WHY ISN'T THE WHOLE OF SPAIN INDUSTRIALIZED?
THE LOCALIZATION OF SPANISH MANUFACTURING DURING THE EARLY
INDUSTRIALIZATION (1797-1910)

Joan R. Rosés ⤵

Abstract
The regions of Spain provide a unique opportunity to study the causes of regional differences in
industrial development because manufacturing concentration and regional specialization rose
substantially during the early phases of industrialization. I employ a model nesting Heckscher-
Ohlin and Economic Geography Frameworks to study that phenomenon, and find that
comparative advantage and increasing returns effects were economically very significant and
practically explained all differences in industrialization levels across regions. The deficit
(surplus) of some regions in industrialization appears to have been largely attributable to their
factor endowments and the absence (presence) of home markets effects for industries with
increasing returns.

Keywords: Heckscher-Ohlin model; Increasing returns; Economic Geography.

† Rosés, Departamento de Historia Económica e Instituciones, Universidad Carlos III de Madrid.
E-mail: jroses@clio.uc3m.es

* I have benefited from the comments of Klaus Demset, Leandro Prados de la Escosura, James
Simpson, and seminar participants at Universidad Carlos III, Universidad Complutense, and
Fourth World Cliometrics Conference held in Montreal. This research has been supported by
Spanish Ministry of Education (Grant PB98-0031).
Economic historians have long been concerned with why within countries there has been regional variation in industrial development and why certain regions and countries have industrialized early in their histories. A substantial literature, for example, has suggested important links between variations in industrialization levels and conditions of local agricultural production. Particularly, scholars have stressed the relationship between the supply of labor in manufacturing and disparities in the size of landholding, in the types of crops, the productivity of females of children relative to that of adult males, the availability of landless workers, and seasonality of labor requirements in agriculture.¹ A second strand of the literature has argued for a strong link between the previous accumulation of human capital and the spread of industrialization across regions and countries.² Others seeking to comprehend why early manufacturing was more likely to surge as a major activity in certain regions rather than in others, tend to highlight how local culture and institutions influenced the paths of manufacturing development. In particular, the economic historian Gregory Clark suggested that workers in poor countries may have been inherently less efficient than their counterparts in rich countries due to local environmental and cultural forces.³ In a similar vein, Brian A’Hearn explained the failure of Italian South to industrialize in terms of cultural forces and institutions; so, he argued that “(Italian South) efforts to mobilize local capital in

² See, for example, Easterlin, “Why isn’t the Whole World Developed” and Sandberg “The case of the Impoverished Sophisticate”.
³ Clark “Why isn’t the Whole World Developed”.
support of industry were undermined by a lack of trust and an inability to cooperate." More recently, a new generation of economic historians has suggested explanations based on trade theory. Sukko Kim, for example, finds that the neoclassical Heckscher-Ohlin model of trade and production scale economies explain long-run trends in U.S. industrial regional structures.  

The regions of Spain provide a unique opportunity to study the causes of regional differences in industrial development during the early phases of industrialization. During the nineteenth century, the spatial concentration of Spanish manufacturing increased steadily. The contrast was particularly stark between the industrializing northeast and the agricultural center and northwest. By 1910 two northern regions, Catalonia and the Basque Country, concentrated a large part of Spain's industry. With only a sixth of Spanish workforce, these two relatively small regions were home to one third Spanish industrial workers and two-thirds of the employment in the modern industries (metallurgy, engineering and textiles). Simultaneously, price gaps among Spanish regions declined as the cost of transporting commodities fell and institutional barriers of home trade were eliminated; so, Spain progressed from being a set of regional economies to becoming an integrated national economy. Dramatic changes in the distribution of income among the Spanish regions accompanied this tendency toward industrial agglomeration and market integration. For instance, by 1800 Catalonia had nearly the average Spanish per capita income while in 1860 was the richest region exceeding in about the 20 percent the national average. The next 70

---

4 A'Hearn, "Institutions", p. 756.
5 Kim, "Geographic Distribution" and "Economic Integration".
6 Note that regional histories of Spain and Italy have close parallels (see, for example, on Italian experience A'Hearn, "Institutions"; and Zamagni, Economic History).
7 The data came from the Spain's Population Census (see appendix 1).
years saw continued divergence, with Catalonia reaching a peak of about 190 percent of national per capita income by 1930.\(^8\)

The question is, then, to explain how the integration of markets can generate simultaneously the concentration of manufacturing production and increasing income per capita divergence across regions. The constant-returns trade models, that is the Heckscher-Ohlin models,\(^9\) would explain the trends in the localization of production but it does not fit with some important details of Spanish economic history. They predict that, when trade cost decreased and factor markets integrated, factor-price-equalization forces should promote income per capita convergence across regions.\(^10\) Perhaps more interestingly, the new economic geography models appear to a way to explain these apparently contradictory historical facts.\(^11\) The rough intuition behind these models runs as follows. In the case of high transport costs (e.g., Spain during the eighteenth century), there was little interregional trade; so, manufacturing establishments were not concentrated. When the transport costs decreased, the interregional trade increased. In presence of imperfect competition, firms located in the region with the larger market. A regional division of labor spontaneously arises through a process of uneven development. Manufacturing concentrated in the regions with higher demand, while the rest of the country suffered deindustrialization. Low wages in the poor regions were not enough to attract manufacturing because of the lack of sufficient backward

---

\(^8\) Estimates of per capita regional GDP came from Alvarez Llano, “Estructura económica regional”.

\(^9\) See Flam and Flanders, Heckscher-Ohlin for a presentation of the theory. See also O’Rourke and Williamson Globalization and History for an application of Heckscher-Ohlin framework to history.

\(^10\) Note that, according to Heckscher-Ohlin Theory, only shocks in relative prices that favored some specialized regions over others might disrupt any income convergence process. However, through the nineteenth century, price shocks favored mainly agrarian regions because the price of manufactures experienced a downward trend compared with the price of agricultural goods (see Prados de la Escosura, “Output and Expenditure”).

and forward linkages. Therefore, economic geography models predict with market openness manufacturing concentration, and income divergence.

This paper employs these two models to account for the structure of manufacturing in Spain during the early industrialization. Following the recent works of Donald Davis and David Weinstein, a model that combines the Heckscher-Ohlin framework with a simple model of economic geography featuring "home market effects" is computed. Estimations suggest that comparative advantage and increasing returns effects were economically very significant and practically explained the localization of Spanish manufacturing during the early industrialization. Moreover, the results give support for the existence of economic geography effects in modern manufacturing industries. In other words, this paper finds strong support for a straightforward economic explanation of differences in industrialization levels among Spanish regions.

The remaining article is organized as follows. The first section of the article offers a brief discussion of the trends in market integration and regional specialization in Spain during the nineteenth century. The next section provides an overview of the interpretations given by the literature to several processes of regional industrialization in Spain. The model and data

12 See Davis and Weinstein "Market Access" and "Economic Geography"; and Davis, Weinstein, Bradford and Shimpo "Factor Abundance Theory". For theoretical justification of a model nesting Heckscher-Ohlin and increasing returns frameworks see Krugman "Increasing Returns", pp. 1245-1251.

13 One should expect to see much bigger effects of Heckscher-Ohlin and Economic Geography frameworks in Spain during the nineteenth century for three reasons. First, Ricardian (technological, institutional) differences among Spanish regions were likely to have been smaller than among countries. Second, Paul Krugman and Anthony Venables ("Globalization") have shown that the impact on localization of agglomeration forces is much bigger in absence of migrations. Precisely, in Spain during the nineteenth century internal and external migrations were smaller than in other European countries and urban and agrarian labor markets were segmented (see below). Third, the relative isolation of the Spanish economy from foreign shocks due to high tariffs reinforced the local forces easing manufacturing concentration.
issues are discussed in section III. Sections IV and V explore, respectively, the impact of factor endowments and economic geography forces on industrial localization. Conclusions and implications for further research are drawn in a final section.

I

During the course of the nineteenth century, Spanish regions went from a set of relatively independent regional economies to an integrated national economy. However, a detailed inspection of regional convergence in prices suggests that regional integration occurred at different rates for commodities and factor markets. Commodity market integration appears to have realized by the 1850s, capital markets by the 1880s, and labor markets by the early twentieth century. In other words, integration of factor markets progressed slower than integration of commodity markets.

The liberal reforms of the first half of the nineteenth-century laid a firm political foundation for economic integration of Spain by eliminating tariffs and local restrictions on home commerce and by ensuring free mobility of people and capital. This institutional progress was accompanied by major improvements in transport and communication systems. For example, the extension of paved roads increased exponentially from 2000 kilometers to 19,815 kilometers between 1800 and 1868. As a consequence of these improvements, Spain’s transport system changed from being based on pack animals to one using carts. At

---

14 Note that the regional interdependence of local prices of commodities was not a nineteenth-century novelty because during the eighteenth-century some market integration existed. See, for example, Hamilton War and Prices; and Ringrose Spanish Miracle.

15 On these liberal reforms see Tedde de Lorca, “Cambio Institucional”; and Simpson, Spanish Agriculture, pp. 84-87. However, this liberalization of the home market was not accompanied by a simultaneous liberalization of foreign imports because Spanish government used tariffs to mute the impact of foreign competition over the nineteenth century.

16 Madrazo, Sistema de transportes, pp. 163-179.

17 Madrazo, Sistema de transportes; and Simpson, Spanish Agriculture, pp. 80-87.
the same time, coastal shipping experienced major advances. So, in the pre-railway age, the signs of the emergence of a national transport system were clearly identifiable as regions were connected to some extent by roads, canals and coastal shipping. Between 1860 and 1890, national transportation and communications system strengthened as the railroads network were completed and telegraph mileage increased exponentially. With the railways, unit transport costs fell, permitting a widening of the market, growth in urbanization, and an increase in agricultural specialization. Both market liberalization and transport improvements induced price convergence among Spanish regions. Recent research on Spanish market integration tends to concur that the Spanish regions were integrated into a national market for basic foodstuffs by the 1850s, and almost all price convergence took place before 1850; that is earlier than the completion of Spain’s railways network.

An examination of regional convergence in short-term interest rates of commercial paper suggests that integration of capital markets seem to have been realized by the latter half of the nineteenth century. More specifically, commercial paper showed rapid convergence in prices across regions after 1850. By 1885 the Bank of Spain completed this process of integration when established the first nationwide branching allowing movements of capital among towns at constant and cheap rates.

The conclusion of the recent research that the Spanish labor market became better integrated from 1860 to 1914 represents somewhat a rupture from the previous literature. The few historians who previously have examined Spanish labor market before 1914 generally were struck by its apparent poor performance. These studies leave the strong notion, if only implicitly, that even if was some migrations, opportunities for arbitrage were not fully

---

18 Frax, *Comercio de Cabotaje.*
19 Gómez Mendoza, *Ferrocarriles.*
exploited. However, analyses of the pattern and extent of migration movements shed little light on the issue of integration. Markets could be perfectly integrated but exhibit little migration or they could exhibit high rates of migration but be poorly integrated. Evidence on agricultural wages clearly identifies a process of integration from 1854 to 1914 when internal and international migration was comparatively low in Spain. Coefficients of variation in agricultural wages declined from 0.26 to 0.18 exhibiting the typical behavior of convergence processes. By contrast, there would seem a widening wage-gap between urban and rural wages between 1860 and 1896, which then begins closing gradually. This would indicate that growing urban demand for labor was not met by appropriate rural migration. Thus, the integration of agrarian and urban labor markets at national level was far from complete.

How the economic structure of Spanish regions responded to this process of progressive market integration? To answer to that question, I assemble Krugman’s index of regional specialization that had been computed using eight macro-regions and one-digit employment levels (agriculture, industry and mining, and services). It shows that, with the reduction of transport costs and the progressive integration of the home market over the nineteenth-century, regional specialization rose substantially in Spain. The index was 0.204 in 1797, fell slightly to 0.200 in 1860 before rose steadily to 0.274 in 1910. Note that the movements in the aggregate index cannot be attributed to changes in a small amount of regions. If one looks in detail at table 1, it can be observed how the aggregate pattern is replicated in most biregional comparisons. In general, each region becomes more specialized compared with any other region between 1860 and 1910.

21 Castañeda and Tafunell, “Las letras de cambio”.
22 See, for example, Mikelarena “Los movimientos migratorios”.
23 See, on the causes of low migrations Sánchez-Alonso “Emigration from the Regions of Spain”.
Table 1

**KRUGMAN'S INDEX OF SPECIALIZATION, 1797-1910**

<table>
<thead>
<tr>
<th>Year</th>
<th>AND</th>
<th>ARA</th>
<th>CAT</th>
<th>NCA</th>
<th>SCA</th>
<th>MED</th>
<th>BAS</th>
<th>NOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1797</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AND</td>
<td>0.043</td>
<td>0.193</td>
<td>0.164</td>
<td>0.062</td>
<td>0.038</td>
<td>0.300</td>
<td>0.253</td>
<td></td>
</tr>
<tr>
<td>ARA</td>
<td>0.229</td>
<td>0.128</td>
<td>0.100</td>
<td>0.035</td>
<td>0.264</td>
<td>0.217</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAT</td>
<td>0.357</td>
<td>0.180</td>
<td>0.194</td>
<td>0.493</td>
<td>0.446</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCA</td>
<td>0.225</td>
<td>0.163</td>
<td>0.136</td>
<td>0.089</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCA</td>
<td>0.095</td>
<td>0.362</td>
<td>0.315</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MED</td>
<td>0.299</td>
<td>0.252</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAS</td>
<td>0.084</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average = 0.204</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>AND</th>
<th>ARA</th>
<th>CAT</th>
<th>NCA</th>
<th>SCA</th>
<th>MED</th>
<th>BAS</th>
<th>NOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1860</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AND</td>
<td>0.162</td>
<td>0.171</td>
<td>0.124</td>
<td>0.153</td>
<td>0.071</td>
<td>0.054</td>
<td>0.312</td>
<td></td>
</tr>
<tr>
<td>ARA</td>
<td>0.315</td>
<td>0.052</td>
<td>0.261</td>
<td>0.142</td>
<td>0.156</td>
<td>0.185</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAT</td>
<td>0.277</td>
<td>0.172</td>
<td>0.242</td>
<td>0.178</td>
<td>0.483</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCA</td>
<td>0.209</td>
<td>0.104</td>
<td>0.104</td>
<td>0.238</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCA</td>
<td>0.205</td>
<td>0.105</td>
<td>0.447</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MED</td>
<td>0.104</td>
<td>0.241</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAS</td>
<td>0.341</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average = 0.200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>AND</th>
<th>ARA</th>
<th>CAT</th>
<th>NCA</th>
<th>SCA</th>
<th>MED</th>
<th>BAS</th>
<th>NOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1877</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AND</td>
<td>0.151</td>
<td>0.127</td>
<td>0.201</td>
<td>0.108</td>
<td>0.074</td>
<td>0.111</td>
<td>0.410</td>
<td></td>
</tr>
<tr>
<td>ARA</td>
<td>0.278</td>
<td>0.051</td>
<td>0.157</td>
<td>0.077</td>
<td>0.206</td>
<td>0.259</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAT</td>
<td>0.329</td>
<td>0.217</td>
<td>0.201</td>
<td>0.170</td>
<td>0.537</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCA</td>
<td>0.208</td>
<td>0.128</td>
<td>0.257</td>
<td>0.208</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCA</td>
<td>0.139</td>
<td>0.049</td>
<td>0.416</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MED</td>
<td>0.142</td>
<td>0.336</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAS</td>
<td>0.465</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average = 0.215</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>AND</th>
<th>ARA</th>
<th>CAT</th>
<th>NCA</th>
<th>SCA</th>
<th>MED</th>
<th>BAS</th>
<th>NOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1887</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AND</td>
<td>0.168</td>
<td>0.294</td>
<td>0.186</td>
<td>0.066</td>
<td>0.064</td>
<td>0.090</td>
<td>0.407</td>
<td></td>
</tr>
<tr>
<td>ARA</td>
<td>0.463</td>
<td>0.041</td>
<td>0.155</td>
<td>0.105</td>
<td>0.259</td>
<td>0.239</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAT</td>
<td>0.481</td>
<td>0.352</td>
<td>0.358</td>
<td>0.226</td>
<td>0.702</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCA</td>
<td>0.173</td>
<td>0.139</td>
<td>0.277</td>
<td>0.221</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCA</td>
<td>0.103</td>
<td>0.126</td>
<td>0.394</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MED</td>
<td>0.154</td>
<td>0.344</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAS</td>
<td>0.498</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average = 0.253</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>AND</th>
<th>ARA</th>
<th>CAT</th>
<th>NCA</th>
<th>SCA</th>
<th>MED</th>
<th>BAS</th>
<th>NOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AND</td>
<td>0.067</td>
<td>0.353</td>
<td>0.171</td>
<td>0.098</td>
<td>0.043</td>
<td>0.264</td>
<td>0.306</td>
<td></td>
</tr>
<tr>
<td>ARA</td>
<td>0.419</td>
<td>0.105</td>
<td>0.132</td>
<td>0.041</td>
<td>0.330</td>
<td>0.240</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAT</td>
<td>0.525</td>
<td>0.294</td>
<td>0.396</td>
<td>0.121</td>
<td>0.659</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCA</td>
<td>0.237</td>
<td>0.129</td>
<td>0.435</td>
<td>0.134</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCA</td>
<td>0.114</td>
<td>0.198</td>
<td>0.371</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MED</td>
<td>0.306</td>
<td>0.263</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAS</td>
<td>0.570</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average = 0.261</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>AND</th>
<th>ARA</th>
<th>CAT</th>
<th>NCA</th>
<th>SCA</th>
<th>MED</th>
<th>BAS</th>
<th>NOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1910</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AND</td>
<td>0.145</td>
<td>0.386</td>
<td>0.151</td>
<td>0.091</td>
<td>0.054</td>
<td>0.283</td>
<td>0.298</td>
<td></td>
</tr>
<tr>
<td>ARA</td>
<td>0.530</td>
<td>0.031</td>
<td>0.161</td>
<td>0.091</td>
<td>0.428</td>
<td>0.154</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAT</td>
<td>0.537</td>
<td>0.369</td>
<td>0.440</td>
<td>0.120</td>
<td>0.684</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCA</td>
<td>0.167</td>
<td>0.114</td>
<td>0.434</td>
<td>0.147</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCA</td>
<td>0.115</td>
<td>0.267</td>
<td>0.315</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MED</td>
<td>0.338</td>
<td>0.244</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAS</td>
<td>0.582</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average = 0.274</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes and Sources: See appendix 1. AND: Andalusia; ARA: Aragon; CAT: Catalonia; NCA: Northern Castilia; SCA: Southern Castilia; MED: Mediterranean; BAS: Basque Country; NOW: Northwest.
It is also interesting to study how manufacturing responded to the integration and specialization of Spanish regions. This can be addressed by estimating Hoover's coefficients of localization. Hoover's coefficients are calculated using employment for nine pseudo-two digit manufacturing sectors and three benchmark years (1797, 1860, and 1910).^6

### Table 2

<table>
<thead>
<tr>
<th>Sector</th>
<th>1797</th>
<th>1860</th>
<th>1910</th>
</tr>
</thead>
<tbody>
<tr>
<td>Textiles</td>
<td>0.340</td>
<td>0.638</td>
<td>0.596</td>
</tr>
<tr>
<td>Cotton</td>
<td>0.343</td>
<td>0.733</td>
<td></td>
</tr>
<tr>
<td>Wool</td>
<td>0.445</td>
<td>0.582</td>
<td></td>
</tr>
<tr>
<td>Silk</td>
<td>0.486</td>
<td>0.711</td>
<td></td>
</tr>
<tr>
<td>Leather</td>
<td>0.255</td>
<td></td>
<td>0.328</td>
</tr>
<tr>
<td>Wood and furniture</td>
<td>0.246</td>
<td>0.264</td>
<td></td>
</tr>
<tr>
<td>Metal industry</td>
<td>0.155</td>
<td>0.268</td>
<td></td>
</tr>
<tr>
<td>Stone, Clay &amp; Glass</td>
<td>0.277</td>
<td>0.291</td>
<td></td>
</tr>
<tr>
<td>Chemicals</td>
<td>0.503</td>
<td>0.430</td>
<td>0.502</td>
</tr>
<tr>
<td>Food</td>
<td>0.296</td>
<td>0.176</td>
<td></td>
</tr>
<tr>
<td>Liquors</td>
<td>0.431</td>
<td>0.505</td>
<td></td>
</tr>
<tr>
<td>Apparel</td>
<td>0.239</td>
<td>0.167</td>
<td></td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>0.264</td>
<td>0.302</td>
<td></td>
</tr>
<tr>
<td>Paper</td>
<td>0.545</td>
<td>0.549</td>
<td></td>
</tr>
<tr>
<td>Unweighted average</td>
<td>0.286</td>
<td>n.a.</td>
<td>0.322</td>
</tr>
<tr>
<td>Weighted average</td>
<td>0.276</td>
<td>n.a.</td>
<td>0.283</td>
</tr>
</tbody>
</table>

**Notes and sources:** See appendix 2.

The unweighted and weighted average of Hoover's coefficient indicates that manufacturing became more localized during the nineteenth century; that is, early industrialization in Spain was associated with a process of industrial concentration. I also find that industries, in the aggregate, became more localized as regions became more specialized. However, the Hoover's coefficients at industry level illustrate major disparity in long-run trends across industries. Textile, metal, miscellaneous and leather industries became more regionally localized throughout the entire period. Other industries, such as food and apparel, became more regionally dispersed from 1797 to 1910. Still other industries, such as chemicals, stone, clay and glass, and wood and furniture, exhibited little change in localization during the nineteenth-century. Note that the two leading sectors of

---

^6 The choice of these years is determined by data availability.
industrialization (metal and textiles) were among those industries increasing their concentration levels. In the case of textiles, greater concentration appears to have been especially strong between 1797 and 1860 during early phases of the transition from cottage to factory production, and early phases of market integration.

II

In recent years there has been considerable empirical work on regional industrialization patterns in Spain, although this has been couched in term of region-based studies, rather than overall interpretations of the process. Interpretations on evolution of particular regions have been proposed based on social attitudes, geography, factor endowments, culture, entrepreneurship, governments’ industrial and commercial policies, demography, infrastructure, income distribution, capital scarcity, education, and agricultural institutions. These explanations are not in general mutually exclusive, the typical regional account combining several, and all are apparently plausible. Therefore, to get a better understanding of what drove historical interpretations, it seems necessary to examine more closely at some of the country’s regional histories. To do this, I will look at Catalonia, Castile, Andalusia and Galicia (in the Northwest), which have been chosen because of their different industrialization experiences. I shall examine each in turn.

As it has been mentioned above, Catalonia was the most successful industrialization in Spain during the nineteenth-century. Economic historians of the Catalan industrialization have in mind a combination of endowments and externalities models to explain this development. Most of them suggest a strong connection between agricultural growth and industrial growth in the area, a cumulative effect, growth of external economies and the subsequent development of Catalonia into an industrialized region. Thus, their view is that agriculture change made an important contribution to Catalan industrialization. The linkages include: transferring labor, making a net contribution to the capital for initial industrial
investment, helping the development of commercial networks and finally, providing a market for consumer goods for the emerging industrial sector. However, most of these functions were not performed by Catalan agriculture. Movements of labor from agriculture to modern industry were relatively small. Also, employment in the agriculture did not peak in absolute, and relative, terms until 1910, when it still accounted for about 50 percent of active population. Finally, it seems that Catalan agriculture did not liberate capital by reducing its demand for investment and, hence, did not make a net contribution of capital for initial industrialization.

The question of why there was so little industry in Castile is still subject to debate. In a provocative article, Nicolás Sánchez-Albornoz has alleged a generally conservative attitude to all change and improvement in Castile. Thus, he has argued implicitly that Castilians during the nineteenth century opted not to industrialize. More recently, Antonio Gómez Mendoza has developed a different view rejecting this argument explaining the failure to industrialize in terms of factor endowments and market integration. According to this author, giving the geographic and climatic conditions and the technology available in the nineteenth century, extensive dry farming was the most efficient type of cultivation in Castile. Under the influence of a rapid integration within the new trade circuits created by the improvement in transports and communications, the region’s comparative advantage increasingly laid with cereals, especially wheat, abandoning manufacturing. This natural tendency toward cereal

---

27 See, for example, Carreras, “Cataluña, primera”; Maluquer, “Revolución Industrial”; and Nadal, Cataluña, la fábrica de España.
28 Camps, Mercado de trabajo.
29 See, for example, Thompson (Cotton in Barcelona, pp. 96-145) for a discussion on the merchant origins of industrial capital.
farming was reinforced by the high tariffs protecting local producers from foreign competition.31

For the past generation, the dominant tendency in literature on nineteenth-century Andalusia has been to treat it in the context of such characterizations as “deindustrialization” and “failure”. Indeed, Jordi Nadal went so far as to title his two most famous articles on the subject “Industrialization and Deindustrialization in the Southeast of Spain” and “The Two Failures of the Industrial Revolution in Andalusia”.32 This is perhaps an extreme posture as the most recent treatments suggest a movement from “absolute failure” to “relative backwardness”. New quantitative evidence tends to support the view that Andalusia was more industrialized than the national average until the early twenty-century, but her industry grew slower than in Catalonia and the Basque Country.33 What diagnosis can be offered for the declining evolution of industry in Andalusia? Until a few years ago, most economic historians would probably have agreed with an explanation combining the inequitable land distribution and foreign dependency. According to the traditional account, as a consequence of the extreme economic inequity, the aggregate purchasing power of regional economy was extremely low and, despite a sizable population, the region presented a very limited market for any industry. Furthermore, the region depended on exports of agrarian commodities, which prices were declining, and foreign entrepreneurs dominated her major resources (mining and export agriculture). Economic historians labeled this state of affairs as a dependency trap.34 Research demonstrating that agrarian terms-of-trade improved during the

32 Nadal, “Industrialización y desindustrialización” and “Los dos abortos”; see, also, Morilla, “la industria andaluza”.
33 See, Martín Rodríguez, “Industrialización interrumpida” and Parejo, “La producción industrial de Andalucía”.
34 See, for example, Delgado Cabeza, Dependencia y marginación.
In consequence, recent interpretations tend to support the view that local forces are to be made responsible for Andalusia’s deindustrialization. Explanations have been based on superabundance of labor, uneven income distribution, poor regional integration, the negative outcomes of protectionism, a relative scarcity of human capital and entrepreneurs for industry, and low agrarian productivity.36

In the second half of the eighteenth century Galicia enjoyed an important textile industry.37 The character of this industry observed the main rules of Frank Mendels’s proto-industrialization model.38 It was predominantly a rural industry in which industrial work represented by-employment (although there were some full-time participants), production was for distant markets, there were some accumulation of capital, the commercial (but not the production) side of the trade was dominated by merchants and the existence of industrial employment stimulated population growth. However, there was not a direct movement from proto-industry to factory-based industry. During the first half of the nineteenth century, Galicia lost their export markets to foreign competition, and then factory-made textiles from Catalonia and Britain took over the local market. Economic historians explained the crisis of local manufacture as a combination of the absence of agricultural change, the organizational limitations of proto-industry and entrepreneurial failure. In particular, local entrepreneurs have been blamed because they preferred to invest capital in agriculture rather than to invest in factories and regional industrialization.39

35 This evidence was collected by Prados de la Escosura, De Imperio a Nación, Ch. 5; and Simpson, Spanish Agriculture, pp. 90-98.
36 See, Martín Rodríguez, “Industrialización interrumpida”; and Tedde de Lorca, “Subdesarrollo andaluz”.
38 Mendels, “Proto-industrialization”.
39 Carmona, Atraso industrial.
Several lessons can be obtained from this brief survey. Each of these regional histories provides a structured way of thinking about the issue, and so an interesting window on the determinants of industrialization. Incrementally they help to narrow the range of alternatives that may reasonably be contemplated. In spite of this, most of these studies suffered from two major shortcomings. First, they are adhered to the idea that a regional agricultural revolution should come before any successful regional industrialization and that agriculture should contribute to industrial development. However, the processes of structural transformation and resource transfer were more intricate and less strictly unidirectional than is habitually depicted by these economic historians. More to the point, agriculture in many parts of Spain experienced an extraordinary variety of advances in the direction of intensification, mixed farming and specialization but not all of these regions industrialized. Second, they also assumed that each region and each industrial sector had an independent history that could not be simplified by application of economic theory. One consequence is that it is impracticable to link all potentially relevant hypotheses for a single critical test. On the other hand, scholars wishing to employ more general explanations have also been guilty of a certain laxity with regards the proof of their case often insisting that what was observed must have been optimal, and any observable differences in industrialization levels were due to differences in economic forces at work. Therefore, no study has been made to compute the impact on the localization of nineteenth-century Spanish industry of factor endowments and increasing returns.

III

As mentioned above, there are two principal theories of why manufacturing activities concentrate: comparative advantage (Heckscher-Ohlin models) and increasing returns to scale (economic geography models). Comparative advantage holds that specialization occurs to take advantage of inherent differences. Increasing returns says that industry concentration
arises to take advantage of scale and variety gains from specialization. Perhaps most interestingly, these two apparently competing theories can be linked to obtain robust empirical calibration of the determinants of industrial localization.

**Methodology**

Under the usual "2 x 2" Heckscher-Ohlin model, with equal numbers of goods and factors, the output $X$ of the good $g$ in industry $n$ in province $p$ is given by: $^{40}$

$$X_{np}^g = \Omega_p^a V^p,$$

where $\Omega_p^a$ is the inverse of the technology matrix mapping output into factors, and $V^p$ is the vector of endowments of province $p$. In this framework, endowments will be sufficient to decide the structure of goods production.

However, under a model of monopolistic competition, endowments will not suffice to determine the output of goods within industries. $^{41}$ Consequently, Donald Davis and David Weinstein assume output structure is determined in two stages. The endowments (Heckscher-Ohlin framework) determine the broad industrial structure of a province (e.g., if it produces rails or shoes), but they tell us nothing about the composition of production across the goods within an industry (e.g., if it produce cars or trucks). Thus, the localization of goods production within industries is determined by the economic geography specification. $^{42}$ They define two variables to deal with this economic geography specification, SHARE and IDIODEM. The first variable measures overall commitment of the province to the

$^{40}$ Davis and Weinstein ("Market Access", pp. 14-15) give the following definition of goods and industries: "Under the hypothesis of increasing returns, a good is a collection of a large number of varieties produced under monopolistic competition (....) By contrast, under the hypothesis of comparative advantage, a good is a traditional homogeneous commodity. Industries, in both frameworks, consist of a collection of goods produced using a common technology".

$^{41}$ Krugman, "Increasing Returns".

$^{42}$ See Davis and Weinstein ("Economic Geography") for a more detailed discussion of the analytics.
encompassing industry and to the importance of that good in the aggregate within that industry. The second variable, IDIODEM, measures demand deviation (idiosyncratic demand) for a good in a province relative to all provinces together. Algebraically:

\[
SHARE^{np}_g = \frac{X^{nRSP}_g}{X^{nRSP}} X^{np},
\]

\[
IDIODEM^{np}_g = \left(\frac{D^{np}_g}{D^{np}} - \frac{D^{nRSP}_g}{D^{nRSP}}\right) X^{np},
\]

where \(D\) denotes absorption in the province, \(P\), or the rest of Spain, \(RSP\). At this point, many readers have probably noted that the variable IDIODEM is central in the issue of this paper.

In a world with decreasing returns, strong domestic demand for a good will tend to make it an import rather than an export.\(^{43}\) Instead, in a world with increasing returns and transport costs local demand should be important. There are several plausible explanations to this phenomenon but all lie in the realm of increasing returns world. One is directly related to the pure effect of market size. Due to economies of scale, each differentiated good is produced in only one place, and put on the market in both. If the cost of production is equal, then the deciding factor in localization is transport costs because, obviously, total transport costs are lower if production takes place in the region with the larger market. Inversely, if transport costs are high or if the markets of both regions had the same size, full concentration of the increasing-returns industry in the larger region will no longer take place.\(^{44}\) A straightforward extension of this model is that, with market integration, small countries will lose their industry producing differentiated goods.\(^{45}\) However, Donald Davis shows that, in

\(^{43}\) Krugman, "Scale Economies", p. 955.

\(^{44}\) Helpman and Krugman, Foreign Trade.

\(^{45}\) This is the so-called Linder hypothesis (see Linder, Trade and Transformation). This does not imply that welfare might decline if a region saw its increasing-returns sectors shrink as a result of trade. Krugman ("Increasing Returns", pp. 1254-1255) demonstrates that the gains from trade come up both
the case in which differentiated and homogeneous goods have identical transport costs, the home market effect disappears. Since it seems difficult to argue that transport costs may be unusually high for differentiated goods, the pure market size effect seems quite implausible. Therefore, a second explanation is that nonconventional transport costs may be higher for differentiated goods than homogeneous goods. If these costs are large and do show such a bias, the home market effect may come back.\textsuperscript{46} Finally, a third explanation suggests that quasi-Ricardian technical differences based on market size may arise if there are increasing returns in the production of intermediaries (backward and forward linkages). If these exhibit a sufficient strong bias towards production of differentiated final goods, then markets integration may yet lead to concentrate manufacturing.\textsuperscript{47}

However, one cannot postulate that endowments play no role in the location of goods production and, hence, it is necessary to estimate a model linking Heckscher-Ohlin framework (endowments) and economic geography specifications (SHARE and IDIODEM). Thus, Donald Davis and David Weinstein propose the following system of equations: \textsuperscript{48}

\[
X_{np} = \alpha_{g} + \Omega_{g} V + \varepsilon_{np} \\
X_{np} = \alpha_{g} + \beta_{SHARE_{g}} + \beta_{2.IDIODEM_{np}} + \varepsilon_{g}
\]

The key to establish the economic geography effects is the coefficient of IDIODEM ($\beta_{2}$), for which they identify three hypotheses. In a comparative advantage world without

\textsuperscript{46} Davis, “The Home Market”.

\textsuperscript{47} Krugman and Venables, “Globalization”. This is the same argument of the Marshallian externalities (labor thin markets and technological spillovers). There is a large literature on this issue see, among others, Marshall, \textit{Principles}; Pred, \textit{Spatial Dynamics}; and Henderson, \textit{Urban Development}.

\textsuperscript{48} It should be noted that this model simplify the geography implicit in Krugman (“Scale economies”) since it assumes that all regions had the same market access. However, this assumption has plausible small effects on my calculations because the differential market access is less pronounced for regions in a single country than for countries across the globe.
transaction costs, where factor endowments suffice to decide production, the localization of demand should have no effect on production structure, so $\beta_2 = 0$. Instead, in a friction world with comparative advantage, the geographical localization of demand does be relevant. If the local response to idiosyncratic components of demand is at most one-to-one (so $0 < \beta_2 < 1$), we are in a comparative advantage world with transaction costs. Finally, in case of economic geography, the response of local producers to idiosyncratic components of demand should be more than one-to-one, so $\beta_2 > 1$.

**Data Issues**

This article takes provinces in the European Spain as geographic unit of analysis. Spanish provinces are mid-sized because they are bigger than French Departments or British Counties, but smaller than US States. It should be noted that the definition of Spanish provinces in 1833 was not based on geographic criteria rather the Spanish government created provinces following pre-existing historical divisions and the areas of influence of major towns. For that reason, Spanish provinces adjust quite well to the functional integration principle that defines regions by the presence of a nucleus and the corresponding area of influence.\textsuperscript{49}

\textsuperscript{49} If one wants to test a model of increasing returns, the unit-of-analysis should be defined in such way. Instead, these kind of analysis units poses problems to test the Heckscher-Ohlin framework, where factors should be mobile within regions but less so across regions (see, for a discussion of the question Kim, “Expansion of markets”, p. 884). However, in mid-nineteenth century, it seems a minor problem since capital and labor movements were overwhelmingly more important within than among Spanish provinces.
The data required for the analysis include sectoral manufacturing output, endowments, technology, and consumption data for 43 of 48 European provinces in Spain for 1861. Five provinces (Alava, Guipuzcoa, Navarra, Orense and Vizcaya) due to the incompleteness of their data have been eliminated from the sample. The choice of year was mainly driven by data availability but this does not appear to be a major problem since, as table 2 shows, manufacturing concentration was quite high in 1861, and a rapid decrease in conventional and nonconventional transport costs happened in the previous thirty years.

Provincial-level data on sectoral output of 16 manufacturing sectors in 1861 was taken from Gimenez Guited's book. The numbers of different categories of labor were entered by province directly from the Population Census of 1860 and then summed to get labor type totals: skilled (clerks, public servants, professions, and commerce), artisans, unskilled (laborers, building, transport, miners, poor, servants, and factory workers), and agrarian labor. The capital stocks by province were derived from income taxes in 1860. Spain's statistical yearbook gives taxes paid by rents of housing, agrarian equipment, livestock,

---

50 This is a concept closely allied to gross output since it is the gross output less intra-industry transactions. See, Ciccone and Hall ("Productivity and Density", p. 60) for a detailed discussion of the advantages of that output concept in regional studies.

51 To avoid simultaneity, 1859-1860 data had been used to compute endowments and other right-hand side variables.

52 Gimenez Guited, Gula Fabril.

53 In empirical literature, some authors employed education-based classifications of labor while others preferred occupational-based classifications (see, for example, Trefler, "Missing Trade"). This occupational-based classification is probably preferable to the educational-based classification for two reasons. The first is that the 1860 census did not recorded education levels, but literacy. Second, as a large literature has pointed out, literacy is a very poor proxy for skills in mid-nineteenth century. However, it should be noted that occupational-based classifications are likely to be less exogenous with respect to output shares than occupational classifications.

54 I do not claim that my measure of capital and land reflects exactly the real amount of capital and land by 1860. However, I am convinced that they are efficient proxies; that is, errors are normally distributed and they are orthogonal to the "real" (unknown) variable.
commerce and industry for each province in 1860. These taxes were used to assign capital rents for each province and kind of capital good in 1860 using tax rates. Then, capital rents were used to impute capital stock levels using interest rate (6 percent), provincial capital goods price deflators and rates of depreciation, which are different for each kind of capital-good. The next problem was how to estimate the amount of land in each province. There are two alternatives to estimate land: (1) to use the quantity of hectares of cropland and pasture and (2) to use land taxes to impute land values. It is hard to see how the gross amount of land could be used for any purpose in Spanish economic history because soil quality, rainfall, crops, and productivity differed strongly from province to province. For that reason, the choice was for using land rents. Spain's statistical yearbook provides information on the amount of provincial land rents. These rents were used to impute land stocks for each province in 1860, using interest rate (6 percent) and provincial land quality-adjusted deflators.

The construction of the consumption data was quite complex. In broad terms, three types of consumption goods were used in this study: intermediate (cotton yarn, wool yarn, silk yarn and metal goods), taxed (olive oil, liquors, cork, paper and soap) and duty-free goods (flour, leather, and textiles). The consumption of intermediate goods for each province was computed using data from Gimenez Guited's book. However, given the absence of direct figures in the provincial consumption of metal goods, this was imputed from the figures on machinery stocks by province. For taxed goods, Spain's statistical yearbook provides the

55 Anuario estadístico de España.
56 The provincial deflators for housing and livestock were taken from Rosés, Carmona and Sánchez-Alonso "Wage Convergence".
57 Anuario estadístico de España.
58 These deflators were taken from Rosés and Carmona "Land prices".
59 Gimenez Guited, Guia Fabril.
quantities consumed by each province capital. Then, provincial totals were computed under
the assumption that province per capita consumption of each good corresponded to the pattern
of her respective capital. Finally, provincial data on household consumption, which broke
household consumption in several categories, was enough to compute figures for the
provincial consumption of flour, leather, paper and textiles. It should be noted that all these
commodities (intermediate, taxed and duty-free) were valued at producer prices because
without this adjustment the data would have greatly underestimated final consumption of
goods. Finally, all provincial totals were scaled so that the 49-provinces total for each
commodity exactly matched the total Spanish consumption.

A proxy for the Provincial Gross Domestic Product was estimated for all provinces in
the sample. The provincial GDP was computed by adding up land and capital rents,
depreciation, and wages (including the remuneration of autonomous labor). Complete
provincial data on unskilled wages was combined with data on the amount of workforce and
skills premium by workers category to get employees remuneration. Moreover, the
remuneration of labor for entrepreneurs was imputed as the remuneration of skilled labor.
Land, livestock and capital rents were imputed from taxes paid on land, livestock, industry
and commerce rents, respectively. Depreciation rates for each type of good were also
imputed from contemporaneous references. Finally, an PPP-adjusted deflator deflated the
resulting current GDP figures.

60 Anuario estadístico de España.
61 Rosés, Carmona and Sánchez-Alonso, “Wage Convergence” furnishes that data on provincial
household consumption.
62 Data on Spanish consumption from Prados de la Escosura, “Output and Expenditure”.
63 The data in unskilled wages came from Rosés, Carmona and Sánchez-Alonso, “Wage convergence”
and data on skills premium came from US Congress, Labor in Europe, pp. 1345-1441.
64 Anuario estadístico de España.
65 The deflator came from Rosés, Carmona and Sánchez-Alonso, “Wage Convergence”.

21
IV

In order to investigate the impact of Heckscher-Ohlin framework on the localization of the production let me to compute several regressions using production as dependent variable and endowments as independent variables. There are, however, several caveats that need to be addressed before to proceed further. First, according to the literature, it should be controlled for differences in technology considering that endowments do not matter at the finer levels of aggregation. Consequently, the Heckscher-Ohlin framework for the whole manufacturing sector and separately by three different aggregates, which pooled several manufacturing sectors, was estimated. To combine the different sectors, the choice was for the pragmatic method of considering aggregation based on capital-labor ratios. The resulting aggregates are: (1) Metal industry, Cotton Spinning, Wool Spinning, Silk Spinning and Paper; (2) Flour mills, Textiles, and Leather; and (3) Olive Oil Refining, Liquors Distilling, Soap and Cork Manufacturing. Second, we are confronted with the problem that there are two sources of heteroskedasticity in these data: the size of both provinces and industries. Hence, errors are likely to be correlated with the size of both regions and industries. There are several alternatives to correct this problem but this article follows Edward Learner's method. Thus, it was postulated the form of heteroskedasticity as arising from the following stochastic process:

\[
\text{var}(e^{\alpha p}) = v_g \text{GDP}_p
\]  

---

66 See Trefler, “Factor Price Differences” and “Missing Trade”.
67 The source for capital-labor intensity ratios is Gimenez Guited, Guia fabril.
68 Learner, Comparative Advantage. Furthermore, I test some alternative procedures for correcting heteroskedasticity without significant changes in my results. Note that estimation in logs is not possible because some observations are zero.
where \( v \) is a parameter. More specifically, first, it was estimated the equations by OLS and generated the squared residuals; second it was regressed that series on the provincial GDP and then employed the fitted values to form my weighting series for heteroskedasticity correction.

| Table 3 |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Dependent Variable | (1) Manufacturing | (2) Aggregate 1 | (3) Aggregate 2 | (4) Aggregate 3 |
| Constant | 16830 (19973.) | -5635 (10189.) | 9507 (13003.) | 7256 (3566.3) |
| Capital | 0.1853 (0.0459) | 0.0027 (0.0274) | 0.1547 (0.0290) | 0.0653 (0.0223) |
| Land | -0.0792 (0.0297) | -0.0229 (0.0163) | -0.0581 (0.0190) | -0.0039 (0.0061) |
| Skilled | -17.8323 (3.9299) | -6.2651 (2.3135) | -11.9509 (2.4876) | -4.7513 (1.7426) |
| Artisans | 3.0440 (1.5370) | 1.1024 (0.8620) | 1.7188 (0.9805) | 1.2992 (0.3333) |
| Unskilled | 2.0048 (0.6596) | 1.0963 (0.3698) | 1.4007 (0.4213) | 0.2020 (0.1587) |
| Agrarian | -0.2861 (0.1374) | -0.0378 (0.0779) | -0.3105 (0.0878) | -0.1284 (0.0395) |
| F-statistic | 47.63 | 11.86 | 62.59 | 9.04 |
| Adjusted R² | 0.8695 | 0.6081 | 0.9001 |
| # Observations | 43 | 43 | 43 | 43 (41) |
| Method | WLS | WLS | WLS | RREG |

Notes: WLS: Weighted Least Squares. RREG: Robust regression. Standard errors are in parenthesis. Adjusted R² is not plausible in RREG.

Table 3 reports the results of the different estimates. They are highly significant and reasonable. The adjusted R² and F-statistic are relatively high in manufacturing and aggregates 1 and 2, but not in aggregate 3. This suggests that even after controlling for size based variation, endowments explain almost the 85 percent of the variance of aggregate manufacturing output. They also explain about the 60 percent in the highest capital-labor ratios industries, which formed the aggregated 1, and about the 90 percent in the case of the high consumption sectors, which formed the aggregated 2. Instead, HO framework fails to explain the localization of the traditional industries with the lowest capital-labor ratio, which formed aggregated 3. However, these industries are quite reasonable candidates for an explanation based on natural advantage rather than comparative advantage. For example, olive oil refining industry was surely affected by the suitability of provinces’ climates for growing olives. The sign and coefficients of the different variables also lie in what one can expect. Provinces well endowed with capital, artisans and unskilled labor had comparative advantage in manufacturing, while the contrary holds for provinces well endowed in land,
agrarian and skilled labor. It should be noted that the sign of agrarian labor confirms the idea of studies on labor market integration that agrarian workers were not recruited in factories during the early industrialization.\textsuperscript{69} Perhaps, the sign of skill labor is puzzling for many readers but it should be noted that during early industrialization very few formal educated workers were recruited to work into the new factories.\textsuperscript{70} Instead the availability of artisans was crucial for their development.

For seven of eight macro-regions (the Basque Country is excluded due to data incompleteness), table 4 reports the impact on industrialization levels of factor endowments. To assist interpretation, the last column of the table indicates the predicted industrialization levels and the panel B the deviations from the Spanish norm. Some regions have industrialization levels quite similar to that of Spain as a whole, and the exercise here usually yields little information. The regions with industrialization levels far above or below the national are of course of most interest.

\textsuperscript{69} Simpson, \textit{Spanish Agriculture}, pp. 195-201; and Camps, \textit{Mercado de trabajo}.

\textsuperscript{70} It is also likely that the skills coefficient includes some congestion forces.
The exercise in table 4 is designed to reassess previous interpretations on the process of regional industrialization. It reveals that Catalonia had a relatively high industrialization level due to the relatively scarcity of land, skilled labor, agrarian labor, and the relative abundance of artisans. Instead, the alleged two major contributions of agriculture to industrialization (unskilled labor and capital) were minor players in explaining the Catalan exceptionality. In a sharp contrast with Catalonia, the relative low levels of Castilia are largely explained by abundance of land and skilled labor. The evidence on Andalusia also contrasts strongly with the previous literature since she was relatively poorly endowed in unskilled and artisan labor and not particularly well endowed in agrarian labor; therefore, her relatively low industrialization levels were mainly consequence of the absence of unskilled and artisan workforce. Finally, the evolution of Galicia in the Northwest can be explained by
appealing to her capital scarcity and superabundance of agrarian labor; consequently, it is
difficult to argue for entrepreneurial failure.  

V  

As mentioned above, to compute economic geography effects one can estimate
equation 3. This equation can be estimated at various levels of aggregation, as linear
specification or as a system of seemingly unrelated regressions (SUR), by separating
endowments in one equation and economic geography specifications in other. In both cases,
one should use endowments as instruments and correct for heteroskedasticity (see above).

| TABLE 5 |  
| ECONOMIC GEOGRAPHY EFFECTS ON PRODUCTION:AGGREGATES | 
| Dependent Variable | Manufacturing | Aggregate 1 | Aggregate 2 | Aggregate 3 | 
| IDIODEM | 1.3241 | 1.1910 | 1.9701 | 1.7594 | 1.3748 | 1.2144 | 0.2626 | 0.2230 | 
| SHARE | 0.8918 | 0.8628 | 0.3439 | 0.5963 | 1.0744 | 0.9639 | 1.1024 | 1.0066 | 
| | (0.0207) | (0.0173) | (0.0299) | (0.0294) | (0.0453) | (0.0301) | (0.0263) | (0.0244) | 
| Endowments | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| F-statistic | 448.66 | 1697.74 | 141.22 | 652.94 | 174.82 | 680.71 | 229.64 | 859.51 |
| Adjusted R2 | 0.9089 | 0.8828 | 0.8933 | 0.8752 | 0.9157 | 0.9221 | 0.9145 | 0.9181 |
| # Observations | 516 | 516 | 215 | 215 | 129 | 129 | 172 | 172 |
| Method | WLS | SUR | WLS | SUR | WLS | SUR | WLS | SUR |

Notes and sources: See text. Standard errors are in parenthesis. In SUR estimation statistics
corresponded to the economic geography equation.

Table 5 reports estimates at the higher levels of aggregation. What is striking in these
estimates is the fact that fits of the regressions (in any method) are quite high. Indeed, it
shows that at the end of the early phase of Spanish industrialization, factor endowments with
economic geography specifications can explain about the 90 percent of the variation in these
aggregates. The results also bolster our confidence on the economic geography effects. Of the
four aggregates computed, idiosyncratic demand is significantly larger than one in three,

Note that the regions of Galicia and Asturias compose Northwest region, the later highly
industrialized by 1860. For that reason, aggregation of both regions produces high levels of
industrialization.
whereas closed to zero only in one (aggregate 3), which corresponds to traditional industries where natural advantages were very important. This implies that the movements in local demand produce more than proportionate movements in production.

One can run an additional robustness check to confirm that it is identifying economic geography effects here. The results above are, obviously, very sensitive to the aggregation scheme. For example, one may dilute the real effects if there are combined industries with constant returns with industries with increasing returns. One simple solution to that problem is to compute the equation for each single industry. The table 6 shows the results from that kind of exercise.

TABLE 6

<table>
<thead>
<tr>
<th>Industry</th>
<th>Economic Geography</th>
<th>IDIODEM</th>
<th>SHARE</th>
<th>F-statistic</th>
<th>Adj. R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cork manufacturing</td>
<td></td>
<td>1.3335 (1.0684)</td>
<td>0.2380 (0.9053)</td>
<td>0.43</td>
<td>0.1224</td>
</tr>
<tr>
<td>Cotton Spinning</td>
<td>Yes</td>
<td>2.1025 (0.0877)</td>
<td>1.2721 (0.0809)</td>
<td>721.96</td>
<td>0.9928</td>
</tr>
<tr>
<td>Flourmills</td>
<td>Yes</td>
<td>1.8567 (0.1675)</td>
<td>1.4305 (0.1064)</td>
<td>33.21</td>
<td>0.8599</td>
</tr>
<tr>
<td>Leather</td>
<td></td>
<td>0.4959 (0.2173)</td>
<td>0.1171 (0.1965)</td>
<td>22.04</td>
<td>0.8003</td>
</tr>
<tr>
<td>Metal Industry</td>
<td>Yes</td>
<td>2.4371 (0.1245)</td>
<td>0.5010 (0.0583)</td>
<td>218.65</td>
<td>0.9764</td>
</tr>
<tr>
<td>Olive Oil</td>
<td></td>
<td>0.3316 (0.2719)</td>
<td>1.1771 (0.0460)</td>
<td>93.45</td>
<td>0.9463</td>
</tr>
<tr>
<td>Paper</td>
<td></td>
<td>-0.1675 (0.1308)</td>
<td>-0.0009 (0.1410)</td>
<td>5.41</td>
<td>0.4567</td>
</tr>
<tr>
<td>Silk Spinning</td>
<td></td>
<td>0.5447 (0.0712)</td>
<td>0.6653 (0.1436)</td>
<td>64.40</td>
<td>0.9252</td>
</tr>
<tr>
<td>Soap</td>
<td></td>
<td>0.1606 (0.1929)</td>
<td>0.3133 (0.1894)</td>
<td>2.80</td>
<td>0.2554</td>
</tr>
<tr>
<td>Spirits</td>
<td></td>
<td>-0.2076 (0.1703)</td>
<td>0.4346 (0.1715)</td>
<td>2.66</td>
<td>0.2404</td>
</tr>
<tr>
<td>Textiles</td>
<td>Yes</td>
<td>1.4967 (0.0893)</td>
<td>0.3981 (0.1362)</td>
<td>614.15</td>
<td>0.9915</td>
</tr>
<tr>
<td>Wool Spinning</td>
<td>Yes</td>
<td>1.8690 (0.2848)</td>
<td>1.1322 (0.1681)</td>
<td>75.12</td>
<td>0.9339</td>
</tr>
</tbody>
</table>

Notes and sources: The number of observations is 43 in all industries. Standard errors are in parenthesis. Estimations performed by WLS. Results obtained by SUR (not reported in the table) are practically identical.

The economic geography effects are significantly larger than unity for five of twelve sectors: cotton spinning, flour mills, metal industry, textiles, and wool spinning. Furthermore, the results are robust to whichever method of estimation one chooses. Moreover, with

The β-coefficients of IDIODEM (not reported in the table) for these five industries also show very statistically significant economic geography effects. They are typically over the 0.8 range (3.45 in flour mills, 0.91 in cotton spinning, 0.84 in metal industry, 0.81 in textiles, and 0.81 in wool spinning). A one standard deviation movement in idiosyncratic demand on average moves production by more

27
perhaps the exception of flourmills, all of the industries with significant economic geography effects seem like plausible candidates for monopolistic competition. For example, Paul Krugman identified metal industry and textiles as canonical examples of industries where backward and forward linkages are important. It is also interesting to note the similitude among my estimates and the estimates of Donald Davis and David Weinstein for Japan and OECD countries. They obtain significant economic geography in about the 42 percent of Japan industries and in about the 34 percent of OECD industries whereas I obtain this result in about the 42 percent of cases. These coincidences obviously bolster the confidence on my results.

Another way to obtain a sense of how economic geography is to industry production in Spain during the early industrialization is to examine the relative size of the sectors for which I computed larger than one IDIODEM coefficients. Using the value-added shares computed by Leandro Prados de la Escosura, it can be estimated that my 12 sectors corresponded to about the 82 percent of value added in industry by 1860. Moreover, the sectors with coefficients on IDIODEM exceeding unity account for 55 percent of the value added in industry. This obviously show that the sectors that seem to have home market effects account for a majority of industrial output in Spain during the early industrialization.

VI

This article offers new insights concerning the causes of the industrialization of Spain's regions by taking a firmly-based theoretical approach. The fortunes of a region are than 0.8 standard deviation. In other words, observed fluctuations in idiosyncratic demand provide a lot of information on production patterns.

73 Krugman, Geography and Trade.

74 Davis and Weinstein "Market Access" and "Economic Geography".

75 Prados de la Escosura, De Imperio a nación,
assumed to depend not only upon its own endowments but also on the market size effects. By contrast, many economic historians take a different approach and study the fortunes of a region by assuming that they cannot be explained with a general law. From the previous pages, a common pattern emerges: regions industrialized or failed to do so according to their comparative advantage (Heckscher-Ohlin framework). More specifically, the excess of some regions in manufactures appears to have been largely attributable to the relative high levels of artisans, unskilled labor and capital. Instead, the deficit of the remaining regions may be explained by the relative abundance of land, skilled labor, and agrarian labor.

However, endowments are not enough to explain the full history. Why did modern manufacturing concentrate in some regions while in others, in spite of their comparative advantage in manufacturing, have few modern factories? The explanation lies in the fact that modern industries that produced heterogeneous goods experienced monopolistic competition and increasing returns. Consequently, they tended to be concentrated in regions in which the home market effects were larger.

What really caused these home market effects? It is premature to attempt to answer this here, but a few observations might be ventured. Home market effects can arise from simple market-size scale-economies or the much more sophisticated Marshallian externalities (backward and forward linkages). Many manufacturing sectors with increasing returns are the typical sectors with backward and forward linkages. Consequently, it is much more plausible that Marshallian externalities were more important for provoking home market effects than pure market-size effects. Anyway, these are simple observations that require much greater analysis and empirical study.

These findings immediately suggest the appeal of some sort of evolutionary interpretation of the concentration of the Spanish manufacturing during the nineteenth century. In the eighteenth century, due to high transport and transaction costs, there is little
interregional trade. Thus, regions with high population densities had lower wages and, hence, lower income per capita. In the first half of the nineteenth century, transport cost and transaction cost decreased. Trade among regions increased and regional goods markets integrated in Spain. At this point, manufacturing became increasingly concentrated. This concentration of manufacturing production arose both from comparative advantage and from the additional external economies. Increasing returns were highly relevant in the new modern manufacturing industries, which produced heterogeneous goods, although negligible in traditional industries. Few regions with a large comparative advantage in manufacturing, like Catalonia, also benefited from gains from external economies. Poor regions did not converged in income per capita with these industrialized regions by two reasons. First, because increasing-returns industries did not flow to poor regions. Second, because their agrarian labor did not migrate to the new rich regions and, therefore, their production structures did not converge with those of rich regions.

Finally, I would like to underline four broad conclusions. First, the article’s findings suggest the need for rethinking the relationship between agricultural revolution and regional industrialization. I believe that the simple conventional view of successful industrialization as mainly determined by the success of a previous agricultural revolution is a misleading point of reference. Second, the result of this study also calls into question the often-assumed links between capital accumulation and industrialization. Third, from the point of view of economic theory, this paper has checked the explanatory power of the Hecksher-Ohlin theory and economic geography in economic history. In particular, the hypothesis of the importance of home market effects in determining the localization of production is fully confirmed by this research. Fourth, I would argue that the spatial concentration of modern manufacturing production in a few regions throughout early industrialization did not imply a loss of per capita income in the rest of Spain. Quite the contrary, I subscribe a rather benign view of
trade under comparative advantage and increasing returns. Spain gained as long as the world output of each external-economy commodity, wherever it was located, exceeded the autarky output. Therefore, it was good for Spain, and the World, that a Catalonia or Lancashire existed. Modern industry needed to be concentrated, and where it was found was really not important.

Appendix 1: Krugman’s Index of Regional Specialization

This appendix describes the methods and data used to create Krugman’s indices of regional specialization from 1797 to 1910. Paul Krugman defined the following index of regional specialization. Algebraically:

$$RS_{CM} = \sum_{i=1}^{n} \left| \frac{L_{iC}}{L_C} - \frac{L_{iM}}{L_M} \right|,$$

where $L_{iC}$ is the amount of employment in industry $i=1,\ldots,n$ for region C and $L_C$ is total employment in region C and similarly for region M. This index moves from zero to two. If the index is equal to two, the regions are completely specialized. If the index is then equal to zero both regions (C and M) are not specialized. Krugman’s indexes of regional specialization are computed for each of the 28-biregional comparisons of eight macro regions and these indices are averaged to calculate an overall measure of regional specialization.

Calculations are based on data from Population Census for respective years. Note that by-employment is not recorded in Spanish censuses and female employment is likely to be understated. 1797 census provides incomplete data for the province of Vizcaya so the data has been completed using weights from adjacent provinces. 1887 census provides no data on regional distribution of military personnel, so they are excluded for 1887. We divide Spain in eight macro-regions by similarity of characteristics (the so-called homogeneity principle). The resulting macro-regions are Andalusia, Aragon, Basque Country, Northern Castile, Southern Castile, Catalonia, Mediterranean and Northwest. The provinces of Almeria, Cadiz, Cordoba, Granada, Jaen, Huelva, Malaga, and Sevilla compose Andalusia. The provinces of Huesca, Teruel, and Zaragoza compose Aragon. Alava, Guipuzcoa, Navarra and Vizcaya compose the Basque Country. The provinces of Barcelona, Gerona, Lerida and Tarragona compose Catalonia. The provinces of Avila, Burgos, Leon, Logroño, Palencia, Salamanca, Santander, Segovia, Soria, Valladolid, and Zamora compose Northern Castile. The provinces of Badajoz, Caceres, Ciudad Real, Cuenca, Guadalajara, Madrid, and Toledo compose Southern Castile. The provinces of Albacete, Alicante, Baleares, Castellon, Murcia, and Valencia

---

76 Krugman, *Geography and Trade.*
compose the Mediterranean region. Finally, Coruña, Lugo, Orense, Oviedo, and Pontevedra compose the Northwest region.

**Appendix 2: Hoover’s Coefficient of Localization**

The simplest and most commonly used measure of industrial localization is the Hoover’s coefficient of localization. This is similar to the Gini coefficient and, hence, it should be interpreted similarly. Thus, if the coefficient is equal to zero, then the industry is completely dispersed across the regions. If it is equal to one, then the industry is completely localized in one region. To compute the Hoover’s coefficient first estimate the following location quotient \( LO_{ij} \):

\[
LO_{ij} = \frac{L_{ij}}{L_{ij}} / \frac{L_{ij}}{L_{ij}},
\]

where \( L_{ij} \) is employment in industry \( i \) for province \( j \), \( L_{ij} \) is the total industrial employment for province \( j \), and \( L_{ij} \) is employment in industry \( i \), and \( L_{ij} \) is total employment in Spain. Then place the provinces by their location quotients in decreasing order, and compute the cumulative percentage of employment in industry \( i \) over the provinces (y-axis). Finally, compute the cumulative percentage of employment in total manufacturing over the regions (x-axis).

The Hoover’s index is computed using 27 pseudo-provinces. This number of provinces is consequence of the modification of Spanish provinces in 1833. For example, in 1797 Catalonia was composed by one province while in 1833 was divided into four provinces (Barcelona, Gerona, Lerida, and Tarragona). Therefore, to allow comparison between censuses these four provinces have been split in one across indexes. The resulting provinces are Madrid, Basque Country, Aragon, Asturias, Avila, Burgos, Catalonia, Córdoba, Cuenca, Extremadura, Galicia, Granada, Guadalajara, Jaén, León, Ciudad Real, Murcia, Palencia, Salamanca, Segovia, Sevilla, Soria, Toledo, Valencia, Valladolid, Zamora and Baleares. Data for 1797 and 1910 indices are drawn from the census of population for the respective years while the data for 1861 is drawn from Gimenez Guited, *Guia Fabril*.

REFERENCES


---

77 Ellison and Glaeser ("A dartboard approach") pointed out the restrictions of that type of index.


Rosés, Joan R., Juan Carmona, and Blanca Sánchez-Alonso. “Regional wage convergence in


