Institut d'Economia de Barcelona

Document de treball de l'IEB 2013/13

MARKET POTENTIAL AND CITY GROWTH: SPAIN 1860-1960

Rafael González-Val, Daniel A. Tirado-Fabregat, Elisabet Viladecans-Marsal

Cities and Innovation

MARKET POTENTIAL AND CITY GROWTH: SPAIN 1860-1960

Rafael González-Val, Daniel A. Tirado-Fabregat, Elisabet Viladecans-Marsal

The **IEB** research program in **Cities and Innovation** aims at promoting research in the Economics of Cities and Regions. The main objective of this program is to contribute to a better understanding of agglomeration economies and 'knowledge spillovers'. The effects of agglomeration economies and 'knowledge spillovers' on the Location of economic Activities, Innovation, the Labor Market and the Role of Universities in the transfer of Knowledge and Human Capital are particularly relevant to the program. The effects of Public Policy on the Economics of Cities are also considered to be of interest. This program puts special emphasis on applied research and on work that sheds light on policy-design issues. Research that is particularly policy-relevant from a Spanish perspective is given special consideration. Disseminating research findings to a broader audience is also an aim of the program. The program enjoys the support from the **IEB-Foundation.**

The **Barcelona Institute of Economics (IEB)** is a research centre at the University of Barcelona which specializes in the field of applied economics. Through the **IEB-Foundation**, several private institutions (Applus, Abertis, Ajuntament de Barcelona, Diputació de Barcelona, Gas Natural and La Caixa) support several research programs.

Postal Address:
Institut d'Economia de Barcelona
Facultat d'Economia i Empresa
Universitat de Barcelona
C/ Tinent Coronel Valenzuela, 1-11
(08034) Barcelona, Spain
Tel.: + 34 93 403 46 46
Fax: + 34 93 403 98 32
ieb@ub.edu
http://www.ieb.ub.edu

MARKET POTENTIAL AND CITY GROWTH: SPAIN 1860-1960 *

Rafael González-Val, Daniel A. Tirado-Fabregat, Elisabet Viladecans-Marsal

ABSTRACT: A few attempts have been made to analyse whether market potential might also have an impact on urban structures. In this paper we employ parametric and non-parametric techniques to analyse the effect of market potential on the growth of Spanish cities during the period 1860-1960. This period is especially interesting because it is characterized by the advance in the economic integration of the national market together with an intense process of industrialization. Our results show a clear positive influence of market potential on city growth.

JEL Codes: R0, N9, O18, N64, F14

Keywords: Market potential, urban structure, city growth, economic history.

Rafael González-Val
Universidad de Zaragoza & IEB
Fac.de Economía y Empresa
Gran Vía, 2
50005 Zaragoza, Spain
E-mail: rafaelg@unizar.es

Daniel A. Tirado-Fabregat
Universitat de València
Av. Blasco Ibáñez, 13
46010 Valencia, Spain
E-mail: Daniel.Tirado@uv.es

Elisabet Viladecans-Marsal
Universitat de Barcelona & IEB
Fac.de Economia y Empresa
Avda. Diagonal 690
08034 Barcelona, Spain
E-mail: eviladecans@ub.edu

* This research has received financial support from projects ECO2010-16934, ECO2012-39169-C03-02 and ECO2009-09332 (Ministerio de Educación y Ciencia), 2009SGR102 (Generalitat de Catalunya), DGA (ADETRE research group) and FEDER. We acknowledge helpful comments from Jeffrey Lin and Gabriel Ahlfeldt. We are also grateful to participants at the 7th Meeting of the Urban Economics Association at the 59th Annual North American Meetings of the Regional Science Association International (Ottawa, 2012) and at the XXXVII Symposium of Economic Analysis (Vigo, 2012). We thank F. J. Goerlich and Julio Martinez-Galarraga for the provision of some unpublished data.

1. Introduction

There is a growing body of empirical studies that analyze the incidence of market potential on the geographical distribution of population and economic activity, migratory patterns, wage levels and differences in regional income[1]. But, to date, few studies have examined the relation between market potential and patterns of city size growth. However, recently a number of papers have introduced market size effects in their explanations of the geographical distribution of cities and of their relative sizes. Indeed, in a recent survey, Redding (2010) points out that it might be interesting to reconcile new economic geography (NEG) models with the findings of the urban economics literature regarding the distribution of population size and population growth patterns. As Krugman (1991) claims, there are two types of factors that can be considered determinants of city growth: *first nature* factors, which are related principally with geography (climate, costal location, access to natural resources, etc.) and which influence city growth in their early stages; and, *second nature* factors, which are related to agglomeration economies and increasing returns of scale.

Here, a part of the literature has considered the change in market potential a good proxy for agglomeration economies. However, the direction of the influence of market potential on city growth is unclear. Trade theory literature (Davis and Weinstein, 2002 and Hanson 2005) concludes that greater market potential should foster growth, the rationale being that nearby cities offer a larger market and, hence, more possibilities of selling products. By contrast, location theory (Fujita et al. 1999) and hierarchy models (Dobkins and Ioannides, 2001) suggest that increasing market potential could affect city growth negatively, the rationale being that the forces of spatial competition separate the larger cities from each other, so the bigger a city grows the smaller its neighboring cities will be. Finally, it is interesting to note that the effects of market potential on city size may differ depending on the initial size of the city.

Although there is a sizeable body of theoretical research developing models around these factors, the empirical evidence remains limited. In recent years, various papers have specifically analyzed the incidence of market potential on city growth. For US cities, Black and Henderson (2003) analyze the determinants of population growth from a long-term perspective (1900-1990), finding clear influence of market potential on city growth.

[1] See Redding (2010) for a survey on this literature.

However, the effect becomes negative when they introduce the squared market potential variable. Henderson and Wang (2007) analyze the influence of market potential on population growth for metropolitan areas with a population over 100,000 inhabitants across 142 different countries between 1960 and 2000. They also report a positive effect and a value that is not negligible: a 1% increase in a city's market potential increases city growth by 0.9% over a decade. Au and Henderson (2006) adopt a similar approach in their analysis of 225 Chinese cities during the nineties but choose to measure city growth in value-added terms. Their results are much more modest indicating that a 1% increase in market potential leads to a 0.16% increase in value added for the city. Finally, da Mata et al. (2007) analyze the determinants of city growth for Brazilian cities from 1970 to 2000 and report smaller values than the two studies discussed above.

This paper is conducted in line with the preceding studies and specifically seeks to analyze the effects of the economic integration and industrialization of the Spanish economy on the evolution of urban system during the period 1860-1960. The hypothesis we test is the following: the geographical distribution and the relative size of the Spanish cities were historically determined by the location fundamentals of each territory. However, when the Spanish market began to be integrated, there was an increase in the concentration of population in a small number of cities. This concentration could explain the increase in the inequalities in the relative city sizes. In other words, in a context in which manufacturing activities were acquiring greater weight in the economy, the construction of the transport network and the integration of the markets would have favored, from the middle of the 19th century, the agglomeration of economic activity and this could have been the basis for the changes in the long-run city patterns. Basically, the urban systems that prevailed before and after market integration and industrialization were quite distinct. Our results lend support to these hypotheses and confirm that the market potential had a clear influence on these processes.

The main contributions of this paper are the following. First, we exploit the long-term historical episode of growth and economic integration that took place in Spain from the middle of the 19th century until the 1960s y. Thus, we are able to study the determinants of city growth throughout the whole industrialization process of the Spanish economy and at a time when the Spanish urban system was undergoing an intense transformation characterized by the concentration of the population in a small number of cities and with a clear impact on their relative sizes. Second, following Black and Henderson (2003), Ioannides and Overman (2004) and Bosker et al. (2008), we test the importance of market potential on city growth. In this respect, we do not use the distance-weighted sum of population of all other existing cities as a proxy of city market potential. We present an empirical measure of city market access that considers a new set of historical estimates of regional GDPs for Spanish regions and the historical transport costs among regions, as well as the changes they underwent during the process of economic integration of the Spanish economy

The remainder of the paper is organised as follows. In section 2 we describe the economic integration process that took place in Spain from 1860 to 1960 and we review the main evidence from the literature regarding its economic effects. In section 3 we analyze the evolution of the Spanish city size distribution from 1860 to 1960. In section 4 we present our data. In section 5 we describe the empirical analysis and we present and discuss the results obtained. Section 6 concludes.

2. Market integration and economic agglomeration in Spain, 1860-1960

From a long-term perspective, Spain's internal market integration received a major push in the middle of the 19th century. Prior to this date, the Spanish regions had relatively independent economies. Barriers to interregional trade and the movement of capital and labour were ubiquitous: local tariffs and regulations on domestic commerce were widespread; weights and measures differed across regions; transport costs were very high due to low public investment in transport infrastructures and the particular geography of Spain, which lacked an extensive water transport system; economic information moved slowly across regions; the banking system was underdeveloped; and many regions had their own currencies (although they were all based on a bimetallic monetary system). As a consequence, regional commodity markets were scarcely integrated and the prices of production factors differed markedly from one region to another. [2]

Market liberalization and improvements in transport systems, particularly the completion of Spain's railway network, prompted the creation of a national market for most major commodities during the second half of the 19th century. According to calculations reported

[2] See, for example, Ringrose (1996) for further details.

by Herranz (2006), the introduction of the railway in 1878 meant a massive 86 per cent reduction in transport costs. In addition to these two factors, the successive political reforms of the 19th century upheld property rights, eliminated tariffs and local restrictions on home commerce and safeguarded the mobility of people and capital. These measures were implemented in three main waves: during the Liberal Revolution (1836-1840), the two-year Progressive period (1854-1856) and the six-year Democratic period (1868-1874).
The Spanish Civil War and the first years of Franco's regime acted as a brake on Spanish growth and its national economic integration. The regulation of markets for goods and factors of production combined with government control of the prices and quantities of final and intermediate goods, energy, capital markets and wages reduced the mobility of factors and resources. The movement of capital across regions slowed and labour migration came to a halt after an initial period of growth in the 1920s (Silvestre, 2005). Likewise, the absence of investment in infrastructure did little to reduce transport costs during the 1940s and early 1950s. The economic liberalisation and stabilisation measures introduced during the decade of the fifties however favoured the transition of the Spanish economy toward a new phase of economic development that would last until the oil crisis..

Recent studies have attempted to analyse the effects of this process of long-term economic integration and growth on the distribution of industry across Spanish regions[3]. Rosés (2003) and Tirado et al. (2002) provide new empirical evidence confirming that, from the second half of the 19th century until the outbreak of the Spanish Civil War, there was a marked increase in the geographical concentration of industry. In addition, both studies stress that this long-term evolution was in line with predictions emanating from NEG models. This strand of the literature suggests that the reduction in transport costs in the presence of scale economies in industrial activities results in the geographical concentration of industry and that production agglomerates in locations that enjoy the best access to markets. In other words, new evidence regarding the evolution of the geographical concentration of industry in Spain in the period that extends from 1860 to 1960 points to the fact that the relative market access of Spanish regions could act as an important explanatory factor of industrial agglomeration geography.

[3] Most of the empirical contributions to the Spanish case discussed below adopt the standard empirical methodologies developed for the analysis of NEG models. Redding (2010) offers a recent survey of these empirical methods.

Besides, a recent set of empirical papers has explored this hypothesis in depth. On the one hand, Martínez-Galarraga (2012), adopting the empirical proposal in Midelfart-Knarvik et al. (2002), demonstrates that, during the first long phase of Spanish industrialization (1860-1930), industrial sectors characterized by the existence of scale economies tended to concentrate in regions with higher market access. This result is in line with findings reported in studies analyzing historical integration processes in other countries (see, for example, Wolf (2007) for Poland, and Crafts and Mulatu (2005) for the United Kingdom). On the other hand, Paluzie et al. (2009a), adopting the empirical framework developed in Hanson (2005), verify, in concordance with the wage equation derived in the Helpman-Krugman model, the existence of a spatial wage structure that relates industrial nominal wages with the relative market potential of regions at different points in time. Other analyses based on this methodology can be found in Brakman et al. (2004) for the case of Germany and Mion (2004) for that of Italy.

Spanish economic growth and market integration also favored the increasing concentration of population across regions. In fact, the Gini index for regional population (at a NUTS3 scale) grew steadily from 0.266 in 1860 to a value of 0.402 in 1960.[4] Several studies have also explored the economic factors underpinning this process. First, in line with the empirical proposals made in the NEG literature, Paluzie et al. (2009b) followed Crozet (2004) to demonstrate the existence of a direct relationship between the location decisions of migrants and the market potential of the host regions during the two main waves of internal migrations in Spain in the 1920s and the 1960s.[5] Second, Ayuda et al. (2010) analyzed the patterns of the geographical distribution of population in Spain from the 18th century onward. They report that in the pre-industrial period, when agriculture was the predominant activity, *first nature* advantages determined the distribution of population across Spanish provinces as climatic and topographic conditions had a direct impact on agrarian productivity. As a result, these natural conditions provided some locations with an initial advantage. However, the authors conclude that market integration in a context of industrialization strengthened this pattern. From 1900 onwards, *second nature* geography, linked to increasing returns and relative access to regional markets, reinforced the process of spatial agglomeration of population.[6]

[4] Ayuda et al. (2010).

[5] Kancs (2005) also makes use of this approach in analyzing the determinants of migratory flows in the European Union.

[6] In a similar vein, Goerlich and Mas (2009) also study the long-term determinants of the agglomeration of population in Spain.

Summing up, this empirical literature records that industrial production and population in Spain agglomerated parallel to the long-term process of development and market integration. Moreover, in line with the hypotheses derived from NEG literature, the differences in regional market access acted as a key factor in explaining the geography of this increasingly agglomerated economy. In line with these conclusions, the sections that follow are devoted, first, to presenting new evidence regarding the changes experienced in the Spanish urban system during this long-term process of economic development and market integration; and, second, to analyzing the role played by differences in the market access of Spanish cities as a factor that accounts for their relative growth.

3. Changes in the Spanish urban system: The evolution of the city size distribution

This section analyses the evolution of Spain's city size distribution from 1860 to 1960. Other studies have examined this distribution, above all during the twentieth century identifying a divergent pattern of growth in city sizes during the period 1900–1970 (see Lanaspa et al., 2003, for a good example of this). Here we seek to add to this literature by offering new empirical evidence dating back to 1860. We estimate Pareto exponents and empirical density functions. Our results show that from 1860 to the beginning of the twentieth century the city size distribution remained stable, but after that date a process of divergent growth that increased inequality within the distribution is identified.

Our geographical unit of reference is the municipality (local government areas), the smallest spatial subdivisions in Spain's administrative system, which cover the whole territory and include all the country's population. Our population data are drawn from the 1860 census and thereafter from the decennial censuses conducted since 1900. Reher (1994) provides population data for 1860, while for all the other decades we use data from the Spanish official statistics institute, the censuses of the *Instituto Nacional de Estadística* (INE - www.ine.es).

Herein Table 1

Table 1 shows the number of municipalities by period and their corresponding descriptive statistics. We impose a minimum population threshold of 5,000 inhabitants in each period since the smallest municipalities can hardly be considered urban (one of the particular features of the Spanish city system is the high number of small rural towns). Furthermore, until the middle of the twentieth century a considerable part of the country's employment was concentrated in the agriculture sector (38.7% in 1960 according to OECD data), so metropolitan structures only really began to emerge in the second half of the century. Figure 1, which plots two maps showing the spatial distribution of the municipalities in our samples in 1860 and 1930, shows that there was a sizable entry of new cities in the distribution by this later date. In 1860 most of the cities were located in Andalusia, the southernmost region of Spain, but several decades after, new cities had emerged in the centre and in the northwestern regions of Spain.

Herein Figure 1

A standard way to analyse the evolution of the city size distribution involves fitting a Pareto distribution to the data (Cheshire, 1999; Gabaix and Ioannides, 2004). Let S_i be the size (population) of city i and R_i its corresponding rank (1 for the largest, 2 for the second largest and so on). We define the relative size of the *i*th city, s_i, as the quotient between the city's population and the contemporary average,

$$s_i = \frac{S_i}{\overline{S}} = \frac{S_i}{\frac{1}{N}\sum_{i=1}^{N} S_i}, \tag{1}$$

where N is the sample size. A power law (Pareto distribution) links city size and rank as follows: $R_i(S_i) = AS_i^{-a}$, where A is a constant and a is the Pareto exponent. Zipf's law is an empirical regularity, appearing when the Pareto exponent of the distribution is equal to unity ($\hat{a} = 1$) and which means, when ordered from largest to smallest, the size of the second city is half that of the first, the size of the third is a third of the first, and so on. Moreover, the greater the coefficient, the more homogeneous the city sizes.

Herein Figure 2

Taking logs we obtain the logarithmic version usually estimated by OLS. We apply the specification proposed by Gabaix and Ibragimov (2011), subtracting $1/2$ from the rank to obtain an unbiased estimation of a:

$$\ln\left(R_i - \frac{1}{2}\right) = b - a\ln S_i + \varepsilon_i\,, \qquad (2)$$

where ε_i is the error term. We estimate Equation (1) by OLS for our sample of cities in the different periods from 1860 to 1960. Graph (a) in Figure 2 shows the results[7], which demonstrate that the distribution remained stable from 1860 to 1900, and the estimated coefficients are greater than one, indicating that city sizes were homogeneous. Since the beginning of the twentieth century the estimated values of the Pareto exponent tend to fall, indicating a process of divergent growth in Spanish cities (Lanaspa et al., 2003). However, the exponent is always greater than one, rejecting Zipf's law. Graph (b) shows the results considering a balanced panel of the 262 municipalities existing in 1860 with population above the minimum population threshold, not allowing the entry or the exit of cities in the sample. The pattern for these cities is similar to that of the whole sample: the Pareto exponent is stable until 1900, when it begins to decrease. The only difference is that, for this sample of cities, the estimated exponent at the end of the period is close to one.

We also estimated the Gini coefficients for each period, which have the advantage of not imposing a specific size distribution (Pareto for rank-size coefficients)[8]. The results are similar; throughout the whole period the evolution of the distribution indicates a divergent pattern as the coefficient rises from 0.45 in 1860 to 0.61 in 1960, with the coefficient growing particularly fast since 1930. Finally, Graph (c) in Figure 2 shows the empirical density functions for the four periods (estimated using adaptive kernels). It can be seen that the distribution remained almost static from 1860 to 1900. Since then, from a highly leptokurtic distribution with much of the density concentrated in the mean value of the distribution, the distribution has lost kurtosis and the concentration has decreased. This evolution is more pronounced for the sample of 262 largest municipalities in 1860 (Graph d), indicating that the initially largest cities were especially involved in the divergence process. Thus, once more, our results point to increasing inequality within the distribution.

Both analyses the parametric and the nonparametric one, show that the distribution remained stable until 1900, when a process of divergent growth started. Our results are robust to the entry of new cities in the sample, because when we consider a balanced panel

[7] We also estimated the Pareto exponent using simple OLS regressions and the Hill estimator, and the results were quite similar.

[8] Although there is a statistical relationship between Zipf's law and the main concentration indices: Gini, Bonferroni, Amato, and the Hirschman–Herfindahl index (Naldi, 2003).

of cities we obtain similar patterns. The hypothesis we test in the following sections is that the factor driving this change in the distribution of city sizes is the economic integration process that took place during the period 1860 to 1960, the effects of which were particularly marked after 1900.

4. Data

To analyze the growth in Spanish cities we use, as in the previous section, official city population data from the decennial censuses. Population data for 1860 are from Reher (1994), and for all the other decades our data source is the census conducted by the National Statistics Institute (*Instituto Nacional de Estadística).* Our main hypothesis is that the domestic market integration that took place between 1860 and 1960, under the presence of agglomeration economies, is a fundamental cause of the change in the structure of Spanish cities. Therefore, our main explanatory variable is the market potential, which reflects the market access of each city. This variable has been extensively used in recent studies focusing on the determinants of growth and the spatial distribution of cities, including Black and Henderson (2003), Ioannides and Overman (2004) and Bosker et al. (2008). However, one of our empirical contributions is that we do not use the distance- weighted sum of population of all other existing cities as a proxy of a city's market potential. Instead of this common option, we derive the market potential variable from a retrospective estimate of regional market potential that is distributed across cities based on the relative size of the cities in each region. The regional market potential is the so-called 'nominal market potential' or the Harris market potential equation, defined as:

$$MP_i = \sum_{j-1}^{j-n} \frac{M_j}{d_{ij}}, \qquad (3)$$

where M_j is a measure of the size of province[9] j (GDP) and d_{ij} is the distance, or in this case, the bilateral transport costs between i and j.

Yet, this measure of market access, proposed by geographers and widely adopted by economists, could be considered an *ad hoc* indicator considering that it is neither built upon a solid theoretical foundation nor it is derived from a structural estimate[10]. However, the advances made by NEG models help overcome this lack of theoretical foundation for this

[9] Provinces are Spanish NUTS3 regions.

[10] Notwithstanding, from an empirical perspective, when compared to structural estimates of market potential, Head and Mayer (2004) conclude that the Harris equation performs well.

empirical measure of market potential. Specifically, Combes et al. (2008) derive an expression for the real market potential (*RMP*) that establishes a relationship with the Harris (1954) equation. We adapt that specification to account for the relative size of cities within each region, where $RMP_{j,r}$ is the real market potential of city *j* in region *r*. Thus, we can establish the following expression:

$$RMP_{j,r} = s_j RMP_r \equiv \frac{Pop_j}{\sum_{j=1}^{n_r} Pop_j} \cdot \sum \phi_{rs} \mu_s Y_s P_s^{\sigma-1} \tag{4}$$

where RMP_r is the real market potential of region *r*, s_j measures the relative size of city *j* in region *r* and the term ϕ_{rs} measures the accessibility of the goods from *r* into market *s* as a function of transport costs, which are represented by $\phi_{rs} \equiv \tau_{rs}^{-(\sigma-1)}$.

Once this expression of real market potential for regions (RMP_r) has been derived from an NEG model, it is possible to establish a relationship between the latter and the market potential equation defined by Harris. To do this, three assumptions have to be made. First, we accept that $\phi_{rs} = d_{rs}^{-\delta}$, where d_{rs} is the distance between locations *r* and *s*, and the exponent δ corresponds to the estimated parameter for distance in the gravity equations that analyze the determinants of the volume of bilateral trade. Second, it is assumed that the share of each good within the total consumption does not vary between regions, so that $\mu=1$[11]. Finally, a key element is the inclusion of the price index $P_s^{\sigma-1}$ in the real market potential (RMP_r), an element that is absent from the Harris equation. However, it allows us to assume that there is no variation in the price indices from one region to another. Bearing these three assumptions in mind, we can thus obtain the Harris (1954) equation from our expression of real market potential.

Employing this expression, Martínez-Galarraga (2010) offers a measure of Spanish NUTS3 market potential for the years 1860, 1900 and 1930 based on Crafts' study (2005)[12]. The author obtains historical market potential figures for Spanish NUTS3 regions as follows. First, he considers that market potential can be divided into two main components. Thus, he calculates the domestic market potential (DMP_r), which includes all Spanish provinces

[11] "This simplifying assumption may be deemed acceptable when working with the consumption of final goods" Combes, Mayer and Thisse (2008), p. 305.
[12] See Martínez-Galarraga (2010) for a detailed description.

including each province's self-potential (SP_r), and the foreign market potential (FMP_{rf}) between the provincial and the international node *f*. In particular, the market potential of a province *r* (MP_r) is calculated as the sum of the domestic and foreign market potential: $MP_r = DMP_r + FMP_{rf}$. Following this expression, the domestic market potential for each one of the 47 provinces *r* is calculated as $DMP_r = \sum_{1}^{s=46} \frac{M_s}{d_{rs}} + SP_r$, being $SP_r = \frac{M_r}{d_{rr}}$ the measure of the self-potential of each province *r*, where d_{rr} is calculated taking a distance θ_{rr} equivalent to a third of the radius of a circle with an area equal to that of the province: $\theta_{rr} = 0.333\sqrt{\frac{(areaoftheprovince_r)}{\pi}}$.

The size of the provincial markets (M_r) is measured in terms of aggregate income. GDP data at the NUTS3 levels are obtained from Rosés et al. (2010). The internal distance between regions d_{rs} is calculated including transport costs, which requires access to data on distances and average transport rates for commodities. Internal transport is assumed to be by railway and coastal shipping. For railway distances, the Ministry of Public Works (*Ministerio de Obras Públicas)* (1902) and Wais (1987) were consulted. For distances between ports, electronic atlases supply information on the length of sea journeys[13]. For transport costs, data on railway rates were obtained from Herranz (2006) and coastal shipping rates in 1865 were obtained from Nadal (1975). In order to consider the reduction in sea transport costs, the data were corrected with the freight rate indices calculated by Mohammed and Williamson (2004). However, in 1860, our first benchmark year, only 32 of the 47 provinces considered were connected to the railway network. For this reason, road transport was also included in the internal market potential estimates for this year[14]. Distances by road were taken from the General Directorate of Public Works (*Dirección General de Obras Públicas)* (1861). For road transport prices, the information provided by Barquín (1999) was used. Finally, the relative weight of each transport mode in the coastal provinces was obtained from Frax (1981).

[13] www.dataloy.com and www.distances.com.

[14] In 1930, however, road transport was not yet playing an important role, and therefore, it was not considered (Herranz, 2006).

Foreign markets were added to the internal market potential[15]. The construction of the external market potential used here is based on the gravity equation for international trade estimated by Estevadeordal et al. (2003). The elasticities obtained for distance and tariffs are used here to reduce the size of foreign markets. The GDP of the main trading partners of Spain was obtained from Crafts (2005) based on the estimates of Prados de la Escosura (2000). Prevailing exchange rates were applied to convert the GDP figures from pounds to pesetas. Maritime distances were once again obtained from an electronic atlas and tariffs come from O'Rourke (2000) and Mitchell (1998a, 1998b). So, the foreign market potential of province *r* $\left(FMP_{rf}\right)$ is obtained according to the next expression, where d_{rp} captures the distance from the inland provincial node to the nearest Spanish port: $FMP_{rf} = \sum_{1}^{f=4} \frac{M_f}{d_{rp}} \cdot Distance^{-0.8} \cdot Tariff^{-1.0}$, with $d_{rp} = 1$ if *r* is a coastal province, and $d_{rp} = d_{rs}$ if *r* is an inland province.

Hence, the total market potential of province *r* $\left(MP_r\right)$ is obtained as the sum of the following terms, the first two corresponding to the domestic market potential (including the self-potential of province *r*) and the last one capturing the foreign market potential:

$$MP_r = \sum_{1}^{s=46} \frac{M_r}{d_{rs}} + SP_r + \left[\sum_{1}^{f=4} \frac{M_f}{d_{rp}} \cdot Distance^{-0.8} \cdot Tariff^{-1.0}\right], \qquad (5)$$

with d_{rp} conditioned to the coastal or inland nature of province *r*. Therefore, the market potential of city *j* in region *r* is $MP_{j,r} = s_j MP_r$.

Finally, we also introduced several geographical variables in the estimations to control for *first nature* causes. Altitude and ruggedness data by municipality were obtained from Azagra et al. (2006) and Goerlich and Cantarino (2010), respectively.

5. Empirical analysis

First, we conduct a nonparametric analysis of the effects of market potential on urban growth. To do this, we estimate the nonlinear relationship between initial market potential

[15] Differences with Crafts' (2005) study are explained in Martínez-Galarraga (2010).

and growth using a local polynomial smoothing[16] for any cross-section in our sample. Figure 3 shows the results, including the 95% confidence bands. We can observe a clearly positive relationship between market potential and city size in the three periods, although there is a temporal evolution pointing to the decreasing influence of market potential over time. Thus, when focusing on cities with a low initial market potential, we can observe that the effect on mean population growth ranges from 0.2 to 0.4 between 1860 and 1900, around 0.2 between 1900 and 1930, and from 0 to 0.2 between 1930 and 1960. A similar pattern is observed for cities with a high initial market potential. Although the effect on mean population growth basically ranges from 0.4 to 0.6 in the three periods, in the last period (1930-1960) a negative effect on population growth can even be identified for top market potential cities.

Herein Figure 3

Second, conducting a parametric analysis, we want to exploit the panel structure of our data. Therefore, we study the period 1860–1960 using panel data and consider three homogeneous periods: 1860–1900, 1900–1930 and 1930–1960. The initial and final sets of cities are those plotted in Figure 1. Our baseline equation is similar to that proposed by Black and Henderson (2003) and Henderson and Wang (2007):

$$g_{it} = \alpha + \beta \ln mp_{it-1} + \gamma(\ln mp_{it-1})^2 + \phi g_{it-1} + \delta X_{it} + \varepsilon_{it}, \qquad (6)$$

where the independent variable is the logarithmic growth rate, $g_{it} = (\ln S_{it} - \ln S_{it-1})$. The main explanatory variable is the city relative market potential (mp_{it}), defined as $mp_{it} = \frac{MP_{it}}{\frac{1}{n_t}\sum_{1}^{n_t} MP_{jt}}$, and X_{it} is a vector of both time variant and time invariant variables.

The cross-sectional measure of market potential is normalized by the contemporaneous average market potential to avoid that later periods can overpower effects in earlier ones through absolute growth in market potential (Black and Henderson, 2003). Furthermore, we use a measure of market potential excluding each province's self-potential to avoid endogeneity concerns (more on this below).

16 The local polynomial smoother fits the growth rate $g_{it} = (\ln S_{it+1} - \ln S_{it})$ to a polynomial form of $\ln MP_{it-1}$ via locally weighted least squares. We use the lpolyci command in STATA with the following options: local mean smoothing, a Gaussian kernel function to calculate the locally weighted polynomial regression and a bandwidth determined using Silverman's (1986) 'rule-of-thumb'.

The time invariant variables represent the locational fundamentals (*first nature* factors) of each location. They include a coastal dummy indicating if the city has access to the sea, the city's altitude and a measure of ruggedness taken at the city level. Provincial fixed-effects are also included to control for other local characteristics for which we have no data. We also introduce the squared market potential variable to capture the possible nonlinear behaviour detected previously with nonparametric analysis. A positive β coefficient is expected, as city growth increases in those locations with greatest market potential. However, this positive effect is expected to fall with the market potential size, since the marginal effect of market potential should be stronger for cities with low market potential (Black and Henderson, 2003). Thus, a γ negative coefficient is expected. Finally, we add the lagged growth rate to the specification to control for the persistence in growth rates.

Herein Table 2

Table 2 shows our first set of results: the OLS estimates of Eq. (6). The first column reports a simple OLS regression with only one explanatory variable, the market potential. The coefficient is clearly positive and significant. In column 2 we add the squared market potential variable, with an estimated negative sign indicating significant quadratic effects. These results remain the same even when we add the geographical variables (column 3) and the set of provincial and time fixed-effects (columns 4 to 6). However, estimated coefficients for market potential slowly increase until a value of 0.067 in column 6. Figure 4 confirms that the relationship between city growth and the initial market potential is positive, and also reveals a slight nonlinear relationship.

Herein Figure 4

Nevertheless, there could be problems with the OLS estimation of Eq. (6). As Henderson and Wang (2007) highlight, there are unobservable effects related to individual cities that may operate at a regional level and be correlated with city growth and market potential: some geographical characteristics, local culture, business climate or institutions. Furthermore, in any city growth estimation there could be an issue with the time persistence in the error structure. Moreover, there might be potential spatial correlation between cities. Due that we consider that the infrastructures are a key element to explain the changes in the market potential for cities, our main concern relates with the role played by these infrastructures. Policy makers tend to improve infrastructures in the most populated cities, but these infrastructures (roads, railways, etc.) undoubtedly also increase the market access of these locations (Holl, 2012), generating problems with our specification. To deal with this last issue we use a measure of market potential excluding each province's self-potential; thus, in our specification changes in local infrastructures can affect population growth in city *i* (Garcia-López et al., 2013), but we exclude the possible effect on city *i* market potential. The policy decision process and the construction of these infrastructures often take several periods, so the past growth rate we introduce in the specification can incorporate the forecasting of these infrastructures, and hence this dynamic model could alleviate the possible endogeneity problem.

We need to instrument the market potential variable (and its square) in the first stage regressions of the IV estimation. We use three instruments: city population as a proportion of the total provincial population, its square and a Madrid dummy variable that takes a value of 1 for the city of Madrid and 0 for all others. Population can serve as a good measure of market potential, and in some papers it is used directly instead of GDP (Black and Henderson, 2003; Ioannides and Overman, 2004; Henderson and Wang, 2007). To be cautious, we use the lagged values of these shares; thus, values from 1787 are used to estimate market potential in 1860, and so on.[17] The Madrid dummy reflects that the capital city tends to be more dominant the more political instability there is in a country and the more authoritarian is its regime (Ades and Glaeser, 1995), given that Spain suffered military dictatorships during the twentieth century (1923-1931 and 1939-1975). Ayuda et al. (2010) also introduce this Madrid dummy variable and find it significant.

It could be argued that using city population as a proportion of the total provincial population might not be a good instrument, because it is probably highly correlated with city growth, our endogenous variable in the second stage regression. However, one of the particular features of the Spanish urban system was that internal migrations were not statistically related to differential city growth for the whole sample of cities in 1930. Silvestre (2005) shows that, although Spanish internal migrations grew significantly during the 1920s, these movements were limited to just a few big cities. In fact, Silvestre calculates

[17] Population data for 1787 come from Reher (1994).

that by 1930, two provinces, Madrid and Barcelona, accounted for 45.97 per cent of the total stock of Spanish migrants.

Herein Table 3

Table 3 shows the IV results. We estimate the second stage regressions by GMM; some statistics from the first stage regressions are also reported.[18] Our instruments seem to perform well, as the R^2 in the first stage regressions exceeds 0.7 in most of the specifications, the weak instruments hypothesis is always rejected using the Stock-Yogo test (results not shown), and all the models pass the overidentification test (Hansen J statistic) for any significance level, except the model in column (1). Here again, when the only explanatory variable is the market potential (column 1) we obtain a clearly positive and significant coefficient, although higher than in the OLS regression. In general, coefficients for market potential and its square tend to be higher than in the OLS estimations in all the specifications. When we add the squared market potential variable (column 2) both coefficients are significant. Once more the effect of market potential is positive and significant, and the squared market potential variable has a significant negative sign revealing that the impact on market potential growth has a quadratic shape. These results remain similar in all the models (column 2 to 6), when we add all the sets of controls, provincial and time fixed-effects. In the last specification (column 6), the estimated effect of market potential on growth is almost 0.071, a slightly higher value to the OLS estimate.[19]

The number of cities in our sample grows from 266 in 1860 to 877 in 1930, so it might be argued that our results are driven by new cities entering the distribution rather than by the effects of market integration. To address this concern we estimate Eq. (6) using data only from those cities that were created before1860, and do not include any of the new cities.[20] As such, our sample comprises the set of cities plotted in map (a) in Figure 1. Table 4 shows the IV results of this robustness check. The instruments used are the same as above. The coefficients for the market potential maintain the sign and significance as in the estimations for the whole sample, confirming that our results are robust to the entry of

[18] The complete results of the reduced regressions, first stage regressions and all the tests, not shown for size restrictions, are available from the authors on request.
[19] Ahlfeldt and Feddersen (2010) examine the incidence of speed rail on market access in Germany. As in our analysis, they obtain similar coefficient estimates both for the OLS and IV estimation. They conclude that the endogeneity concerns can be rejected in that subject case.
[20] As Black and Henderson (2003), we also estimate Eq. (6) including a variable measuring the number of cities in each period, and results do not vary.

new cities. Nevertheless, there are two interesting differences. First, the estimated coefficients for the market potential tend to be smaller, meaning that the influence of market potential on growth is weaker in these cities. Thus, as these cities tend to be large cities this result supports the decreasing influence of market potential over time. Second, the past population growth is always negative and significant, while in the estimations using all cities this variable was only significant in one case (column 5 in Table 3), confirming that the initially largest cities were especially involved in the divergence process.

6. Conclusions

In this paper we have analyzed the determinants of city growth, focusing primarily on the factors that determine the configuration of the urban system. Our hypothesis has been that, in the context of the growing economic integration and industrialization experienced by Spain from the second half of the 19th century onwards, the presence of agglomeration economies implied that the access to markets was an important factor explaining the relative growth of Spanish cities and so affecting the country's urban pattern. In order to test this hypothesis, in the empirical analysis both the location fundamentals (relating to Spain's geography) and the city market size (a variable introduced by theoretical approaches conducted within NEG) have been considered as determinants of city growth. Our results show that Spain's urban growth was related initially to *first nature* characteristics but that it was also affected by the forces of the agglomeration economies measured in terms of home market size. All these elements had an impact on Spanish urban design, which experienced a growth in its city size inequality due to the fact that, throughout the process of economic growth and integration, the cities with greatest market potential benefited most from the presence of agglomeration economies.

The present analysis confirms the results obtained in some works devoted to the study of determinants of population growth in US, Brazilian or Chinese cities during different periods comprising the second half of the XXth century. In this respect our main contribution has been the use of a direct measure of market access. It has been constructed making use of regional GDPs and transport costs at the time of each one of the benchmark years. Results point out to the presence of a positive effect of market potential on population growth of Spanish cities.

Our results also fall in line with those that have pointed to the importance of agglomeration forces in the definition of Spanish economic geography. A bulk of recent NEG empirical literature devoted to the analysis of the Spanish historical experience has showed that market access acted as a relevant factor explaining regional industry location, the upsurge of wage gradients or the direction and intensity of internal migratory flows along the process of national market integration and industrialization. In this paper it has been shown that the relative market potential of cities was also a key factor in explaining their relative growth during the period 1860-1960, and specially from the very beginning of the XXth century.

References

Ades, A. F. and E. L. Glaeser (1995), Trade and circuses: explaining urban giants. *The Quarterly Journal of Economics*, 110, 195–227.

Ahlfeldt, G. M. and A. Feddersen (2010), From periphery to core: economic adjustments to high speed rail. Working Paper, London School of Economics & University of Hamburg.

Au, C. C. and J. V. Henderson (2006), Are Chinese Cities Too Small. *Review of Economic Studies*, 73, 549-576

Ayuda, M.I., F. Collantes, and V. Pinilla (2010), From locational fundamentals to increasing returns: The spatial concentration of population in Spain, 1787-2000. *Journal of Geographic Systems* 12, 25–50

Azagra, J., P. Chorén, F. J. Goerlich, and M. Mas (2006), *La localización de la población española sobre el territorio. Un siglo de cambios: un estudio basado en series homogéneas (1900–2001),* Fundación BBVA.

Barquín, R. (1999), El ahorro social: una perspectiva dinámica, in Muñoz, Sanz and Vidal (eds.), *Siglo y medio del ferrocarril en España 1848-1998. Economía, Industria y Sociedad,* Fundación de los Ferrocarriles Españoles, Madrid, 337-354.

Black, D. and V. Henderson (2003), Urban evolution in the USA. *Journal of Economic Geography*, 3 343-372.

Bosker, E. M., S. Brakman, H. Garretsen, H. de Jong and M. Schramm (2008), Ports, Plagues and Politics: Explaining Italian city growth 1300-1861. *European Review of Economic History,* 12: 97-131.

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

Cheshire, P. (1999), Trends in sizes and structure of urban areas. *Handbook of Regional and Urban Economics*, Vol. 3, P. Cheshire and E. S. Mills, eds. Amsterdam: Elsevier Science, Chapter 35, 1339–1373.

Combes P. P., T. Mayer, J. F. and Thisse (2008), *Economic Geography: The integration of regions and nations*, Princeton University Press.

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

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

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

Da Mata, D., U. Deichmann, J. V. Henderson, S. V. Lall, and H. G. Wang (2007), Determinants of city growth in Brazil. *Journal of Urban Economics*, 62(2), 252-272

Davis, D. R., and D. E. Weinstein (2002), Bones, Bombs, and Break Points: The Geography of Economic Activity. *The American Economic Review*, 92, 1269-1289.

Dirección General de Obras Públicas, (1861), Memoria sobre el progreso de las Obras Públicas en España, en los años 1859 y 1860, Madrid.

Dobkins, L. and Y.M. Ioannides (2001), Spatial Interactions among US cities: 1900–1990. *Regional Science and Urban Economics*, 31, 701–731.

Estevadeordal, A., B. Frantz, and A. M. Taylor (2003), The rise and fall of world trade, 1870-1939. *Quarterly Journal of Economics*, 118(2): 359-407.

Frax, E., (1981), Puertos y comercio de cabotaje en España, 1857-1934, Banco de España, Madrid.

Fujita, M., P. Krugman, and A. J. Venables (1999), *The spatial economy: cities, regions and international trade*, The MIT Press, Cambridge.

Gabaix, X., and Y. M. Ioannides (2004), The evolution of city size distributions. *Handbook of urban and regional economics*, Vol. 4, J. V. Henderson and J. F. Thisse, eds. Amsterdam: Elsevier Science, 2341–2378

Gabaix, X., and R. Ibragimov, (2011), Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents. *Journal of Business Economics and Statistics,* vol. 29(1), p. 24–39.

Garcia-López, M.-A., A. Holl, E. Viladecans-Marsal (2013), Suburbanization and highways: when the Romans, the Bourbons and the first cars still shape Spanish cities. Working papers 2013/5, Institut d'Economia de Barcelona (IEB).

Goerlich, F., and I. Cantarino (2010), Un índice de rugosidad del terreno a escala municipal a partir de modelos de elevación digital de acceso público. Documento de Trabajo. Julio 2010. FBBVA

Goerlich, F. J., and M. Mas (2009), Drivers of Agglomeration: Geography vs. History. *The Open Urban Studies Journal*, 2, pp. 28-42.

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(4), 315-348.

Head, K., and T. Mayer (2004), The empirics of agglomeration and trade, in *Handbook of Urban and Regional Economics: Cities and Geography*, 4, 2609-2669, Elsevier.

Henderson, J. V., and H. G. Wang (2007), Urbanization and city growth: The role of institutions. *Regional Science and Urban Economics*, 37, 283–313.

Herranz, A., (2006), Railroad impact in backward economies: Spain 1850-1913. *Journal of Economic History*, 66, 853-881.

Holl, A., (2012), Market potential and firm-level productivity in Spain. *Journal of Economic Geography*, Vol. 12(6): 1191–1215.

Ioannides, Y., and H. Overman (2004), Spatial evolution of the US urban system. *Journal of Economic Geography*, 4: 131-156.

Kancs, D. A., (2005), Can we use NEG models to predict migration flows? An example of CEE accession countries. *Migration Letters* 2:32–63

Krugman, P., (1991), Increasing returns and economic geography. *Journal of Political Economy*, 93, 483-499.

Lanaspa, L., F. Pueyo, and F. Sanz (2003), The Evolution of Spanish Urban Structure during the Twentieth Century. *Urban Studies*, Vol. 40(3): 567–580.

Martínez-Galarraga, J., (2010), *Market integration and regional inequality in Spain, 1860-1930*, PhD dissertation, Universitat de Barcelona.

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

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

Midelfart-Knarvik, K. H., H. Overman, S. Redding, and A. J. Venables (2002) The location of European industry. In: European Commission (Ed.), *European economy: European integration and the functioning of product markets*. Directorate-General for Economic and Financial Affairs, Brussels, pp. 213–269.

Ministerio de Obras Públicas, (1902), Estadísticas de Obras Públicas, años 1899 y 1900, Dirección General de Obras Públicas, Madrid.

Mion, G., (2004), Spatial externalities and empirical analysis: the case of Italy. *Journal of Urban Economics*, 56, 97–118.

Mitchell, B. R., (1998a), *International historical statistics. Europe 1750-1993*, Macmillan, London.

Mitchell, B. R., (1998b), *International historical statistics. The Americas 1750-1993,* Macmillan, London.

Mohammed, S. I. S., and J. G. Williamson (2004), Freight rates and productivity gains in British tramp shipping 1869–1950. *Explorations in Economic History*, 41, 172-203.

Nadal, J., (1975), El fracaso de la revolución industrial en España, 1814-1913. Ariel, Barcelona.

Naldi, M., (2003), Concentration indices and Zipf's law. *Economics Letters*, 78, 329–334.

O'Rourke, K. H., (2000), Tariffs and growth in the late 19th century. *The Economic Journal*, 110, 456-483.

Paluzie, E., J. Pons, and D. A. Tirado (2009a), A test of the market potential equation in Spain. *Applied Economics*, 41, 1487-1493.

Paluzie, E, J. Silvestre, J. Pons, and D. A. Tirado (2009b), Migrants and market potential in Spain over the 20th century: a test of the New Economic Geography. *Spanish Economic Review*, 11(4): 243-265.

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. J., (2010), The empirics of new economic geography. *Journal of Regional Science,* 50, 297-311.

Reher, D., (1994), Ciudades, procesos de urbanización y sistemas urbanos en la península ibérica, 1550-1991. In: Guàrdia, M., Monclús, F. J., Oyón, J.L. (Eds.), *Atlas histórico de las ciudades europeas*. Península Ibérica. Salvat - Centre de Cultura Contemporània de Barcelona, Barcelona, 1-30.

Riley, S. J., S. D. De Gloria, R. Elliot (1999), A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Science* 5, 23-27.

Ringrose, D., (1996), Spain, Europe and the "Spanish Miracle", 1700-1990. Cambridge University Press, Cambridge (MA)

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

Rosés, J. R., J. Martínez-Galarraga, and D. A. Tirado (2010), The Upswing of Regional Income Inequality in Spain (1860-1930). *Explorations in Economic History*, 47, 244-257.

Silverman, B. W., (1986), Density Estimation for Statistics and Data Analysis, 1st Edition. Chapman and Hall/CRC, Cleveland.

Silvestre, J., (2005), Internal migrations in Spain, 1877-1930. *European Review of Economic History*, 9, 233-265.

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

Wais, F., (1987), Historia de los ferrocarriles españoles, Fundación de los Ferrocarriles Españoles, Madrid.

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

Table 1. Number of cities and descriptive statistics by year

Year	Cities	Mean population	Standard deviation	Minimum	Maximum
1860	266	13,267.34	24,712.85	5,004	279,379
1900	657	13,720.86	32,795.34	5,001	539,835
1930	877	15,684.87	50,554.06	5,000	1,005,565
1960	1,030	20,878.10	91,372.98	5,004	2,259,931

Data sources: Reher (1994) and Instituto Nacional de Estadística, www.ine.es.

Table 2. Spanish city size growth: Panel 1860-1960, OLS

	(1)	(2)	(3)	(4)	(5)	(6)
Relative MP excluding self-potential	0.029***	0.057***	0.058***	0.052***	0.071***	0.067***
	(0.004)	(0.007)	(0.009)	(0.008)	(0.011)	(0.011)
Relative MP square excluding self-potential		-0.002***	-0.002***	-0.002***	-0.003***	-0.003***
		(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Population growth (t-1)			-0.011	0.014	-0.096*	-0.067
			(0.051)	(0.048)	(0.053)	(0.050)
Coastal dummy			-0.036	-0.050	-0.045	-0.057
			(0.036)	(0.034)	(0.039)	(0.036)
ln(Altitude)			-0.016	-0.022**	-0.042***	-0.044***
			(0.009)	(0.009)	(0.013)	(0.012)
ln(Ruggedness)			-0.048***	-0.028**	-0.033*	-0.024
			(0.012)	(0.012)	(0.018)	(0.017)
Provincial fixed-effects	N	N	N	N	Y	Y
Time fixed-effects	N	N	N	Y	N	Y
R^2	0.040	0.050	0.089	0.170	0.207	0.266
Observations	1661	1661	1010	1010	1010	1010

Notes: Coefficient (robust standard errors). Significant at the *10%, **5%, ***1% level.

Table 3. Spanish city size growth: Panel 1860-1960, IV GMM

	(1)	(2)	(3)	(4)	(5)	(6)
Relative MP excluding self-potential	0.056***	0.102***	0.098***	0.071***	0.097***	0.071***
	(0.006)	(0.014)	(0.013)	(0.009)	(0.014)	(0.011)
Relative MP square excluding self-potential		-0.004***	-0.004***	-0.002***	-0.004***	-0.003***
		(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Population growth (t-1)			-0.051	-0.012	-0.122**	-0.075
			(0.052)	(0.049)	(0.053)	(0.049)
Coastal dummy			-0.046	-0.053	-0.048	-0.057
			(0.037)	(0.034)	(0.038)	(0.035)
ln(Altitude)			-0.007	-0.016*	-0.038***	-0.041***
			(0.009)	(0.009)	(0.012)	(0.011)
ln(Ruggedness)			-0.046***	-0.028**	-0.028	-0.023
			(0.012)	(0.012)	(0.017)	(0.017)
Provincial fixed-effects	N	N	N	N	Y	Y
Time fixed-effects	N	N	N	Y	N	Y
First Stage statistics						
FS (Relative MP)): F test, (p-value)	64.69	64.69	82.92	89.63	95.78	107.82
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FS (Relative MP): Uncentered R^2	0.764	0.764	0.844	0.854	0.884	0.893
FS (Relative MP square): F test, (p-value)		19.98	20.78	21.00	27.08	28.13
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FS (Relative MP square): Uncentered R^2		0.661	0.695	0.699	0.736	0.740
Second Stage (GMM) statistics						
Uncentered R^2	0.403	0.426	0.439	0.496	0.518	0.556
Hansen J statistic, p-value	0.046	0.391	0.421	0.879	0.520	0.994
Observations	1010	1010	1010	1010	1010	1010

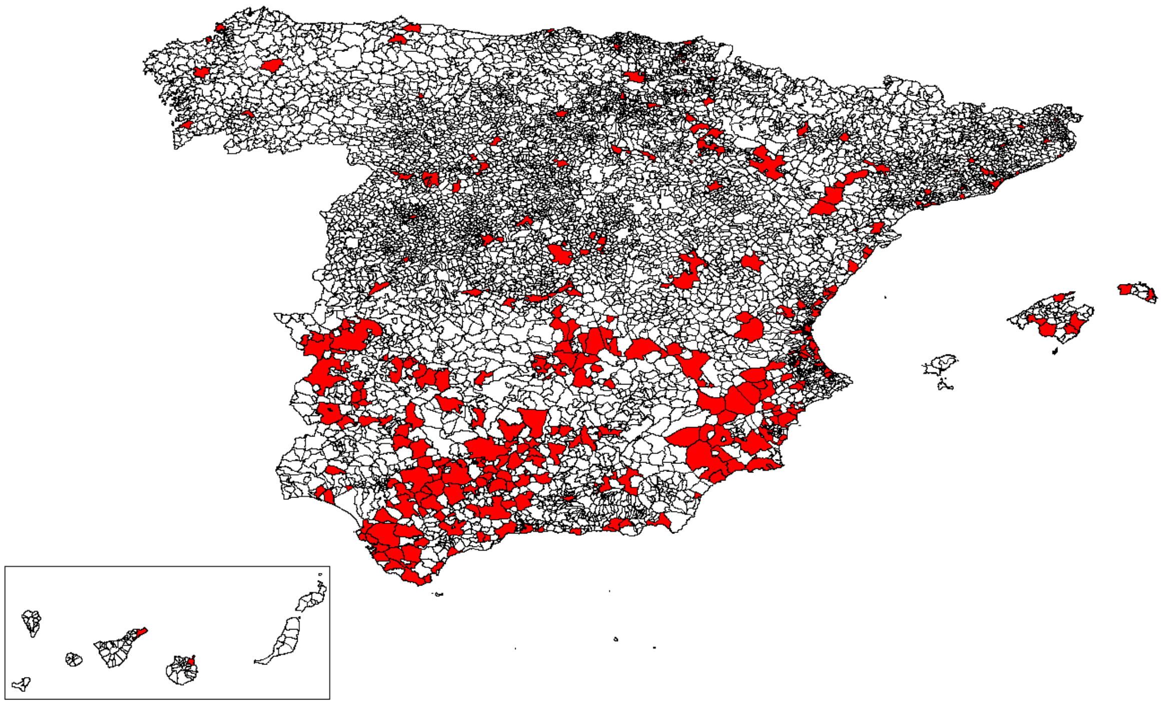
Notes: Coefficient (robust standard errors). Significant at the *10%, **5%, ***1% level.

Table 4. Spanish city size growth: Robustness checks, panel with no entry, IV GMM results

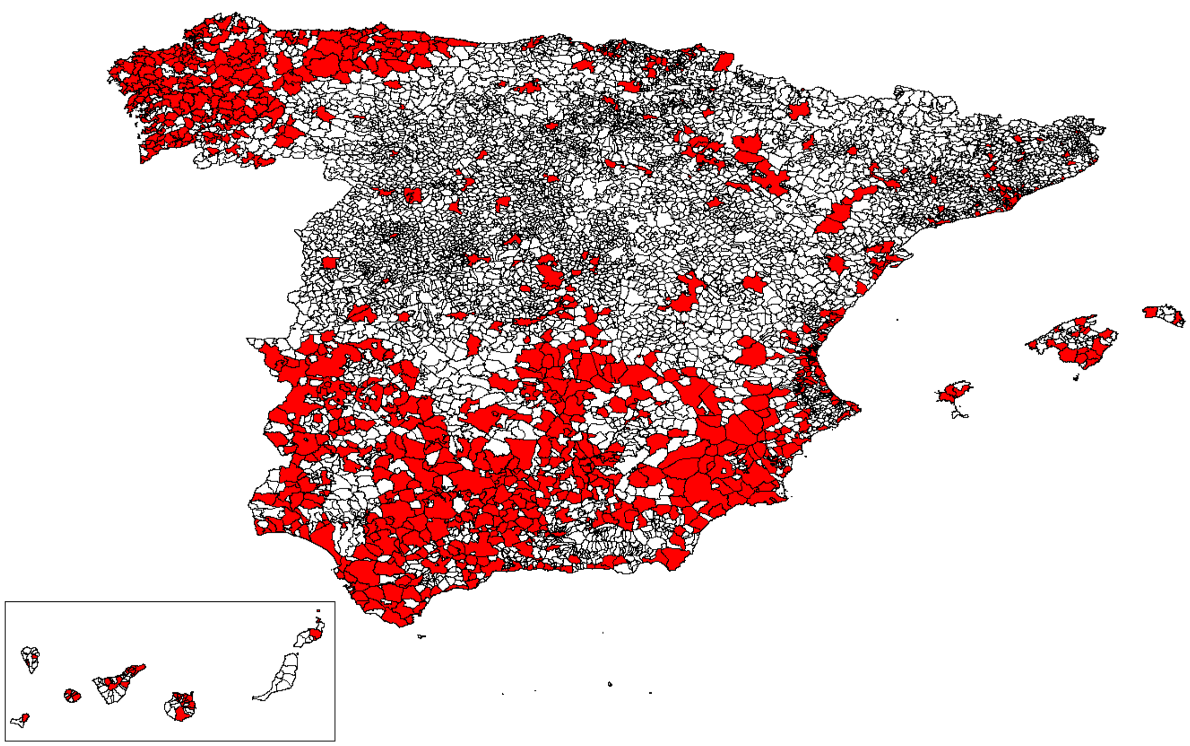
	(1)	(2)	(3)	(4)	(5)	(6)
Relative MP excluding self-potential	0.043***	0.078***	0.083***	0.070***	0.071***	0.058***
	(0.005)	(0.011)	(0.010)	(0.008)	(0.012)	(0.011)
Relative MP square excluding self-potential		-0.003***	-0.003***	-0.002***	-0.003***	-0.002***
		(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Population growth (t-1)			-0.201***	-0.145***	-0.276***	-0.224***
			(0.055)	(0.054)	(0.054)	(0.053)
Coastal dummy			-0.049	-0.053	-0.051	-0.059
			(0.052)	(0.049)	(0.058)	(0.055)
ln(Altitude)			-0.009	-0.013	-0.038**	-0.041**
			(0.013)	(0.013)	(0.018)	(0.017)
ln(Ruggedness)			0.026	0.025	0.005	0.007
			(0.017)	(0.016)	(0.022)	(0.022)
Provincial fixed-effects	N	N	N	N	Y	Y
Time fixed-effects	N	N	N	Y	N	Y
First Stage statistics						
FS (Relative MP)): F test, (p-value)	60.16	60.16	79.49	84.81	86.32	94.24
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FS (Relative MP): Uncentered R^2	0.766	0.766	0.853	0.861	0.908	0.917
FS (Relative MP square): F test, (p-value)		19.32	20.34	20.48	25.42	25.79
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FS (Relative MP square): Uncentered R^2		0.661	0.705	0.708	0.787	0.791
Second Stage (GMM) statistics						
Uncentered R^2	0.511	0.524	0.543	0.571	0.621	0.642
Hansen J statistic, p-value	0.097	0.888	0.621	0.474	0.896	0.746
Observations	673	673	673	673	673	673

Notes: Coefficient (robust standard errors). Significant at the *10%, **5%, ***1% level.

Figure 1. Cities in the sample, 1860 and 1930



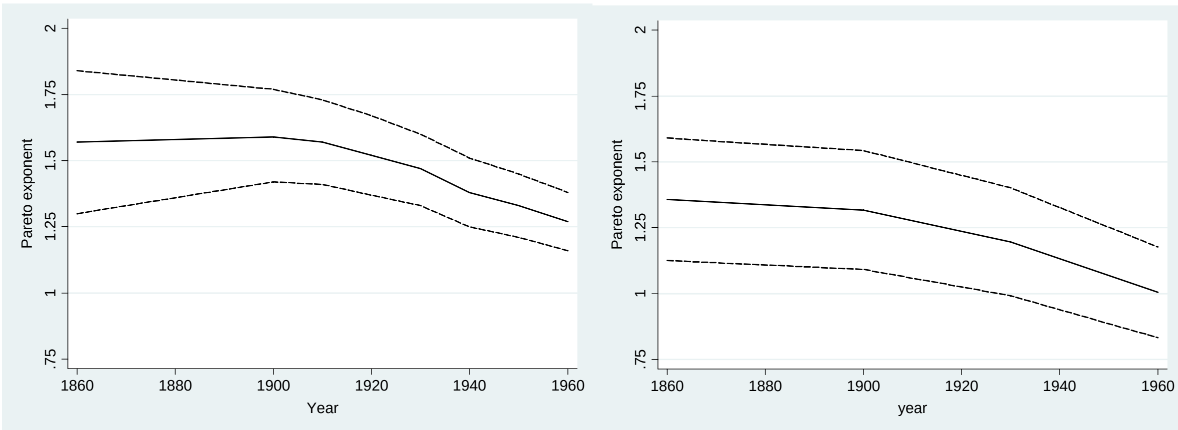
(a) 1860



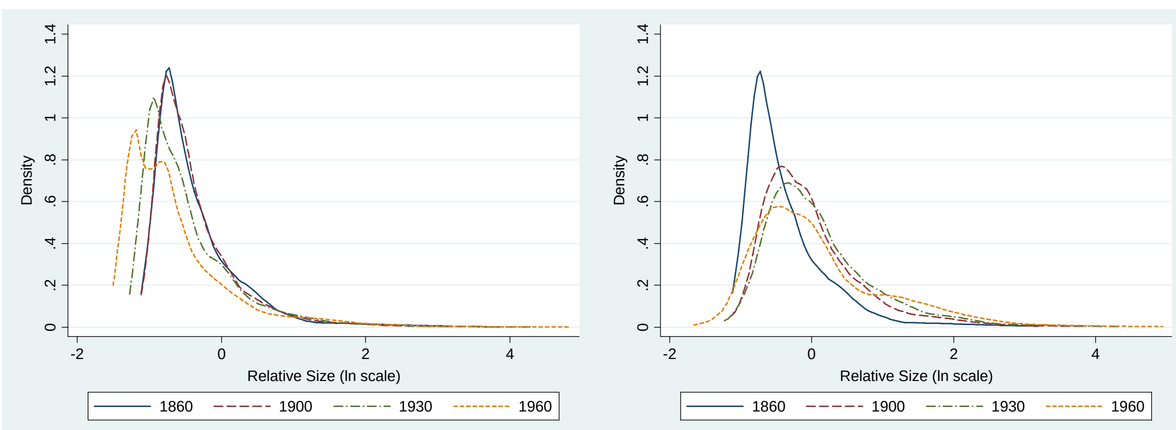
(b) 1930

Notes: Geographical boundaries defined according to census in 2000.

Figure 2. Evolution of the Spanish city size distribution



(a) Pareto: All sample (b) Pareto: Balanced panel (262 municipalities)



(c) Empirical pdfs: All sample (d) Empirical pdfs: Balanced panel (262 municipalities)

Notes: The Pareto exponents are estimated using Gabaix and Ibragimov's Rank-$1/2$ estimator. Dashed lines represent the standard errors calculated applying Gabaix and Ioannides's (2004) corrected standard errors: $\text{GI s.e.} = \hat{a} \cdot (2/N)^{1/2}$, where N is the sample size.

Figure 3. Growth (ln scale) by initial market potential

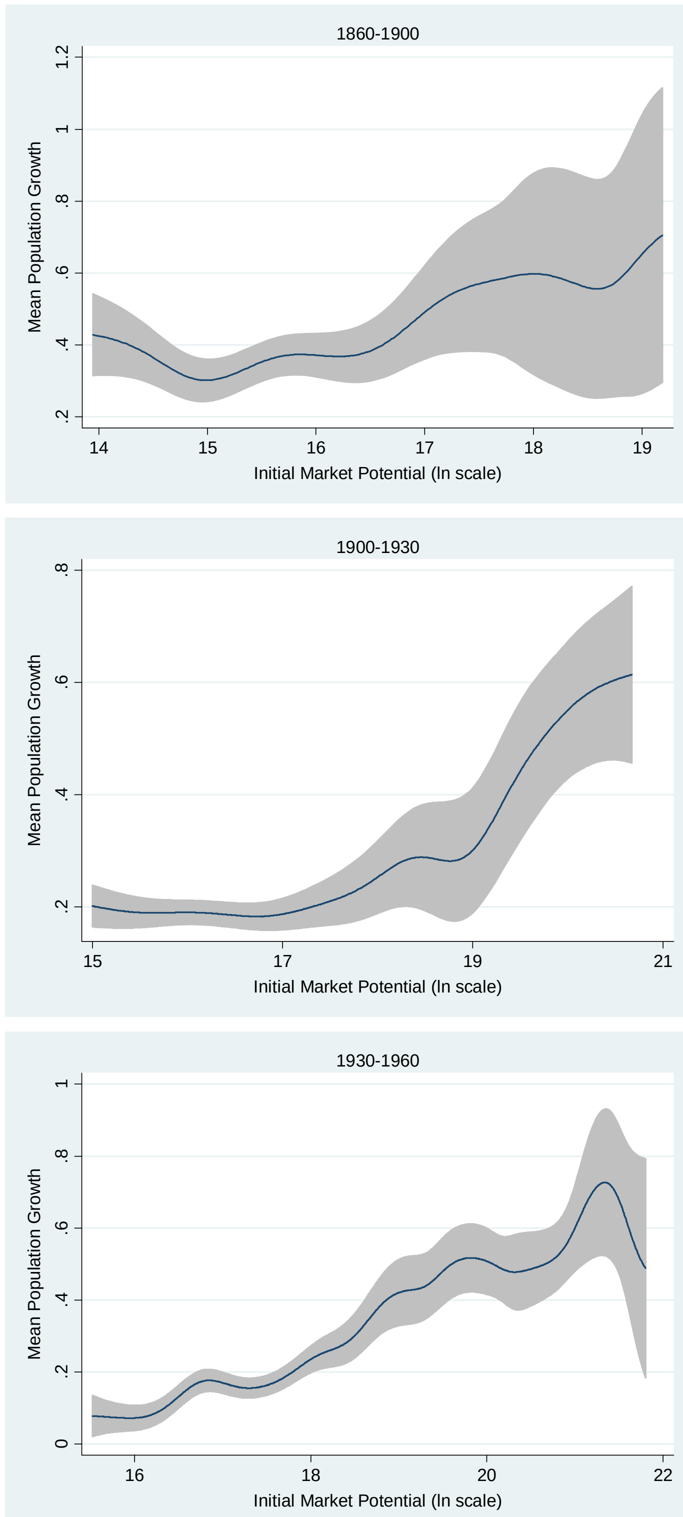
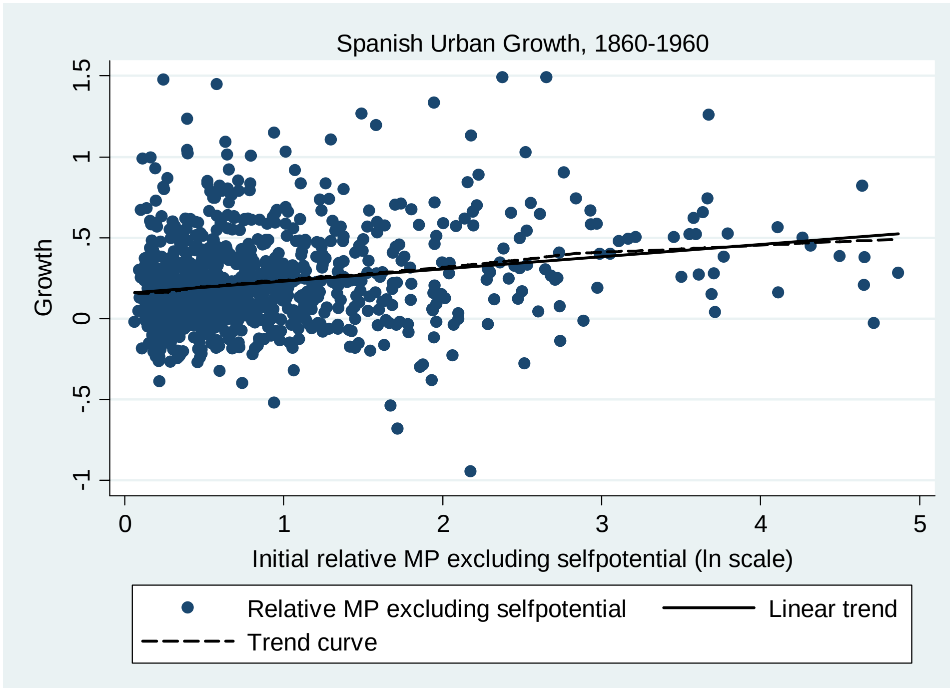


Figure 4. Scatter plot of city growth (ln scale) against initial relative market potential (excluding selfpotential), Panel 1860–1960



Notes: The values are shown until relative market potential of 5, because the few extreme highest values distort the graph. Curve fitted using LOcally WEighted Scatter plot Smoothing.

2011

2011/1, Oppedisano, V; Turati, G.: "What are the causes of educational inequalities and of their evolution over time in Europe? Evidence from PISA"
2011/2, Dahlberg, M; Edmark, K; Lundqvist, H.: "Ethnic diversity and preferences for redistribution "
2011/3, Canova, L.; Vaglio, A.: "Why do educated mothers matter? A model of parental help”
2011/4, Delgado, F.J.; Lago-Peñas, S.; Mayor, M.: “On the determinants of local tax rates: new evidence from Spain”
2011/5, Piolatto, A.; Schuett, F.: “A model of music piracy with popularity-dependent copying costs”
2011/6, Duch, N.; García-Estévez, J.; Parellada, M.: “Universities and regional economic growth in Spanish regions”
2011/7, Duch, N.; García-Estévez, J.: “Do universities affect firms’ location decisions? Evidence from Spain”
2011/8, Dahlberg, M.; Mörk, E.: “Is there an election cycle in public employment? Separating time effects from election year effects”
2011/9, Costas-Pérez, E.; Solé-Ollé, A.; Sorribas-Navarro, P.: “Corruption scandals, press reporting, and accountability. Evidence from Spanish mayors”
2011/10, Choi, A.; Calero, J.; Escardíbul, J.O.: “Hell to touch the sky? private tutoring and academic achievement in Korea”
2011/11, Mira Godinho, M.; Cartaxo, R.: “University patenting, licensing and technology transfer: how organizational context and available resources determine performance”
2011/12, Duch-Brown, N.; García-Quevedo, J.; Montolio, D.: “The link between public support and private R&D effort: What is the optimal subsidy?”
2011/13, Breuillé, M.L.; Duran-Vigneron, P.; Samson, A.L.: “To assemble to resemble? A study of tax disparities among French municipalities”
2011/14, McCann, P.; Ortega-Argilés, R.: “Smart specialisation, regional growth and applications to EU cohesion policy”
2011/15, Montolio, D.; Trillas, F.: “Regulatory federalism and industrial policy in broadband telecommunications”
2011/16, Pelegrín, A.; Bolancé, C.: “Offshoring and company characteristics: some evidence from the analysis of Spanish firm data”
2011/17, Lin, C.: “Give me your wired and your highly skilled: measuring the impact of immigration policy on employers and shareholders”
2011/18, Bianchini, L.; Revelli, F.: “Green polities: urban environmental performance and government popularity”
2011/19, López Real, J.: “Family reunification or point-based immigration system? The case of the U.S. and Mexico”
2011/20, Bogliacino, F.; Piva, M.; Vivarelli, M.: “The impact of R&D on employment in Europe: a firm-level analysis”
2011/21, Tonello, M.: “Mechanisms of peer interactions between native and non-native students: rejection or integration?”
2011/22, García-Quevedo, J.; Mas-Verdú, F.; Montolio, D.: “What type of innovative firms acquire knowledge intensive services and from which suppliers?”
2011/23, Banal-Estañol, A.; Macho-Stadler, I.; Pérez-Castrillo, D.: “Research output from university-industry collaborative projects”
2011/24, Ligthart, J.E.; Van Oudheusden, P.: “In government we trust: the role of fiscal decentralization”
2011/25, Mongrain, S.; Wilson, J.D.: “Tax competition with heterogeneous capital mobility”
2011/26, Caruso, R.; Costa, J.; Ricciuti, R.: “The probability of military rule in Africa, 1970-2007”
2011/27, Solé-Ollé, A.; Viladecans-Marsal, E.: “Local spending and the housing boom”
2011/28, Simón, H.; Ramos, R.; Sanromá, E.: “Occupational mobility of immigrants in a low skilled economy. The Spanish case”
2011/29, Piolatto, A.; Trotin, G.: “Optimal tax enforcement under prospect theory”
2011/30, Montolio, D; Piolatto, A.: “Financing public education when altruistic agents have retirement concerns”
2011/31, García-Quevedo, J.; Pellegrino, G.; Vivarelli, M.: “The determinants of YICs’ R&D activity”
2011/32, Goodspeed, T.J.: “Corruption, accountability, and decentralization: theory and evidence from Mexico”
2011/33, Pedraja, F.; Cordero, J.M.: “Analysis of alternative proposals to reform the Spanish intergovernmental transfer system for municipalities”
2011/34, Jofre-Monseny, J.; Sorribas-Navarro, P.; Vázquez-Grenno, J.: “Welfare spending and ethnic heterogeneity: evidence from a massive immigration wave”
2011/35, Lyytikäinen, T.: “Tax competition among local governments: evidence from a property tax reform in Finland”
2011/36, Brülhart, M.; Schmidheiny, K.: “Estimating the Rivalness of State-Level Inward FDI”

2011/37, García-Pérez, J.I.; Hidalgo-Hidalgo, M.; Robles-Zurita, J.A.: "Does grade retention affect achievement? Some evidence from Pisa"
2011/38, Boffa, f.; Panzar. J.: "Bottleneck co-ownership as a regulatory alternative"
2011/39, González-Val, R.; Olmo, J.: "Growth in a cross-section of cities: location, increasing returns or random growth?"
2011/40, Anesi, V.; De Donder, P.: "Voting under the threat of secession: accommodation vs. repression"
2011/41, Di Pietro, G.; Mora, T.: "The effect of the l'Aquila earthquake on labour market outcomes"
2011/42, Brueckner, J.K.; Neumark, D.: "Beaches, sunshine, and public-sector pay: theory and evidence on amenities and rent extraction by government workers"
2011/43, Cortés, D.: "Decentralization of government and contracting with the private sector"
2011/44, Turati, G.; Montolio, D.; Piacenza, M.: "Fiscal decentralisation, private school funding, and students' achievements. A tale from two Roman catholic countries"

2012

2012/1, Montolio, D.; Trujillo, E.: "What drives investment in telecommunications? The role of regulation, firms' internationalization and market knowledge"
2012/2, Giesen, K.; Suedekum, J.: "The size distribution across all "cities": a unifying approach"
2012/3, Foremny, D.; Riedel, N.: "Business taxes and the electoral cycle"
2012/4, García-Estévez, J.; Duch-Brown, N.: "Student graduation: to what extent does university expenditure matter?"
2012/5, Durán-Cabré, J.M.; Esteller-Moré, A.; Salvadori, L.: "Empirical evidence on horizontal competition in tax enforcement"
2012/6, Pickering, A.C.; Rockey, J.: "Ideology and the growth of US state government"
2012/7, Vergolini, L.; Zanini, N.: "How does aid matter? The effect of financial aid on university enrolment decisions"
2012/8, Backus, P.: "Gibrat's law and legacy for non-profit organisations: a non-parametric analysis"
2012/9, Jofre-Monseny, J.; Marín-López, R.; Viladecans-Marsal, E.: "What underlies localization and urbanization economies? Evidence from the location of new firms"
2012/10, Mantovani, A.; Vandekerckhove, J.: "The strategic interplay between bundling and merging in complementary markets"
2012/11, Garcia-López, M.A.: "Urban spatial structure, suburbanization and transportation in Barcelona"
2012/12, Revelli, F.: "Business taxation and economic performance in hierarchical government structures"
2012/13, Arqué-Castells, P.; Mohnen, P.: "Sunk costs, extensive R&D subsidies and permanent inducement effects"
2012/14, Boffa, F.; Piolatto, A.; Ponzetto, G.: "Centralization and accountability: theory and evidence from the Clean Air Act"
2012/15, Cheshire, P.C.; Hilber, C.A.L.; Kaplanis, I.: "Land use regulation and productivity – land matters: evidence from a UK supermarket chain"
2012/16, Choi, A.; Calero, J.: "The contribution of the disabled to the attainment of the Europe 2020 strategy headline targets"
2012/17, Silva, J.I.; Vázquez-Grenno, J.: "The ins and outs of unemployment in a two-tier labor market"
2012/18, González-Val, R.; Lanaspa, L.; Sanz, F.: "New evidence on Gibrat's law for cities"
2012/19, Vázquez-Grenno, J.: "Job search methods in times of crisis: native and immigrant strategies in Spain"
2012/20, Lessmann, C.: "Regional inequality and decentralization – an empirical analysis"
2012/21, Nuevo-Chiquero, A.: "Trends in shotgun marriages: the pill, the will or the cost?"
2012/22, Piil Damm, A.: "Neighborhood quality and labor market outcomes: evidence from quasi-random neighborhood assignment of immigrants"
2012/23, Ploeckl, F.: "Space, settlements, towns: the influence of geography and market access on settlement distribution and urbanization"
2012/24, Algan, Y.; Hémet, C.; Laitin, D.: "Diversity and local public goods: a natural experiment with exogenous residential allocation"
2012/25, Martinez, D.; Sjögren, T.: "Vertical externalities with lump-sum taxes: how much difference does unemployment make?"
2012/26, Cubel, M.; Sanchez-Pages, S.: "The effect of within-group inequality in a conflict against a unitary threat"
2012/27, Andini, M.; De Blasio, G.; Duranton, G.; Strange, W.C.: "Marshallian labor market pooling: evidence from Italy"
2012/28, Solé-Ollé, A.; Viladecans-Marsal, E.: "Do political parties matter for local land use policies?"

2012/29, Buonanno, P.; Durante, R.; Prarolo, G.; Vanin, P.: "Poor institutions, rich mines: resource curse and the origins of the Sicilian mafia"
2012/30, Anghel, B.; Cabrales, A.; Carro, J.M.: "Evaluating a bilingual education program in Spain: the impact beyond foreign language learning"
2012/31, Curto-Grau, M.; Solé-Ollé, A.; Sorribas-Navarro, P.: "Partisan targeting of inter-governmental transfers & state interference in local elections: evidence from Spain"
2012/32, Kappeler, A.; Solé-Ollé, A.; Stephan, A.; Välilä, T.: "Does fiscal decentralization foster regional investment in productive infrastructure?"
2012/33, Rizzo, L.; Zanardi, A.: "Single vs double ballot and party coalitions: the impact on fiscal policy. Evidence from Italy"
2012/34, Ramachandran, R.: "Language use in education and primary schooling attainment: evidence from a natural experiment in Ethiopia"
2012/35, Rothstein, J.: "Teacher quality policy when supply matters"
2012/36, Ahlfeldt, G.M.: "The hidden dimensions of urbanity"
2012/37, Mora, T.; Gil, J.; Sicras-Mainar, A.: "The influence of BMI, obesity and overweight on medical costs: a panel data approach"
2012/38, Pelegrín, A.; García-Quevedo, J.: "Which firms are involved in foreign vertical integration?"
2012/39, Agasisti, T.; Longobardi, S.: "Inequality in education: can Italian disadvantaged students close the gap? A focus on resilience in the Italian school system"

2013

2013/1, Sánchez-Vidal, M.; González-Val, R.; Viladecans-Marsal, E.: "Sequential city growth in the US: does age matter?"
2013/2, Hortas Rico, M.: "Sprawl, blight and the role of urban containment policies. Evidence from US cities"
2013/3, Lampón, J.F.; Cabanelas-Lorenzo, P-; Lago-Peñas, S.: "Why firms relocate their production overseas? The answer lies inside: corporate, logistic and technological determinants"
2013/4, Montolio, D.; Planells, S.: "Does tourism boost criminal activity? Evidence from a top touristic country"
2013/5, Garcia-López, M.A.; Holl, A.; Viladecans-Marsal, E.: "Suburbanization and highways: when the Romans, the Bourbons and the first cars still shape Spanish cities"
2013/6, Bosch, N.; Espasa, M.; Montolio, D.: "Should large Spanish municipalities be financially compensated? Costs and benefits of being a capital/central municipality"
2013/7, Escardíbul, J.O.; Mora, T.: "Teacher gender and student performance in mathematics. Evidence from Catalonia"
2013/8, Arqué-Castells, P.; Viladecans-Marsal, E.: "Banking towards development: evidence from the Spanish banking expansion plan"
2013/9, Asensio, J.; Gómez-Lobo, A.; Matas, A.: "How effective are policies to reduce gasoline consumption? Evaluating a quasi-natural experiment in Spain"
2013/10, Jofre-Monseny, J.: "The effects of unemployment benefits on migration in lagging regions"
2013/11, Segarra, A.; García-Quevedo, J.; Teruel, M.: "Financial constraints and the failure of innovation projects"
2013/12, Jerrim, J.; Choi, A.: "The mathematics skills of school children: How does England compare to the high performing East Asian jurisdictions?"

ieb@ub.edu
www.ieb.ub.edu

Cities and Innovation

Document de treball de l'IEB