1986. 'Increasing returns and long-run growth', *Journal of Political Economy* 94 (5), 1002-1037.

Roos, M. 2001. 'Wages and market potential in Germany', *Jahrbuch Für Regionalwissenschaft* 21(2), 171-195.

Rosés, J.R. 2003. 'Why isn't the whole of Spain industrialized? New economic geography and early industrialization, 1797-1910', *Journal of Economic History* 63(4), 995-1022.

Rosés, J.R., Martinez-Galarraga, J. and Tirado, D.A. 2010. 'The upswing of regional income inequality in Spain (1860-1930)', *Explorations in Economic History* 47(2), 244-257.

Sáiz, P. 2005. 'Investigación y desarrollo: patentes'. In A. Carreras and X. Tafunell (dirs.), *Estadísticas Históricas de España, Siglos XIX-XX* (Madrid, FBBVA).

Tirado, D.A., Paluzie, E. and Pons, J. 2002. 'Economic integration and industrial location: the case of Spain before World War I', *Journal of Economic Geography* 2(3), 343-363.

Tirado, D.A., Pons, J., Paluzie, E. and Martinez-Galarraga, J. 2013. 'Trade policy and wage gradients: evidence from a protectionist turn', *Cliometrica* 7(3), 295-318.

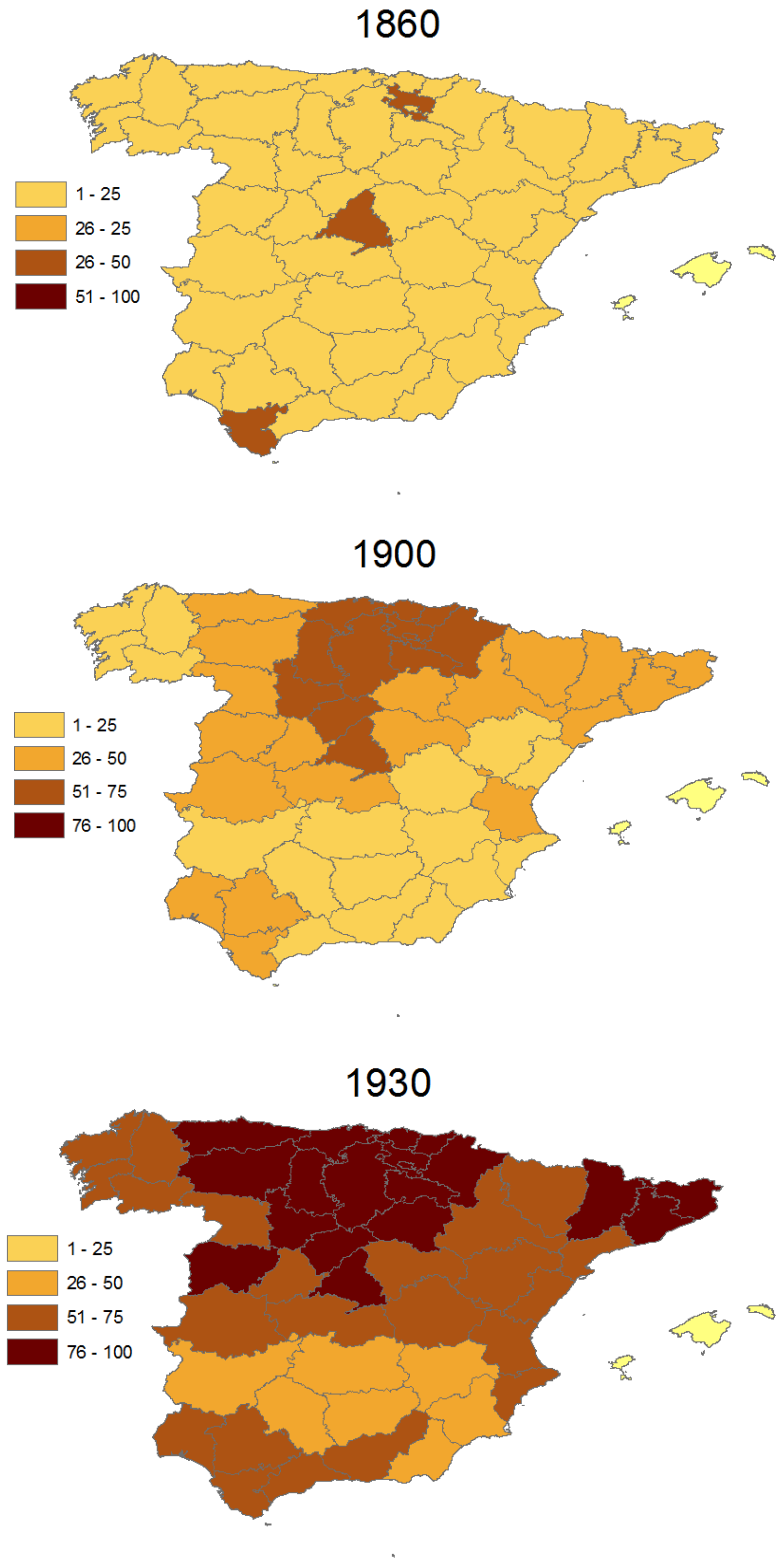
Tollnek, F. and Baten, J. 2016. 'Age-heaping-based human capital estimates'. In C. Diebolt and M. Haupert (eds.), *Handbook of Cliometrics* (London: Springer), 131-154.

Viñao, A. 1990. 'The history of literacy in Spain: evolution, traits, and questions', *History of Education Quarterly* 30(4), 573-599.

Wolf, N. 2007. 'Endowment vs. market potential: what explains the relocation of industry after Polish reunification in 1918?', *Explorations in Economic History* 44(1), 22-42.

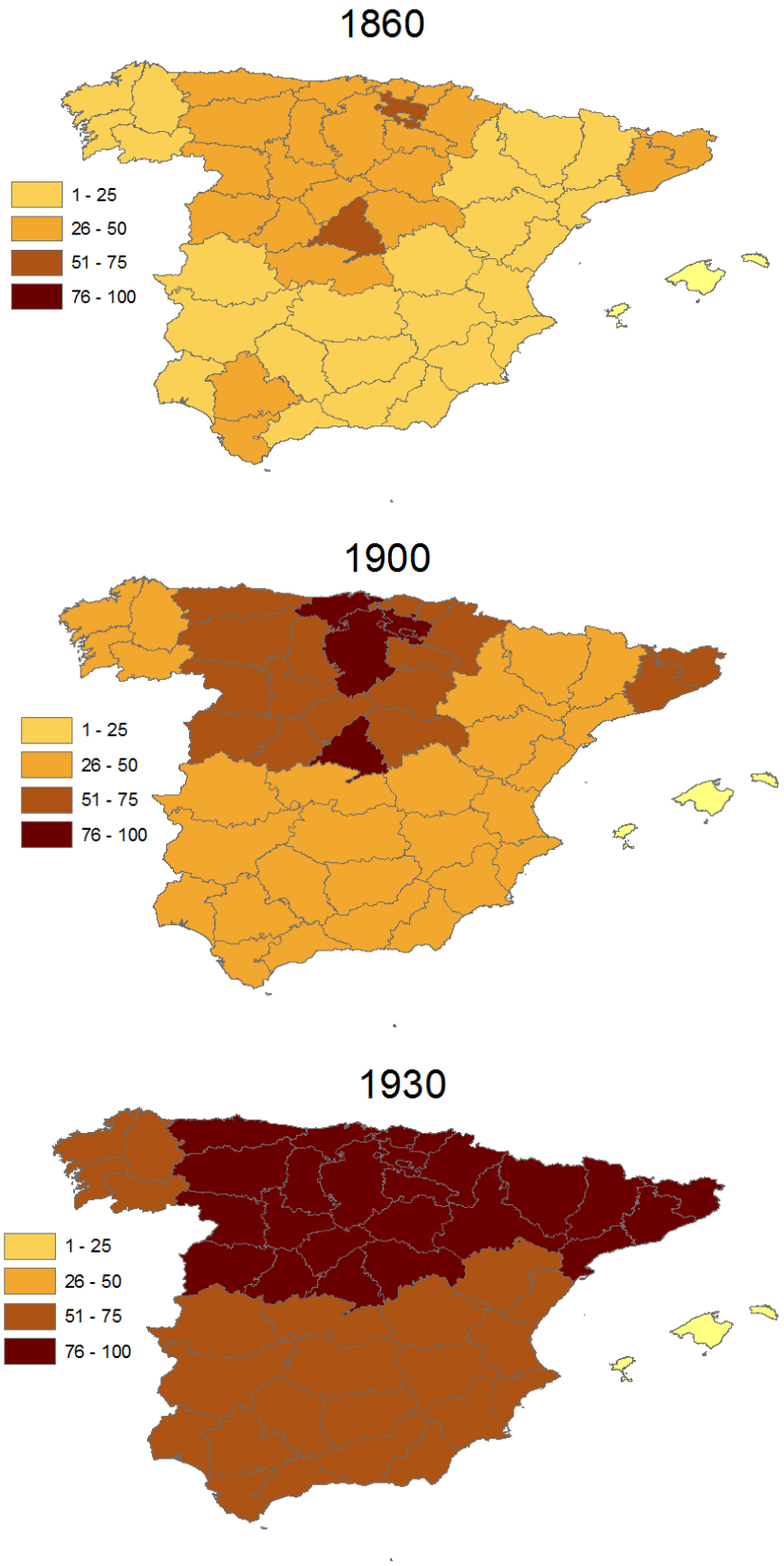
Appendix

Figure A.1. Female literacy rates, 1860-1930



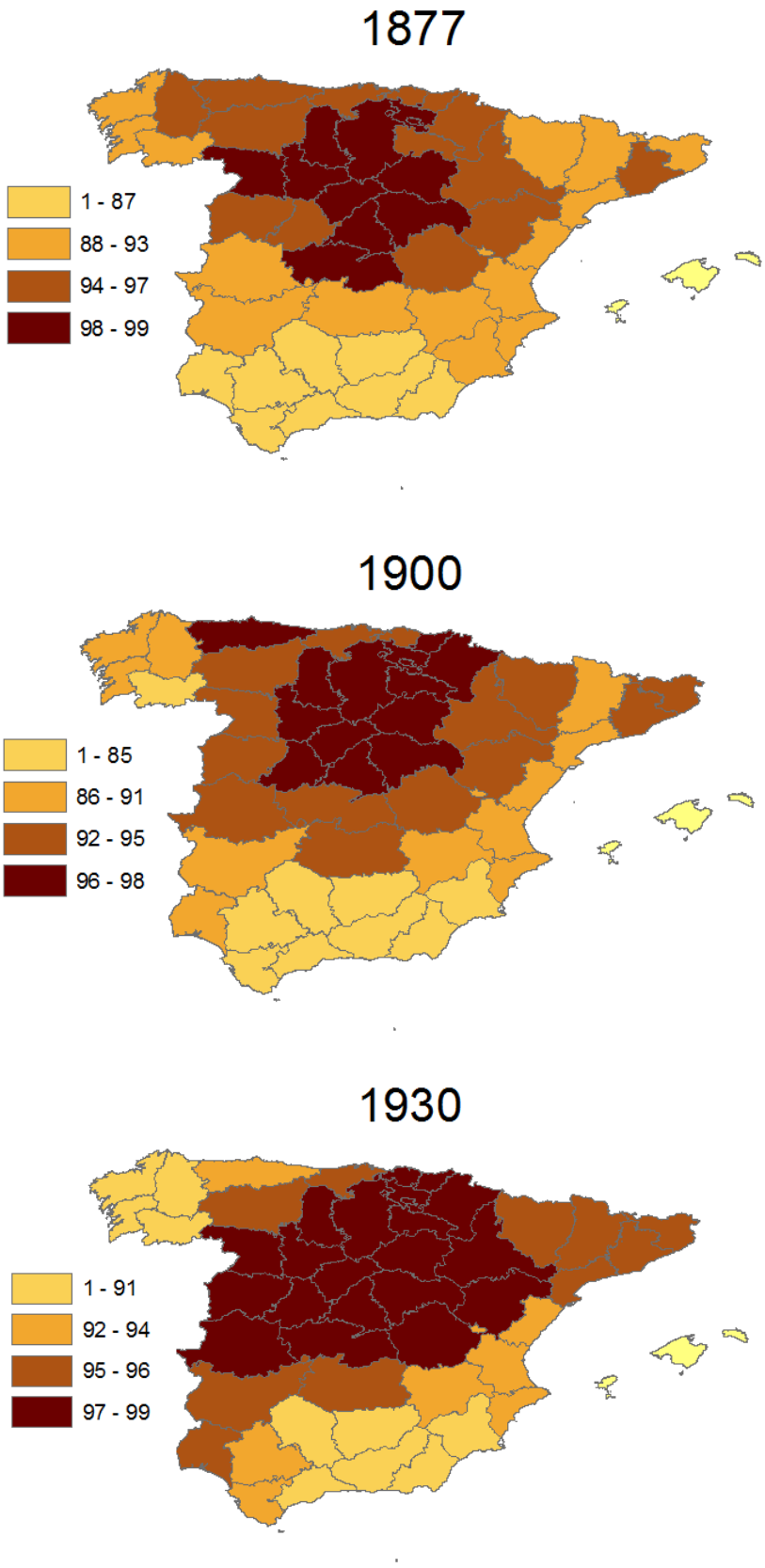
Source: Population censuses.

Figure A.2. Total literacy rates, 1860-1930



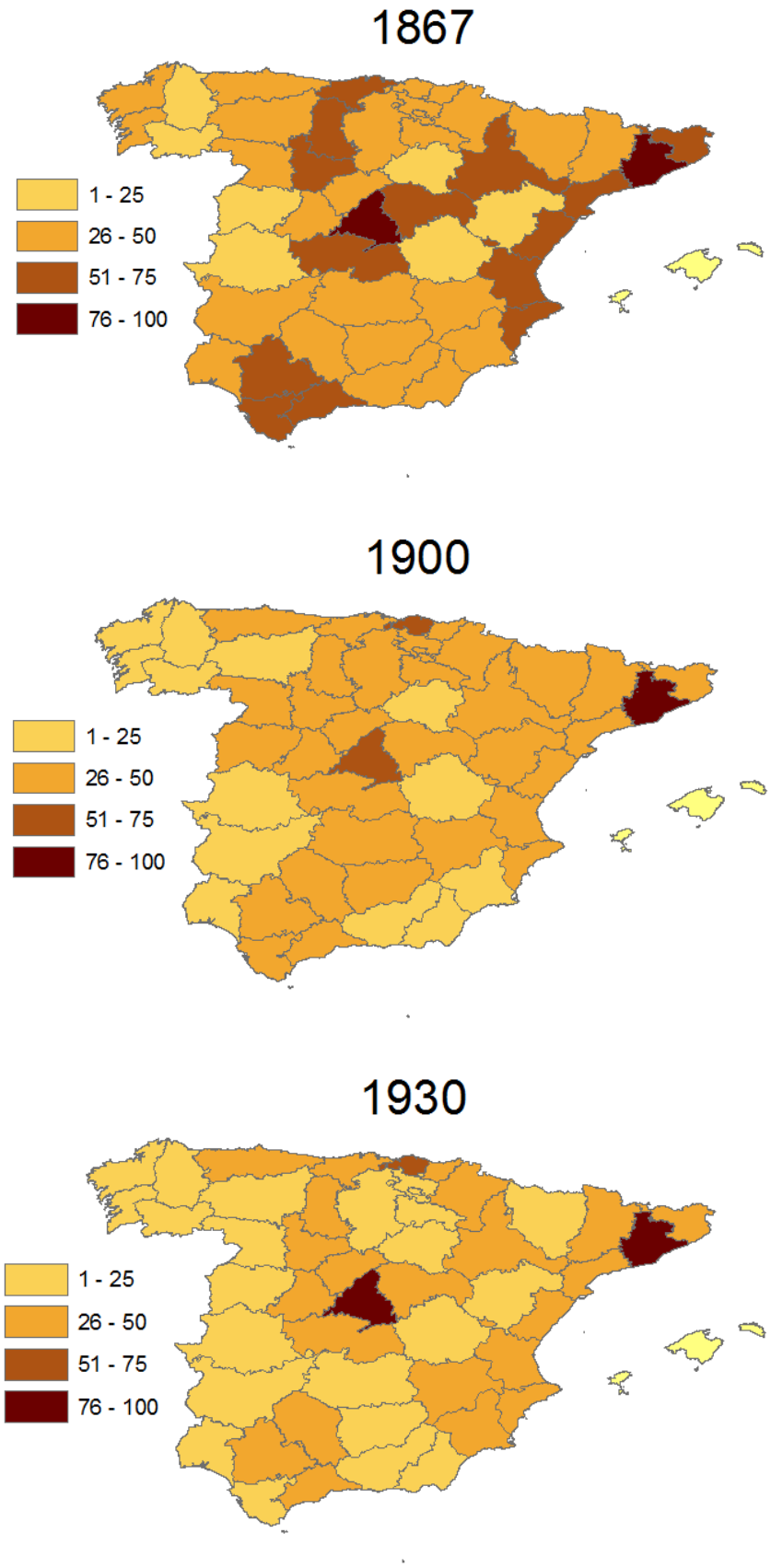
Source: Population censuses.

Figure A.3. Numeracy levels (ABCC index), 1877-1930



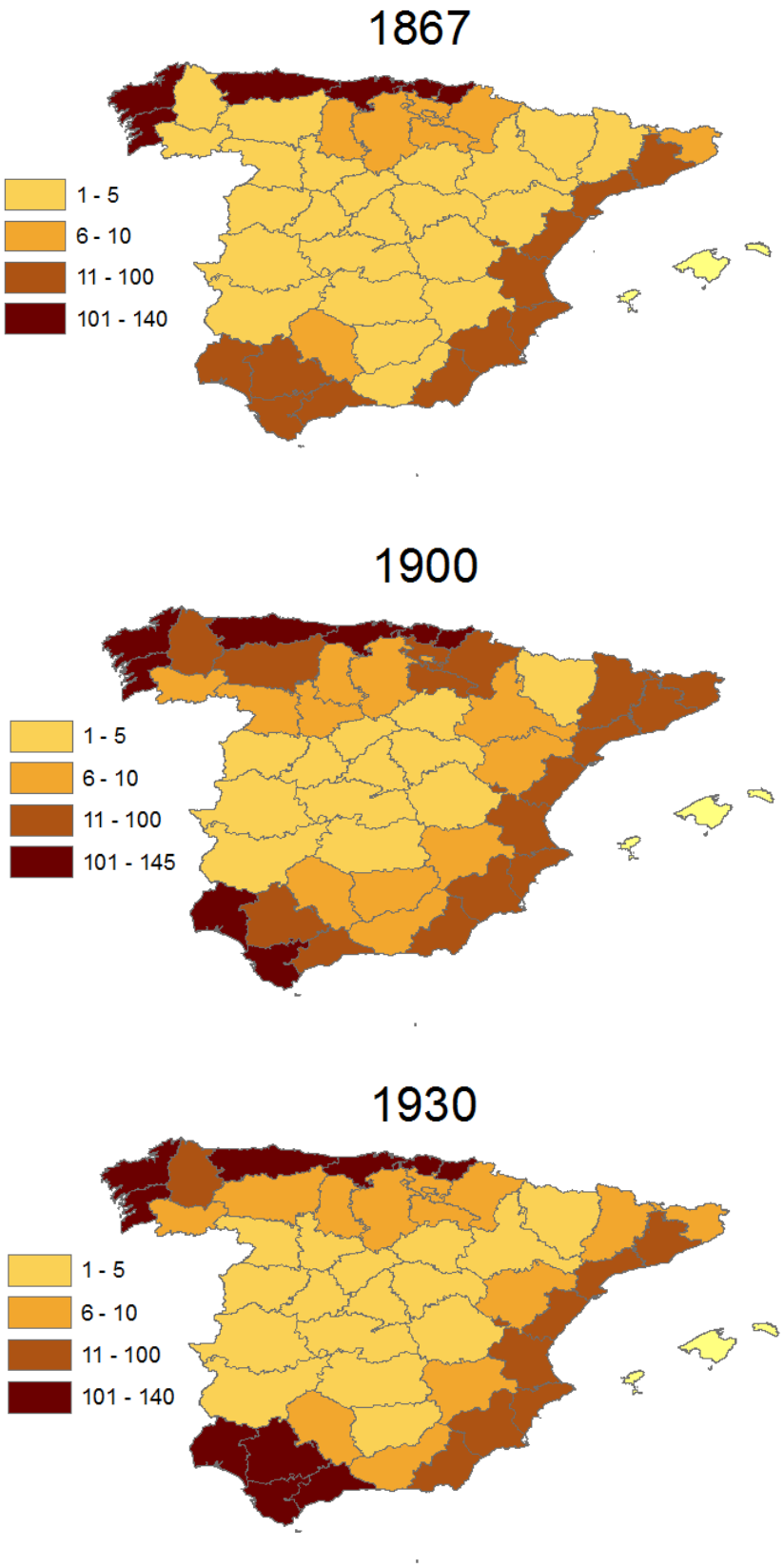
Source: Beltrán Tapia et al. (2021).

Figure A.4. Internal market potential, 1867-1930 (Barcelona=100)



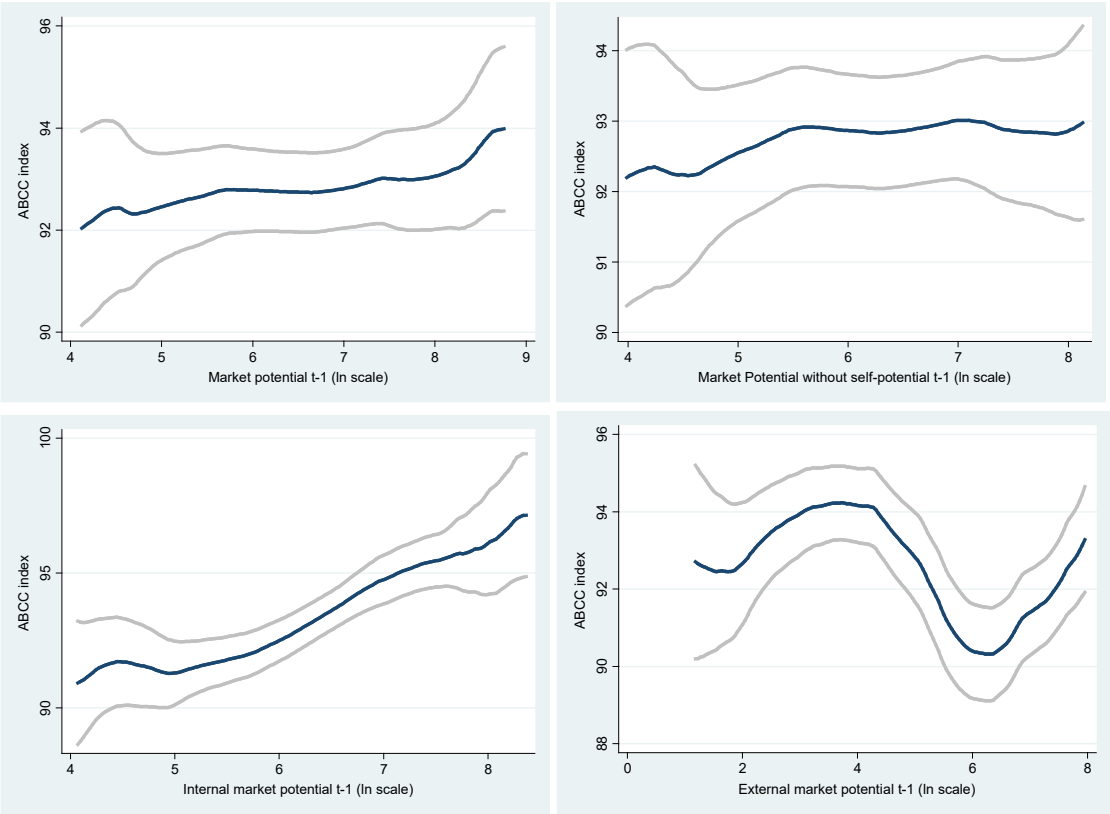
Source: Martinez-Galarraga (2014).

Figure A.5. External market potential, 1867-1930 (Barcelona=100)



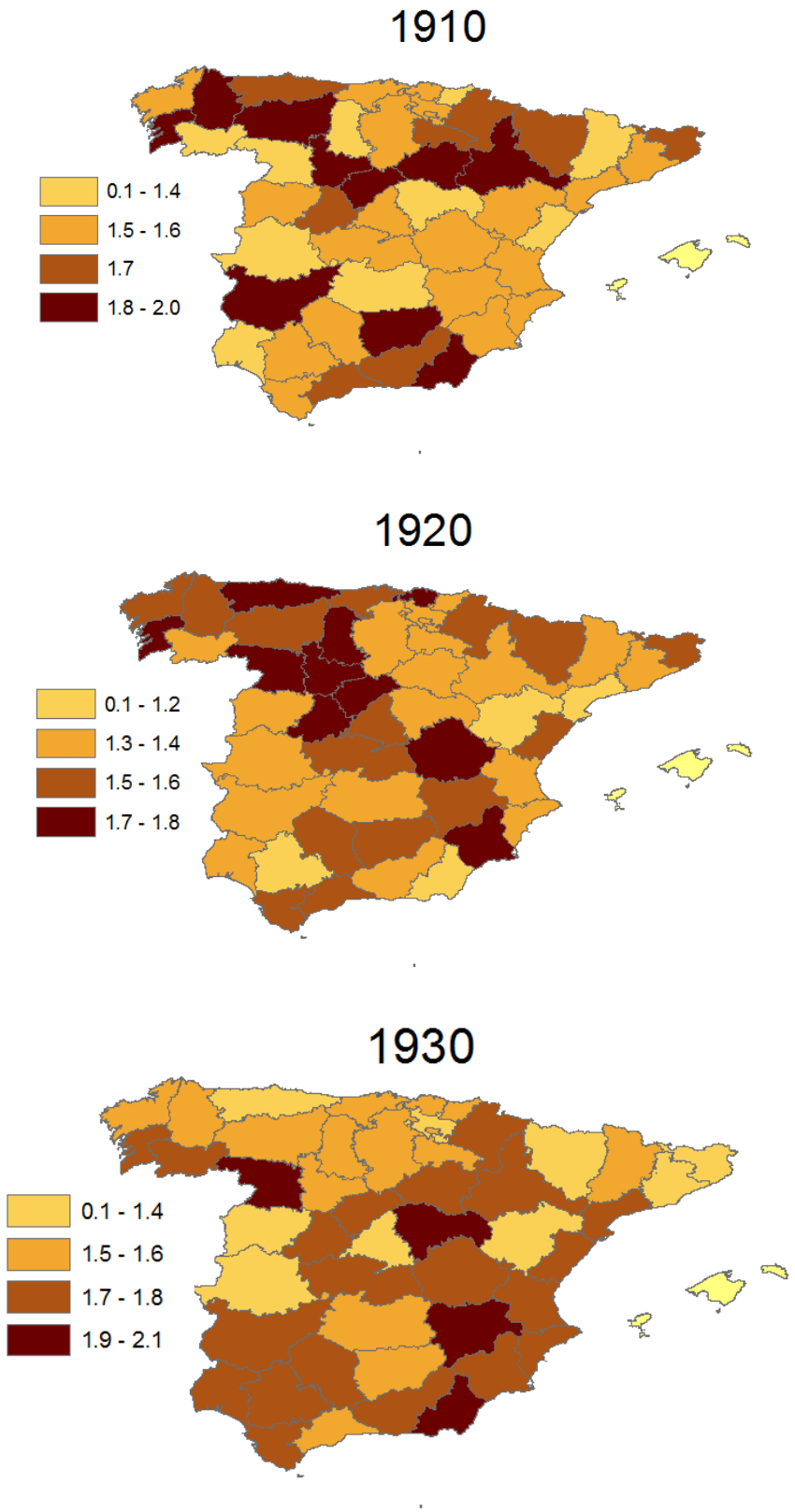
Source: Martinez-Galarraga (2014).

Figure A.6. Non-parametric relationship between the ABCC index and different measures of lagged market potential (1860-1930)



Source: see text.

Figure A.7. Skill premium: skilled industrial workers v. unskilled industrial workers



Source: Ministerio de Trabajo (1927, 1931).

European
Historical
Economics
Society

EHES Working Paper Series

Recent EHES Working Papers

2022

EHES 228 Union, border effects, and market integration in Britain
Daniel Cassidy, Nick Hanley

EHES 227 Welfare Reform and Repression in an Autocracy: Bismarck and the Socialists
Felix Kersting

EHES 226 Harmonious Relations: Quality transmission among composers in the very long run
Karol Jan Borowiecki, Nicholas Ford, Maria Marchenko

EHES 225 Urban Political Structure and Inequality: Political Economy Lessons from Early Modern German Cities
Felix Schaff

EHES 224 Health, Income, and the Preston Curve: A Long View
Leandro Prados de la Escosura

EHES 223 Boomtowns: Local Shocks and Inequality in 1920s California
Sarah Quincy, Rowena Gray

EHES 222 Poverty in early modern Europe: New approaches to old problems
Guido Alfani, Francesco Ammannati, Wouter Ryckbosch

2021

EHES 221 Human Capital and Industrialization: German Settlers in Late Imperial Russia
Viktor Malein

EHES 220 Resilience, Adaptability and Transformability: Danish Butter Factories in the Face of Coal Shortages
Sofia Teives Henriques, Paul Sharp, Xanthi Tsoukli, Christian Vedel

All papers may be downloaded free of charge from: http://www.ehes.org/

The European Historical Economics Society is concerned with advancing education in European economic history through study of European economies and economic history. The society is registered with the Charity Commissioners of England and Wales number: 1052680