Economic Development and Market Potential

European Regional Income Differentials, 1870-1913

Paul Caruana Galizia

Department of Economic History

London School of Economics

A thesis submitted for the degree of

Doctor of Philosophy

2015

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 74,646 words.

Acknowledgements

My first debt is to my supervisor, Max Schulze, for securing the funding for this research, and for his enthusiasm in helping me to pursue it. Mark Schaffer and Harald Tauchmann shared their *Stata* code. Jordi Marti Henneberg shared his railway, waterway, and road maps. Alexis Wegerich shared coal price data. Tomas Cvrcek shared Austro-Hungarian schooling data; Kerstin Enflo supplied the Swedish regional income data; Emanuele Felice supplied Italian regional literacy and schooling data; and Max Schulze supplied the Austro-Hungarian regional income and employment data.

Seminar participants at the Caltech/All-UC Huntington Conference, LSE, Warwick, Lund, and Venice made helpful comments on parts of this dissertation. Steve Broadberry and Oliver Volckart carried out my second-year review, steering me clear of unpromising roads. This dissertation owes much to James Dowey's suggestions and to converstaions we had over the years, as well as this dissertation's examiners Nikolaus Wolf and Kevin O'Rourke who gave invaluable feedback on my work.

My debt to my nuclear and extended families and to Jessica, not only for their patience, counsel and inspiration, is too great to be acknowledged adequately.

This research was funded by an Economic and Social Research Council Nominated +3 Studentship, and a STEPS Doctoral Scholarship from the Maltese Ministry of Education and the European Social Fund.

P.C.G
Berlin
March, 2015

EUROPEAN SOCIAL FUND
MALTA 2007-2013

The research work disclosed in this publication is partially funded by the Strategic Educational Pathways Scholarship (Malta). This Scholarship is part-financed by the European Union – European Social Fund (ESF) under Operational Programme II – Cohesion Policy 2007-2013, “Empowering People for More Jobs and a Better Quality Of Life”.

Operational Programme II – Cohesion Policy 2007-2013
Empowering People for More Jobs and a Better Quality of Life
Scholarship part-financed by the European Union
European Social Fund (ESF)
Co-financing rate: 85% EU Funds; 15%National Funds

Investing in your future

Abstract

This dissertation examines the extent to which proximity to markets - as measured by market potential, the trade cost-weighted sum of surrounding regions' GDP - can explain late-nineteenth century Europe's regional per capita income differentials. The research questions are: (1) was the spatial distribution of regional income random; (2) how helpful are traditional explanations - coal and institutions - of regional income; (3) how helpful is market potential when controlling for traditional explanations; and (4) did market potential have an effect on other determinants of income? This dissertation finds that: (1) the distribution of regional per capita income increasingly concentrated in the northwest; that there was little tendency to income convergence; and regional inequalities were higher within than between countries; (2) while a measure of regional institutions is correlated with income, simple distance-to-coal and a cost-to-coal measures are not; (3) market potential has a significant effect on income; foreign market potential more so than domestic; and increasing core relative to peripheral market potential results in perpherial income losses; and (4) changes in literacy rates, a proxy for human capital, responded to changes in market potential. In conclusion, a new economic geography framework with market potential at its core fits the historical experience well. Certain regions performed better than others generally because they had cheaper access to markets. At the start of the period, trade costs were high, and so economic activity - long concentrated in Britain - was spread out more or less evenly across the Continent. By the end of the period, when trade costs dropped dramatically, economic activity concentrated in the northwest of Europe at the cost of the periphery.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

This dissertation has its roots in the observation that the distribution of European regional income today is not dissimilar from the distribution that existed 140 years ago. Figure 1.1 plots the distribution of European regional GDP per capita relative to its cross-sectional mean from 1870 to 2010.[1] While the right tail of the distribution has grown in length, indicating a growing gap from the mean to the very rich regions, the general shape of European regional inequality has remained comparable over the period. A bi-modal distribution that took form in 1910 has remained present until in 2000 and 2010. Indeed, looking at the simple coefficient of variation on the absolute GDP per capita levels underlying 1.1 we go from 0.29 in 1870 to 0.31 in 1910 and to 0.30 in 2010. While recent papers from the OECD and articles in *the Economist* indicate that regional inequality has become the topic *du jour*, economic historians long ago wrote about Europe's highly unequal distribution of production within and between countries in the ninteenth century (Fredriksen, 2012; Pollard, 1973; The Economist, 2011).

Is something interesting just because it is old? Leaving aside the effects of regional income inequality, looking at historical inequalities shows us that contem-

[1]The EU-15 countries are: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden and the United Kingdom. My sample of countries: Austria-Hungary, Britain (including Ireland), France, Germany, Italy, Spain, and Sweden. There are obvious problems with national border changes and sample inconsistencies, but this a simple illustrative exercise.

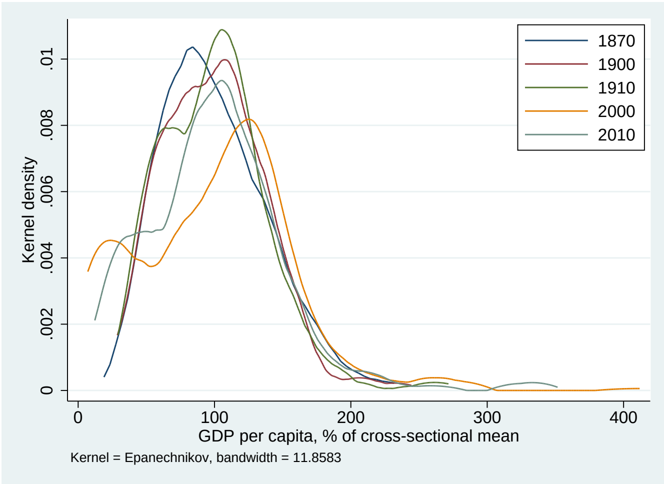

Figure 1.1: Long-run European regional GDP per capita distributions

Notes: Figures are expressed as percentage of cross-sectional mean. The post-1910 countries are: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden and the United Kingdom. The 1870-1910 countries: Austria-Hungary, Britain (including Ireland), France, Germany, Italy, Spain, and Sweden. The pre-1910 currency unit is the 1990 Geary-Khamis dollar; post-1910, purchasing-power parity national currencies. For sources of 1870-1910 data, see Chapter 3; and European Commission (2011) for post-1910 data.

porary explanations - often based on relatively recent European Union structural policies or exogenous shocks - can be ahistorical and so miss the true cause of inequality.

Initially, mainstream economists and policymakers based their arguments for European regional integration on classical trade and growth theory. They argued that a low level of competitiveness - an absolute disadvantage - does not prevent regions from enjoying the wider gains from trade that are based on comparative advantage. Balassa (1961) theorised that the expansion of the European Common Market would lead to overall productivity gains. Aitken (1973) and Wang and Winters (1992) provided empirical support for this theory. Growth theorists predicted capital market integration would lead to convergence because of decreasing returns to scale to capital: regions with low capital stocks and per capita incomes should have higher marginal products and returns to capital. This, argued Barro and Sala-i Martin (2004), is what explains (slow) per capita income convergence in late twentieth century Europe, when core capital flowed into the periphery.

Almost 60 years after the European common market was established, and 140 years after our earliest measures of regional income, we are, however, still living with high regional per capita inequalities. A helpful explanation of this inequality must account for its persistence and for the fact that ever-deeper economic integration has not produced the results predicted by classical theory. If achieved, such an explanation might also provide an antidote to presentist accounts of European economic inequalities, increasingly seen as the root of the current crisis, as having been primarily caused by the European Union and monetary union (Ferguson, 2012).

1.2 Theoretical background

In his pioneering new economic geography (NEG) work, *Geography and Trade*, Krugman (1991, p. 95) asked, 'What will happen to [poorer] regions as Europe becomes more closely integrated?' Classical growth and trade theorists presumed that with improved access of low-wage peripheral regions to high-wage core regions, industry will be incentivised to move to the periphery. This happened to some extent, but it was not a necessary effect. Krugman and Venables (1990) showed that improved access to markets can be detrimental to peripheral industry.

Regional integration and peripheral industry

Suppose that an industry can locate in one or both of two regions: a *central* region where wages and hence production costs are high, but it has good access to markets, and a *peripheral* region where wages are low, but market access is poor. It is easy to think that a reduction in trade costs - transport costs plus tariffs - would move industry from the central to peripheral region, but in reality something quite different is likely to happen.

A reduction in trade costs can have two simultaneous effects. First, it allows production to locate where it is cheapest. Second, it allows the concentration of production in one region, enabling economies of scale to be realised. Crucially, this agglomeration might make sense at the region with higher wages, but with better market access.

Table 1.1, from Krugman (1991), offers a practical example of this. Suppose a

Table 1.1: Hypothetical effects of lower trade costs

	Production costs	Shipping costs High	Medium	Low
Produce in Belgium	10	3	1.5	0
Produce in Spain	8	8	4	0
Produce in both	12	0	0	0

Notes: Taken from Krugman (1991, p. 96).

good can be produced in one or both of Belgium, the central location, and Spain, the peripheral location. For simplicity, Krugman (1991) takes total sales as given, ignoring any elasticity of demand. That is, producers choose the location that minimises the sum of production and trade costs. While it is cheaper to produce that good in Spain than in Belgium, since Spanish wages are lower, it is cheaper still to produce the good in *either* location than in both because of economies of scale. On the other hand, producing in both locations would minimise trade costs, and producing in Belgium would involve lower trade costs than producing in Spain.

Table 1.1 shows three cases of trade costs: - high, intermediate, and low (or zero). If costs are high then production will occur in both countries, but if they are low it will occur in low-wage Spain. A 50 per cent reduction in costs from the high case of 3 to 1.5 causes production to move from low-wage Spain to high-wage Belgium. This is because in the intermediate trade cost case, costs are low enough to incentivise agglomeration, but still too high for market access to outweigh production costs (wages in this example) as determinant of industrial location. The implication is that the relationship between trade costs and Spanish production is U-shaped rather than monotonic. Across some range of costs, 'integration actually leads production to move perversely from the point of view of comparative costs' (Krugman, 1991, p. 97).

Which part of the U-shape?

Krugman (1991, p. 97) hypothesised 'that we are now on the good part of the U, not the bad: that railroads and steamships led to deindustrialization of the periphery, but that 1992 will actually favour peripheral manufacturing.' Thinking about economic activity more broadly, Krugman's hopes have not be fulfilled.

That is, regional income inequalities remain high and the *map* of regional income has not changed susbtantially, going by our historical measures and descriptions (Landes, 1969; Pollard, 1973, 1981; Wrigley, 1961).

Cost reductions from "railroads" and "steamships" did indeed facilitate the concentration of production in the northwest of Europe. Figure 1.2 shows, on the left, a familiar picture of declining ton-mile freight rates - average of coal and grain - in global ocean shipping and European railway transport. On the right of the figure, a breakdown of transporting the same "good" from London, the European core, to Kosice (in what is now Slovakia), the eastern periphery, shows an overall drop in trade costs of 38 per cent. It also shows how the arrival of railways obliterated road and waterway transport, shortening the length of the transport network from 1,551 miles in 1870 to 1,389 miles in 1910.[1] The issue, however, is that more dramatic trade cost reductions since - the EU free trade area and monetary union, no less - have failed to move it back.

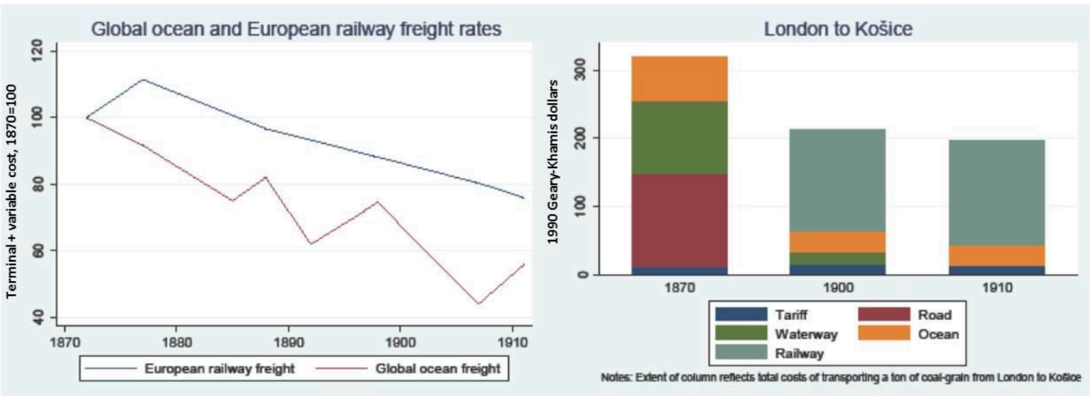

Figure 1.2: Transport cost declines from 1870 to 1910

Notes: Left panel shows terminal (docking) plus variable (per ton-mile) costs of coal-grain freight. Underlying data are in 1990 Geary-Khamis dollars, and 1870 is set to 100. The railway data are from Schulze (2007b) and the shipping data from Kaukiainen (2006). Right panel shows the cost of transporting a ton of coal-grain from London to Kosice (Austria-Hungary), broken down into five components. The data re in 1990 Geary-Khamis dollars. See Chapter 4 for more details on sources and calculations.

Through forward and backward linkages, a region that accumulated physical

[1]See Chapter 4 for all the sources and explanations of these data. The network is my own GIS of roads, waterways, railways, shipping lines, ports, and regional cities. The costs are in 1990 Geary-Khamis dollars, and average the freight rates of coal and grain, as representative goods.

and human capital will have a higher - rather than lower - rate of return on investment than a region where physical and human capital are scarce. If the rate of capital accumulation depends on the rate of return, then an 'unequalizing spiral' will ensue, where Europe becomes differentiated into richer and poorer regions (Krugman, 1991, p. 94).

These increasing returns do not only feed off themselves and low trade costs: they depend on a 'demand externality' (Krugman, 1991, p. 20). Industry will locate where the market is largest, and the market is largest where industry locates. If this cicularity is strong enough, then historical patterns of economic activity can persist over long horizons. Interestingly, the idea of the market as a determinant of industrial location is as old as Marshall (1920). Economists did not take much notice of it, but some geographers in the mid-twentieth century did, before it was forgotten again for another three decades. Looking at retailers, Harris (1954) measured the pull of markets - "market potential" - through distance-weighted regional populations, as a proxy for market size weighted by trade costs. His idea has been revived in NEG, and plays a central role in my dissertation (Crafts, 2005a; Hanson, 2005; Head and Meyer, 2004; Martínez-Galarraga, 2014). It captures in a simple fashion the idea that producers want to minimise the sum of production and trade costs, and maximise their market size.

1.3 Market Potential: Theory and Mechanism

Recent theoretical work in new economic geography explains spatial agglomeration as a function of product-market linkages between regions. As mentioned above, the market potential function in Harris (1954) is a precursor to this recent work. His function equates the potential demand for goods and services produced in a region with that region's proximity to consumer markets. Following the notation in Hanson (2005):

$$MP_j = \sum_{k \varepsilon K} Y_k e^{-djk} \tag{1.1}$$

where MP_j is the market potential for region j, Y_k is income in region k, and d_{jk} is the distance between j and k. Early work asserted the existence of equation 1.1 with little theoretical foundation. More recently, researchers have derived a structural relationship similar in form to 1.1 from general-equilibrium spatial models.

It is useful to start with the basic framework of Krugman (1991), and extend it along the lines of Helpman (1998) which provides a tractable form for empirical work. The difference between the two comes down to Helpman (1998) replacing the agricultural sector in Krugman (1991) with a housing sector thus creating a nontraded good, and more realistic distribution of production. All consumers have identical Cobb-Douglas preferences over two bundles of goods, traded manufacturing goods and housing services,

$$U = C_m^{\mu} C_h^{1-\mu} \tag{1.2}$$

μ is the share of expenditure on manufactured goods, C_h is the quantity of housing services consumed and C_m is a composite of symmetric manufacturing good varieties given by

$$C_m = \left[\sum_i^n c_i^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{1.3}$$

where σ is the elasticity of substitution between any pair of product varieities and n is the number of product varieties. The production of each manufactured variety has increasing returns, so that

$$L_{im} = a + bx_i \tag{1.4}$$

where a and b are constants, L_{im} is labour used in the production of variety i and x_i is the quantity of i produced. When in equilibrium, each variety of manufactured good is produced by a single monopolistically competitive firm and the foreign onboard price (f.o.b.) price for each variety is a constant markup over its marginal cost, which in turn depends on the wage, w.

Labourers, L, are perfectly mobile between J regions. The housing stock in region j is normally assumed to be fixed at H_j. Housing stock ownership is then assumed to be symmetric across individuals so that each labourer owns $1/L$ of the housing stock in each one of the J regions. With iceberg trasnport costs in trading goods between regions, the cost, insurance, freight (c.i.f.) price of good i produced by region j and sold in region k is

$$P_{ijk} = P_{ij} e^{\tau djk} \tag{1.5}$$

where P_{ijk} is the f.o.b. price of good i produced in region j, τ is the unit transport costs and d_{jk} is the distance between j and k. The solution of this model, given in Helpman (1998) and Fujita et al. (1999a), is well known. Under certain parameters, the manufacturing sector spatially concentrates. Firms have an incentive to locate in a region with employment to serve a large regional consumer market at low trade costs while saving on fixed production costs. The costs associated with location in a large market are higher wages, resulting from high housing costs, in turn the result of regional congestion.

As shown in Hanson (2005), it is helpful to derive the demand for traded goods produced in region j to show the intuition behind the market potential function. Let C_{ijk} be the quantity of good i so that region k buy from region j. Given constant elasticity substitution (CES) over traded goods, the symmetry of traded goods in both preferences and technology, along with the equilibrium condition on the constant markup pf prices on top of marginal cost ($P_{ij} = \frac{\sigma}{\sigma-1} bw_j$), total sales of manufactured goods by region j are defined as

$$\sum_k \sum_i P_{ijk} C_{ijk} = n_j \sum_k \mu Y_k \left[\frac{\sigma}{\sigma - 1} bw_j e^{\tau djk} \right]^{1-\sigma} T_k^{\sigma-1} \tag{1.6}$$

where T_k is the CES price index for manufactured goods available in region k. When profits are zero, sales of manufactured goods in region j equal wages paid to labour in region j. The wage cost in region j, therefore, equals $w_j n_j a\sigma$. We can now follow Hanson (2005) to arrive at a modified market potential function, defined as

$$w_j = \theta \left[\sum_k Y_k e^{-\tau(\sigma-1)d_{jk}} T_k^{\sigma-1} \right]^{\frac{1}{1-\sigma}} \tag{1.7}$$

where σ is a function of fixed parameters. Wages in a region increase in the the income of surrounding regions, decrease in trade costs to these regions as well as increase the price of competing traded goods in these regions.

Following from this, the price index for traded goods in region j can be written as

$$T_j = \left[\sum_k n_k \left(\frac{\sigma}{\sigma - 1} b w_k e^{\tau djk} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{1.8}$$

Equation 1.8 defines the market equilibrium for traded goods. Their price index is higher where a larger fraction of goods is imported from distant regions.

There are three additional equilibrium conditions to equations 1.7 and 1.8. First, real wages are equalised across regions so that

$$\frac{w_j}{P_j^{1-\mu} s T_j^{\mu}} = \frac{w_k}{P_k^{1-\mu} s T_k^{\mu}}, \forall j \neq k \tag{1.9}$$

where P_j is the housing price in j. The second additional equilibrium condition is regional income equals income derived from labour and housing,

$$Y_j = n_j w_j a\sigma + \frac{1-\mu}{\mu} \frac{n_j a\sigma}{L} \sum_k n_k w_k a\sigma, \forall j. \tag{1.10}$$

The final additional equilibrium condition is that housing payments equal housing expenditure,

$$P_j H_j = (1 - \mu) Y_j, \forall j. \tag{1.11}$$

By now it is clear how the above extension of Krugman (1991) resembles a spatial labour-demand function: nominal wages increase in consumer income in surrounding regions, and decrease in trade costs between these regions. This is why even Krugman's initial framework focused on the mobility of labour rather than capital in driving agglomeration and regional divergence. When workers move from a peripheral to core region, drawn by initial levels of manufacturing in the core, they spend their income in their new host region, thus increasing demand there, but decreasing demand in their region of origin. As Combes et al. (2008, p. 131) write, with emphasis,

the migration of workers, because it sparks the combined move of pro-

> *duction and consumption capacities, modifies the relative size of markets, thus generating new agglomeration forces.*

The migration of workers, as stated in the model, also increases congestion costs in the core region. This implies that the migration decision is based on the core-periphery nominal wage differential as well the cost of living differential; in short, the real wage differential. If wages tend to increase in regions with greater market potential, then we should expect this to trigger migratory movement. The extent of this movement depends on the degree of workers's spatial mobility.

This not to say that the firm re-location channel is irrelevant. While high market potential makes a region more attractive to firms as well as individuals, it attracts firms through higher profits, but individuals through higher wages. In practice, the estimating equations look similar as wages are invariably proxied using per capita GDP levels, which can also be used to measure economic productivity more generally (Breinlich, 2006; Combes et al., 2008; Redding and Venables, 2004; **?**). Indeed, the historical literature indicates that distinguishing between the two mechanisms might be difficult. First, before the industrial and transport revolutions that took hold of Europe in the mid- to late-nineteenth century, **?**, p. 201 tells us that

> *the gaps between different parts of Europe were much smaller than they were to become later and some industrial activity not unlike that in Inner Europe was to be found almost everywhere.*

After the Industrial Revolution and after trade costs declined, **?**, p. 11 writes that

> *the industrial regions colonize their agricultural neighbours [and take] from them some of their most active and adaptable labour, and they encourage them to specialise in the supply of agricultural produces, sometimes at the expense of some preexisting industry, running the risk thereby that this specialisation would permanently divert the colonized areas from becoming industrial themselves.*

In short, there was a simultaneous migration of workers *and* firms towards the new industrial regions, particularly Germany. In the 1911 to 1915 period alone, Germany received 434,489 immigrants (Wilcox and Fercenzi, 1929a, p. 239). Italian emigration to other European countries went from 75,065 in 1878 to 313,032 in 1913 (Wilcox and Fercenzi, 1929b, p. 124). As for firms, Bairoch (1997, p. 116-7) echoes **?** in writing of the period,

> *There began a massive sales flow of manufactured articles toward what was gradually becoming the Third World, with the notable appearance of one of its characteristics: the more or less complete and rapid disappearance of all its industries. As a counterpart to these sales of manufactured articles, a massive flow of untreated products (tropical goods and raw materials) went to the West, which had more and more means to absorb them.*

Disentangling precisely which mechanism was at work, or which mechanism was more influential than the other, is beyond what the available data allow for. My main dependent variable is per capita GDP, as I go on to explain in the coming chapters, which has an interpretation as a broader measure of industrial activity. My empirical strategy throughout this dissertation is simply to examine whether a simple new economic geography model can explain Europe's changing spatial distribution of regional income in the 1870 to 1910 period.

1.4 Research questions

NEG theory has nothing to say about initial conditions apart from 'history matters' (Krugman, 1991, p. 20). Some times in can be pure accident. Dalton, Georgia emerged as America's "carpet capital" after a teenaged Catherine Evans discovered a better way of tufting bedspreads that was then mechanised (Krugman, 1991, p. 20). In other cases it can be due to fixed inputs. The north of England emerged as an industrial region after technological innovations that enabled and required the use of coal. What is important to me are not the initial conditions - the debate on that is large, long-running and I am ill-equipped to enter it - but 'the cumulative process that allowed such accidents to have such large and long-lasting effects' (Krugman, 1991, p. 62). I look for evidence that these processes were at work in late-nineteenth century Europe.

While it would be ideal to go further back in time than 1870, the data simply do not allow for this. The conventional pre-WWI ending avoids the border changes after the War that make this analysis intractable, and marks the limit of what a single research student can do in four years. In any case, the 1870 to 1910 period - the "highwater mark" of the "first globalisation" - in Europe - a collection of different states and regions - provides a good opportunity to understand the distribution of economic activity within and between countries (Daudin et al., 2010; Epstein et al., 2003; O'Rourke and Williamson, 1999).

My approach to this broader theme is to ask the following sub-questions, each one informed by the historical context: (1) was the spatial distribution of regional economic activity random; (2) how helpful are traditional, non-NEG explanations of regional income; (3) how helpful is a market potential explanation when controlling for traditional explanations; and (4) did market potential have an effect on other determinants of income? The answers to these questions together address the main concern of my dissertation: is the spatial distribution of European regional income a result of NEG forces? If I succeed in providing a satisfactory answer, then future research can draw a link from here to current patterns of European economic activity.

In answering these questions, I find that: (1) the distribution of European regional per capita income increasingly concentrated in the northwest, there was little tendency to income convergence, and regional inequalities were higher within than between countries; (2) while a measure of regional *de facto* institutions is correlated with income, simple distance-to-coal and a cost-to-coal measures are not; (3) market potential has a significant effect on income, foreign market potential more so than domestic, and increasing core relative to peripheral market potential results in perpherial income losses; and (4) changes in literacy rates, a proxy for human capital that was also spatially concentrated, responded to changes in market potential.

In conclusion, an NEG framework fits the historical experience well. Certain regions performed better than others generally because they had cheaper access to markets. At the start of the period, trade costs were high, and so economic activity - long concentrated in Britain - was spread out more or less evenly across the Continent. By the end of the period, when trade costs dropped dramatically, economic activity concentrated in the northwest of Europe, being better placed to serve the larger markets there, to serve more distant markets at lower trade costs, and reap the gains from agglomeration. The point at which trade costs drop so much that economic activity is free to locate anywhere - as it would be equally cheap to serve any market - had not been reached by this point. The resulting pattern was the consolidated "Golden Triangle" that persists until today.[1]

1.5 Methods and data

Empirical tests are organised around explaining my dissertation's main dependent variable, regional real GDP per capita. In this period of European industrialisation, productivity growth was important, but economic growth was primarily driven by re-allocating resources out of agriculture and into industry (Broadberry et al., 2010). Income (or GDP per capita), however, captures variation in economic activity and progress beyond industrial employment shares or industrial value-added alone - London's high-income economy, for example, was driven by services. Income is also more easily comparable across time and place, avoiding issues that come with changing employment categories over time or different sectoral definitions across countries. Lastly, the Geary-Khamis dollar unit I use throughout this dissertation is perhaps the most widely used income unit in the economic history literature, making comparisons with other work much simpler.

Analysing the role of market potential, the main independent variable of interest, required the construction of a detailed trade network, using a geographical information system (GIS). The network was built up from railway, roads, waterways, and shipping lines, and ports and regional nodes. I routed through the network using Dijkstra's lowest-cost search algorithm (Yan, 2014). This required data on transport costs, transhipment costs, and tariffs. As a cross-check on this construction of market potential, I estimated an alternative version of market potential, using international bilateral trade flows (**?**). Controlling for competing explanations of income required the estimation of a novel measure of *regional* institutions, based on regional literacy data and the ideas in Acemoglu et al. (2005),

[1]The geographical region roughly delimited by London, Paris, and Berlin. The term was first used by geographers, as in Hall (1993), but quickly entered popular discourse (Buck, 2006).

and of proximity to coal deposits, based on the writings of a number of historians (Fernihough and O'Rourke, 2014; Pollard, 1981; Wrigley, 2010).

The principal methodology I employed in this dissertation is the quantitative analysis of a panel of 199 regions from Austria-Hungary, Britain, France, Germany, Italy, Spain, and Sweden over the benchmark years 1870, 1900, and 1910. Coverage is wide, as regions are from all corners of the continent, as well as deep, given the countries accounted for 93 per cent of European GDP during the period (Bolt and Van Zanden, 2013). This regional panel approach has a number of advantages. The first and most obvious is greater historical accuracy. If we are interested in the spatial distribution of income, then a smaller geographical scale (compared to the standard national unit), is more historically accurate. Industrialisation, or economic development more generally, varies between as well as *within* countries. A regional panel contains this information. Nations, on the other hand, aggregate industrial and agricultural regions, and ignore spatial clusters of economic activity. Second, a regional panel allows us to narrow in on the variation that mattered for economic development. A national unit, aggregating internal differences in economic activity, prioritises noise over signal. This matters in a causal study where we ask the fundamental question, why did some places achieve higher income levels than others? Lastly, a panel is comparative - spatially and temporally - in nature. Disjointed regional economic histories are of the same use as disjointed national economic histories: they contribute to the accumulation of evidence. Taking nations or regions as standalone units, however, is saying that they do not form part of a wider economy, and it also makes generalisations less reliable. My panel draws on a large body of work by country specialists, allowing me to extract comparative and generalisable insights.[1]

Broadly speaking, the quantitative techniques are designed around two themes. The first is to approach a *causal* explanation of European regional income. While I used exploratory techniques to define patterns and trends, I then use a series

[1] The literature now provides estimates for Austria-Hungary, Britain, Italy, Sweden, and Spain (Crafts, 2005b; Enflo et al., 2010; Felice, 2009; Roses et al., 2010; Schulze, 2007b). Most of these estimates, for the post-1900 period, have been incorporated into he European Science Foundation-funded project, *the Historical Economic Geography of Europe 1900-2000*, coordinated by Joan Roses and Nikolaus Wolf. In this dissertation, I provide estimates for France and Germany. Elswhere I provide estimates for China and India (Caruana Galizia, 2013; Caruana Galizia and Ma, 2014).

of instrunmental variable (IV) regression analyses to test causal explanations put forward by the literature, and the market potential explanation. As a study of economic history covering a broad topic, I did not have the luxury of controlled labratory or even "natural" experiments (Diamond and Robinson, 2010). IV analyses have their critics, but they allow us to get as close to a causal story as possible, given the data and historical constraints (Chang, 2010). The second underlying theme is spatial analysis: a dimension often implicitly ignored in the literature. An a-spatial analysis is, for example, to ignore Europe's clear clusters of income or its core-periphery income structure. To take this into account, I make extensive use of a GIS. Some variables, for example a distance or cost to coal deposit, are constructed using GIS. More complex calculations, like the lowest-cost route algorithms underlying the market potential variable, would have been impossible without GIS. In his latest presidential address of the 2013 *Economic History Association* annual conference, Atack (2013, p. 332) took the opportunity to discuss the use of GIS in economic history. Speaking of his ongoing work on American railways in the nineteenth century, he concluded, 'it is not too soon to claim that historical GIS transportation databases will change our interpretation of American economic history.' The GIS dataset here is large in scale and scope, and this dissertation can only be a first pass at exploiting its full potential.

1.6 Chapter outline

This dissertation is composed of four linked analytical chapters (Chapters 3, 4, 5, and 6). The following contextual chapter (Chapter 2) and the concluding chapter (Chapter 7) bring the four analytical chapters together in addressing the bigger question of regional income differentials. The appendices (Appendix A, B and C) contain additional notes and tests on the data.

Chapter 2: Context: Historiographical Background
Recent NEG-work in economic history has provided us with a generalisable explanation of European development - the limitation with older historical work like Pollard (1981) - but is restricted to national studies, as in Roses et al. (2010)

- the advance of of that same historical work. We have taken two steps forward, and one step back to William N. Parker's question, on 'why European history is not the history of a continent' (Rhode, 2008, p. 192). The U-shape of spatial economic development uncovered by Combes et al. (2011) for France, for example, and posited generally by NEG theory, did not only exist within France, but was occurring *across* European national borders. NEG can thus explain not just individual within-country patterns, as recent authors have, but also the broader European patterns of regional industrialisation and income. This is where my dissertation comes in. I collect all available data, fill in the gaps, and explain the resulting dataset using NEG.

Chapter 3: Explicandum: the Distribution of European Regional Income

Was the spatial and temporal distribution of late-nineteenth century European income random? I estimate the regional incomes of France and Germany, and standardise them along with all the other estimates from the literature. I then conduct a number of exploratory empirical exercises on the income data, detecting non-random spatial patterns, or clusters, and how they change over time. I find that income inequality was higher within than between countries; that contrary to the historiography, there is no discernible trend of income convergence; that the income distribution went from normal in 1870 to bi-modal in 1910; and that the spatial distribution went from high income clusters in Britain and northern France with fragmented clusters elsewhere, to the consolidated "Golden Triangle" that persists until today.

Chapter 4: Traditional Explanations: Coal and Institutions

How helpful are traditional explanations of late-nineteenth century European regional income? In this chapter, I explore the extent to which regional income differentials were due to either institutions or natural endowments, specifically coal. After estimating a novel measure of regional institutional efficiency, I find that coal access - as measured by distance and cost to the nearest deposit - was not an important determinant of income levels. Regional institutional efficiency has a strong positive effect on income levels, when controlling for coal access, and country and year fixed effects. At least for the late-nineteenth century period, this calls into question a long line of work of "coal-men" from Pounds and Parker (1957) to Wrigley (2010). It also lends support, in a regionally-modified form, to the conceptual arguments in Acemoglu et al. (2005).

Chapter 5: An Alternative Explanation: Market Potential

In this chapter, I provide an alternative explanation of regional income; one that can account for the spatial and temporal distribution of regional income and fits with the broader historical context of falling barriers to trade. First, I examine the neo-mercantilist arguments put forward in the historiography (Pollard, 1981). I then describe my measures of market potential and the sources used, giving further details in *Appendix C*. I then set out my empirical strategy, which is geared towards uncovering baseline effects of market potential on income, differences between foreign and domestic market potential, and whether the relationship is uniform across regions. The final section concludes with a discussion of the issues raised. The general findings of this chapter are: (1) market potential has an economically meaningful and statistically significant effect on income levels; (2) foreign markets have a larger effect on income levels than domestic markets; (3) increasing core market potential relative to peripheral market potential results in an absolute decline in peripheral per capita income levels and a decline in peripheral share's of total GDP; and (4) the relationship between market potential and per capita income is not uniform. Residuals suggest the influence of increasing returns, and a smaller role for factor endowments and locational fundamentals.

Chapter 6: Addendum: Market Potential and Human Capital

Did market potential have a direct effect on other determinants of income, namely human capital? My hypothesis here is that, generally, regions far from the north-western European core had low stocks of human capital, as measured by literacy, but as economic distance, as measured by market potential, between the core and periphery dropped over the late-nineteenth century, the periphery's incentives to invest in human capital increased. This produced an *international* convergence in literacy (Cipolla, 1969; Crafts, 1997). National law that made primary enrolment compulsory is an alternative explanation for this convergence but, as I show in this chapter, variation in literacy was higher within than between countries. Further, regional literacy rates were spatially correlated. These features in the data are consistent with my hypothesis: market potential, itself spatially correlated, varied by region causing the incentives for human capital investment to also vary by region. The emphasis here is on market potential providing an incentive on which agents can decide to act: it reflects the demand for, rather than supply of (as with national law), human capital. I show that changes in literacy rates were responsive to changes in market potential, as people sought to capitalise on higher returns to education. To the extent that technological and industrial progress depended on a region's human capital, the implications of these results are that market potential shaped a region's ability to industrialise in other ways than through the trade cost channel.

Chapter 7: Conclusion

In this chapter, I conclude the dissertation by reviewing the main findings of each analytical chapter. I summarise these findings in light of the literature, asking, how has this dissertation changed our understanding of European economic history? I re-assert the usefulness of a regional approach, and of explicitly considering the spatial dimension of economic activity. I then cautiously relate the historical context to current topical debates on European regional income inequality.

Chapter 2

Context: Historiographical Background

2.1 Introduction

The literature provides a number of different ways of thinking about Europe in the late-nineteenth century (circa 1870 to 1913). Table 2.1 summarises some influential ones. Different categorisations can be the result of different methods or different research aims. Hobsbawm (1987) and Hobson (1902), for example, were both interested in Europe's unprecedented colonial expansion and the "global imperialism" that it created. O'Rourke and Williamson (1999) and Williamson (1998) define the period as one of unprecedented economic globalisation, when commodity prices converged rapidly thanks to improvements in transport technology and, in the latter case, created a protectionist policy backlash. Daudin et al. (2010) make the same argument, with a clearer focus on Europe. The authors see economic rather than colonial forces as the defining feature of the time. Also focusing directly on Europe, Pollard (1981) writes in a similar vein about the policy backlash, when newly-unified European states hiked tariff rates on agricultural goods and manufactures in response to an influx of cheap North American commodities. Kaplan (1957), writing from a political history perspective, sees late-nineteenth century Europe as having been in a functioning and informal "balance of power" until the very-late nineteenth century, when national political interests - Germany's annexation of Alsace-Lorraine, in particular - took precedence. While there may have been a breakdown in trade and political relations, Bordo (1981) reminds us that this period was that of the "Classical Gold Standard." During this time, the majority of gold standard members adhered to fixed-rate convertibility and international capital flows reached levels that were not seen again until the end of the twentieth century.

This was, in short, an exciting period as the literature makes clear. While many of the themes will emerge in my analytical chapters, of all these "eras", the two most relevant for my purposes are those from Mokyr (1998) and Foreman-Peck (1999): the "Second Industrial Revolution" and the "Zenith of European Power."

Table 2.1: Historiographical eras in late-nineteenth century Europe

Era	Source
Age of Empire	Hobsbawm (1987)
Age of Neo-Mercantilism	Pollard (1981)
Classical Gold Standard Era	Bordo (1981)
European Balance/Alliances	Kaplan (1957)
First Globalization	O'Rourke and Williamson (1999)
Globalization Backlash	Williamson (1998)
High water mark of 19th century globalization	Daudin et al. (2010)
La Belle Epoque	Esteves (2011)
New Imperialism	Hobson (1902)
Second Industrial Revolution	Mokyr (1998)
Zenith of European Power	Foreman-Peck (1999)

Starting at the end of the nineteenth century, industrialisation took a new form. As Mokyr (1998, p.2) writes, the period 'witnessed the growth in some industries of huge economies of scale...Some vast concerns emerged, far larger than anything seen before. This change occurred because of ever more important economies of scale in manufacturing.' This first important feature is follwed by a second: these 'vast concerns', as Mokyr (1998, p.2) put it, emerged in Continental Europe, especially Germany. No longer was modern industrialisation limited to Britain. While it spread across Europe in varying degrees, the change was enough, Foreman-Peck (1999) argues, to take Europe to its economic "zenith." These changes have expectedly attracted a lot of attention in the historiography, much of it of course preceding Mokyr (1998) and Foreman-Peck (1999).

2.2 Historiographical background

One of the first "explanations" for Europe's economic history during this period was simply description - and not always accurate, as Foreman-Peck (1995) writes. In the "leader-follower" scheme, Britain industrialised first, which raised its productivity in manufacturing and its per capita income levels above European levels. Then, in varying degrees, European countries are argued to have adopted Britain's productive technology and economic organisation. The pace at which they did this, the scheme goes on, determines the extent of their income lag behind Britain. Rostow (1960) provided the GDP figures, a novelty, and one of the scheme's keywords: 'discontinuity', which described the movement from one of his stages of growth to another. Gerschenkron (1962), focusing on industrial production instead, wrote of economic 'backwardness' in peripheral Europe, giving the scheme its next keyword. Landes (1969) and Pollard (1973) wrote of the 'diffusion' of industrialisation across Europe (Pollard, 1973, 1981). Following this scheme, explanations of income differentials would be the determinants of the speed of adoption of the leader's technology. The authors assumed that new technology is adopted only if more capital is available, and so European economic history was one of differential rates of capital accumulation. Gerschenkron (1962), for example, spent a lot of time on the role of banks and states in overcoming underinvestment. This was seemingly validated by the Solow-Swan growth models of the 1950s and 1960s, which attempted to explain economic development through, mainly, capital accumulation (Solow, 1956; Swan, 1956).

This first wave was undermined by the production of quantitative evidence - ironically, given this was Rostow's aim - which was occurring in parallel. Rostow's "discontinuities" and "take-offs" in European economic development proved to be hard to pin down. Gerschenkron's higher shares of production-goods in the output of backward economies, and the importance of investment banks, did not show up in the data underlying the comprehensive book of Milward and Saul (1973). Nor did the theory hold up well to newer economic ideas (Sylla and Toniolo, 1992). The jump in capital accumulation was not necessitated by manufacturing, but railways. In Britain, railways accounted for a much higher proportion of the economy's capital stock until the interwar years (Feinstein and Pollard, 1988).

So-called "alternative" work, starting in the 1970s, showed that the "leader" was not so great after all, eroding the credibility of the first wave. O'Brien and Keyder (1978) showed that France did not lag as far behind as Britain, as others argued (Landes, 1969). Further, it avoided some of the welfare costs that came with Britain's urbanisation for industrial development. Crafts (1984) showed that Britain's reliance on coal was, in fact, unusual throughout Europe's industrialisation. He argued that while coal shaped the possibilities of an economy's industrialisation, human capital could act as a substitute input. This is what happened in coal-poor, highly literate, and fast-growth Sweden (Cameron, 1985; Sandberg, 1979). Morris and Adelman (1988) provided what is perhaps the most "alternative" approach: a principal component analysis of 35 macro-economic variables measured for 23 - not just European - countries. Their components tell them that there existed five development paths: two industrial, two agricultural, and one balanced. As North (1989, p. 90) wrote in his review of the book, 'the theories that are the building blocks, the inconventional statistical techniques they employ, and the quality of the data they use all make the study vulnerable to...criticism.' Foreman-Peck (1995, p. 444) calls their 'selection of indicators is somewhat restricted and arbitray.'

Another group of researchers avoided quantitative methods. As with traditional economic history, they focused on the role of institutions, broadly conceived. Their geographical concern, however, was not exclusively European. Chandler (1990) wrote about British and German business organisation, with a focus on how it competed with and learned from the United States. Kennedy (1986) argued that the structure of London's capital market distorted Britain's economic development at large. North (1981), and to some extent Olson (1982), used the Western European experience to create a theory of institutions that proved to be useful in current research (for example, Acemoglu et al. (2005)). Jones (1981) wrote about the rise of Europe - the 'European Miracle' - in a global comparative perspective. Indeed, institutions are central to most general economic histories that feature Europe. Rosenberg and Birdzell (1987) argued that Western Europe's economic success was a function of its loosening political and religious controls, and that economically-useful innovation is a result of competition among politically independent units. Jones's explanation is based on the universal propensity to accumulate wealth, and that in the West, governments did not get in the way of this - through wars, invasions, taxation, or regulation - too much.

This wave of institutional work offers compelling, but not rigorous, accounts of European development. Chandler's industry-level analysis of British and German firms fails to deal with the fact that Ford, until the 1920s, was a family-run company - something Chandler (1990) blames for business inefficiencies in Europe. Olson's influential theory of distributional coalitions is also liable to criticism when it comes to nineteenth century Europe. In his account, coalition within national economies become stronger over time in the absence of shocks. As they grow stronger, they re-distribute income to themselves, slowing national economic development as a result. The formation of Germany's customs union in 1834, the Zollverein, may have acted as a powerful economic shock, but the unification of Italy, as Foreman-Peck (1995, p. 445) points out, 'does not appear to have worked a similar miracle.' France went through several institutional upheavals while Britain's constitutional change took comparatively less to effect. Still, France's income per capita, even under the most optimistic scenario in O'Brien and Keyder (1978), never reached Britain's. Spain went through 58 governments and 83 ministers of finance between 1868 and 1915, but these shocks were in no way positive influences on Spanish economic development (Platt, 1984, p. 107).

A strand of research to have come off this institutional work is based on the theme of "path dependence." In short, the idea that a unique event or shock - say, a policy or institution - moves an economy into another equilibrium, in which it remains until the event of another shock. The narrow implication here is that development is a function of specific events in previous periods, leaving no room for general explanations of initial economic development other than "history matters" (Nunn, 2014). An early example from the historiography of this work was the traditional explanation of nineteenth-century French economic development, as in Caron (1979). It was argued that the French Revolution led to what would be an adverse re-distribution of land. A resultant more egalitarian income distribution, and political represention among agriculturists, led to savings patterns and industrial policy that was detrimental to industrialisation, compared with Britain. Levy-Leboyer and Bourguignon (1990), however, cast doubt on the contingency of the French Revolution. Their econometric analysis showed that, in fact, agriculture was much more dynamic than industry, contributing most of French growth between 1825 and 1859.

Having accumulated enough quantiative data, starting in the late-1980s, research followed Baumol (1986) and Abramovitz (1986) in trying to show a general tendency to convergence since 1870 of GNP per capita among industrial countries. The underlying thesis was that backward economies have higher growth potential from unrealised re-allocation of labour from agriculture into industry, and by being able to adopt existing frontier technologies. DeLong (1989), however, extended Baumol's original sample to include economies like Chile and Argentina - high incomes in the nineteenth century, but poor development thereafter - finding little evidence of income convergence. Barro (1991), focusing on the post-1960, *did* find evidence of per capita income convergence, and in later work with Sala-i-Martin, gave it a more rigorous theoretical foundation (Sala-i-Martin, X., 1997). This neo-classical model, rooted in Solow (1956), held that convergence was a function of markets being free to allocate factors of production to where they were most needed: poor countries. This implied a negative correlation between initial income and future growth. Their reasoning behind the strong post-1960 convergence is that it captures the trade and financial liberalisation that followed Bretton Woods. The economic history literature tells us, however that the 1870 to 1914 period was an era of rapid globalisation, particularly in the goods and capital markets. Daudin et al. (2010) called in the "highwater mark" of nineteenth century globalisation, O'Rourke and Williamson (1999) and O'Rourke and Williamson (1997) found fast commodity price and real wage convergence, and Bordo (1981) called the "classical gold standard" era, when capital mobility was high.

North and Thomas (1973), in their institutional account of the rise of the West, expressed frustration with the original Solow (1956) model, claiming that factor accumultion is not a cause of growth, but is growth itself. The criticism applies to the more recent work, and has recently been taken up by empirical economists - also institutionalists. Moving away from all-encompassing definitions, recent work emphasises clearly defined property rights and low risks of state-led expropriation drive economic performance (Acemoglu and Dell, 2010;

Acemoglu and Johnson, 2005; Acemoglu et al., 2001). There is a lot of empirical weight behind this formulation of institutions, but it can criticised on the grounds that institutions can be *de facto* and these are rarely measurable, and that institutions are endogenous to economic activity (Chang, 2010). Another potentially serious criticism comes from Redding and Strum (2008), who argue that property-rights institutions are national and so cannot account for subnational income differentials. Not dealing with subnational income differentials in a rigorous way is not only an issue in this branch of the literature. It has, in fact, been a feature for much - not all - of the historiography until recently.

2.3 The modifiable unit area problem

In 1991, the economic historian William N. Parker asked, 'Why isn't European history the history of a continent? Why do we keep all such heavy emphasis on national histories?' (Rhode, 2008, p. 192). While, for example, Cipolla's *Fontana Economic History of Europe* covers specific places and engages in comparative assessments, the underlying assumption is that economic change is best explained by looking at nations (Cipolla, 1976). Even Landes (1969), one of the first economic historians to break the nation's stranglehold in his *Unbound Prometheus*, took in other work for the *Cambridge Economic History of Europe* the British model as a national yardstick with which to measure Continental progress (Landes, 1965). This is not to say that national units are always problematic, but they do lack a spatial perspective that is necessary to understand European economic development during this period.

The implications of using individual nations are considerable. First, using a single country case-study, we cannot be sure that the insights it provides are generalisable. Second, different levels of spatial aggregation can lead to different, possibly incorrect, conclusions - the modifiable area unit problem (Opensahw, 1984). The second is more serious than the first, since the purpose of case studies is the accumulation of evidence. Researchers are aware of the trade-off between internal and external validity, and do not draw general conclusions based on, in effect, one observation. The importance of the second point is highlighted by the fact that many authors, from Gerschenkron (1962) to Acemoglu et al. (2005), using national units of analysis, concluded that national policies were instrumental in economic development. As forcefully argued by Pollard (1973, p. 636-7), however, 'the spread of industrialization' was scattered across nineteenth century Europe 'with little reference to political boundaries.'

Pollard, first with a 1973 article in the *Economic History Review* and then with his 1981 book *Peaceful Conquest*, made the case for disaggregating nations into subnational regions, and for looking at Europe as a whole. Of course, the modifiable area unit problem means one can always argue for a different scale, but subnational regions - like the Rhineland or Piedmont - are a compromise between the ever smaller regions and ever more precision, and data constraints. As Aldcroft (1994, p. 3) wrote, Pollard 'was one of the few writers to acknowledge the extent of "pan-Europeanism" in the development of the European continent.' Unlike the writers who came before him or his contemporaries, Pollard analysed regions of modernisation which were connected to one another by "transmission paths", but were surrounded by underdeveloped, agricultural regions. In short, he argued that nations were not particularly helpful since in the nineteenth century all or most European nations were composed of industrial, agricultural, and stagnant regions. National incomes grew when countries had more of the first than second and third components.

However, Pollard (1981) ultimately remained traditional in using the British model as a benchmark, which the Continent emulated - "diffusion", as I mentioned earlier, being the keyword. In his scheme, explanations for modern regions' successes range from the location coalfields to idiosyncratic events, like wars or government policies. Like much of the "old economic history", it is rich in detail and ideas, but weak in analytical rigour. Further, the last third of Pollard (1981) consists of conventional country-by-country studies, detailing developments within countries with frameworks specific to each one. Pollard justifies this dual treatment - first regions, then nations - by arguing that states were, after 1870, the drivers of economic change, while earlier the development was regional and spontaneous. Problematically, however, regional development continued to characterise European development after 1870 more deeply than before. The rapid industrialisation of Germany's Ruhr area in the late-nineteenth century is clear testament to this. So, even if we accept that states were the main drivers of economic change post-1870, how do we explain the state's differential effects *within* the same countries? While Pollard (1981) recognises regional clusters of industry within and across nations, he fails to account for them in a satisfactory way. Ten years later, we got a theoretical framework that can pick up where Pollard (1981) left off.

2.4 The new economic geography

Krugman (1991) provided an alternative to Hecksher-Ohlin trade theory in forming NEG (see also Fujita et al. (1999b); Krugman (1991, 1993); Krugman and Venables (1995)). The typical NEG framework depends on the interaction between market size, economies of scale, and trade costs. The three interact to form a U-shaped curve, relating trade costs (inclusive of transport costs) on one axis to industrial location and real wages on the other. As there are economies of scale in industry, there is an incentive for production to concentrate in one region and reap those benefits. If trade costs were very high, shipping manufactures between markets will be expensive and so production will be scattered among the core and periphery. However, if trade costs are at an intermediate level, it would be too expensive for the periphery to produce for the larger core markets, but cheap for the industrial core to produce for the small peripheral market. If trade costs were very low, then production would again be scattered among the core and periphery since costs to market would not figure in location choices. This is what produces the U-shaped pattern.

The implications for regional economic activity are that high trade costs first lead to falling incomes in the periphery, as industry moves to the core, before leading to rising incomes when trade costs fall, as industry moves back to the periphery. Further, since the move of industry to the core market increases the size of that market, there are *increasing* returns to concentration. Marshall (1920), who is the source of these ideas, outlines another two reasons for increasing returns. There is also a concentration of useful knowledge that can be implemented in production (think Silicon Valley today), and there will be a concomitant concentration of suppliers to industry.

NEG has proved empirically successful in explaining contemporary international differences in per capita income, European regional differences, and differences between US counties (Hanson, 2005; Redding and Venables, 2004; ?). The key variable in these studies is market potential, which is ignored in most historian's accounts of European regional development (Berend, 2013; Pollard, 1981). That is, the level of economic activity in a region is affected by its potential access to demand markets; the size of demand markets is usually proxied with GDP, the numerator, and access, the denominator, is determined by trade costs (mainly tariffs and transport costs, but some times cultural variables like language). As the numerator increases relative to its denominator, that region becomes a more attractive location for industrial production. While economic historians are usually too data-constrained to measure Marshall's second and third reasons, NEG has still enabled them to succsessfully explain patterns of economic activity within countries using market potential.

Why isn't the whole of Europe industrialised?[1]

Roses (2003) is an early example of the application of NEG in the economic history literature. Looking at Spain between 1797 and 1910, he shows that as transport costs decreased and internal barriers to domestic trade were eliminated, Spanish industry became increasingly concentrated in a few north-eastern regions, Catalonia and the Basque Country. Roses (2003), however, also finds an important role for relative factor endowments, concluding that differences in industrialisation between Spanish regions were totally accounted for by increasing returns *and* comparative advantages. Very much in the same vein, Wolf (2007) shows that after Polish reunification in 1918, when internal tariffs were removed and infrastructure improved, Polish industry re-located based on both comparative advantage - in skilled labour - as well as access to markets. Henning et al. (2011), looking at Sweden between 1860 and 2009, finds the same patterns playing out, but gives more weight to markets during the earlier years of Sweden's industrial-

[1]This is adapted from Roses (2003), a study on Spanish industrialisation, titled 'Why isn't the whole of Spain industrialized?: a new economic geography and early industrialization, 1797-1910'.

isation.

Crafts (2005a) and Schulze (2007b) both shy away from a formal empirical analysis, but reveal the importance of market potential in determining patterns of economic activity. Looking a Britain between 1871 and 1931, Crafts shows that the North, Scotland, and Wales were much more economically peripheral before WWI than in 1985. The main reason, he argues, for their relative GDP decline was changing transport costs that came with the demise of coastal shipping and the rise of road haulage. Schulze (2007b, p. 1), looking at late-nineteenth century Austria-Hungary, shows that being economically peripheral was associated with having lower GDP per capita levels, but that 'there was no uniform relationship between changes in regions' relative GDP position and their market potential.'

Crafts and Mulatu (2006) provide a more formal empirical analysis for nineteenth century Britain, in line with Roses (2003) and Wolf (2007). Broadly speaking, the authors show that spatial patterns of industry during this period were persistent, with regional specialisation changing little. Their econometric analysis shows that relative factor endowments mattered more than market access, although market access did matter for industries with large plant size. In conclusion, they show that transport costs had a negligible effect on industrial location during a period when proximity to coal mattered most. This result is unsurprising. First, their data start in 1841 by when Britain's industrial pattern, as they concede, had long consolidated. Second, British industry, as argued earlier in this review, was unusual in its concentration around coal deposits. Indeed, Wrigley (2010) tried generalising the British experience across Europe, but Spain and Austria-Hunagry, for example, both had a number of coal deposits that either went unexploited or were far from industrial centres. The case of British industrialisation rarely fits Europe's (Pollard, 1981).

In another interesting example, Combes et al. (2011) show that over 1860 to 2000, the patterns of economic activity within France formed the U-shape hypothesised by NEG. That is, their data on departmental value-added per capita confirms the U-shaped evolution of spatial concentration of manufacturing and services. Further, they find that for the 1860 to 1930 period, market potential was the main determinant of location patterns. In the 1930 to 2000 period, when economic activity became more advanced, human capital became the main determinant.

2.5 Conclusion

Recent NEG-work in economic history has provided us with a generalisable explanation of European development, which is lacking in the broad-sweep historiography (for example, Pollard (1981)), but NEG work has so far been restricted to national studies (for example, Roses (2003)). We have taken two steps forward, and one step back to William N. Parker's question, quoted earlier, on why European history 'is not the history of a continent.' The U-shape of spatial economic development uncovered by Combes et al. (2011) for France, and posited generally by NEG theory, did not only exist within France, but was occurring *across* European national borders.

The consolidation of the European "Golden Triangle" marks the point on the *European* U-curve where trade costs were at the intermediate point, making it too expensive for the periphery to produce for core markets, but cheap enough for the core to produce for peripheral markets, further concentrating industry in the core and hence enlarging its market. This typology also fits with the description of continental industrialisation in Mokyr (1998, p. 2) as the culimination of 'economies of scale in manufacturing' - a concept central to NEG. That at the start of the period regional economic activity was more scattered around Europe in what Pollard (1981, p. 123) called 'provincial markets' also lends support to these ideas. It was the point on the U-curve at which trade costs were much higher, causing producers to locate between the core and periphery since shipping to markets was expensive. The dramatic transport cost reductions, that more than made up for increases in tariffs, over the late-nineteenth century are what moved regions along the U-curve (O'Rourke, 1997).

NEG can thus explain not just individual within-country patterns, as recent authors have, but also the broader European patterns of regional industrialisation and income. This is where my dissertation comes in. I collect all available data, fill in the gaps, and explain the resulting dataset using NEG. The end-product, hopefully readers will agree, is a coherent account of Europe's late-nineteenth century regional economic development.

Chapter 3

Explicandum: the Distribution of European Regional Income

3.1 Introduction

Was the distribution of European regional income random? Asking this question is the first step towards an understanding of the causes of European regional development.

Economic historians tend to characterise this period favourably. In his discussion of factor price convergence among the OECD club, Williamson (1996) characterised the 1870 to 1913 period as one of fast growth, globalisation, and convergence. Similarly, Maddison (1995, p. 59,87) characterised it as a 'relatively peaceful and prosperous era' where 'per capita growth accelerated in all regions and in most countries.' These generalisations are helpful in setting the broader context of growth, but obscure too much of the variety of economic experiences within and between countries. Growth was particularly slow in Britain, for example, which is what allowed other countries to converge on it. By contrast, the newly unified Germany was growing fast, but then it contained stark internal differences. While Saxony and the Rhineland were powering ahead, East and West Prussia were left behind, comparable in their income levels to the far-flung regions of the Austro-Hungarian Empire.

Establishing the geographical distribution of income over time is not just about greater historical accuracy. It is a necessary first step towards a causal explanation. Looking at regions, we can narrow down variation to precise patterns and locations, and then ask why economic activity was found where it was. This excludes all the potential confounding variation that might come with, say, looking at a single time-series or using larger geographical units of analysis. Looking at the timing of economic development, we can then find common temporal patterns or shocks, and ask why they happened when they did.

The rest of this admittedly dry chapter is organised as follows. In the following section, I develop a technique to estimate the regional incomes of Germany and France, which are missing from the literature. I then standardise the estimates, and test them for robustness. The subsequent section features what are essentially descriptive exercises on the data. I look at the regional income distributions across the sample and how they change over time. The final section concludes with a discussion of patterns and trends, relating them to the literature.

3.2 Scope and Data

GDP per capita is my measure of income. While this was the age of European industrialisation, income captures variation in economic activity and progress beyond industrial employment shares alone. Even then, in this period, variation in income was highly correlated with that in industrial employment (Broadberry et al., 2010). Income also makes for a measure of economic activity that is comparable across time and place. There are often issues with comparing sectoral employment shares across countries and over time, as categories and industries change, that do not feature as much in international income comparisons. Further, the Geary-Khamis dollars unit I use throughout my dissertation is perhaps the most widely used in the literature, making comparisons with research much simpler.

The income data I use are from the work of a number of country specialists who over the past 10 years have produced a flurry of research on regional GDP per capita. Most of these estimates, for the post-1900 period, have been incorporated into the European Science Foundation-funded project, *the Historical Economic Geography of Europe, 1900-2000*, coordinated by Joan Roses (of Roses et al. (2010)) and Nikolaus Wolf (of Wolf (2007)). The literature now provides regional income estimates for Austria-Hungary (22 regions); Britain (12); Italy (18); Spain (17); and Sweden (24) (Crafts, 2005b; Enflo et al., 2010; Felice, 2009; Roses et al., 2010; Schulze, 2007b, 2011).[1] The research done on other countries does not feature enough data within my period to merit inclusion. The Badia-Miro et al. (2012) estimates for Portugal start in 1890, while Buyst (2009) on Belgium starts in 1896. These countries, along with France and Germany whose incomes I estimate in the following section, make up 93 per cent of European GDP over 1870 to 1910 according to the latest Maddison data base (Bolt and Van Zanden, 2013).[2] Coverage is also reliable in a geographical sense. Regions are not concentrated in particular area of Europe, but are spread from the centre to every corner. This reduces the potential for sample selection bias, since we know that some determinants of income are correlated with geographical variables (Sachs et al., 1999).

After discussing my income estimation method for France and Germany in the following sub-section, I review them along with all the other income estimates. The task here is to ensure the consistency of currencies, prices, borders, and method.

3.2.1 French and German Regional Income Estimation

Given their size and importance, it is perhaps surprising that regional income estimates for France and Germany during this period are scarce. Using this fragmentary evidence, I develop a novel income estimation model, where income is specified as a function of shifts in sectoral employment structure. Its concep-

[1] There has also been a lot of recent work on Asian and Latin America regional income estimation for the period: China, India, and Mexico have all been covered (Aguilar-Retureta, 2014; Caruana Galizia, 2013; Caruana Galizia and Ma, 2014).

[2] Bolt and Van Zanden (2013) is the latest update of the Maddison (2007) version of the data. By 'Europe', I refer to Maddison's Western Europe category plus his Eastern Europe one.

tual basis is uncontroversial, its data requirements are low, and it withstands robustness checks.

Higher income levels in one economy compared to another can indicate two things. First, they can indicate higher labour productivity due to higher capital and labour ratios or better technology. Second, they can indicate a more efficient allocation of labour among economic activities. It is, I maintain throughout this section, the latter that really counts. European economic history has shown that countries that remained heavily committed to agriculture remained relatively poor, while those that reallocated labour to the industrial and services sectors became relatively wealthy. In fact, Broadberry et al. (2010) observe a negative relationship between the level of per capita incomes and the share of the labour force in agriculture for a sample of 14 European countries between 1870 and 1913.

Relationships between sectoral distributions of labour and incomes are hardly news. Studies on the subject have a long history, going back to, for example Fabricant (1942) and Good and Ma (1998). Another strand of research has used correlates of GDP to produce estimates for international samples (Chenery and Syrquin, 1975; Crafts, 1984).

More recently, in the economic history literature, the relationship has been exploited to derive income estimates for a number of European countries, starting with Geary and Stark (2002), which put forward a short-cut method for estimating regional GDP based on sectoral employment and wages. The method uses national GDP estimates, breaking them down according to regional employment structure and corresponding regional-sectoral wages. This is the estimation method underlying most of sample drawn from the literature. However, as Wolf (2010) writes of Germany, such specific data are hard to come by. The approximations required (say, using national wages or city wages to represent regions) make it not much more rigourous than the method I present here. Furthermore, my method is particularly relevant for France because, as O'Brien and Keyder (1978, p. 98) point out, French backwardness was a 'failure to realise re-allocation of labour from agriculture to industry.'

As we do have census book data on regional population and sectoral employment for both countries, as well as a single regional cross-section of income (from Frank (1994) for Germany and Combes et al. (2011) for France), an empirical implementation of the concept that structural change determines income level would be

$$\begin{aligned}\ln GDP_i = & \alpha + \beta_1 \ln P_i + \beta_2 \ln \frac{A_i}{LF_i} + \beta_3 \ln \frac{I_i}{LF_i} \\ & + \beta_4 \ln \frac{S_i}{LF_i} + \varepsilon_i\end{aligned} \tag{3.1}$$

where subscript i indexes regions, α is a constant term, and ε is a random error term. The dependent variable is regional GDP, and independent variables are population (P), agricultural (A), industrial (I), and services (S) share of the labour force (LF) all at the regional-level. Taking logarithms on both sides, this model will produce the elasticities at which GDP responds to structural shifts and population change. As in Crafts (1984), I include P to control for regional size effects that would not be captured by sectoral labour force *shares* alone.

Necessarily, I had to drop one of the terms of the right of the model. This was because of multicolinearity issues: the three sectors, which would sum to the total labour force, are a linear combination of each other. I dropped $\frac{A}{LF}$: a backward stepwise regression procedure showed it to be statistically insignificant for both countries. The refined implementation is as follows

$$\ln GDP_i = \alpha + \beta_1 \ln P_i + \beta_2 \ln \frac{I_i}{LF_i} + \beta_3 \ln \frac{S_i}{LF_i} + \varepsilon_i \tag{3.2}$$

where all variables are previously defined. I expect the signs on the coefficients to be positive and significant. Population growth boosts aggregate demand and allows for the division of labour. Industry was the main driver of growth in income, so growth in the industrial labour force should be associated with increasing income levels. The service sector is composed of high value-added sub-sectors, such as the legal profession, as well as low value-added ones like the army, navy, and domestic and catering subsectors. While positive, I expect its magnitude to be smaller than the coefficient on industry. Applying the estimated elasticities $\beta_{1...3}$ in a linear transformation to the independent variable from one year to the next allowed me to project GDP from 1860 (benchmark of available regional GDP cross-section) to 1911 for France and, following the same process, from 1907 (benchmark of available regional GDP cross-section) back to 1871 for Germany.

I explain the procedure in more detail in the following sections.

Data

German census books provide sectoral employment by *Land* (or state; the regional unit I use) for the benchmark years 1871, 1882, 1895, and 1907 (Deutsch Statistischen Bureau, 1871, 1882, 1895, 1910, 1912). Frank (1994, p. XXX) provides a per capita GDP cross-section for 1907/13 in constant 1913 marks by district (subdivision of *Länder*). French census books provide sectoral employment by department (the regional unit I use) for the benchmark years 1872, 1886, 1901, and 1911 (Statistique de la France, 1872, 1886, 1901, 1911). Combes et al. (2011, p. 22) provide sectoral value-added cross-sections for 1860 in current francs. These data alone, following some adjustments that I explain in detail below, allow me to exploit the structural change-GDP relationship.

Germany The first step was defining a region or *Land*. I multiplied the district GDP per capita series in Frank (1994) by population by district (taken from the censuses), to produce total GDP by district. I then aggregated these districts into *Länder* according to what the *Länder* administrative boundaries of the time were. Frank's series misses a few free states and minor principalities. I either left these as standalone units, as with Brunswick and Hamburg, or else put them into the nearest *Land*, as with Lubeck in Schleswig-Holstein. I aggregated in this way as the method is not detailed enough to deal with rapid shifts or fluctuations in employment structure that you so often get with very small regions, as the result of temporary internal migrations or boundary changes. Some *Länder*, and their districts, are altogether missing from Frank's series, like Alsace-Lorraine. I included these in the present series, and to get their GDP levels, since Frank does not provide them, I scaled GDP as a function of population, according to the sample average. Scaling according to population is in fact reliable as population showed a consistent and strong positive correlation with income across all benchmark years in all *Länder*. Furthermore, the sum of GDP of this new *Länder* series (including my additions) is only around 5 per cent higher than the widely used national level GDP for 1907/13 from Hoffman (1965) . Still, to ensure accuracy in the empirical stage, I scaled this new series according to the Hoffman data average for 1907/13, yielding a *Land* cross-section of GDP for 1907/13 in constant 1913 marks.

Frank's GDP per capita series for the remaining years (1882, 1895) proved to be too unreliable for use here.[1] By Frank's own admission, the estimates are clearly too high.[2]The GDP levels for these years were derived by keeping wages constant at their 1913 level, and simply multiplying them by their corresponding employment figures for the remaining three years. This overstates output. When I multiplied Frank's district GDP per capita estimates by district population, the resulting total GDP estimates summed to a figure far higher than existing national GDP estimates (Hoffman, 1965; Maddison, 2007). This can be seen in table 3.1, which compares Hoffman's estimates for national GDP, and national GDP estimates derived by multiplying Frank's GDP per capita estimates by population, and then summing.

Table 3.1: Reliability of Frank's income figures for years other than 1907/13.

	1882	1895
Frank (derived)	26,136,719,191	32,812,766,867
Hoffman	18,441,000,000	27,621,000,000
Difference %	41.2	18.8

Source: own calculations, based on census books, Frank (1994), and Hoffman (1965).

The next step was collecting and aggregating sectoral employment data for the corresponding *Länder*. For the years 1907, 1895, and 1882 this was simple. The census books list six-sectors: (1) agriculture (including fishing, forestry and related industries), (2) industry (including manufacturing like textiles and chemicals), (3) trade and commerce (including transport and communication), (4) professional workers and the civil service (including lawyers and government bureaucrats), (5) army and navy, and a (6) residuary category of "other occupations." Conventionally, I grouped sectors (3) to (6), to produce a single services sector in the contemporary sense (Broadberry et al., 2010). The year 1871 proved a little trickier. The book for this year listed seven-sectors: (1) agriculture, (2) industry, (3) trade and commerce, (4) wage labourers (including workers like farmhands), (4) army and navy, (5) professional workers and the civil service, (6)

[1]Frank provides estimates for 1849 and 1939 as well, but besides also being unreliable, these figures are beyond my period . He provides no figures for 1871.

[2]See Frank (1994, p. XXX).

the same residuary sector of “other occupations”, and (7) a sector listing what should be translated as unemployed persons. Following Schulze (2007a, p. 212), the first necessary adjustment was grouping (1) and (4). The sum of persons in (1) makes clear that census enumerators disaggregated the agricultural labour force into permanent workers, and the daily wage labourers listed in (4). Sector (7) was altogether dropped: there is no comparable data of unemployed persons for the other years. The proportion of unemployed as part of the national labour force was six per cent. The last adjustment was grouping (3), (4), (5), and (6), to produce a service sector comparable to the one found in later benchmark years.

These adjustments and groupings yield intuitive results, as table 3.2 shows. Expectedly, agriculture shows a very large drop in its share of the labour force, from 50 to 37-per-cent, while industry and services are making rapid gains. Services grew by 40-per cent.

Table 3.2: National sectoral percentage shares in national labour force.

Year	Agriculture	Industry	Services
1907	37	42	21
1895	40	40	20
1882	47	36	17
1871	50	35	15

Source: own calculations, based on census books.

France Here the process is more straightforward. Combes et al. (2011) provide cross-sections of gross value-added (GVA) in current francs for the agricultural, industrial, and services sectors for 1860. These series are largely drawn from the painstaking work of Toutain (1981, 1987), who has spent a number of years building up very detailed production estimates of the French economy. I summed the three sectoral cross-sections to produce a single GVA series for 1860. Combes et al. (2011) also present value-added series for 1895, but leave out a services series due to data unavailability. This made the calculation of total GVA series for that year impossible.

The arguments that GVA is incomparable to GDP are unconvincing here. The relationship between the two is defined as *GVA+(taxes on products - subsidies on products)=GDP*. The reason why Toutain (1981) used value-added to build up departmental estimates is that the total aggregates of product taxes and subsidies are only available on the national economy level. Applying taxes and subsidies would not affect cross-sectional differences, as taxes are mostly, if not all, national. It does not, during this period, substantially affect the estimates in absolute terms either: summing the national 1860 GVA cross-section closely matched Toutain's GDP (or *product interieur brut*) for that year. From here on I shall refer to the series as GDP to avoid confusion. Conveniently, Toutain (1987) also provides a GDP deflator with a 1905/13 base year, which I used to deflate the GDP series after estimation.

Combes et al. (2011) provide data on population and sectoral employment for 1860. I took these data for the remaining benchmark years from the census books (Statistique de la France, 1872, 1886, 1901, 1911). The books covered agriculture (*professions agricoles*) and industry (*professions industrielles*) as two broad categories over all benchmark years. As they contained all the standard sub-divisions (for example, forestry under agriculture and chemicals under industry), I stuck with these aggregations. Services were slightly more complicated. Until the turn of the century, services were split into the following: *transports*, *commerce* (retail and wholesale trade), *professions libérales* (lawyers and doctors), *force publique* (army and navy), and *administration publique* (civil service). These categories together form what we nowadays think of as the services sector. I aggregated them under the latter. After the turn of the century, the categories were as follows: *professions commerciales* (same as *commerce*), *professions libérales* (same as before), *service domestique* (domestic workers such as maids), *services publics administratifs* (civil services and public employees), and *personnes non classes suivant la position* (a residual services category that includes the army and navy, as well as professions like merchant seamen). These categories are in essence different groupings of the previous categories, and so I aggregated them under the services sector. In all the sectors in all the census books, I ignored the field that records unemployed persons numbers (*sans emploi*), as I did with Germany. I also ignored a field that recorded all the dependants of workers in each sector, that is, the families and whoever else might be living off the workers' wages (*familie des précédents. Parents à tous degrés et autres vivant avec les précédents*). Both of these fields inflate the labour force of department with persons who are not economically active.

While census enumeration strategies were constant across departments, they were not constant across years. For example, in some years, labourers' dependents were counted along with labourers, but counted separately in other years. Whatever the case, the same strategy held across all departments within the same year. That is, census data are reliable in the cross-section, but not in the time-series. Marchand and Thelot (1991) correct this time-inconsistency for national-level data. For example, to estimate the "true" agricultural labour force, the authors extrapolated the number of workers from the rural working-age population as a constant ratio of the agricultural labour force. This ratio survived a battery of statistical tests, and they used similar methods to arrive at new numbers for industrial and services employment, as well as population. Their work is widely accepted as the standard reference for this sort of data. To apply their correction to departments, I expressed census departmental figures as proportions of census national figures, and then multiplied these proportions by the Marchand and Thelot (1991) national figures. This straightforward ratio scaling thus preserves the original cross-sectional differences (found in the census books), and makes the cross-sections comparable over time (by taking into account Marchand and Thelot's statistical corrections). The importance of correcting and scaling in this way diminishes as census books near the second half of the 20th century. Already for 1930, for example, the difference between the census national totals and Marchand and Thelot's are only: 0.72 per cent (population), -5.80 per cent (agriculture), 2.78 per cent (industry), 8.22 per cent(services). In later census books the differences become unsubstantial.

It is important to point out here that I excluded Corsica from the analysis, as Combes et al. (2011) do not cover it. Generally, it is not normally included in studies on France, as technically speaking it is a collectivite territoriale rather than a department. Bas-Rhin and Haut-Rhin were annexed by Germany between 1871 and 1918, which is why I drop them after 1860. A small part of Haut-Rhin remained French, as Territoire Belfort. I grouped this relatively small piece of land under Haut-Saone, its nearest neighbour. Combes et al. (2011) group Meurthe and Moselle into contemporary $Meurthe-et-Moselle$ for 1860. From 1871 to 1918, both of these regions were annexed by Germany. About a fifth of Moselle - the iron rich part, regrettably - and about two-thirds of Meurthe were left behind. I maintained the contemporary $Meurthe - et - Moselle$ grouping as in Combes et al. (2011).

Empirical results

The results of model 3.2 are displayed in table 3.3. It is reassuring that the model produced very similar results for both countries even though the cross-section years are different and the sample size for Germany is much smaller. This indicates that model is applicable across samples (countries). Secondly, if we are to accept this (so far limited) applicability, it indicates that the sizes of elasticities do not change radically over the couple of decades I cover. There is a difference of 0.2 between the industrial elasticites ($\frac{I}{LF}$), which is expected given German industry was among the most productive industry on the continent at the time.

Table 3.3: Regional GDP estimation model

	France, 1860	Germany, 1907
P	1.046***	1.038***
	(0.050)	(0.026)
$\frac{I}{LF}$	0.325***	0.527***
	(0.053)	(0.057)
$\frac{S}{LF}$	0.151**	0.156**
	(0.063)	(0.067)
$Constant$	6.405***	6.662***
	(0.706)	(0.387)
$\hat{R}^2$	0.92	0.98
F-Stat.	332.1***	672.6***
N	87	23

*Notes: Statistical significance: *** 1 per cent, ** 5 per cent. Standard errors in brackets. A Breusch-Pagan test for heteroskedasticity failed to reject the null hypothesis of constant variance, with a χ^2 value of 2.7 and a probability value of 0.101 for Germany, and a χ^2 value of 0.33 and a probability value of 0.566 for France.*

In both countries, the sizes and significance of the elasticities are intuitive. Population was growing fast during this period, and showed a strong positive correlation with income, as in previous historical studies Williamson (1998). The mechanisms behind are likely to be population growth both boosting aggregate demand and allowing for the division of labour. Chenery and Syrquin (1975, p. 17) argue that the mechanisms captured by the elasticity are likely to be economies of scale and transport costs. The elasticities on industrial employment and commercial employment shares are also positive. The industrial elasticity indicates a substantial effect on income, which is also expected during this time of industrialisation. The size of the services elasticity is much lower, at least in part, because the sector includes unproductive sub-sectors such as domestic services and the army.

The model shows high explanatory power, but with an $\hat{R}^2$ of 0.98, the result for Germany looks suspicious. The usual reason for this is multicolinearity, but as table 3.4 shows, there is no correlation among the independent variables. France's $\hat{R}^2$ of 0.92 is high, but not high enough to indicate any statistical issues.

Table 3.4: Correlation matrix of German-model independent variables.

	P	$(\frac{I}{LF})$	$(\frac{S}{LF})$
P	1		
$\frac{I}{LF}$	0.235 (0.281)	1	
$\frac{S}{LF}$	-0.335 (0.118)	0.051 (0.818)	1

Notes: p-values in brackets.

It is worth discussing an early paper by Orsagh (1968) here. The paper uses a similar top-down method to the one I propose here, to provide regional GDP estimates for Germany between 1882 and 1963. There are, however, a number of differences. First, Orsagh leaves all three sectoral terms in his regression equation. While the sectoral terms are expressed as proportions of the national - rather than regional - labour force, this is still a serious mis-specification error. The sectoral terms are correlated. I produced a correlation matrix (table 3.5) of the variables as defined by Orsagh and the correlation coefficients are strong and significant. As we cannot ascertain the correct GDP elasticity of each sectoral term because of this multicolinearity, we cannot reliably project GDP. Secondly, Orsagh does not include a "size" control in his specification (I used population), but concedes that it is necessary to go beyond shares of employment alone. He resorts to a complicated counter-factual exercise on the elasticities. The exercise, though, is pointless: the elasticities have already been derived from an equation that leaves out a "size" control and so they are likely to suffer from an over or under-estimation bias.

As the model worked equally well on a sample size of 87 (France) as well as one of 23 (Germany), I expect that these new estimates are open to three main criti-

Table 3.5: Correlation matrix of Orsagh's independent variables.

	Services	Industry	Agriculture
Serives	1		
Industry	0.878 *(0.000)*	1	
Agriculture	0.565 *(0.004)*	0.538 *(0.008)*	1

p-values are in brackets.
Source: own calculations, based on Orsagh (1968).

cisms, though I am sure readers will think of a few more. First, I am venturing too far out of the range of evidence and the GDP estimates themselves can be either too high or too low. Second, using elasticities estimated from a cross-sectional model to create a panel dataset (one with a time dimension) is unreliable, and keeping productivity elasticities fixed over time is unreliable. Lastly, the model can be criticised for having limited applicability, that is, we need a replication test.

Testing levels

That the specific GDP levels I present here are simply too high or low is perhaps the most fundamental criticism that can be leveled. Besides going on intuition, and the estimates *do* work intuitively (generally, agricultural regions have lower incomes and industrial regions have the highest incomes), the only possible check is to see whether the sum of projected regional GDP estimates add up to a widely used and carefully calculated national GDP estimate.

Table 3.6 shows the sum of my *Länder* GDP estimates and the corresponding national estimates from Hoffman (1965). It is important to point out here that the Hoffman figure for 1907 is actually an average for the years 1913 to 1907, as Frank used such an average to construct his 1907/13 marks per capita by district cross-section used for this paper's estimates.

Table 3.6: German projected and observed national GDP levels.

	1871	1882	1895
Hoffman	14,653,000,000	18,441,000,000	27,621,000,000
Own	15,899,601,418	19,907,875,916	29,687,580,840
Difference %	8.51	7.95	7.48

Source: own calculations and Hoffman (1965). figures are in 1913 marks.

It is clear that there is no substantial difference at the national level: the biggest deviation from Hoffman's estimates is for 1871 at 8.51 per cent, which is encouraging. In their seminal "short-cut" regional income estimations for the UK and Ireland, Geary and Stark report that their "best" specification estimates deviate from official estimates by a maximum of 7.5 per cent for one region.

Table 3.7 displays the results of the same exercise for France. Toutain (1987) provides a national GDP series in 1905/13 francs, as well as the GDP deflator that I used deflate my own estimates. In the case of France, the deviations are all still below 10-per cent, making them tolerable deviations by national accounting standards.

Table 3.7: French projected and observed national GDP levels.

	1872	1886	1901	1911
Toutain	24,603,300,000	27,984,300,000	36,090,300,000	42,159,700,000
Own	25,171,854,313	30,505,442,294	33,606,543,691	36,605,587,402
Difference %	2.31	9.01	-2.03	-9.33

Source: own calculations and Toutain (1987). Figures in 1905/13 francs.

This test, of course, does not tell us much about the *distribution* of national GDP among regions. It is impossible to check that because no similar regional GDP estimates exist. The test does show, however, that the regional estimates are well within the frame of a national total and so are unlikely to be far off from their "true" regional values.

Testing parameter reliability

Are elasticities derived from a cross-sectional model useful when it comes to projecting backwards and forwards in time? The assumption here is that the elasticities will be the same, or very similar, in both the cross-section and at different points in time. To test this, I used the elasticities to re-create the widely-used annual national GDP time-series in Hoffman (1965) for Germany, and in Toutain (1987) for France. The same method of projection is employed: I apply the elasticities to national level data on population and employment. If the model's elasticities are reliable in the cross-section, they should also be reasonably reliable in a time series. This test also reveals whether it is reliable to keep the elasticities fixed, that is to assume constant productivity, over the period.[1]

[1]Why not use a time-series model on Hoffman or Toutain's data and apply those coefficients

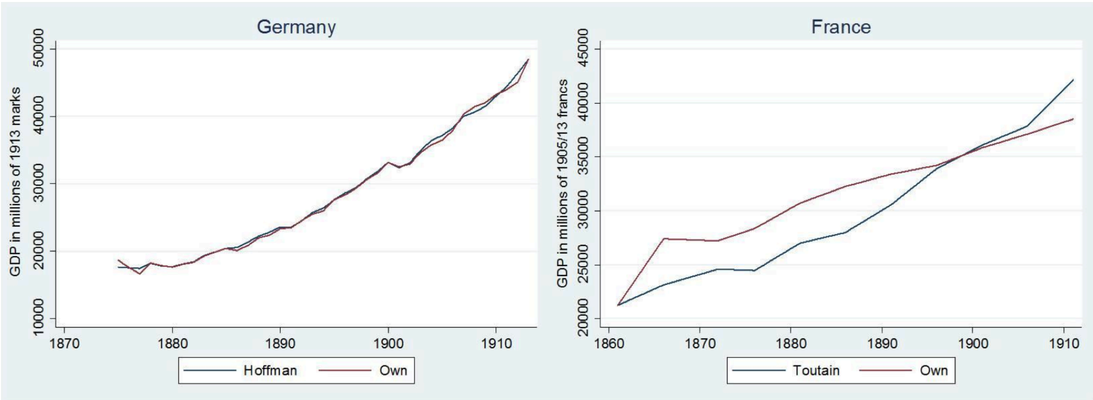

Figure 3.1: Estimated versus observed national time series for France and Germany

Notes: Own calculations based on Toutain (1987), Marchand and Thelot (1991), and Hoffman (1965). All data are in real terms. "Own" refers to the estimates; "Toutain" refers to the "observed" GDP series for France from Toutain (1987); and Hoffman refers to the "observed GDP series for Germany from Hoffman (1965). Germany: starting year is 1875 and ending year, from which the projections were made, is 1913. The time-series is annual. France: starting year, from which the projections were made, is 1861 and ending year is 1911. The benchmark years are: 1861, 1866, 1872, 1876, 1881, 1886, 1891, 1896, 1901, 1906, 1911.

Figure 3.1 displays the results for both countries. For Germany, there is a gap between 1875 and 1871 in the series in Hoffman (1965). In projecting backwards, I therefore stopped at 1875. Data on population and employment is also taken from Hoffman (1965). For France, I took data on national population and employment from Marchand and Thelot (1991), and data on national GDP from Toutain. Marchand and Thelot are only able to provide employment data for 11 benchmark years between the period, so the exercise is limited to those 11 years. It is impossible to re-create an annual time-series as with Germany, but the point of the exercise can still be executed. The starting year is 1861 and the end year is 1911.

The average deviation between the projected and the Hoffman series is 0.67 per cent. The correlation coefficient is 0.998, significant at one per cent. The average deviation between the French projected series and the Toutain series is 6.15 per cent. The correlation coefficient is 0.942, significant at one per cent.

to the cross-section? Due to unsurprising problems of autocorrelation, the time-series model was impossible to estimate. Even when correcting using the Cochrane-Orcutt procedure, the model produced spurious, unreliable results that did not withstand robustness tests.

While figure 3.1 shows that the method at times under and over-estimates national GDP levels, it manages to replicate GDP levels within tolerable standards. These results show that the elasticities used here are reliable in capturing the time dimension, and so the cross-sectional model is useful in projecting GDP estimates backwards in time. This supports the assumption of stable sectoral productivity over the period.

Replicating existing regional income estimates

Here I use the method to replicate Sweden and Austria-Hungary's regional GDP levels, using the data underlying Enflo et al. (2010), and Schulze (2007b) (Austria) and Schulze (2011) (Hungary). The countries are a useful choice for three reasons. First, they are composed of 24 and 22 regions respectively, which make for a smaller number of observations than France but a similar one to Germany, and larger still than other countries in the literature. For example, Britain is composed of 12 regions and Italy of 15 regions (Crafts, 2005b; Felice, 2009). Second, Enflo et al. (2010) estimated regional GDP levels using the Geary-Stark method, which is the most popular "top-down" estimation method in the literature Geary and Stark (2002). Schulze used a "top-down" approach for the Hungarian regions, but a "bottom-up" one for the Austrian regions. This makes for a useful comparison of methods. Third, the authors estimate regional GDP cross-sections for every decade in the period I am analysing here.

Sticking to the same approach, I ran model 3.2 on the 1860 cross-section of Swedish regional GDP in 1910/12 kronor and on the 1910 cross-section of Austro-Hungarian regional GDP in 1990 dollars, and then used these elasticities to project regional GDP levels for every decade up until 1910 for Sweden and back to 1870 for Austria-Hungary. The choice of the estimation benchmark year does not change the results. I am working with an early-nineteenth century (Sweden) and late-nineteenth century (Austria-Hungary) to show the technique's flexibility.

The model results for this experiment are in table 3.8. For Sweden, the sizes of the elasticities are very similar to those for France and Germany, presented in table 3.3. For Austria-Hungary, the size of the population coefficient is smaller, but the coefficients on industrial and services employment are in the same orders of magnitude as in the estimations of the other countries. Across all countries, the hierarchy of coefficients, with the largest size going to population, then industry, and lastly services, is the same. Sweden's $\hat{R}^2$, at 0.85, is slightly lower than France's 0.92, and Austria-Hungary's is a large 0.98, similar to Germany's. These are of course all high figures, and the overall similarity of the model's results for these countries lends support to the technique's wider applicability. There is, however, one difference. Sweden's services elasticity has a p-value of 0.207. Given that this elasticity was also the "weakest" - but still significant - for Austria-Hungary, France and Germany, this result for Sweden is unsurprising. Perhaps it is telling us something about the construction of the services sector variable; that it mixes too much of the high value-added with the low. This is the first time my method has been applied to secondary data, so it is difficult to judge whether this one p-value invalidates the method or is just a "Swedish fluke" - it is more likely to be the latter, given the coefficient comes out as statistical significant for the other three countries. One way to find out if it makes regional GDP estimation unreliable, seeing as we have the (Enflo et al., 2010) constructed estimates, is to go ahead with the estimation procedure.

Table 3.8: GDP estimation model replication test

	Sweden, 1860	Austria-Hungary, 1910
P	0.901***	0.446***
	(0.095)	(0.072)
$\frac{I}{LF}$	0.321**	0.444***
	(0.174)	(0.036)
$\frac{S}{LF}$	0.241	0.089**
	(0.185)	(0.039)
$Constant$	7.78***	8.930***
	(1.203)	(0.461)
$\hat{R}^2$	0.85	0.98
F-stat.	43.97***	457.3***
Breusch-Pagan	0.09	0.62
N	24	22

*Notes: Statistical significance: *** one per cent, ** five per cent. Standard errors in brackets. A Breusch-Pagan test for heteroskedasticity failed to reject the null hypothesis of constant variance, with a χ^2 value of 0.09 and a probability value of 0.8 for Sweden; 0.62 and 0.43 for Austria-Hungary.*

Table 3.9 shows the ratios of my estimates to those of Enflo et al. (2010). The cross-sectional average ratios are as follows: 0.91 (1870); 0.99 (1880); 1.02 (1890); 0.89 (1900); and 0.81 (1910). While there are regional ratios that are far from these averages, the general result is a reassuring one. Most ratios are close to 1. As can also be seen in table 3.9, we see similarly positive results on the national aggregate level. The ratios of my national aggregates (sum of all regional estimates) to those of (Enflo et al., 2010) are as follows: 0.95 (1870); 1.03 (1880); 1.07(1890); 0.96 (1900); and 0.90 (1910). These differences on the national level are similarly small to the those reported for France and Germany in table 3.6.

Table 3.9: Ratios of own regional GDP to Enflo et al.

Region	1870	1880	1890	1900	1910
Stockholms lan	1.27	1.24	1.27	1.37	1.34
Uppsala lan	1.16	1.36	1.45	1.19	1.00
Sodermanlands lan	1.11	1.17	1.15	1.13	0.95
Ostergtlands lan	1.15	1.27	1.13	1.04	0.89
Jonkopings lan	0.96	0.82	0.97	0.93	0.76
Kronobergs lan	0.77	0.73	0.72	0.68	0.55
Kalmar lan	0.97	1.02	0.89	0.71	0.61
Gotlands lan	0.79	0.90	0.55	0.51	0.39
Blekinge lan	0.94	0.82	0.81	0.68	0.63
Kristianstads lan	1.09	1.17	1.04	0.91	0.77
Malmohus lan	0.87	0.89	0.98	0.95	0.84
Hallands lan	0.89	1.03	1.25	0.93	0.86
Goteborgoch Bohus lan	0.97	1.29	1.45	1.22	1.29
Alvsborgs lan	0.94	1.05	1.14	0.91	0.86
Skaraborgs lan	0.86	0.89	0.98	0.75	0.65
Varmlands lan	0.96	1.06	1.05	0.97	0.89
Orebro lan	0.78	1.04	1.04	0.90	0.86
Vastmanlands lan	1.01	1.06	1.13	0.88	0.91
Dalarnas lan	0.56	0.69	0.67	0.64	0.54
Gavleborgs lan	0.86	0.97	1.06	0.94	0.87
Vasternorrlands lan	0.96	1.07	1.21	0.98	1.02
Jamtlands lan	0.84	0.82	1.05	0.92	0.82
Vasterbottens lan	0.70	0.84	0.77	0.59	0.57
Norrbottens lan	0.51	0.62	0.61	0.55	0.59
Sweden	0.95	1.03	1.07	0.96	0.90

Table 3.10 shows the ratios of my estimates to those of Schulze (2007b). The cross-sectional average ratios are as follows: 1.29 (1870); 1.28 (1880); 1.17 (1890); 1.06 (1900); and 1 (1910). While there are regional ratios that are far from these averages, the general result is a reassuring one. Most ratios are close to 1. As can also be seen in table 3.9, we see similarly positive results on the national aggregate level. The ratios of my national aggregates (sum of all regional estimates) to those of (Enflo et al., 2010) are as follows: 1.3 (1870); 1.27 (1880); 1.16 (1890); 1.06

(1900); and 0.99 (1910). These differences on the national level are similarly small to the those reported for France and Germany in table 3.6, but the gap widens with the distance from the estimation year in this case.

Table 3.10: Ratios of own regional GDP to Schulze

Region	1870	1880	1890	1900	1910
Bohemia	1.43	1.34	1.27	1.14	1.04
Bukovina	1.13	1.25	1.13	1.04	1.01
Carinthia	1.33	1.42	1.31	1.05	1.02
Carniola	1.50	1.51	1.37	1.13	1.09
Croatia-Slavonia	1.29	1.37	1.31	1.14	1.12
Dalmatia	1.16	1.11	1.18	1.21	1.07
Danube Left Bank	1.28	1.33	1.16	1.15	1.10
Danube Right Bank	1.40	1.39	1.12	1.07	0.99
Danube-Tisza Basin	1.38	1.33	1.15	1.16	1.09
Galicia	1.13	1.12	1.07	1.00	0.94
Littoral	1.39	1.41	1.37	1.18	1.08
Lower Austria	1.18	1.14	1.08	0.93	0.88
Moravia	1.48	1.39	1.31	1.12	1.06
Salzburg	1.06	1.07	1.03	0.88	0.84
Silesia	1.69	1.62	1.48	1.35	1.21
Styria	1.20	1.17	1.07	0.92	0.91
Tisza Left Bank	1.22	1.12	0.98	0.97	0.91
Tisza Right Bank	1.29	1.30	1.08	1.06	0.98
Tisza-Maros Basin	1.25	1.20	1.00	1.02	0.95
Transylvaina	1.15	1.13	0.98	0.99	0.95
Tyrol/Voralbg.	1.28	1.23	1.27	1.03	0.98
Upper Austria	1.20	1.15	1.03	0.92	0.87
Austria-Hungary	1.30	1.27	1.16	1.06	0.99

As this method is meant to approximate regional income rankings, rather than provide precise point estimates of income for each region, what really matters is here are the correlations between the original and replicated series. Leaving aside the starting year of 1860, the two Swedish national series are correlated at 0.99 as are the Austro-Hungarian ones. The average deviation of the estimated Swedish series from the observed is 7.31 per cent; 15 per cent for Austria-Hungary, which is slightly above the usual 10 per cent threshold for national accounts. More importantly, the technique managed to preserve cross-sectional differences, which is, again, its main purpose. The Spearman rank correlations between my estimated Swedish cross-sections and those of the authors are all significant at the 1 per cent level: 0.92 (1870); 0.94 (1880); 0.89 (1890); 0.89 (1900); and 0.90

(1910). For Austria-Hungary, the coefficients are also all significant at the 1 per cent level: 0.99 (1870); 0.98 (1880); 0.98 (1890); 0.99 (1900); and 0.98 (1910).

This replication test highlights the robustness and applicability of the technique. Using a single cross-section of regional GDP from 1860 for Sweden and 1910 for Austria-Hungary, along with regional employment shares and population levels, the technique has replicated the authors' estimates, within a reasonable degree of accuracy. The authors themselves arrived at their estimates using the Geary and Stark procedure as well as a "bottom-up" approach in the case of Schulze (2007b). Even Geary and Stark's "top-down" procedure requires more data: national GDP, regional sectoral employment, regional population, and, most problematically (as they are hardest to find) and distinct from this technique, regional sectoral wages and value-added Geary and Stark (2002). Moreover, in the case of Sweden it yielded robust results despite the services coefficient p-value of 0.207. Whether this means that such a level of statistical significance is tolerable for this method or that this Swedish case is a "fluke" is up to future work to decide. It is only by applying the method to a number of different countries that we can get a really solid idea of what the bounds of the coefficients and their significance should look like.

Finally, *Appendix A* includes a comparison of my French and German regional GDP estimates to those kindly supplied by Nikolaus Wolf, which were not available to me at the time of writing. The comparison shows, despite different estimation methods, that both sets of estimates are similar.

3.2.2 Standardisation

Table 3.11 summarises my regional GDP dataset. The sample consists of 199 regions from seven countries. The benchmark years do not match perfectly, but are never too far off one another to make general cross-sectional comparisons difficult. Some estimates come in real terms, making their conversion to a standard unit straightforward. For the others, deflation is required before conversion.

Deflation of nominal estimates

Ideally, I would have regional price data with which to deflate regional GDP es-

Table 3.11: Regional income dataset.

Country	Years	Method	Currency	Regions
France	1911, 1901, 1886, 1872	TD	Francs	85
Germany	1907, 1895, 1882, 1871	TD	1913 Marks	23
Spain	1910, 1900, 1860	TD	Pesetas	17
Britain	1911, 1901, 1891, 1881, 1871	TD	Pounds	12
Italy	1911, 1901, 1891, 1881, 1871	TD	2001 Euros	16
Austria-Hungary	1910, 1900, 1890, 1880, 1870	BU/TD	1990 Dollars	22
Sweden	1910, 1900, 1890, 1880, 1870, 1860	TD	1910/12 Kronor	24

Notes: 'Years' column shows benchmark years at which GDP data are available; 'Method' is GDP estimation method, where 'TD' is top-down and 'BU' is bottom-up. Britain includes Ireland. For sources, see references in text.

timates. There are two problems with this. First, collecting regional price and quantity data for my sample would be a slightly unrealistic endeavour, even if we are sure that the data exist to begin with. Second, there is a more fundamental methodological concern. Apart from the estimates in Schulze (2007b), most estimates are derived using the top-down Geary-Stark method, which scales regional estimates to a given *national* total. Applying regional price indices would invalidate the scale of regional estimates. Furthermore, authors often proxy regional wages using neighbours' wage levels or national averages, making the use of a regionally-specific price index inconsistent.

These points aside, I am unaware of any evidence that would have us believe regional price variation within European countries during this period was so great that it substantially affected cross-sectional regional GDP differences. As Felice (2009, p. 4) writes,

> *… looking at some specific price data, house rents, some basic foods, in Giusti (1914) and Maic (various years)*[1] *respectively, in a few towns, it seems plausible that in the years previous to World War I consumer price levels were not so different across the Italian regions, for sure not clearly higher in the North.*

Felice mentions the North because it was much wealthier than the South of Italy. Its relative wealth was not, as Felice indicates, the result of relatively lower

[1]This refers to the numerous publications of the Ministry of Agriculture, Industry, and Commerce. For example, Ministero di Agricoltura (1887)

prices. It was the result of there being more industry in the north. Felice's argument is backed up by the sporadic data we have on regional prices. Between 1870 and 1877, the coefficient of variation on wheat prices between Brescia, Padua, and Rome was an average of 0.06 (Jacks, 2005). Looking beyond Italy, wheat price variation across Sweden between 1870 and 1914 is slightly higher at 0.07 (Jorberg, 1973). Flour price data covering 12 British cities in 1872 show variation was 0.06 (Ward and Devereux, 2003). While these are all the data we have to go on, they do indicate that regional price variation was on average low; at least not high enough to re-order the rankings of relative regional GDP levels, which is the potential issue. Given these constraints, I deflated the regional estimates for Spain (to 1958 pesetas) and Britain (to 1913 pounds), using the GDP deflators in Smits et al. (2009) dataset, and those for France (to 19005/13 francs) using the deflator in Toutain (1987).

Conversion to standard unit

Using the latest version of the widely used Maddison data base (Bolt and Van Zanden, 2013), I converted my sample of now real GDP estimates into 1990 Geary-Khamis dollars ($GK). Maddison's data are the 'best estimates' of this kind available for the period, and make comparisons of my work with the rest of the literature much simpler (Prados de la Escosura, 2000, p. 2). For consistency, I always derived the conversion rate as the period starting year Maddison national GDP per capita divided by the starting year national GDP per capita. For example, for Italy, which had a starting year of 1871, $CR_Italy = GDPpc_Maddison1871/GDPpc_Felice1871$. I then used this same conversion rate (CR) to convert all real regional GDP per capita cross-sections.

As different countries have slightly different starting years, does this variability affect conversion rates and ultimately rankings of relative income? I re-converted regional GDP per capita estimates for all regions for 1870, 1900, and 1910, using different exchange rate benchmark years - a midpoint (1900) and endpoint (1910) - rather than the country starting year, which hovers around 1870. The Spearman rank correlations are displayed in table 3.12.

The near-perfect correlations show that the choice of exchange rate benchmark year is of little consequence in the grand scheme of things. This small degree of

Table 3.12: Spearman correlations of exchange rate benchmarks 1900 and [1910]

	1870	1900	1910
1870	0.986 [0.987]		
1900		0.979 [0.980]	
1910			0.973 [0.974]

Notes: All coefficients are significant at 1 per cent. Coefficients in square brackets are for 1910; those not in brackets are for 1900. They show the rank correlation between regional GDP per capita cross-sections standardised using the method outlined earlier (with country starting years) and cross-sections converted using 1900 and 1910 benchmarks.

measurement error is tolerable, and is unlikely to affect general patterns when it comes to the following empirical analysis.

Border Changes

Before carrying this out, I checked whether the territorial borders in my sample matched up with those in Bolt and Van Zanden (2013). There are no border changes for Italy, Spain, and Sweden during the period, and the units match those in Maddison (2001). France lost some territory to Germany, most notably Alsace-Lorraine, during the Franco-Prussian War, but this happened in 1871. So, using 1860 as a starting year for $GK conversions was safe: for that year, there were no changes and my data matched that in Maddison (2001, p. 27). It is pertinent to point out here that Maddison based his French estimates on Toutain (1987). My estimates for Germany start in 1871, by when the Germans had acquired Alsace-Lorraine and some other parts of France. Maddison (1995, p. 131) adjusts his data for these changes, and as there were no other changes until 1918, there are no other changes that affect my estimates. I should also point out here that Maddison based his German figures on Hoffman (1965).

While there were no border changes for Britain during the period, its conversion required some more adjustments. As Ireland was not part of Britain proper during this period, Crafts (2005b, p. 59) does not list Ireland in his table (Table 4) of regional GDP per capita for Britain. Crafts (2005b, p. 56) does, however, list total GDP for Ireland in another table (Table 3). I calculated Irish GDP per capita using population data from (Bolt and Van Zanden, 2013) and the data in Table 3 of Crafts (2005b). After applying the deflator from Smits et al. (2009), I derived an *CR* for Ireland from Geary and Stark (2002), which provides an 1871

$GK GDP per capita figure for Ireland alone, and one for the rest of Britain from Maddison (2007). I did not use Maddison's "UK" figure for both as it excludes Ireland, and he provides no Irish-only figure for 1871.

Different estimation methods

One last potential concern is the mixture of top-down and bottom-up estimation strategies used in the sample. Does the use of different regional GDP estimation methods introduce a systematic bias? This would matter if, say, regional GDP estimates derived using the Geary-Stark top-down method were persistently lower or higher than those derived using a national accounting approach. This is unlikely. Admittedly, short-cut methods are less accurate, but as Geary and Stark (2002), Enflo et al. (2010), and Buyst (2009)'s checks show, the margin of error is tolerable according to national accounting standards. Comparing their estimates to official ones from 1971, Geary and Stark (2002, p. 11) conclude that 'the likely error in the estimates of GDP generated using [their] method...are plus or minus 10 per cent of the correct total.' Checking against official figures for 2000 and 2007, Enflo et al. (2010, p. 16-7) find an even smaller error of 5 and 4 per cent of total GDP respectively. Furthermore, they show that the error is distributed randomly across regions - differences exceed 10 per cent in three regions in 2000, for example. The correlation coefficient between regional errors in 2000 and 2007 is a mere 0.08, suggesting that measurement errors for individual regions are not the product of a systematic bias. In sum, the error is random and its margin tolerable.

My own technique differs from the Geary-Stark method so I subjected it to additional robustness tests in the previous section to ensure that it also does not introduce an estimation bias. As the tests show, much like the Geary-Stark error magnitudes, the sum of my regional GDP estimates as deviations from given national estimates range from 3.74 to 6.75 per cent for Germany; and for France the range is -8.39 to 9.27 per cent. I also use my method to estimate Enflo et al. and Schulze's regional numbers for Sweden. The differences are very small, with the replicated and given national level series correlated at 0.99, and cross-sectional Spearman correlations ranging from 0.89 to 0.99.

Perhaps more basically, the potential influence of such a systematic error, even if it existed, would be small. It is only Schulze's estimates for Austria - and the mining and iron industry for Hungary (a small part of a small economy) - that were derived using a bottom-up approach. Austrian regions form seven per cent of the total sample. Further, table 3.10 shows no clear difference between the replication results for Austrian and Hungarian regions.

Following all the adjustments, table 3.13 summarises the regional per capita income data.

Table 3.13: Regional GDP per capita.

Year	Mean	Maximum	Minimum	Standard Deviation
1870	1,778	4,282	606	526
1900	2,425	5,877	771	694
1910	2,874	8,109	873	891
Sample	2,543	8,109	606	1,065

Notes: Regional GDP per capita in 1990 Geary-Khamis dollars.

3.3 Empirical Analysis

Figure 3.2 shows what needs to be explained. During the 1870 period, industrialisation had peaked in Britain and was spread in quite equal measure around the rest of Europe. The leaders were first Britain and then France, as O'Brien and Keyder (1978) wrote. The historical literature, however, does not give us an idea of the extent of internal per capita income inequality. Central France, for example, bears more in common with Sweden than it does with northern France. That is, a large part of the country - a "leader" - is as poor as Europe's periphery. The literature does not tell us how these patterns changed over time either. We can be more systematic about this and calculate a Theil index for the whole sample at each year.[1]

Following Combes et al. (2011), I decompose the Theil index into within-country (T_w) and between-country (T_b) components:

$$T = T_w + T_b. \tag{3.3}$$

[1]As market potential data are only available for 1870, 1900, and 1910, for the remainder of my dissertation these will be my benchmark years.

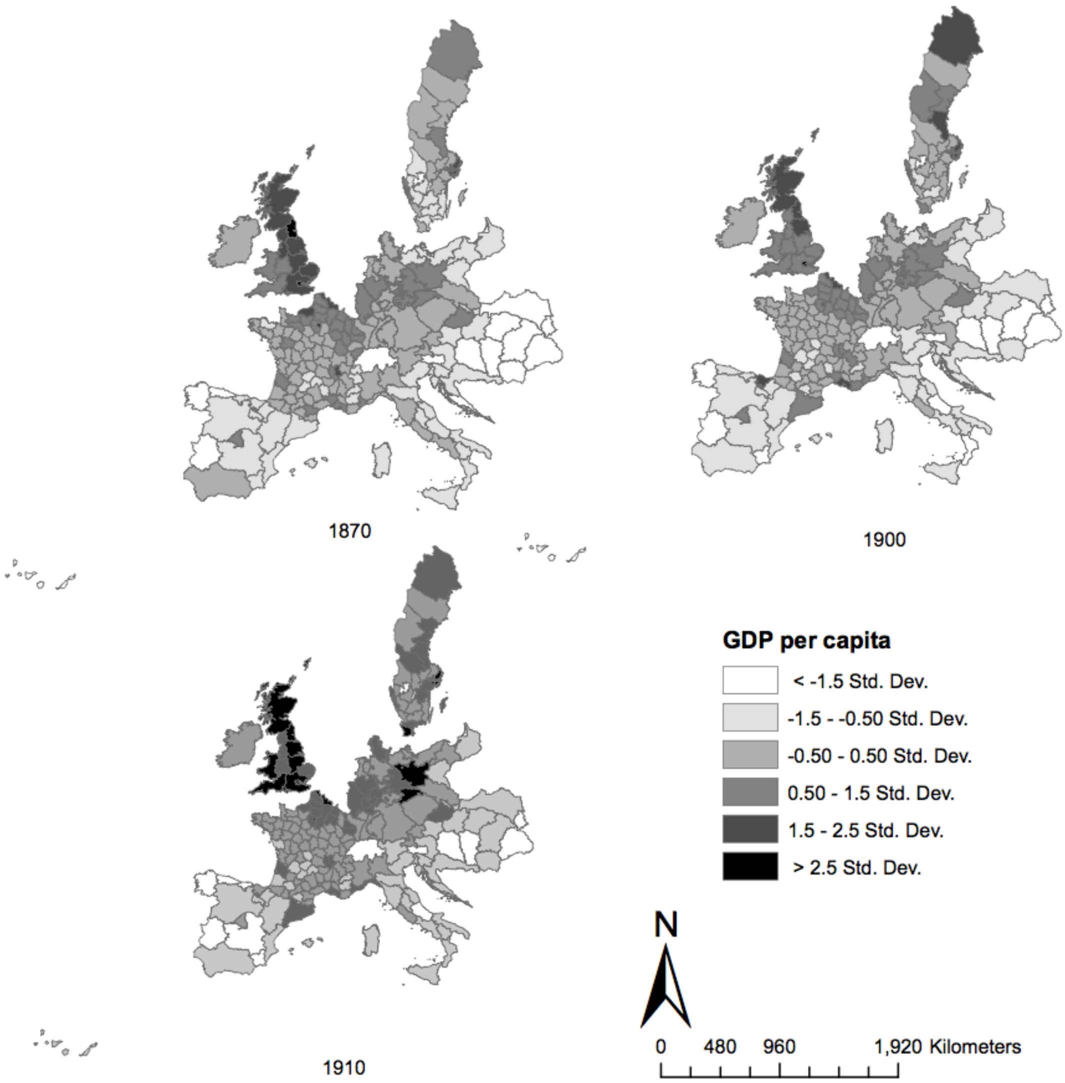

Figure 3.2: Regional GDP per capita in 1990 Geary-Khamis dollars.

Notes: See text for underlying data; maps are own.

The T_w captures the weighted average of Theil indices within region r, T_r:

$$T_w = \sum_{r=1}^{R} \frac{A_r}{A} T_r \tag{3.4}$$

where R is the number of regions, and $A_r = \sum_{d=1}^{D_r} A_d$ the level of income in country r, which includes D_r regions. The Theil index for country r is given by the same expression as T, but applied to the regions belonging to country r:

$$T_r = \sum_{d=1}^{D_r} \frac{A_d}{A_r} \ln \frac{A_d}{A_r/D_r}. \tag{3.5}$$

The T_b term corresponds to the between-country Theil index:

$$T_b = \sum_{r=1}^{R} \frac{A_r}{A} \ln \frac{A_r/D_r}{A/D}. \tag{3.6}$$

Table 3.14 shows the index values obtained for GDP per capita. The last row shows the percentage difference between the within- (T_w) and between-country (T_b) values.

Table 3.14: Theil indices for income.

	1870	1900	1910
T_w	0.0375	0.0406	0.0412
T_b	0.0330	0.0278	0.0298
% Difference	14	46	38

Notes: All per capita income data are in 1990 Geary-Khamis dollars. T_w refers to inequality in regional GDP per capita within countries. T_b refers to inequalities between countries.

The indices show that GDP per capita inequalities within countries are much higher than inequalities between countries. Indeed, Bourguignon and Morrisson (2002) find that world income inequality between 1820 and 1929 was mainly accounted for by within country inequalities; but that between country inequalities came to dominate in the mid-twentieth century. The present GDP per capita data show that the difference between the within and between inequality components grew over time, having peaked in 1900. While inequalities between countries narrowed over the period, which would explain the literature's convergence findings, inequalities within countries grew (Williamson, 1996). These within-country differentials run contrary to the standard assumptions in Solowian growth theory that physical capital, its main variable, is freely mobile within countries. We know that labour was highly, and freely, mobile within countries, but quantitative analyses of internal regional emigration rates have 'failed to find any systematic relationship between emigration rates and economic conditions' (Baines, 1994, p. 42). High within-country inequalities also pose a problem to the newer empirical tradition in economics, which finds, based on national units of observation, that institutions trump all other explanations of income differentials (Acemoglu et al., 2005). These within-country inequalities show us that the popular focus on institutions is either misguided or that institutions operate very differently within the same national environment.[1]

A traditional (neo-classical) empirical prediction in comparative income studies is that of convergence, as production factors flow to where they are needed most. For a group of economies, σ-convergence implies that over time incomes relative to the sample mean will decline (Barro and Sala-i Martin, 2004). That is, dispersion in income drops. This concept is most often measured as the coefficient of variation on income (Williamson, 1995). Table 3.15 shows the coefficient of variation on GDP per capita by country and year for my sample. In panel A, the coefficients are unweighted. They show generally high levels on inequality - a value of zero would represent perfect equality - which is to be expected after the Theil indices in table 3.14. Another notable feature is that only three countries - Austria-Hungary, Germany, and Sweden - experienced a decline in their regional inequality from 1870 to 1910. In the remaining four countries, which comprise the bulk of the sample, inequality increased; quite dramatically in the case of Italy. In panel B, I weighted regional GDP per capita values by regional population shares in the total sample population. Milanovic (2007) and Kenny (2005), among others, both show us how weighting by population can change the picture. Kenny (2005) explains its relevance by rhetorically asking whether it would, in today's world, make sense to assign equal weight to European countries as to India and China, which account for a large share of the world's population and saw dramatic growth in the late-twentieth century. The results paint an even less opti-

[1] In the following chapter, I expand these conceptual arguments and empirically test them.

mistic picture of income convergence: only Spain experiences a decline in income dispersion and, even then, a fairly modest one of eight per cent. Italy's increase in income dispersion is now less pronounced, and Austria-Hungary registers a marginal increase in dispersion. Sweden, famed for its income equality past and present, sees a large increase in income dispersion over the period (Henning et al., 2011). The largest increase comes from France, where the population-weighted coefficient of variation went up by 36 per cent from 1870 to 1910. Rather than the dramatic convergence discussed in the literature, these numbers point towards persistent income inequality (Epstein et al., 2003; Williamson, 1996).

Table 3.15: Coefficients of variation by country and year

Panel A	1870	1900	1910	Change (%)
Austria-Hungary	0.346	0.329	0.310	-10
Germany	0.211	0.194	0.186	-12
Spain	0.300	0.427	0.377	25
France	0.302	0.314	0.344	14
Britain	0.210	0.246	0.281	34
Italy	0.167	0.241	0.266	59
Sweden	0.248	0.209	0.146	-41
Panel B				
Austria-Hungary	0.978	0.969	0.984	1
Germany	0.753	0.757	0.822	9
Spain	1.063	0.919	0.981	-8
France	0.961	1.243	1.304	36
Britain	0.492	0.606	0.629	28
Italy	0.622	0.674	0.681	9
Sweden	0.529	0.624	0.673	27

Notes: Panel A contains the unweighted data; Panel B contains the population-weighted data. All per capita income data are in 1990 Geary-Khamis dollars. 'Change' is percentage change in coefficient from 1870 to 1910. Coefficient of variation is the standard deviation over the mean for each country at each year. A value of zero indicates perfect equality.

To explore these trends further, in figure 3.3, I show a box-plot of logged regional GDP per capita levels. We can see that as the median income is increasing, the tendency is, again, towards stratification or slight divergence. These results are more in line with Epstein et al. (2003), which shows that for an OECD sample of countries, the interquartile range of income actually increased between 1873 and 1893, and stabilised by 1913. By 1910, the distance between the median and and the upper-qaurtile closed, implying that the per capita income distribution was shifting to the right, but still leaving a large number of regions behind.

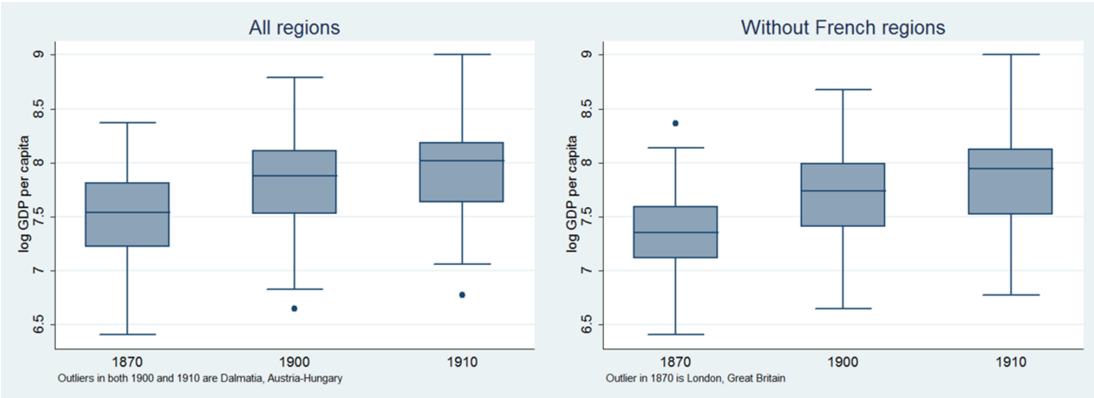

Figure 3.3: Box plot of regional income

Notes: Per capita income in 1990 Geary-Khamis dollars: 1870 (left), 1900 (middle), and 1910 (right). See text for underlying sources.

I included a box plot that excludes French regions, which in table 3.15 showed the most dramatic increase in weighted dispersion. France also accounts for a large share of my sample. The plot shows the same pattern as that of the full sample: increasing median income along with increasing inter-quartile distance. Still, to deal with the concern that French regions make up too large a share of the sample, *Appendix A* re-runs all the exploratory exercies in this chapter for a larger aggregation of French regions. The new aggregation brings the number of French regions down to its 19 NUTS-2 regions. In short, I find no difference in measures of sample-wide within- and between- country inequality, no changes to the general dispersion of regional income or to divergence rates, no changes to the distribution of regional income over time, and no changes to a global measure of spatial autocorrelation. For these reasons, the remainder of my dissertation proceeds with the use of *departments*.

Young et al. (2008) shows that it is possible for economies to exhibit β-convergence - a negative correlation between initial income levels and growth rates - even if income dispersion - σ-convergence - is high. An initially narrow distribution of income relative to the distribution of growth paths can give β-convergence, but σ-divergence. Figure 3.4 shows the standard representation of the β-convergence relationship: per capita income growth regressed on log initial per capita income level. A negative relationship indicates convergence and vice-versa. In the unweighted version, we can see a slow annual rate of income convergence - 0.61 per cent - but one that is highly statistically significant, with a t-statistic of -6.66. At this rate of convergence, it would take 128 years for regions to fill half the initial gap of income inequalities; not something people would have been holding their breath for.[1] In the population-weighted version, we can see a dramatic reversal of the relationship. The coefficient is positive and highly significant, with a much a larger t-statistic of 16.55. It implies a very slow rate of *divergence* of 0.03 per cent per year. This version supports the previous results of persistent income dispersion in table 3.15. The unweighted version, even if we accept it, offers very weak evidence of convergence.

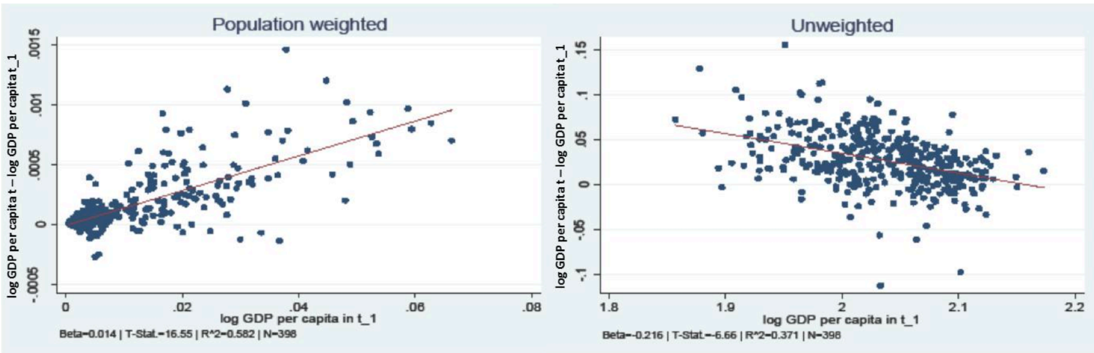

Figure 3.4: β-convergence of pooled sample

Notes: Y-axis is the log difference between GDP per capita in t and $t-1$; x-axis is the log GDP per capita level in t. Population weights are regional shares in sample total population at each year.

One issue with these measures of income dynamics is that they obscure intra-distribution movement. Following Epstein et al. (2003) we can get a better idea of income mobility by computing transition probabilities for the regions between each year. To do this, I placed regions in quartiles at each cross-section, 1870, 1900 and 1910. I then *counted* regions' transitions from one quartile to another between 1870 and 1900 and between 1900 and 1910. These count transitions are displayed in the matrices in table 3.16.

They show the movements of individual regions across income states between the years. For example, between 1870 and 1900, 10 regions went from the lowest quartile to the second quartile, while 14 regions went from the second quartile to

[1] Following Abreu et al. (2005), the convergence rate for panel model is $-ln(1+\beta)/T$, where T is the number of years in the period. Half-life is calculated as $ln(2)/ln(1+\beta/T)$.

Table 3.16: Regional income count transition matrix.

	1870-1900			1900-1910		
	p25	p50	p75	p25	p50	p75
p25	37	10	3	43	6	1
p50	10	26	14	6	28	16
p75	3	14	82	1	16	82

Notes: All income per capita data are in 1990 Geary-Khamis dollars. Figures are simple counts of transitions. p25 refers to the 25^{th} *percentile; p50 is the median; and p75 is the* 75^{th} *percentile.*

the top quartile. Equally interesting is that a number of regions fell back during this period: 10 regions went from the second to first quartile, and three from the top to first quartile. The diagonal cells represent the regions that did not move states, and they make up 74 per cent of the sample. During the much shorter 1900 to 1910 period, 77 per cent of regions did not move income states. Only six regions made it from the first to second quartile. More made it from the second to top quartile - 16 regions - but overall mobility was constrained. These findings for regional economies do not fit with the broader historical literature, but do match the analysis in Epstein et al. (2003, p. 84), who write: 'despite the strong case made in the historical literature for the significance of the forces of convergence between 1870 and 1913, the empirical distributions suggest that mobility was less common between those two dates than between 1913 and 1950 or 1950 and 1992.' Still, more regions were moving up the income scale than were falling back.

We can examine the changing shape of these distributions through a graphical shape of an Epanechnikov kernel density estimator, as in Epstein et al. (2003). The advantage of this representation is that it is smoother than a simple histogram, allowing us to discern trends over time. Figure 3.5 shows the kernel estimates for 1870, 1900, and 1910. The final year is clearly the peak. Regional per capita income levels moved from a relatively even distribution to an increasingly peaked and skewed distribution. It is also possible to see an almost bi-modal distribution emerge by 1910, with a much lower secondary peak farther to the left of the distribution. This observation, along with the limited mobility and stratification seen in table 3.16, finds more in common with Quah's concepts of convergence clubs and twin peaks of rich and poor economies (Quah, 1996). Interestingly, it also resonates with contemporary debates of a "two-speed" Europe, with low growth and high unemployment in the periphery and the reverse in the core (de Grauwe, 2012). I have again included a separate plot that excludes French regions. This results in a slighlty lower mean income in 1870 - France was at this point the second most successful economy - and a clearer bi-modal distribution of income in the following two years.

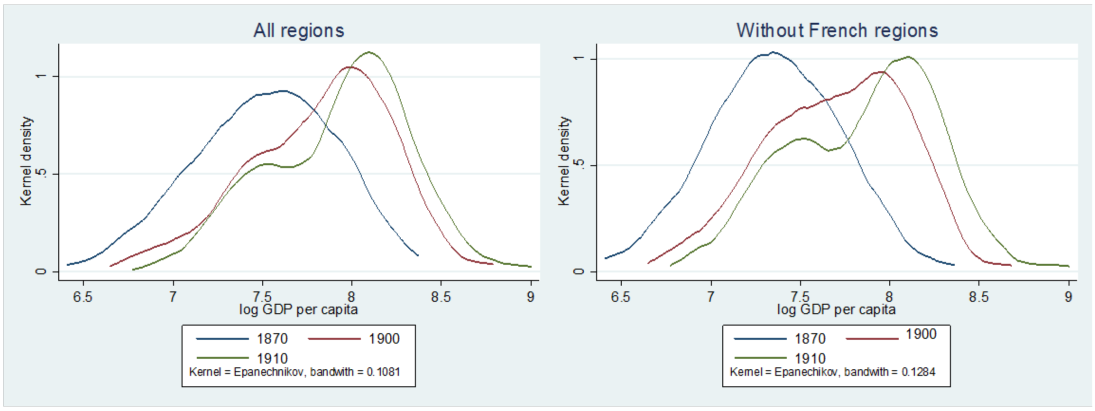

Figure 3.5: Kernel density plots of regional income

Notes: Per capita income in 1990 Geary-Khamis dollars.See text for underlying data. Estimated using Epanechnikov kernels.

Epstein et al. (2003) caution against placing too much weight on the transition matrices and kernel density plots, as they may be affected by short-term shocks or the atypical years. This is a fair point, but taking spaced-out benchmark years as I do here reduces this risk. A more important interpretative issue is that these exercises are a-spatial. Development does not only spread over time, but across space. It is harder to express this more eloquently than Lösch (1939), a founder of location theory:

> *[I]f everything occurred at the same time, there would be no development. If everything existed in the same place, there would no particularity. Only space makes possible the particular, which then unfolds in time.*

Understanding the geographical spread of development allows us to ask why certain areas developed while others did not. We can then look at the characteristics of these areas to arrive at a causal explanation. As a first step, I calculate Local

Moran's I statistics of spatial association for the regions. Anselin (1995) defines it as:

$$I_i = \frac{x_i - \overline{X}}{S_i^2} \sum_{j=1, j \neq i}^{n} w_{i,j}(x_j - \overline{X}) \tag{3.7}$$

where in my case x_i is the GDP per capita of region i, $\bar{X}$ is the cross-sectional mean of regional GDP per capita, w_i, j is an squared inverse distance spatial weight between regions i and j, where Euclidean distance is measured between regional nodes[1] , and:

$$S_i^2 = \frac{\sum_{j=1, j \neq i}^{n} (x_j - \overline{X})^2}{n - 1} - \overline{X}^2 \tag{3.8}$$

where n is the total number of regions. The index identifies spatial clusters of regions with similar per capita income levels. It also identifies spatial outliers, that is, high (low) income regions located near low (high) regions. The index can be tested for statistical significance, by calculating a z-score, as in Anselin (1995). Statistically significant positive index values indicate that a regiona has neighbouring regions with similarly high or low income levels. Negative values indicate a region has income levels dissimilar to those of its neighbours - it is an outlier. Figure 3.6 shows the results. Only regional index values significant at 5 per cent are displayed, rejecting the null that the spatial distribution of income is random, that is, there is no clustering of income.

The patterns are, perhaps, unsurprising. High income regions are clustered in the north of Europe: primarily in Britain (excluding Ireland), and the north of France. So high are British regional incomes relative to the rest of the sample, that the whole country is, in effect, a high income cluster. The parts of France closest to Britain also make up a high income cluster, but then they are co-located with some low income regions, forming a few low-high income clusters. Southern Italy and most of Austria-Hungary make a broad low income cluster. Moving into the middle of the period, in 1900, a new high income cluster emerges in

[1] The conceptual basis for this matrix is straightforward. It avoids an arbitrary fixed distance band or threshold, and assumes that all regions are associated, but the farther away regions are, the weaker is the association. As it is *squared* distance, the slope is sharper, so association drops off quicker and only a region's closer neighbours will be substantially associated with it.

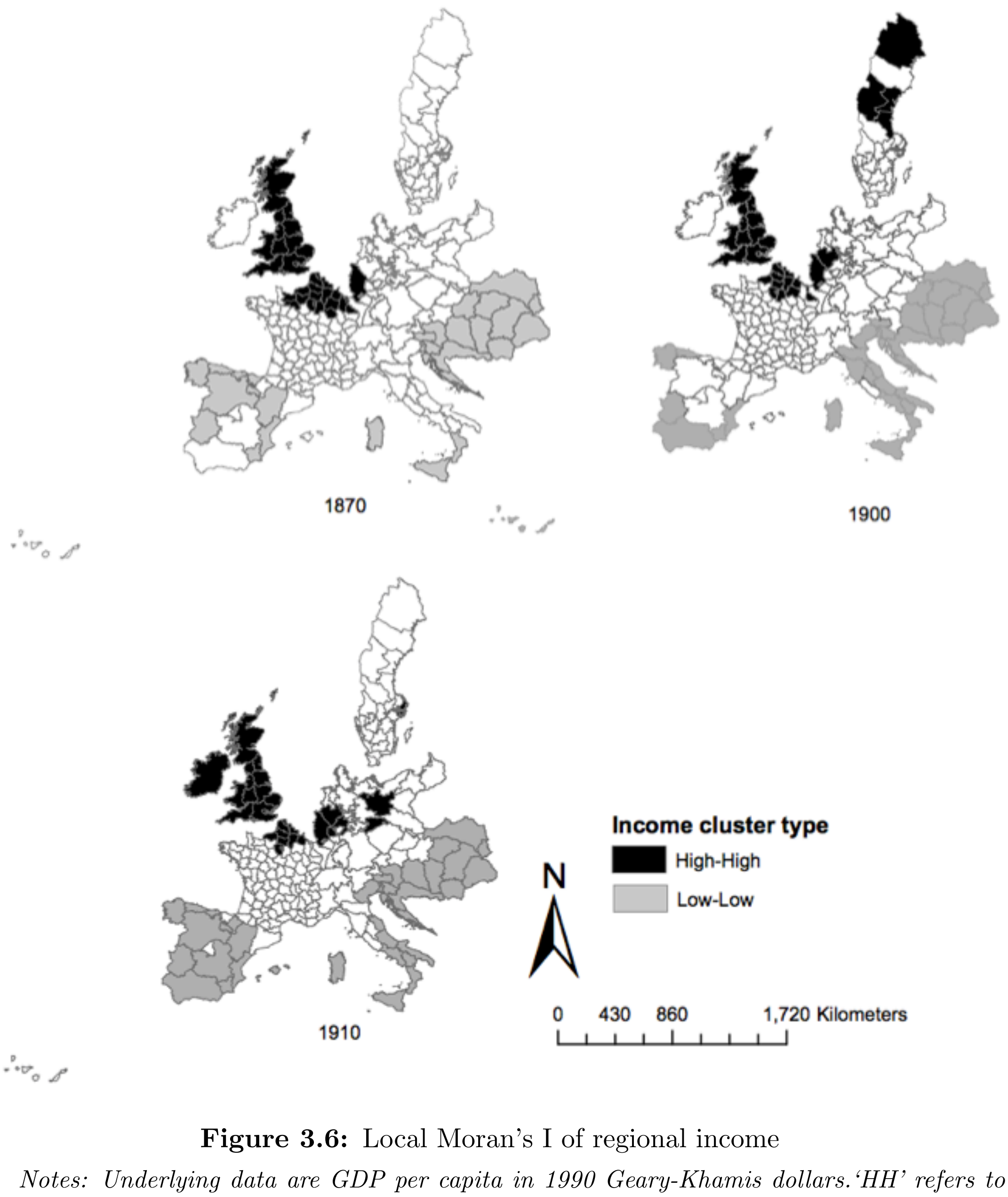

Figure 3.6: Local Moran's I of regional income

Notes: Underlying data are GDP per capita in 1990 Geary-Khamis dollars.'HH' refers to a high income cluster; 'LL' to a low income cluster; and 'LH' to a low-high cluster, that is, outliers. All other regional indices (in white here) indicate a random distribution of income.

western Germany. Other clusters of high income emerge in the south of France, but wealth is still largely concentrated in the north. Less happily, central Italy has now joined the low income cluster made up of the south of the country, and extending all the way into Austria-Hungary. These spatial patterns corroborate the emerging bi-modal distributions of per capita income in figure 3.5. At the close of the period, France's high income regions were entirely clustered in the north, running from the coast down to the German border. The high income cluster in Germany has spread outwards, and new ones formed slightly farther afield in the centre and north of the country. What has formed by 1910 is clearly the "Golden Triangle:" the developed core of Europe, roughly delimited by Paris, London, Amsterdam, and Frankfurt, which is still the most prosperous European macro-region today. Writing almost 50 years ago, Landes (1965, p. 456) claimed that 'By 1870 the industrial map of Europe was substantially what it is today.' If industry was driving these income clusters, then Landes was off by some 40 years. Western German industry was not yet "on the map" by 1870.

Broadly speaking, figure 3.6 shows us that while the core's of income clusters display persistence over time, they also shed and gain regions in a process of consolidation. Pollard (1981, p. 112) makes the point that regions as a unit of significance in European industrialisation have a 'historic dimension.' That is, clusters and concentrations of income were preceded by, and correlated with, older 'conglomerations'. The initial conditions for the first clusters can be fixed resources like coal or simply historical accident, as argued by Krugman (1991). Indeed, the appeal of coal based explanations of European industrialisation is that high-income clusters were often found around deposits (northern Britain and northeastern France), but then so were low-income clusters (in southern Spain or eastern Hungary).

Even if we do accept these explanations, it still leaves us with the fact that some regions dropped in and out of clusters. How do we, for example, explain the regions on the French western coast that were part of a high-income cluster in 1870, but not by 1910? What do we make of the central German regions that did not feature in a cluster up until 1900, but joined a high-income one in 1910? The economic structure of a region changes in relation to that of its surrounding regions. Heywood (1995), for example, argued that regional inequalities in France were lower at the start than at the end of the nineteenth century, when railway expansion formed a national market as opposed to an even distribution of regional ones. Improved links among regions mean industry has no reason to distribute itself evenly so as to reach scattered markets. The implication, however, is that peripheral regions, with their own provincial industrial centres, become exposed to increasing competition from more advanced regions, with their bigger markets. In Pollard (1981, p.115), we see the example of Dijon in the Cote-d'Or region of France. The initial phase of French railway expansion benefited local brewers by extending their regional markets, but by the end of it these producers suffered from the even more effective producers of beer in the east, close to the industrial core.

3.4 Discussion

The general patterns uncovered in this chapter are as follows: (1) per capita income inequality was higher within countries than between them; (2) there is no clear trend of regional per capita income convergence, but one of stratification; (3) while more regions were moving up rather than falling down the income scale, most regions were static; (4) the overall regional per capita income distribution went from a normal to a bi-modal distribution; (5) the spatial distribution went from high income clusters in Britain and northern France, and fragmented clusters elsewhere, to the consolidated "Golden Triangle." The *geographical* periphery (Iberian peninsula, central and southern Italy, Sweden, and Austria-Hungary) was the *economic* periphery.

Taking each in turn, the first point may be unsurprising to an earlier generation of historians. Pollard (1973) argued that the nation is not an appropriate unit of analysis in explaining the spread and emergence of industrialisation during this period. We cannot really understand economic development on this level because industrialisation does not, as much of the literature and models imply, hop from one country to another. The literature is full of explanations as to why Britain was more successful than France or why Britain was "first," but these explanations do not account for the dramatic regional variation in development within countries (Allen, 2009; O'Brien and Keyder, 1978). In his prescient paper, Pollard (1973, p. 636-7) writes,

> *On a map of Europe in which industrialization was coloured, say, red, it would by no means be the case that an area corresponding to a country within its boundary would turn uniformly pale pink, dark pink, and so on to deepest crimson. On the contrary, industrialization would appear as red dots, surrounded by areas of lighter red diminishing to white, and with the spread of industrialization these dots would scatter across the map with little reference to political boundaries.*

This may seem fairly obvious once stated, but it is still not how a lot of the relevant research is done. Some 40 years after Pollard's paper, most researchers still use the nation as their unit of analysis. Often data do not allow anything else, but the implications of using nations are considerable. Acemoglu et al. (2001), in an well-cited paper, take nations and national institutions as their observational unit to arrive at the empirical finding that institutions trump all other explanations for relative economic performance. It is national institutions, they argue, that provide the legal and social framework necessary for a freely-functioning free enterprise environment. However, as Pollard writes, European development proceeded with little reference to political boundaries. So how is it that political or institutional explanations of development have been so dominant for so long (Acemoglu et al., 2001, 2002; Gerschenkron, 1962; North and Weingast, 1989)? Further to this point, how can national institutions explain intra-national differences in economic performance?

This is not to say that national institutions are irrelevant. Gerschenkron (1962) used nations as his units of analysis, and perhaps much of the literature's current dependency on the national unit stems from his seminal work. As Sylla and Toniolo (1992) argue in reference to his work, nations historically tend to have common languages, laws, customs, public institutions, currencies, bureaucracies, economic and social policies, and the elites who affect all these characteristics. According to the authors, 'these variables have a profound influence on the industrialisation process and vary much more between nations than within them' (Sylla and Toniolo, 1992, p. 15). This is a fair point, but it is easy to disagree with it on factual and conceptual grounds. First, in a Europe where nations were only just forming, languages, customs and many other the other variables crossed national borders or varied within them (Schulze and Wolf, 2012). In an unofficial sense, is Alsace German or French, for example? How do we define Swiss cantons; are the French areas more similar to France than to their German neighbours? Indeed, what is interesting about figure 3.6 is that high and low income clusters seem to form among regions within countries. That is, income is not delimited by national borders. Second, if we do accept that national institutions and policies mattered, then we must consider the possibility that these variables operated differently in different settings, even within the same country. Sometimes this could be due to a difference between national *de jure* and regional *de facto* institutions (Acemoglu and Dell, 2010). Third, this example makes clear that institutional explanations are not *a priori* spatial in nature: there is nothing in them that leads us to expect a spatial concentration of "good" institutions, and hence of income. Weak southern Italian institutions lived side by side with their northern neighbours. This is problematic, given the broader spatial structure of European regional income that existed in the nineteenth century - and still exists today. Either way, using nations obscures much of what is interesting about Europe's economic development.

The next two points concern per capita income mobility and stratification. The economic history literature covering this period has been largely influenced by the findings in Williamson (1996) and O'Rourke and Williamson (1999). Both papers argue that open economy forces, mainly mass migration and trade, made for fast real wage convergence in the pre-1914 period - an era of rapid globalisation. They also review later eras, finding that when open economy forces are hindered, so too are rates of convergence. As Williamson (1996, p. 277) boldly put it, 'history offers an unambiguous positive correlation between globalisation and convergence.' Unfortunately, the evidence I presented here does not give us such great cause for optimism. Regional per capita income levels showed no obvious trend of convergence, and while the median income was rising, most regions remained firmly placed in the hierarchy, as can be seen in table 3.16. My analysis finds much more in common with Epstein et al. (2003, p. 95), who write that 'the data we have analysed do not suggest that the regime prevalent before 1914 was consistent with strong convergence.' The authors find, instead, a low degree of distributional mobility, and forces of persistence and stratification. Rather than "unambiguous" convergence then, we see here a different kind of distribution.

Much like Quah (1996) writing on income distributional dynamics, the pattern I have uncovered here is one of stratification. Figures 3.5 and 3.6 show clearly the development of what Quah (1996) would call "convergence clubs" - the income distribution polarising into twin peaks of rich and poor. The economic history literature cited above misses this characteristic of the distribution, and is perhaps why it concludes in favour of convergence. As Quah (1996, p. 16) writes,

> *...because in [the] traditional approach, the researcher only estimates a cross-section regression, he [sic.] sees only the behaviour of the (conditional) representative economy. He will never detect the multi-peakedness that arises in the cross-country distribution.*

The implication is that causal explanations based on empirics that miss an important characteristic of the data are likely to be misleading. If the main drivers of per capita income were global economic forces, as Williamson (1996) claims, then why did they affect different parts of Europe so differently? The finding that this twin-peak distribution also has a clear spatial regime - that stratification was also occurring spatially; something that would be hard to detect using national units - raises additional questions, which brings us to the last general point.

Understanding the spatial patterns observed here is central to understanding Europe's economic development. Allen (2009), for example, does not simply try to explain the Industrial Revolution, but tries to explain why it was *British*. Similarly, we can ask here, what was it about the north of France and west of Germany that allowed them to developed? Of course, the reverse is equally interesting: why did Europe's geographical periphery, the Iberian peninsula, south of Italy, and Austria-Hungary, fail to develop? In the language of Quah (1996), we need to explain this twin-peak distribution. Bringing in the spatial dimension is more likely to produce a reliable explanation. Spatial patterns, though, should always be interpreted with care, as we can see with a brief example from the literature.

Pollard (1981, p. xv) shows a map of Europe with the location of coalfields and industry in 1875. The idea is to display the spatial correlation between the two and, indeed, at first glance there appears to be a reasonable one. Relying on the spatial aspect alone, then, we might conclude that coal was a necessary condition for the development of industry: that income clubs formed around coal deposits. However, as Pollard (1981, p. 121) goes on to discuss, the timing of development differed across space. In Britain, the expansion of coal-mining preceded as well as accompanied industrialisation. On the Continent, he writes, 'coal rarely came first, and it was apparently favourable cost ratios of non-coal methods, such as using water-power for spinning mills...which frequently held back some of the major continental centres...' Pollard (1981, p. 121) concludes the matter by writing 'Thus coal as a determinant developed late.'

What this example shows us is the importance of considering both the geographical and the temporal. Now that I have established the general income patterns, the following chapters extend these themes. The following chapter empirically deals with coal and institutions.

Chapter 4

Traditional Explanations: Coal and Institutions

4.1 Introduction

How helpful are traditional explanations of late-nineteenth century European regional income? This chapter explores the extent to which regional income differentials were due to either institutions or coal endowments. The latter picks up on the work of a number of historians who have emphasised the role of coal in industrialisation (Cameron, 1985; Deane, 1965; Fernihough and O'Rourke, 2014; Pollard, 1981; Pounds and Parker, 1957; Wrigley, 2010). The former explanation is related to the new empirical work on institutions as determinants of income, spearheaded by Acemoglu et al. (2001) but with roots going back to North and Thomas (1973).

To the extent that development depends on coal, or any other unequally distributed natural resource, opportunities for development are not going to be equally available to all regions. In this line of reasoning, coal deposits provide an explanation of spatial income inequality. It has been argued for example that French efforts at industrialisation were hindered by the lack of easily-accessible coal (Heywood, 1995). Writing more generally, Parker (1961, p. 160) claimed, 'Resources, mineral, agricultural, and transport, were largely responsible for the direction and speed of nineteenth century development among western countries.' There are, however, reasons to be sceptical of coal-based explanations. The strength of a coal explanation deteriorates in a comparative analysis. In contrast to France, for example, Japan did not find limited coal and other raw material supplies to be a major obstacle to development (Kenwood and Lougheed, 1982). Crafts (1984) showed that Britain's reliance on coal in industrialisation was unusual, compared to other countries.

North and Thomas (1973) famously argued that the development of "efficient organization" in Western Europe accounts for the rise of the West. By "efficient organization" the authors mean institutions that ensure property rights, which create incentives that bring private rates of return close to the social rate of return. Property rights are taken to be embedded in institutions and, the authors argue, allow people to realise economies of scale, improve factor market efficiency, and encourage investment and innovation. Some thirty years later, these ideas started receiving serious empirical treatment. Emphasising clearly defined property rights and low risks of state-led expropriation, recent work finds support for the idea that "institutions" drive economic performance (Acemoglu and Dell, 2010; Acemoglu and Johnson, 2005; Acemoglu et al., 2001). This empirical literature is for the most part based on national units of observation, since these institutions are typically national. This is where my regional (subnational) units of analysis are particularly useful: they offer an *a priori* challenge to the institutional explanations of income differentials, since we know from the preceding chapter that within-country inequality was higher than between-country inequality.

This chapter is organised as follows. In the next section, I review the coal explanation and construct a more accurate measure of coal access based on a comprehensive GIS. In the subsequent section, I estimate a measure of regional "institutions," to be able to empirically test the institutional explanation. The empirical section brings these two explanations together in a"horse-race" specification, where I also construct an instrumental variable to address the endogeneity of institutions. I conclude this chapter with a discussion of the conceptual and empirical limits of these explanations.

4.2 Coal

Two weeks after I finished a draft of this chapter, Fernihough and O'Rourke (2014) released the most comprehensive treatment of coal in European industrialisation to date. The authors, who find that proximity to coalfields explains a large share of city growth in the late-nineteenth century, make clear an important distinction in the debate. Looking at variation over time, one strand of the argument has it that exploiting coal deposits explains Europe's subsequent development - the 'growth hypothesis' (Deane, 1965; Wrigley, 2010). Looking at spatial variation, the other strand of the argument has it that the location of economic activity was determined by the location of coal deposits - the 'location hypothesis' (Fernihough and O'Rourke, 2014; Mathias, 1983; Pollard, 1981; Wrigley, 2010).

In terms of the growth hypothesis, Wrigley (2010, p. 23) argues that switching to coal was a 'necessary condition for the industrial revolution', but an insufficient one. Industrialisation, he argues, was by definition a move away from an organic economy, where energy could only be drawn from labour, animals, wood, water and wind. These sources of energy placed a tight constraint on development: 'Iron, for instance, has many physical properties that make it of the greatest value to man but as long as the production of 10,000 tons of iron involved the felling of 100,000 acres of woodland, it was inevitable that it was used only where a few hundred- weight or at most a few tons of iron would suffice for the task in hand' (Wrigley, 1988, p. 80). Coal production and its related technologies, when they came, thus freed swathes of land and masses of labour from inefficient energy production and use.

According to the second strand of the debate, regions needed easy access to coal if they were to industrialise. Mathias (1983) sums up the debate: 'The logistics of energy inputs based upon coal, translated against available transport in a pre-railway age, precluded any major industrial complex in heavy industry from developing except where coal and ore were plentiful and adjacent to one another or to water carriage.' Wrigley (1961) calculated substantial cost savings for coal use close to coal mining. Pollard (1981) relates the location of Continental European industry to the location of coalfields: northern France, the Ruhr, and Belgium. Most authors concede that transport cost declines starting in the late-nineteenth century weakened the need to be located near a coalfield (Wright, 1990).

Fernihough and O'Rourke (2014) find that proximity to coalfields explains European city growth, a proxy for economic growth, supporting the location hypothesis. Further, they find that proximity to coal really matters when coal technologies were introduced, supporting the growth hypothesis. Their results are robust to a number of controls, an instrumental variable specification, and different specifications. What can I add to this debate? As with Fernihough and O'Rourke (2014, p. 6), my interest is not in whether coal was needed to *start* industrialisation - my period is too late for that - but whether, once industrialisation was in progress, access to coal 'mattered a lot or a little' for development. In practical terms, I have data on actual regional GDP rather than a proxy, which Fernihough and O'Rourke (2014) lament are missing, and I have new data on coal transport costs as well as simple geodesic distances.

4.2.1 Re-creating Europe's transport network

Since the broader coal explanation rests ultimately on acces to coal, we need to measure "access" as accurately as possible. Most empirical studies use simple geodesic distances, but being physically far from coal does not necessarily imply that access was bad or vice versa. While Pollard (1981, p. 121) is emphatic on the location hypothesis, he does write that industrial regions 'only survived if they had reasonable access by *water* to a supply of good coal [emphasis added].' As Kenwood and Lougheed (1982, p. 112) put it,

> *...investment in cheap transport is often more important than the possession of high-grade minerals for a region's successful industrialization, since it makes possible either the opening-up of local resource supplies or the importation of resources from elsewhere.*

Using a GIS of railways, waterways, steamship lines, and roads, I re-created Europe's late-nineteenth century transport network to measure regions' cheapest possible access to coal. I also collected coal freight rates for this network, which means that the resulting variable on coal access is time-*variant*.

Sea transport

Shipping and port data came from RRG Spatial Planning and Geoinformation at the University of Dortmund which collects and provides European geographical data and digitised maps (RRG, 2012). I used their shipping routes and seaports map. It originally included some 1,820 ports and some 2,110 routes between those ports. I eliminated ports that did not exist during the period or that were irrelevant like fishing ports or marinas, going by Bartholomew's historical map, *Europe and Near East - General Commercial Chart- Suez Canal* (Bartholomew, 1907). I kept routes the same as there is no reason why these would have changed. Ultimately, some 1,210 ports and 1,030 routes remained.

I took ocean coal freight rates from Kaukiainen (2006, p. 54), which is the standard source for this type of work. Kaukiainen conventionally breaks down freight rates into terminal costs - the cost paid at the end of a journey or at transhipment - and variable costs - the costs per ton-mile for coal.

Railways

Railway data came from the Historical GIS of Europe project (HGISE) at the University of Lleida (University of Lleida, 2012). HGISE's database provides information on railways every decade between 1870 and 2000. Corresponding to the GDP and transport costs data, I took data for the years 1870, 1900, and 1910. As the HGISE team explain, European railway data could not be extracted from a single source. Sources that were digitised and standardised include historical maps from Cambridge University and the Cartographic Institute of Catalonia, Thomas Cook publications, specialised magazines and online sources like the *Histoire Chronologiques des Chemins de Fer Europeens.*

Railway coal freight costs come from a variety of sources. The U.S. Bureau of Railway Economics (1915) provides comparative freight rate data for a large number of different-length routes in different countries and for different goods. I decomposed these costs into terminal and variable components, to use the variable cost data in Noyes (1905) and Cain (1980) to extrapolate back across the remaining benchmark years, following the procedure in Schulze (2007b). This yielded railway coal freight data for Austria-Hungary, Britain, France, Germany, Italy, Spain, Sweden, and Russia.

Waterways

Navigable river and canal data (henceforth, waterways) are also from (University of Lleida, 2012). Europe's map of waterways did not, according to the (University of Lleida, 2012) dataset, change substantially between 1870 and 1910. Indeed, Moulton (1914, p. 264) provides figures showing that the total length of Germany's waterway network was the same length in 1885 as it was in 1905, while France's declined by 2.5 per cent. Hadfield (1968, p. 208) shows that the length of England and Wales' canal system plateaud in 1850. Sources that were digitised and standardised include historical maps from the Cartographic Institute of Catalonia, Thomas Cook publications, specialised magazines and online sources like the *Voies navigables d'Europe* and *Histoire et Patrimoine des Rivieres et Canaux*.

Given, as the HGISE team explain, waterway maps are very rare, it is perhaps unsurprising that waterway freight rates are rarer still. I collected what was available in Moulton's epic, *Waterways Versus Railroads* (Moulton, 1914). The author provides waterway coal freight rates for France, Germany and Britain in 1905 (Moulton, 1914, pp. 291, 217-8, 212). The rates include tolls and haulage fees, but the author does not provide terminal costs. The best way of getting around this was to assume that the ratio of terminal to variable costs in sea transport was the same as that in waterway trasnport. For the remainder of my sample, I used the average of these three countries. Since the map of waterways does not change over my period, and since coal freight rates for waterways are unavailable for years other than 1905, the waterway network is fixed. While this is not ideal, it is better than leaving out the waterway network altogether, especially since recent work has argued for its importance (Klemann and Schenk, 2013).

Roads

The base road map is again from (University of Lleida, 2012). One difficulty in analysing historical road maps is that it is rare to find common criteria being used to identify different road categories in different countries. This is a problem that persists even to today with European Union road maps. It is therefore necessary to use the same source when comparing road densities in different countries. The HGISE team used General Dufours maps published between 1835 and 1872, which enabled them to make meaningful comparisons of public highways over time. The map I use here is for 1872. Much like the waterway network, the team found little substantial change in Europe's public highway network between 1870 and 1910. Unfortunately, no maps for Sweden were found. To fill in this gap, I used the *ArcGIS* "Grouping Analysis" tool to create a minimum spanning tree overland network for Sweden. While this "idealises" Sweden's road network, it is the most objective way of handling the gap.

There is a dearth of quantitative historical data road transport. The only road coal freight rate I found is from Van Vleck's study of Britain's internal coal freight road network (Van Vleck, 1997). The author provides a single ton-mile estimate for 1914. For a terminal rate, I assumed the ratio of road terminal to variable rates is the same as that of railway transport. We are unlikely to have anything better than this for some time, but it is arguably better than assuming away the road network, as often happens (Armstrong, 1987).

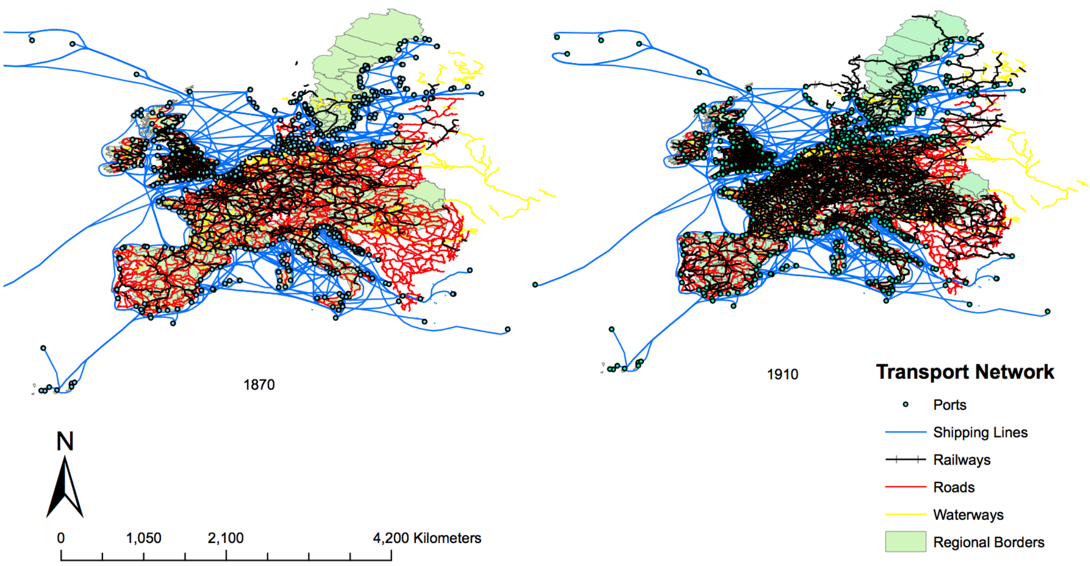

Figure 4.1: Europe's transport network

Notes: Regional nodes and map for 1900 are not shown to offer some clarity. For sources, see text.

Figure 4.1 shows the network in 1870 and in 1910. I left out regional nodes, used in the network routing I explain in the following section, to make the map easier to read. The idea here is to show the detail and extent of the network, and the ability of a GIS to integrate different datasets.

Network routing from Nodes to Coal to Deposits

The first step in calculating the cheapest routes between regions and coal deposits is defining points of connection. Since distance can only be measured from point to point - rather than from and to regional polygons - I attributed a node to every region. The criterion for choosing nodes was economic rather than political importance. It was more often than not that regional (or provincial) capitals were the most economically important, say, Vienna with Lower Austria. For coastal regions, these nodes often corresponded with the RRG (2012) map of ports.

Coal deposits - both anthracite and lignite, as in Fernihough and O'Rourke (2014) - are from the United States Geological Survey (USGS) (USGS, 2013). Coal forms when organic matter is first pressurised into peat, which turns into lignite, then into sub-bituminous coal, after that bituminous coal, and finally anthracite. The last in the order is the most valuable form of coal, since it has compressed over the longest period the most organic matter. It has a heating value of 13,000 to 15,000 Btu/lb (British thermal unit per pound) versus lignite's 4,000 to 8,300 Btu/lb (American Society for Testing and Materials, Subcommittee: D05.18, 1998). According to Fernihough and O'Rourke (2014), both were used in Europe's industrialisation efforts.

These data cover all coal *deposits*, and not just the ones known at the time or mines that were in operation. If the coal explanation holds that coal was required for development, then we would be testing a different (endogenous) explanation if we were to delete coal deposits that were not known. Indeed, there is something to be said of regions that did not use or discover coal deposits in their territory. Fernihough and O'Rourke (2014) use coalfields, but ultimately instrument it with proximity to rock strata from the Carboniferous era - in effect, coal deposits. Another issue with these data is that the location of deposits as measure using centroids: a point place on the geographical centre of each deposit. This means a region may be close to the border of one deposit, but still far from its centroid.

As a cross-check on both the centroid and on the “knowability” of coal deposits, I digitised the ’Outline Map of Europe’s Principal Coalfields, 1910’ from Ralph et al. (1910, p. 259). Crucially, the map is from 1910 and shows coalfields rather than deposits, so these were known and exploited sources of coal at the time. I digitised the coalfields as polygons rather than points to get around the centroid issue.

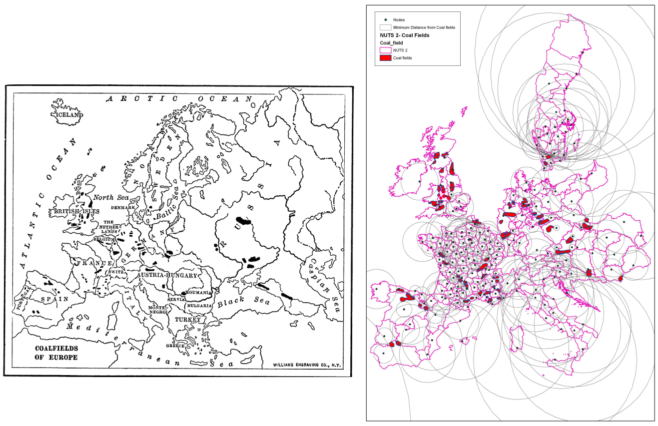

Figure 4.2: Digitising a 1910 coalfield map.

Notes: Map on left is a 1910 map of European coalfields from Ralph et al. (1910); map on right is the digitised version showing distance from each regional node and NUTS-2 borders.

I then measured the distance from each regional node to the nearest coalfield broder. Figure 4.2 shows the original and digitised map, with regional distances and regional borders highlighted in the digitised version. The summary statistics of these two measures shown in table 4.1. The simple distance to deposit measure has a 16 per cent higher mean value. The difference between their standard deviations is equally comparable, at 17 per cent. The coalfield distance measure has a minimum value of 0km since some nodes are directly on coalfields. For the coal deposit measure, the minimum is 1km. The difference between their maximum values is four per cent. Finally, their skewness statistics, at 2.95 and 3.01, show that they follow a very similar distribution.

Are the simple distance to coal deposit centroid (km) and distance to nearest coalfield border (km) correlated? Their Pearson correlation is 0.752, significant at one per cent. Their Spearman rank correlation is 0.673, significant at one per cent. Regressing the coalfield distance on the coal deposit centroid distance

Table 4.1: Distance to Coal Deposit Centroid and Coalfield Border (kms).

	N	Mean	Std. Dev.	Min.	Max.	Skew
Coal_Distance	597	182	161	1	1,423	2.95
Coal_Field	597	157	195	0	1,478	3.01

Notes: Coal_Distance is the distance to coal deposit centroid variable, and Coal_Field is the distance to coalfield border variable.

across all 199 regions, I get a coefficient of 0.911, significant at one per cent, and a constant term of -8.80, which is insignificant. This parametric test indicates no differences in levels (insignificant constant term) and an almost one-for-one correlation (significant coefficient of 0.911). The R^2 is 0.57, indicating a respectable overall correlation.

Given the correlation between the simple distance to centroid and coalfield border, it is unlikely that using the latter in the transport calculations would have produced meaningfully different results. Further, in *Appendix B* I show that neither measure has an effect on regional per capita income, which is what ultimately matters, when controlling for regional institutions.

The second step is connecting the lines of transport with the nodes, using GIS topology tools. To allow transport from one node to another, it must be possible to continuously travel along railways, waterways, roads and shipping lines, through ports, and other nodes. Shipping lines naturally connected directly to ports, so there was no need for adjustment there. Railways did not always connect directly onto nodes or ports. This made it necessary to use a buffer zone around nodes: Any railway that enters that buffer zone is assumed to connect to the node. The buffer was defined with a radius of 40-miles. This choice is based on the following logic. Non-rail land-transport during the period occurred most commonly by horse-drawn carriage. Economic actors are unlikely to ship goods that would have taken longer than a day to make it from the rail station to the regional node, as overnight stays greatly increase costs. So, the question is, how far can the average horse pull the average carriage in an eight-hour day? As horses are unable to gallop, canter or trot for eight straight hours, the horse will walk. The average walking speed for a horse is around three to five to 10 miles-per-hour. That means, in eight-hours, it can travel between 40 and 80-miles, so the lower-bound estimate was taken, since the horse will be drawing a carriage

(Weinstein and Adam, 2008).

The third step is defining connections along the network. It would be unrealistic to allow routes to switch freely from rail to waterway and ship or vice versa. There are obvious transhipment costs. Unfortunately, at the time of writing, no detailed data on the location of rail stations or on direct and regionally specific transhipment costs were available. This made two safe assumptions necessary. First, transhipment was possible whenever a railway connected to a port or vice versa, and possible whenever two railways met at a node. This is sound, as ports were more often than not served directly by rail or had potential to nearby railway stations, and railways crossing at a regional node usually marked the location of stations. Second, as transhipment cost data were unavailable, the terminal cost data discussed above were used in its place. The two are likely to be similar in size, but more to the point, countries' terminal costs were likely to have been proportionate to their transhipment costs. It is difficult to imagine a scenario where London's transhipment cost is five-times its terminal cost, but that Italian Ancona's is only twice as high. Therefore what matters is that there is consistence taking the same measures for the same things across countries and time. For consistency with the GDP data and as required for panel model work, I converted all variable and freight rates into 1990 Geary-Khamis dollars using the exchange rates at Global Financial Data (2014), and the purchasing-power rates implicit in Maddison (2007).

Finally, the GIS allows for transport over two-dimensional space. The network at this stage allows routing along railways, waterways, roads and shipping lines, but travel remains costless. This step was just a matter of applying the variable costs (costs per ton-mile of coal) to the appropriate railways and shipping lines, and the terminal or transhipment costs to the appropriate ports and connections. For transhipments, I applied the terminal cost associated with the current mode of transport; for example, when switching from railways to another mode, I applied the railways terminal cost. With all the costs in place, I used the *ArcGIS* "Network Analyst" tool, which uses Dijkstra's graph search algorithm to solve for the cheapest routes between regional nodes and coal deposits. The algorithm solves for the single-source shortest/lowest-cost route between a group of nodes.

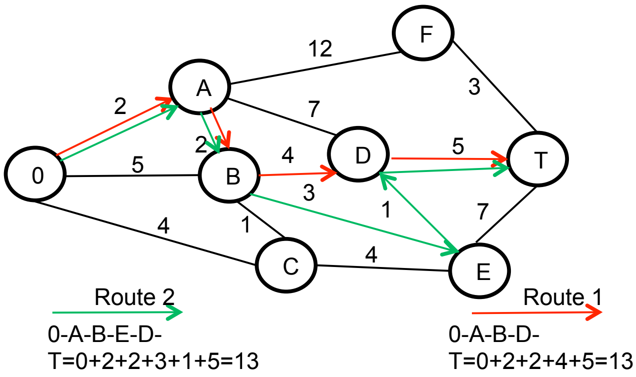

Figure 4.3: Dijkstra's graph search algorithm

Notes: Circled letters represent nodes. Connecting lines are arcs. Figures represent length of arcs. The shortest routes (both of equal length) are highlighted in green and red.

For an illustration, see figure 4.3.[1] It is a network where arcs (lines) are labelled with their lengths (costs). We can use Dijkstra's algorithm to find the shortest route from the origin node 0 to the destination node T. The origin is labelled 0, since it is 0 units from the origin. For each arc connecting the origin node, we calculate the candidate distance to neighbouring nodes. The candidate distance is the distance to, in this step, the origin node plus the length of the arc. So, between 0 and A, it is two $(= 0 + 2)$; between 0 and B it is five $(= 0 + 5)$; and between 0 and C it is four $(= 0 + 4)$. We choose the shortest arc - $0 - A$ - and add it to our set. This process is repeated moving through the nodes, until we reach T. In figure 4.3, there are two equally short routes: $0 - A - B - D - T$ and $0 - A - B - E - D - T$ both "cost" 13 units. This is, of course, a very simple version of the algorithm. In my dataset, the nodes are my 199 regional cities, and my arcs are the combination of railways, waterways, roads, and sea routes. Further, it is not simply the cost along the arc, but the terminal costs that matter.

Results

Figure 4.4 shows the pooled correlation between the resulting estimates of minimum transport costs to coal deposits ($Coal\_Transport$) and GDP per capita, and the pooled correlated with the latter and a geodesic distance to coal deposits ($Coal\_Distance$) measure. All variables are logged.

[1]See Yan (2014) for am introduction of Dijkstra's algorithm and its importance.

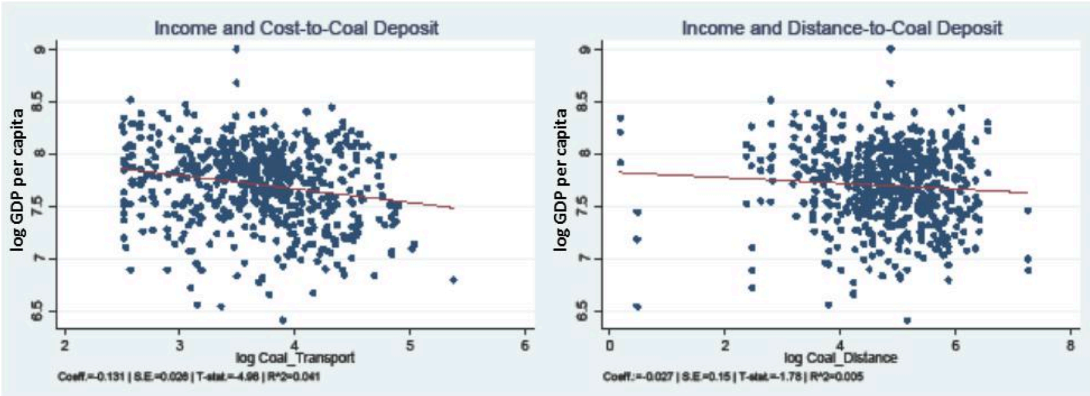

Figure 4.4: Coal and income correlations

Notes: Values on y-axis are log GDP per capita. Coal_Transport is the cost in dollars of transporting a tonne of coal from a region's nearest deposit. $Coal_Distance$ is the geodesic distance in kilometers to the nearest deposit. Includes full sample. For sources, see text.

While neither provide a good fit, the R^2 for the *Coal_Transport* correlation at 0.04 is greater than that for *Coal_Distance* at 0.004. Both coefficients are statistically significant: *Coal_Transport* at one per cent and *Coal_Distance* at 10 per cent. The estimate on the former implies that for every standard deviation increase in a region's cost of transporting coal from the nearest deposit to its node, its regional income drops by 0.203 standard deviations, compared to -0.073 standard deviation effect for *Coal_Distance*. These effects translate into, respectively, seven and three per cent of the sample mean per capita income. In sum, a more accurate measure of coal access shows us that the effect of proximity to coal on income may be larger than the standard geodesic measure. Figure 4.5, which overlays coal deposits on the GDP per capita map, shows that this should not come as a surprise. The simple spatial correlation between income and coal only works in Britain, central Germany, and northeastern France. The coal deposits in southeastern France, Spain, and Hungary seem to have done little for per capita incomes. While both the geodesic and transport-cost measure support the coal explanation, figures 4.5 and 4.4 highlight the importance of measuring access to coal accurately.

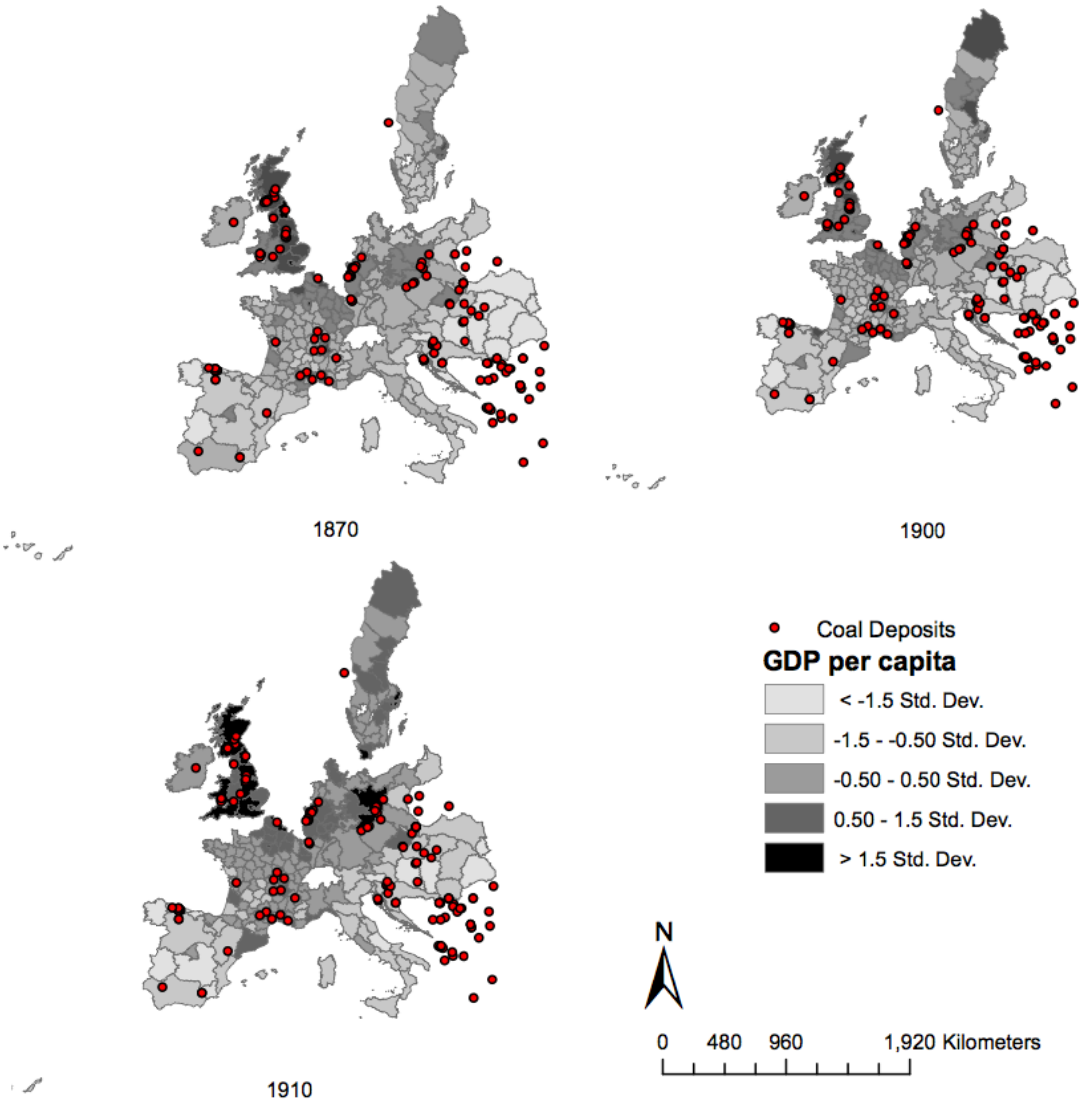

Figure 4.5: Coal and income map

Notes: Coal deposits, from USGS (2013), mark the geographical centres of exploitable deposits. The GDP per capita data is the same as that in Chapter 3.

4.3 Institutions

According to North and Thomas (1973), it is differences in institutions that explain differences in development. Acemoglu et al. (2005), currently the leading proponents of the institutional view, argue that economic institutions are what matter for economic outcomes. These institutions collectively refer to the 'structure of property rights and the presence and perfection of markets' (Acemoglu et al., 2005, p. 389). Put simply, economic institutions matter because they shape the structure of economic incentives in society. These ideas have now received a lot of empirical support - one well-cited instalment in the debate is called 'Institutions Rule: the Primacy of Institutions Over Geography and Integration in Economic Development' (Rodrik et al., 2004). As successful as they have been, however, institutional explanations suffer from a number of issues.

First, the term "institutions" is often all encompassing. As Voigt (2013) writes, it has been used to refer to newspapers, supermarkets and even phone booths. There needs to be clarity on whether we are working with institutional structures (federal states, say) or organisations (say, the Church). Second, we need to be clear on whether it is *de jure* or *de facto* institutions that matter or that we are measuring. While, for example, Italy's institutions were, *de jure*, the same after unification, we know that there were - and still are - *de facto* institutional differences within Italy (Banfield, 1958; Dimico et al., 2012). Third, and perhaps most importantly, it is plausible that institutions are endogenous to economic activity - that economic development leads to good institutions. This is a serious concern, and I come to it in the empirical strategy section after constructing my institutional measure in the coming subsection.

4.3.1 Measuring institutions

Measuring the concepts underlying "institutional" explanations is fraught with controversy, especially when using subnational regional units of observation. I take as my starting point the idea that *de facto* subnational institutions exist and are important determinants of economic performance (Acemoglu and Dell, 2010). This runs contrary to the institutional critique that altogether denies a role for institutions because the explanation does not fit with the fact of subnational income differentials documented in the previous chapter. The argument goes, institutions are national and so cannot account for *sub*national differentials (Redding and Strum, 2008). This straw-man critique ignores the existence of *de facto* regional institutions shown for example in Acemoglu and Dell (2010), insisting that only *de jure* national ones exist. In spite of this, the presence of these regional institutions is something political scientists and historians have long-recognised (Banfield, 1958; Fukayama, 2001; Pollard, 1973). Economic historians, too, pointed out some time ago that *de jure* and *de facto* institutions that can co-exist on different scales. Montinola et al. (1995) explained the puzzle of late-twentieth century China's inefficient institutions and its rapid growth with the matrix in figure 4.6. In the two-by-two matrix, *de jure* institutions can be either present or absent and *de facto* institutions can be either present or absent. Researchers often assume that *de facto* and *de jure* go together - as in cell 1 - but, as the matrix shows, they do not necessarily correspond. Cell 2, in which *de jure* institutions are absent but *de facto* decentralisation is extensive, is what Montinola et al. use to explain the puzzle.

		De jure institutions: Present	De jure institutions: Absent
De facto institutions	Present	1 Ideal institutions	2 Neglected alternative
De facto institutions	Absent	3 Bogus institutions	4 Highly centralised in theory and practice

Figure 4.6: De facto and de jure institutions
Notes: Adapted from Montinola et al. (1995).

As for quantitative measurement, I rely on the popular definition of institutional efficiency from Acemoglu et al. (2005, p. 389-90):

> *[Institutions] not only determine the aggregate economic growth potential of the economy, but also an array of economic outcomes, including the distribution of resources in the future (i.e., the distribution*

of wealth, of physical or human capital). In other words, they influence not only the size of the aggregate pie, but how this pie is divided among different groups and individuals in society.

In a late-twentieth century study, it is possible to proxy this formulation of institutions using indicators on income inequality, education levels, and social mobility. In historical settings, when such indicators are unavailable, measurement is harder. My attempt is based on frontier analysis. If we follow the formulation that institutions determine the distribution of resources, and that efficient institutions distribute resources more evenly, we can measure institutional quality as the efficiency at which material resources (GDP per capita) are converted into resources like human capital.

An efficiency frontier approach

Technical efficiency in a production frontier model refers to the ability of a unit (regional institution, in this case) to achieve maximum potential output (say, human capital) from given amounts of inputs (say, income). Farrel (1957) was perhaps the first to use frontier analysis to measure technical efficiency. Technically efficient units are ones that are on the frontier, while inefficient ones will be located below the frontier as they are achieving less output than is technically possible. Technical efficiency can thus be measured as the relationship between the achieved output (Y) and what would be achieved if the unit were on the frontier (Y^*). That is, $0 \leq Y/Y^* \leq 1$.

The two main methods used in frontier analysis are data envelopment analysis (DEA) and stochastic frontier analysis. The first does not assume a functional form for the frontier (it is non-parametric), and the second does assume a functional form. It is parametric, stochastic, and uses econometric methods. I use this latter method as it allows the frontier model's error term to be decomposed into statistical noise and inefficiency. DEA does not do this, and so attributes deviations from the frontier only to inefficiency. Furthermore, in a cross-section of various regions like mine, this would make it hard to tell whether we are picking up heterogeneity or efficiency-differences between regions.

Stochastic frontier analysis was pioneered by Aigner et al. (1977). Once a functional form has been chosen for the production function, the authors put forward the following model $y = f(x_i\beta) + \varepsilon_i$, where y_i is the output achieved by region i, x_i is the vector of used inputs, β is a vector of parameters to be estimated, and ε_i is an error term composed of two elements, $\varepsilon_i = v_i + u_i$. The component v_i is the symmetric disturbance, capturing random variations in production due to random errors, observation errors and measurement errors, and is assumed to be identically and independently distributed. The component u_i is an asymmetric term, capturing technical inefficiency, and is assumed to be distributed independently of v_i, and to satisfy $u_i \leq 0$. Within this framework, maximum likelihood is used to produce consistent parameter estimates. The values for technical efficiency can be calculated as the expectation of the term u_i conditional on the composed error term ε_i. That is, technical efficiency for the i^{th} region is the ratio of the observed output for the i^{th} region relative to the potential output (frontier function). I use the standard assumption of a half-normal distribution for u_i. Therefore, technical efficiency TE can be measured as $TE_i = exp(-u_i)$, with $0 \leq TE_i \leq 1$ so as to ensure that scores remain either on or below the frontier. TE measures the efficiency of institutions in distributing resources (outputs) based on its inputs.

How do I adapt this model to my case, that is, what do I use as an output and as an input to measure the efficiency of institutions? Acemoglu et al. (2005) mention, in particular, wealth, and physical or human capital. While for my current case, I do not have data on regional physical capital, I do have data on regional literacy rates (a proxy for human capital) and data on regional GDP per capita (a measure of income or economic prosperity). Literacy rates are a useful indicator of how well distributed gains from economic prosperity are.[1] Efficient regional institutions are thus ones whose literacy rates are high relative to their GDP per capita. This formulation is not a causal model, but one that is intended to measure the *association* between the two variables. As such, establishing exogeneity is not a concern here.[2] The relationship between GDP per capita and literacy rates also has a sound economic historical basis. Around this period, governments began public education plans in earnest, often devolved to subnational levels as

[1]See *Appendix 2* for the literacy data sources and *Appendix 3* for robustness tests of these data.

[2]Refer to the discussion around table 4.3 for a test on the usefulness and information content of this relationship.

with Germany, with specific goals to spread and improve literacy. GDP was the taxable base used to fund these plans and achieve these goals. As Engerman and Sokoloff (2000, p. 227-7) write about Latin America,

> *The institution of public primary schools was...the principal vehicle for high rates of literacy attainment and an important contributor to human capital formation. Major investments in primary schooling did not generally occur in any Latin American country until the national governments provided the funds.*

Assuming a Cobb-Douglas production function, using linear presentation, and taking into account my variables, I re-write the function to be estimated as

$$\ln L_{it} = \alpha + \theta \ln Y_{it} + \gamma_i + \theta_t + \mu_{it} \tag{4.1}$$

where i continues to index regions and t years, L is literacy rate, Y is GDP per capita, and γ_i is a countey fixed effects term and θ_t is a year fixed effects term. The last two terms are particularly useful since the first allows me to analyse variation within regions, controlling for potentially confounding effects of national-level institutions, and second controls for the fact that both variables were trending upwards over time. The parameter of interest θ gives the elasticity of output (literacy) to input (per capita income).

Figure 4.7 provides a more intuitive representation of the frontier relationship between literacy rates and GDP per capita. The red line marks the efficiency frontier, at which the "production" of literacy for a given amount of GDP per capita is most efficient. A region's distance from that frontier, with the dashed line providing an example, provides a measure of TE. To derive TE scores, henceforth $INST$, for each region at each year, I ran the model on the entire panel.

Results

The results of model 4.1 are displayed in table 4.2. Estimated values for the parameter θ are positive and highly significant across all years, as expected. The two variances of the two error components indicate that the inefficiency component u is much more statistically significant (T-statistic of -24.56) than the random component v (T-statistic of -12.57). This implies that inefficiency u makes a more

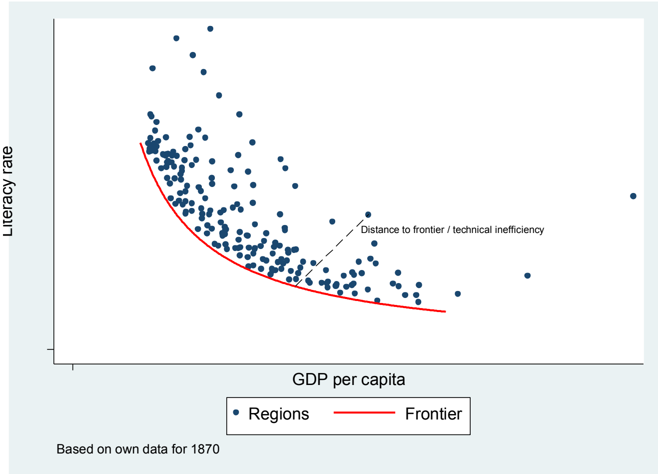

Figure 4.7: Stylised literacy and GDP per capita frontier

Notes: While based on data for 1870, this is a stylised frontier plot, drawn as a visual aid to the discussion on the frontier technique. Harald Tauchmann at the Friedrich-Alexander-Universitaet Erlangen-Nuernberg kindly supplied the Stata code to produce frontier plots.

important contribution to the variability of the total error in the frontier model, and that inefficiency is highly significant across regions and years.

The resulting mean technical efficiency score is 0.79, with a standard deviation of 0.175 and a minimum score of 0.18. The maximum score, which marks the frontier, is always one. The main implication of this result is that the average regional institution could have reduced its input (GDP per capita) by up to 21 per cent without reducing its output (literacy rates) - simply by improving its institutional efficiency. More efficient regional institutions, therefore, could have increased levels of human capital without necessarily increasing their economic prosperity. That idiosyncratic inefficiencies exist tell us the measures of $INST$ are capturing something beyond the simple correlation between literacy and income.

In their evaluation of the human development index (HDI) and its usefulness relative to per capita income, McGillivray and White (1993, p. 187) propose criteria for a variable's redundancy. First, a variable is redundant if the correlation coefficient is higher than 0.90 ('Level 1 Redundancy') or 0.70 ('Level 2 Redundancy'). Second, a variable is redundant if a 'restricted' computation with the relevant component (in this case, GDP per capita) is excluded and remains highly correlated with excluded component. While I cannot test the second cri-

Table 4.2: Frontier model results

	$\ln L$
α	2.684***
	(0.267)
$\ln Y$	0.253***
	(0.037)
$\ln v$	-6.415
	(0.510)
$\ln u$	-1.825
	(0.074)
Log-likelihood	70.319
N	597

*Notes: Estimated on entire panel with country fixed effects, using maximum likelihood. Standard errors are in brackets. *** denotes statistical significance at the one per cent level. Dependent variable is log of regional literacy rates; independent is log of GDP per capita (Y). v is the symmetric (random) disturbance; u is the asymmetric term, capturing idiosyncratic technical inefficiencies.*

terion, since the only component or input is GDP per capita, I can test the first. Table 4.3 shows that while the pair-wise and Spearman rank correlation coefficients are statistically significant, as we would expect them to be, they pass the McGillivray and White (1993) 'Level 1' and 'Level 2' redundancy criteria. That is, $INST$ contains useful information beyond its GDP per capita input.

Table 4.3: Correlations of GDP per capita and $INST$.

	Pearson	Spearman
1870	0.228***	0.223**
1900	0.403***	0.162**
1910	0.420***	0.182**

*Notes: Statistical significance: *** one per cent, ** five per cent.*

Another concern is that $INST$ is not capturing anything beyond its output, literacy. In table 4.4 I carry out a paramteric test of the two variables' independence by regressing $INST$ on literacy rates with both variables expressed in percentage terms for comparability. For evidence of dependence, we would need a coefficient (β) that is significant and close to one, and a constant term (α) that is significant and close to zero. The results of this exercise are in table 4.4.

While the βs are all highly significant, only that for 1870 is near enough to one to indicate substantial dependence. Further, α for that year is 13 per cent,

Table 4.4: Correlations of literacy and $INST$.

	1870	1900	1910
α	12.696	31.043	41.427
	(3.208)**	(2.463)***	(2.677)***
β	0.866	0.625	0.490
	(0.040)***	(0.032)***	(0.038)***
R^2	0.926	0.832	0.646
N	199	199	199
RMSE	5.645	5.605	6.172

*Notes: The dependent variables is $INST$, the measure of regional institutional efficiency, expressed in percentage terms; the independent variable is literacy rates, also in percentage terms. Statistical significance: *** one per cent, ** five per cent. 'RMSE' is the root mean-square error. Robust standard errors clustered by country in brackets.*

which is close to $INST$'s sample standard deviation of 0.17 or minimum value of 0.18. Moving to 1900, the relationship weakens, with a β of 0.63 and an α of 31; almost double the standard deviation of $INST$. In 1910, it weakens further with a β of 0.49 and an α of 41; this time surpassing $INST$'s standard deviation and minimum value. The root mean-square error for all estimations hovers around six per cent. This is large enough to take the sample median of $INST$ (86 per cent) to the top quartile of the distribution (91 per cent). On the balance, the results in table 4.4 indicate that $INST$ contains useful information beyond its literacy rate output measure.

All quantitative measures of institutions are imperfect, and the one I present here is no exception. It is both indirect and incomplete. The challenge I faced here was constructing the same measure for all 199 regions across seven countries at three different points in time within the timeframe of a doctoral project. This excluded a number of options, not least popular national level measures or specific within-country "natural experiments."

One idea for future work would be to use the geographical and temporal variation in the French occupation of European regions in the early nineteenth century. This occupation involved a set of reforms, including the imposition of a civil legal code, the abolition of guilds and feudalism, and the introduction of equality at law, and the undermining of aristocratic privileges. Acemoglu et al. (2011) find for a number of independent German polities that the "spread of the French Revolution" fostered economic progress through these reforms. I have left this for future work because digitising the geographical spread of the reforms, as well as taking qualitative notes of their de facto implementation, is no small task. The analysis of Acemoglu et al. (2011), which is based on one country, is the work of five professors and five research assistants.

4.4 Empirical Strategy

The goal here is to estimate the baseline implementation in model 4.2:

$$\ln Y_{it} = \alpha + \pi \ln Coal\_Transport_{it} + \rho INST_{it} + \gamma_c + \theta_t + \mu_{it} \qquad (4.2)$$

where GDP per capita Y of region i at year t is regressed on my measure of coal transport costs ($Coal\_Transport$), and institutions ($INST$). The terms γ and θ are, respectively, country and year fixed effects. By the start of the period, Europe's regions had formed national units. Using region fixed effects allows me to control for region-specific, time-invariant characteristics like first-nature geography or a region's national institutional context. According to the popular 'executive constraints' measure, these changed very little between 1870 and 1910: no change for Austria-Hungary, France, Spain, Sweden, and Britain; and a one point improvement for Germany from 1900 to 1910, and a two point improvement for Italy between 1870 and 1900 INSCR (2012).[1] Region fixed effects allow me to econometrically exploit spatial variation within regions, filtering out confounding variation. The year fixed effects term controls for shocks common to all regions, but specific to certain years. In practice, these were rare, but it controls for the fact that both Y and $INST$ are variables that were trending up with time, while $Coal\_Transport$ was trending down.[2] The main issue with model 4.2 is, however, the threat of endogeneity between income and institutions. This would result in

[1] "Executive constraints" measures the degree of constraints on a country's political executive. It ranges from zero to seven, with seven being the greatest degree of constraints. Its use in recent years has been popularised by Acemoglu et al. (2001).

[2] On common shocks, there were no major climatic events that affected my sample. The "Long Depression" of 1873 to 1896 misses my benchmark years. The numerous conflicts arising mainly in southern and eastern Europe were bilateral or civil in nature, for example, the Albanian Revolt of 1910, the 1897 Greco-Turkish War or the 1872 to 1876 Third Carlist War.

a positive bias for $INST$ - or the institutional explanation.

4.4.1 Identifying the institutions

Chang (2010) outlines three channels through which incomes can drive institutional development. First, higher incomes may lead to higher demands for higher-quality institutions (for example, demands for institutions with greater transparency). Second, higher incomes make better or more efficient institutions more affordable, as functioning institutions are costly to establish and run. Third, development creates new agents of change who demand new institutions. In the nineteenth and early-twentieth centuries, the working class' growing power led to the rise of the welfare state and proactive labour laws, for example. Chang (2010, fn. 3), in a footnote, writes that these channels are largely ignored by the empirical literature with Acemoglu et al. (2001) being a 'partial exception.' "Partial" because the authors recognise the channels of reverse causality on a conceptual level, but 'go on to conclude, through the use of an instrumental variable, that empirically the causality basically runs from institutions to development.'

While his conceptual critique is robust, Chang (2010) is weaker on the empirical problems and on suggestions. He does not say precisely what the issue with an instrumental variable approach is. An instrumental variable (IV) approach does not totally disregard the two-way relationship, but decomposes it into plausibly exogenous and endogenous variation. Even Chang (2010) admits that some of the variation in development is accounted for by variation in institutions, and that "some" is what gets us one step closer to understanding the causal relationship. Like all econometric estimations, we must interpret it with caution.

The Habsburg Division of 1521

Looking into the patterns of $INST$ gives us an idea of what IV should be used. Table 4.5, which summarises the $INST$ data, shows that the lowest $INST$ values, that is the least efficient institutions, were found in Spain and Italy, with Austria-Hungary much closer to the leaders, but still a laggard. These are the only three countries whose regional mean $INST$ values are below the sample mean. What do these countries have in common? They range from the western to eastern-most extremes of my sample, and are separated first by France and then by Germany. Simple geographical variables often captured with climate, latitude or longitude do not work here (Acemoglu et al., 2001; Rodrik, 2004). While Spain and Italy were both predominantly Catholic countries, Austria-Hungary was home to a large population of Protestants, following the Reformation Berend (2013). Other important cultural variables, like language, were different across and even within the countries (Schulze and Wolf, 2012).

Table 4.5: *INST* summary statistics

Country	Mean	Maximum	Minimum	Standard Deviation
Austria-Hungary	0.768	1	0.193	0.133
Britain	0.785	0.933	0.301	0.179
France	0.847	1	0.347	0.157
Germany	0.946	1	0.693	0.050
Italy	0.445	0.881	0.136	0.208
Spain	0.471	0.838	0.163	0.182
Sweden	0.897	0.957	0.810	0.033
Sample	0.789	1	0.136	0.212

Notes: INST figures of regional institutional efficiency, where a value of one marks the frontier.

What these countries *did* have in common was a Habsburg legacy. Until 1700, before the War of the Spanish Succession, the Habsburg direct-rule dominions - rather than titular sovereignties or vassal states - within Europe included all of Spain, Milan (Lombardy), Sardinia, Sicily and southern Italy, along with the Austrian Habsburg lands.[1] These are the regions, and not just countries, which are afflicted with the lowest *INST* scores in my sample.

Two-sample T-tests of Habsburg direct-rule versus all other regions reject the null of equal *INST* variances with a T-statistic of 7.83 in 1870, 13.85 in 1900, and 13.76 in 1910. The number of Habsburg direct-rule regions is 39 versus 160 all other regions. The degrees of freedom equals 197. The difference in means in

[1]Ceaseless inbreeding meant the Habsburg line died out by the late-seventeenth century. The last Habsburg king, Charles II, designated Philip of Anjou, the grandson of sun-king Louis XIV, as his successor. Fearing a dynastic unification of France and Spain, and a change in Europe's balance of power, the Grand Alliance (consisting, at various times, of Austria, Bavaria, Brandenburg, the Dutch Republic, England, the Holy Roman Empire, Ireland, the Palatinate of the Rhine, Portugal, Savoy, Saxony, Scotland, Spain and Sweden) intervened. The War was concluded with the treaties of Utrecht (1713) and Rastatt (1714), which put the Bourbon Philip V in place as King of Spain, but removed him from the French line of succession.

1870 is 0.291 (standard error of 0.037); 0.347 (0.025) in 1900; and 0.297 (0.022) in 1910.

Given their shared Habsburg legacy, why are $INST$ scores so much higher in Austria-Hungary than in Italy and Spain? On April 21, 1521 King Charles I of Spain (or Emperor Charles V) assigned the Habsburg's Austro-Hungarian possessions to his brother, Ferdinand I. Habsurg Europe thus split into a Spanish branch and an Austrian branch. The first branch ruled over Spain (Castille and Aragon), Sardinia, Sicily, Milan, and southern Italy, along with nominal rule of out-of-sample Flanders. The Austrian branch ruled over the Austrian regions, including Bohemia, and *de jure* presided over Transylvania and the Holy Roman Empire, which was to become the German Empire. In reality, the Holy Roman Empire was a very diverse collection of *de facto* princely states. Parts of Hungary remained under Ottoman rule. Figure 4.8 shows the geographical extent of this division before the War of Spanish Succession.

Two-sample T-tests of *Spanish*-branch Habsburg regions versus all other regions reject the null of equal $INST$ variances with a T-statistic of 8.12 in 1870, 11.47 in 1900, and 8.86 in 1910. the number of Spanish-branch Habsburg regions is 24 versus 175 all other regions. The degrees of freedom is 197. The difference in means in 1870 is 0.319 (standard error of 0.039); 0.265 (0.023) in 1900; and 0.169 (0.019) in 1910. The next question is why.

Mechanisms: 'The laws and statutes of a nation are inherited disease'[1]
Why were Spanish Habsburg institutions less efficient and why did their inefficiency persist until at least 1910? Taking the last point first, historians have long appreciated that 'the legacy of the past posed a heavy burden on the present' (Berend, 2013, p. 320). A number of recent studies have empirically identified institutional persistence over long periods (Acemoglu et al., 2001; Tabellini, 2010). In a relevant example, Becker et al. (2011a, p. 1) show that the institutions of the Austrian Habsburg Empire, which collapsed in 1918, 'still [affect] trust and corruption in local public services in Central and Eastern Europe today.' They

[1] In the words of Goethe's Mephisto, 'The laws and statutes of a nation/Are an inherited disease/From generation unto generation/And place to place they drag on by degrees.'(Goethe, J.W. von, 1961, p. 203)

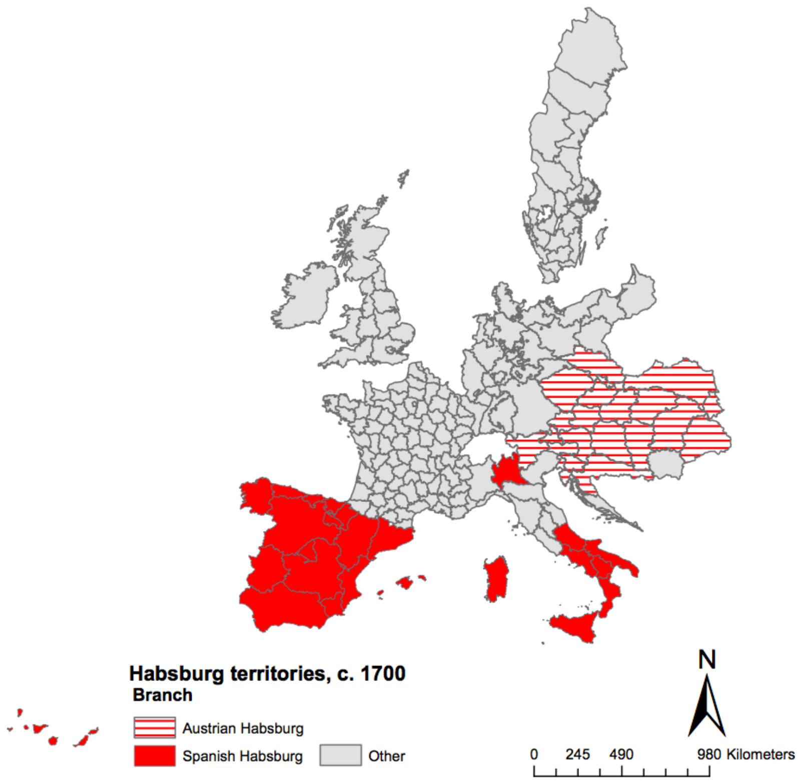

Figure 4.8: Spanish and Austrian Habsburg branches

Notes: This map is imprecise since branches were allocated to regions in their 1913 borders: if a region was for the most part under the rule of one of the branches before 1700, it was coded as belonging to that branch. The sources consulted were: Berenger (1994); Luebke (2014); Palmer et al. (2002); William Ward et al. (1912).

argue that the effects of long-gone formal institutions move through the channels of current cultural norms, values, beliefs, and formal institutions. Emperor Franz Joseph, for example, was known to start his day early, and expected to able to contact his civil servants equally early. In the Czech Republic today, public offices still generally open at seven in the morning (Becker et al., 2011a, p. 8). Using data from the *Life in Transition Survey*, the authors find that respondents from previously Habsburg-ruled areas were more likely to have higher levels of trust in courts and the police, and are less likely to pay bribes. This was a legacy of the trustworthy institutions that came before. As Taylor (1948, p. 44) wrote of the eighteenth and nineteenth centuries, 'the Austrian bureaucracy was fairly honest, quite hard-working, and generally high-minded, it probably did more good than harm.'

The institutions Taylor (1948) discusses and Becker et al. (2011a) analyse have their roots in the sixteenth century. Once Ferdinand I was given reign over the eastern dominions, he immediately sought to establish a common administration and royal authority over what was a fragmented territory. Predictably, he clashed with the kings and princes who came before him; who enjoyed 'personal privileges...and...public liberties, and a long period of omnipotence' Berenger (1994, p. 162). Intelligent and resolute, unlike his profligate brother, Ferdinand I tackled these problems of re-organisation, created a central government, established order in particularly unruly regions like Bohemia, harmonised legal and financial systems, and reduced the Diet's power. By royal decree on 1 January 1527, the central government now consisted of an Aulic Council (a supreme court), Privy Council, a Post Master (a state postal service), a court chancellery and a chamber of accounts (a forebearer to a ministry of finance). As Berenger (1994, p. 162) wrote, 'jurisdiction extended over the whole monarchy, without distinction between countries and particular privileges.'

Similar changes during this period were happening in the rest of my European sample, *excluding* the Spanish Habsburg regions. Even a brief overview of these widely discussed changes provides a contrast to the Spanish Habsburg regions.[1]

In *Britain*, the end of the conflict between Parliament and th Stuarts in 1689

[1]See Craig and Fisher (1997) for a useful overview of European political integration in the nineteenth century.

led to two significant events in econonic history. First, the perpetuation of Parliament's right to approve new taxes. Second, the establishment of a consitutional monarchy, which delineated the interests and responsbilities of the monarch, whose survival depended on the recognition of the Parliament. Rules were now made by a body - Parliament - whose 'interests were best served by private property and elimination of crown monoploies' (North, 1981, p. 156). This institutional set-up guaranteed the property rights of a new commercial class, which in turn invested its finances and human capital in industry, trade, and education. Fiscal consolidation followed the growing tax base, giving the country institutions, like the Bank of England (founded in 1694), which persist to this day.

France made no move towards a constitutional monarchy. Through marriage and conquest, the Bourbons came to control the entire country in 1589, which was previously divided among Brittany, Anjou, Bourbon, Valois, and Burgundy. The Bourbons financed their wars and conquests through local monopolies enforced by guilds in return for taxes. The offices responsible for the collection, disbursement, and borrowing of revenues were all up for sale. Though unthinkable by contemporary standards, this consolidated the French state very early on and created an effective bureaucracy. Some economic historians refer to the period as 'the beautiful sixteenth century' (Goubert, 2002). The well-documented political and social upheaval that marked the end to the Bourbon reign ushered in, starting around 1792, important institutional change: 'the imposition of the civil legal code, the abolition of guilds and the remnants of feudalism, the introduction of equality before the law, and the undermining of aristocratic privileges' (Acemoglu et al., 2011, p. 3286).

Germany was the last to take its modern form. The Catholic-Protestant divided in the Holy Roman Empire during the Reformation delayed the country's political integration. Much of what became Germany was under the titular sovereignty of the efficient Austrian Habsburgs by 1555. In that year, Ferdinand I signed the Peace of Augsburg, which sanctioned the defeat of his brother's aggressively intolerant (anti-Protestant) religious policy. In contrast to his brother's rule, Ferdinand I, much like he did in the Austrian dominions, sought a compromise between the various stakeholders: Catholics, Lutherans and Protestants. He allowed the hundreds of princes to make their own religious policy (*Cuius regio,* *eius religio*), and define the religious confession of their territories. He also recognised the secular status of ecclesiastical property secularised before 1552. This overall peace was interrupted by the Thirty-Years War, which ended in 1648, and was the worst of the religious conflicts. Costly in economic and human terms, it split Germany into more than 300 principalities that would take another 200 years to mould together. Starting in the eighteenth century, the unlikely Hohenzollerns, the ruling House of the Electorate of (then peripheral) Brandenburg, became a force for German integration. Between 1720 and 1772, they conquered West Pomerania, West Prussia, and Silesia, along with many other territories. A consolidated system of excise taxes, evenly distributed across income classes, helped fund their wars. A series of customs unions, of which the Prussian Customs Union (1828) was the first and the Zollverein (1834) the most decisive, set Germany on the path to coherent state.

Sweden between 1387 and 1523 was part of the Kalmar Union of Norway, Denmark, and Sweden. In 1523, Gustavus Vasa (a Swedish noble, and top administrator of the joint kingdom) took Sweden out of the Union, 'and began a series of adventures in Scandinavia and on the Continent that brought Sweden to a 200-year reign as the most powerful economic and military force in the Baltic' (Craig and Fisher, 1997, p. 30). While it gained and lost territories during the Thirty Years War, where it allied with France despite being Protestant, its sovereinty was never under threat. Indeed, even by 1611 Sweden passed a royal charter that limited the powers of the King. The country can claim the world's first central bank - the Riksbank, founded in 1668 - and thanks the famous Church Law of 1686, 'had the most educated population in the whole of Europe' (Craig and Fisher, 1997, p. 31). According to Sandberg (1979, p. 229) 'this statute made the [Lutheran] parish priests responsible for assuring that every young person in their charge learned to read the Gospel and other specified religious works, and even to write.' Helpfully, the Swedish clergy was numerous, university-trained, and reached large areas of the peasantry.

By contrast, the *Spanish Habsburg regions* did not enjoy the same institutional changes - even though some were half-measures or took rocky paths to completion - as did the Austrians, and other Europeans discussed above.[1] Their ruler, Charles

[1] Grafe and Irigoin (2006, p. 41), who throw a more benign light on the Spanish Empire,

I, had an ‘aristocratic...and medieval conception of patrimony...and...matters of government and matters of family for him were closely connected’ (Berenger, 1994, pp. 140-4). It is with this mind-set that he approached governing Spain and much of Italy. Berenger (1994, p. 145) gives evidence of hindered institutional advance in Spain through tax revenues, which in 1523 were five per cent of what they were in France. Further, Spanish revenues showed no real increase from 1504. Charles I was financing his wars with France, as he spent money on little else, through credit from German, Genoese and Antwerp bankers. The uncontrolled accumulation debt throughout Habsburg Spain is a clear indicator of weak institutions, with clear contemporary parallels of Spanish public debt profligacy (de Grauwe, 2010). In sum, Charles I proved to be an incapable ruler, retiring some 30 years after the 1521 Habsburg division to a Hieronymite monastery in Extremadura, Spain. Control of the Spanish Habsburg regions was then assumed by the infamous serial defaulter Philip II (Drelichman and Voth, 2011).

We can see more institutional decay quite dramatically in the Spanish Inquisition, which Philip II greatly expanded in the mid to late-sixteenth century, making Church orthodoxy a goal of public policy. Following this, we see the expulsion of the industrious Moriscos, contrasting greatly with the societal trust between citizens and their public institutions found in the Austrian dominions (Becker et al., 2011a). Indeed, attempts at reforming Spain’s bloated and inefficient bureaucracy running through to the mid seventeenth century were met with staunch resistance (Elliott, 2009). Phillip II also continued his father’s legacy of plundering the state to finance unsuccessful wars, accumulating ‘towering debts while stopping all payments to his lenders four times’ (Drelichman and Voth, 2011, p. 1205). Pointing to inefficient and ineffective rule, historians locate Spain’s long decline between the late-sixteenth and mid-seventeenth centuries (Elliott, 1961; Hamilton, 1938; Thompson and Yun, 1994).

Some interesting variation comes from the prosperous state of Milan (what became Lombardy), which the Spanish Habsburgs also ruled over. We normally think of Lombardy as a developed region not just of Italy, but of Europe.

still show that even in Latin America the ‘Spanish [Habsburg] path to the formation of an empire turned out to be a poor basis for state formation and institution building in the post-independence period.’

While under Spanish rule, Lombardy managed a mean $INST$ score of 0.738: respectable, but still below the sample mean in table 4.5. As Tabellini (2010) documents, even though Charles I had taken over with a new constitution, this legislation was drafted by Lombard jurists on the basis of local legal traditions. The Spanish Habsburgs were in effect 'caretakers' of local traditions (Tabellini, 2010, p. 42). The Lombard Senate had, in the words of Tabellini (2010, p. 42) 'strong powers in implementing the law and the king's pardons, and was able to exert strong influence on the whole legislation. The senate often refused to implement the Governor's deliberations, appealing against them to the king's final decision.' The *de jure* rule of Spanish Habsburgs and *de facto* rule of the Lombard Senate explains the below-mean, but not disastrous, institutional efficiency of this region.

The remaining parts of Habsburg Italy were not so fortunate. Tabellini (2010, p. 45) describes them simply as 'absolutist and autocratic,' giving them the lowest rank in his institutional scoring system. Berend (2013, p. 319) writes that even the reforms that came with Italy's *Risorgimento* in the 1860s could be described as 'a failed revolution' of 'pseudo-reforms' and 'spurious changes.'

Given the historical institutional differences between Spanish Habsburg regions and others, and given the persistence of institutions, I define my IV SH as $SH_i = \{1 \; if \; ever \; under \; Spanish \; Habsburg \; rule \; 0 \; otherwise\}$. Following this definition, $SH = 1$ for 24 (17 Spanish and seven Italian regions) of my 199 regions, that is, 12 per cent of the sample. It is unlikely that SH breaks the exclusion restriction. Spanish Habsburg regions show no obvious geographical pattern (latitude or longitude), with regions in the western-most of my sample; southern Italy, and northern-most Italy. These regions show no co-location patterns with natural resources either, as a look back at map 4.5 shows. Neither am I simply identifying regions of low per capita income levels, since the lowest were in Austria-Hungary and, for example, Lombardy enjoyed high income levels. Religion, or Catholicism, is the only semi-palusible candidate. SH regions were all Catholic. But then so were a number of high-income efficient-institution German regions, and so was France. Secondly and more fundamentally, it is difficult to conceptually disentangle religion from public institutions in an era of divine right monarchs, Papal States, Holy Empires, and people like Cardinal Richelieu. We should not enforce a separation of powers if one did not exist.

There is one final necessary adjustment. As it stands SH is time-invariant, making incompatible with my regional fixed effects implementation. To introduce variation over time, as well as place, in the legacy of Spanish Habsburg rule I interacted the SH dummy variable with the inverse number of years from the Habsburg Division until each benchmark year in the sample. Put together, the components measure the idea that regions under Spanish Habsburg rule needed more time to recover from their institutional legacy.

The empirical implementation

The correlation between SH and $INST$ is -0.456, significant at one per cent. Regressing $INST$ on SH gives a coefficient of -0.001, significant at one per cent, with an R^2 of 0.21. The reduced form estimation - regressing $\ln Y$ on SH - gives a coefficient of -0.001, significant at one per cent, with an R^2 of 0.19. The conceptual as well as empirical basis for SH has been set out, so my IV estimation strategy is as follows:

$$\ln Y_{it} = \alpha + \pi \ln Coal\_Transport_{it} + \rho \widehat{INST}_{it} + \gamma_c + \theta_t + \mu_{it} \tag{4.3}$$

$$INST_{it} = \alpha + \pi \ln Coal\_Transport_{it} + \varnothing SH_i + \gamma_c + \theta_t + \mu_{it} \tag{4.4}$$

where inthe first stage (model 4.4) SH instruments $INST$. In the second stage (model 4.3), the predicted values, $\widehat{INST}$, are used to explain GDP per capita, along with $Coal\_Transport$. This specification represents the clearest framework for estimating the respective effects of coal access and regional institutions. It is general, straightforward, and treats both variables symmetrically, giving either one an equal chance.

4.4.2 Results

In column (1) of table 4.6, both coefficients enter with their expected signs. The more costly it is for coal to reach a region, the lower its per capita income. While highly statistically significant, the effect is small, with a 10 per cent increase in cost leading to a two per cent decrease in income. In contrast, the effect of $INST$ is both large and statistically significant. A 10 per cent increase in a region's $INST$ score results in a 23 per cent increase in its per capita income level. As a proportion of the sample mean income level, this effect equals $607, which is greater than Galicia's (Spain) income in 1870.

In column (2) I swap the cost to coal measure for the simple distance to coal measure. The latter is correctly signed, but insignificant. $INST$ remains significant and in the same order of magnitude. Column (3) introduces year fixed effects. Here the magnitude of $INST$ drops substantial to 0.203 from 0.817 in column (1). Further, the cost to coal measure has been rendered insignificant. One interpretation of the insignificance of $Coal\_Transport$ is that the need to be near coal declined dramatically towards 1910, and this decline is not fully captured by the transport cost measure, but is captured by the year fixed effects (which are statistically significant). In column (4), I introduce region fixed effects. The size of $INST$ increase slightly to 0.297, and remains statistically significant. The coefficient on $Coal\_Transport$ remains insignificant.

While statistically significant in the fixed OLS estimations, $INST$ has modest effect on regional per capita income levels. Is $INST$ robust to an IV estimation?

Table 4.6: Coal and Institutions OLS estimates

	(1) $\ln Y$	(2) $\ln Y$	(3) $\ln Y$	(4) $\ln Y$
$\ln Coal\_Transport$	-0.189***		-0.028	0.000
	(0.034)		(0.028)	(0.044)
$INST$	0.817***	0.920***	0.203***	0.297***
	(0.053)	(0.048)	(0.046)	(0.048)
$\ln Coal\_Distance$		-0.009		
	(0.027)			
Year F.E.	No	No	Yes	Yes
Region F.E.	No	No	No	Yes
$\hat{R}^2$	0.393	0.263	0.868	0.866
N	597	597	597	597

*Notes: *** denotes significance at one per cent. Robust standard errors clustered on regions, and reported in brackets. Y is GDP per capita; $INST$ is a measure of regional institutional efficiency; and $Coal\_Distance$ is distance to nearest coal deposit; $Coal\_Transport$ is transport cost to nearest coal deposit.*

Column (1) in table 4.7 shows the reduced form estimation, where GDP per capita is regressed on the IV SH, controlling for $Coal\_Transport$ and region and year fixed effects. The highly significant and negative coefficient on SH provides support for its use as an instrument. The coefficient on $Coal\_Transport$, as in the first of the OLS estimates, is significant. The first stage results in column (2) show that SH has a large negative effect on $INST$, even when controlling for fixed effects and $Coal\_Transport$, which itself has no effect on $INST$.

Given the estimates in columns (1) and (2), it is perhaps unsurprising that SH passes an (Angrist-Pischke) excluded-IV F-test and an (Anderson) under-idenification test. The second stage results in column (4) show, again, that $Coal\_Transport$ is insignificant. We can also see that correctly identifying $INST$ yields a larger, but less significant, coefficient in comparison to the OLS estimate in column (4). The IV estimate implies that a 10 per cent increase in a region's institutional efficiency results in a 15 per cent increase in its per capita income level; compared to 23 per cent for the OLS estimate.

One important thing to keep in mind on the institutional effect is that while the potential gains of improving institutional efficiency may have been large, it does not mean they were realised. The transition probability of a region's institutional efficiency going from the 25^{th} percentile to the 50^{th} percentile between 1870 and 1900 was 20 per cent; from the 50^{th} to the 75^{th} it was 22.5 per cent. Between 1900 and 1910, a much shorter period, those same probabilities were zero per cent (that is, no region's institutional efficiency went from the 25^{th} to 50^{th} percentile) and 20.4 per cent. In both periods, most regions stayed where they were. Some 68 per cent of the regions in the 25^{th} percentile in 1870 remained in the 25^{th} percentile by 1900. Even more strikingly, 96 per cent of the regions in the 25^{th} percentile in 1900 remained in this percentile in 1910. If frequencies of transitions indicate the ease of that transition, then these figures would show us that income gains from improvements in institutional efficiency, although strong, were unlikely.

While it might be difficult to have full confidence in the point estimates of $INST$, the results in tables 4.6 and 4.7 do show that institutional efficiency mattered for regional per capita income levels. On the other hand, access to coal, whether measured by geodesic distance or transport costs, did not matter. A closer look at history, and indeed the basic correlations in figures 4.4 and 4.5,

Table 4.7: Coal and institutions IV estimates

	(3)	(4)
	OLS	IV-2
	$\ln Y$	$\ln Y$
$INST$	0.297***	0.408*
	(0.045)	(0.222)
$\ln Coal\_Transport$	0.000	-0.007
	(0.045)	(0.034)
Year F.E.	Yes	Yes
Region F.E.	Yes	Yes
$\hat{R}^2$	0.867	0.650
N	597	597
Angrist-Pischke multivariate F-test		12.56
Anderson canon. corr. LM statistic		12.28
	(1)	(2)
	RF	IV-1
	$\ln Y$	$INST$
SH	-0.001***	-0.003***
	(0.000)	(0.000)
$\ln Coal\_Transport$	-0.278***	-0.027
	(0.032)	(0.034)
Year F.E.	Yes	Yes
Region F.E.	Yes	Yes
$\hat{R}^2$	0.653	0.551
N	597	597

*Notes: *** denotes significance at one per cent; ** at five per cent; and * at 10 per cent. Robust standard errors clustered on regions, and reported in brackets. Angrist-Pischke multivariate F-test is on the excluded instrument, SH (Spanish Habsburg Regions). Anderson canon. corr. LM statistic is for under-identification. Y is GDP per capita; $INST$ is a measure of regional institutional efficiency; and and $Coal\_Distance$ is distance to nearest coal deposit; $Coal\_Transport$ is transport cost to nearest coal deposit. Column [1] is the reduced form estimation; [2] is the first stage; [3] is the OLS estimation; and [4] is the second stage. Mark Schaffer at Heriot-Watt University supplied the Stata code for this estimation.*

shows that this should not be much of a surprise.

4.5 Discussion

There are at least three reasons that can explain coal's insignificance in this context. First, transport costs declined so much over the period that they ceased to be important, especially relative to other costs (Jevons, 1915; Wright, 1990). Second, for regions that did not have access to coal, energy substitutes existed (Simpson, 1997). Third, technological advances in industrial production lowered the price of coal relative to other inputs (Simpson, 1997).

Taking the last point first, Simpson (1997) reminds us that the Bessemer process - the first technique for the mass-production of steel from pig iron, patented in 1855 by Henry Bessemer - favoured locating industry near iron ore rather than coal deposits. The process, by heating iron ore more efficiently allowed for 'a dramatic reduction in costs' of coal inputs (Strassman, 1959, p. 343). It also increased the value of iron ore deposits, which were previously used to produce iron only. The Basque region (Spain) 'enjoyed Europe's best resources of high grade non-phosphoric [iron] ore' (Simpson, 1997, p. 353). It had, however, poor access to coal: Bilbao, the main centre, was 209 kilometers from the nearest deposit versus the sample mean of 181 kilometers. Once the Bessmer process took hold commercially, in the late-nineteenth century, the Basque region developed an *international* comparative advantage in iron, exporting a quarter of its output between 1881 and 1910 (Simpson, 1997, p. 353). The same is true of Spain's other affluent, iron-producing region Catalonia, which is also coal-free.

Figure 4.9 shows the correlation between regional GDP per capita and regions' distances to their nearest *iron* ore deposits, from the same USGS (2013) source. The strong negative correlation is immediately clear, but what is important here is that the correlation is stronger than that of GDP per capita and distance to coal, as in figure 4.4. Here we have an R^2 of 0.147 versus 0.005 for coal-distance, and a T-statistic of -10.13 versus -1.78. While the effect of iron-distance ultimately washes out in the same way as coal in the models I used throughout this chapter (see *Appendix B* for some empirical tests), it does seem unusual that the recent literature has focused on coal when iron is more likely to provide a better explanation for the location of industry. This is a point not missed by an earlier generation of economic historians. Landes (1965, p. 456-7) relates the

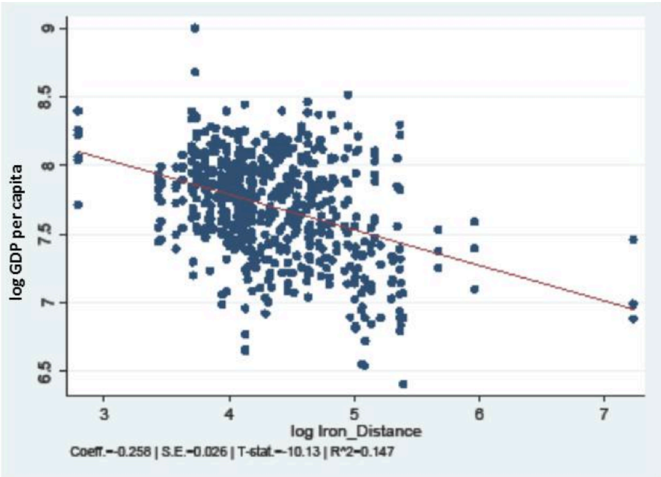

Figure 4.9: Iron ore and income correlation

Notes: Values on y-axis are log GDP per capita. Iron_Distance is the geodesic distance in kilometers to the nearest iron ore deposit. Includes full sample. For sources, see text.

'industrial map of Europe' in 1870 to deposits of *iron* ore - but more recently attention has shifted almost exclusively to coal.

The second point seems to hold some power, given the results in tables 4.6 and 4.7. There is some more evidence I have to back up the idea that coal transport costs, or access costs broadly conceived, faded into significance over the late-nineteenth century. First, Jevons (1915) documents a rapid growth in Britain's coal exports to coal-poor areas over the late-nineteenth century.[1] While greater export demand may reflect a better ability to pay for coal in industrialising Europe, the rate of growth, as can be seen in table 4.8, would be inconsistent with prohibitive or even high transport costs. By 1912, Britain controlled 70.8 per cent of the world's sea-borne coal trade (Jevons, 1915, p. 681). Other coal-rich countries were also exporting coal, however: in 1912, Germany exported 10.36 million tons (four million in 1906) and Belgium, 1.28 million tons (0.69 million in 1906) (Jevons, 1915, p. 681).

The primary reason for this growth cited by Jevons (1915, p. 691) is 'the wonderful fall of freight rates [which] far exceeds the reduction in the cost of railway transport, or in the price of any of the staple commodities of trade.'

[1] This is *not* the same "Jevons" as Jevons (1865) of 'The Coal Question' fame. Jevons (1915) is a far better resource for economic historians.

Table 4.8: Britain's coal exports ('000 tons)

Importers	1887	1912
Baltic and North Sea	7,296	25,271
France and Mediterranean	11,814	31,132
Brazil, Argentine, Uruguay, and Paraguay	1,203	5,879
North and South America and Pacific Coasts	300	575

Notes: From Jevons (1915, p. 683).

The "transport revolution" durinsg this period is understood well by economic historians (O'Rourke and Williamson, 1999). It was also recognised quite clearly by contemporaries. Declining freight rates, according to Jevons (1915, p. 691),

> *...resulted from the manifold improvements in steam navigation... Vessels have been built of steel, and much larger, and with engines more economical in fuel consumption; so that both the initial cost [terminal cost] and the cost of running per ton of carrying capacity [variable cost] have been greatly reduced.*

He goes on to explain that while the carrying capacity of ships increases as the cube of their dimensions (lenght or width), the resistance to water increases approximately only as the square of such dimension. This makes for savings in fuel. Further, the navigating staff increases much less proportionally with the size of a ship. A large steamship can carry much more coal than a small one, and only needs a few, if any, additional officers. Jevons (1915, pp. 692-3) provides a table of 'outward freight rates from Cardiff [a major coal exporting city] to representative foreign ports.' I decimalised the figures, deflated them into 1910 pound sterling, and plotted them in figure 4.10.

As figure 4.10 makes clear, the decline in coal freight rates was universal and fast. The start of the period is characterised by a wide dispersion in freight rates, with all rates at an elevated level. The lowest rate in 1864, for Bordeaux (France) at £1.50 per ton, dropped to less than £0.50 by 1912. Even more dramatic was the decline in rates for destinations that pre-Suez had to be reached via the Cape. Singapore's rate dropped from £4.50 in 1864 to around £0.50 in 1912, and Bombay's from £4.25 to £1.00, as the nautical distances to these locations dropped by 29 per cent and 41 per cent, respectively. This reduction in intercontinental transport costs should put the debate on coal access in Continental Europe in

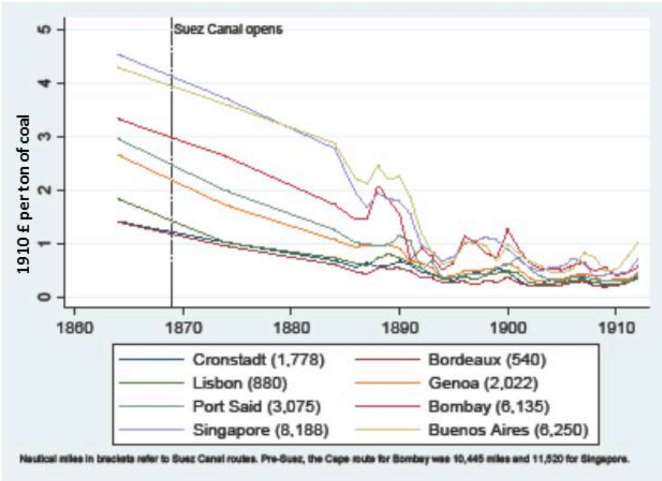

Figure 4.10: Coal freight rates from Cardiff

Notes: Values on y-axis are 1910 £per ton of coal. Data, including nautical miles in brackets, from Jevons (1915, pp. 692-3). Decimalisation: 12 pennies per shilling, and 240 pennies per pound. The numbers for Bordeaux are given in francs, which were converted using Global Financial Data (2014). London and Paris consumer price indices from Allen (2001).

perspective. The average distance between regional nodes and coal deposits is 113 miles (98 nautical miles).

The effects of this drop in transport costs can be seen in the unusually detailed German data on domestically produced versus imported coal prices - which include tariff costs - from the *Die Grosshandelspreise in Deutschland von 1792 bis 1934* (Jacobs and Richter, 1935).[1] The spread or imported and domestic coal prices, displayed in figure 4.11, went from an average of £1.00 between 1850 and 1860 to £0.13 by 1913. That is, Germans were paying a premium of 13 pennies per ton for imported coal. This number represents 21 per cent of the price of a ton of coal in 1913. This is a much lower figure than that given by Mokyr (1983, p. 152-8) (quoted in Fernihough and O'Rourke (2014)) for Ireland's import-premium on British coal of between 100 and 150 per cent. Further, Mokyr (1983)s writes that fuel costs in 'nonmetallurgical industries' were at most four per cent of total costs, implying that Irish costs were pushed up by at most 10 per cent relative to British costs. This brings us to Jevons's second reason for the growth in the coal trade.

[1] Alexis Wegerich at the University of Oxford kindly supplied these data.

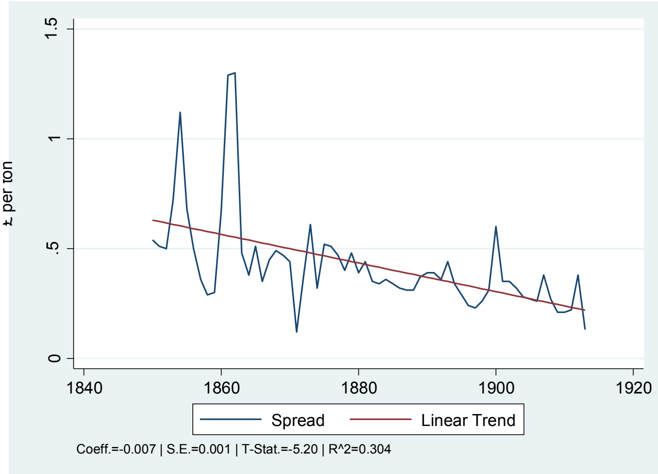

Figure 4.11: German imported and domesitc coal price spread

Notes: Values on y-axis represent the difference in £ between imported and domestically produced coal. Data from Jacobs and Richter (1935). The underlying numbers were converted from American dollars into pound sterling using Global Financial Data (2014), and refer to the price per ton of coal. Alexis Wegerich at the University of Oxford kindly shared these data.

As 'the peoples of Europe have gradually become more wealthy, so have they been able to purchase more coal for domestic as well as industrial purposes' (Jevons, 1915, p. 690). As the cost of producing and transporting coal declined, the purchasing power of Europeans rose. The expenditure proportion of fuel costs dropped. Further, this also matters because, as Allen (2006, p. 10) summarises the argument against his case, '...businesses are only concerned about costs *in toto*-and not about labor costs or energy costs in particular-so all cost reductions are equally welcome.' While businesses in coal-poor regions in, say, Spain or Hungary paid a premium on coal, they enjoyed lower input costs elsewhere, especially with labour. Further, there are substitution possibilities to consider.

Simpson (1997, p. 353) tells us that in Spain, for example, the

> *...cotton textile industry...adapted to high coal prices, with the improvement in turbine technology allowing the industry to relocate away from Barcelona's coastal plain [where imported coal arrived] to the mountainous interior [where rivers flowed fastest] after 1860. By 1914, about 80 per cent of its spindles were water driven, and the hydraulic energy used was equivalent to roughly a quarter of Barcelona's*

coal imports.

Similarly, Zamagni (1993, p. 92) writes that in Italy, a virtually coal-free country, by 1913 coal was rarely used in 'industrial processes.' According to her figures, 'only 20 per cent of industrial energy took the form of [coal-powered] steam power, a further 22 per cent was hydraulic, whereas 48 per cent was...generated by electric motors, which made it possible to decentralise factories and workshops...'

For the 1870 to 1910 period at least, the empirical and historical case for a coal based explanation of European regional income is weak. We have to consider costs of accessing coal, and those costs relative to other inputs, the presence of potential more valuable endowments like iron, and of substitutes like rivers and water power. Furthermore, in some cases, as in northern Britain and central Germany, coal might have been the primary reason for industrial location, but high *income* did not necessarily need coal: think of London and Brandenburg. That institutions trump coal in my empirical results also tell us something about coal's second-order nature.

Spain had numerous coal deposits, as figure 4.5 shows. But as the discussion surrounding 4.5 shows, it also had weak institutions. Indeed, Berend (2013, p. 318) writes of 'a medieval mining law [that] impeded the extraction of Spain's highly abundant natural resources.' When this law was modified in 1825, it was 'deemed that all of the country's natural resources belonged to the crown - thus assuring that all landowners would have no interest in exploring for natural resources on their properties.' With the Spanish Liberal Mining Act of 1868, nationals and foreigners were allowed to obtain mining rights 'as perpetual concessionaires by paying royalties to the State' (Lieberman, 2013, p. 126). The first law outright prevented resource extraction, and the modified versions created monopolies and ensured that large gains go to the crown - hardly what we would call institutional efficiency.

Institutions do not only determine whether a natural resource is exploited. They determine per capita income in their own right, and they do so on a subnational scale. While European economic history has only just started dealing with these issues empirically, Latin America has provided fertile ground for research into subnational institutions and their economic effects. Dell (2010), for example, examines the long-run effects of the *mita*, an extensive forced mining labour system used in colonial Peru and Bolivia. She finds that the *mita* effect lowers household consumption by 32 per cent in subjected districts today. Dell (2010) traces this to the fact that *mita* districts historically had fewer large landowners and lower educational attainment. Today these same districts are poorly integrated into road networks and their residents are more likely to be subsistence farmers. Acemoglu et al. (2008) find a negative association between political inequality in nineteenth century Cundinamarca, Colombia (measured by the lack of turnover mof mayors in the municipalities) and economic outcomes today. They argue, in line with Dell (2010), that the availability of local (subnational) public goods is an important intervening channel. Naritomi et al. (2012) obtain similar results for Brazil.

As Acemoglu and Dell (2010) summarise, the recurring theme in this line of research - what they refer to as "local institutions" and current economic outcomes - is an attempt to isolate a source of historical exogenous variation in future institutions. My attempt at this was to use institutional legacies from the sixteenth century. If it stands up to future research, then the implication is that the 1521 Habsburg Division has had important implications for regional income differentials in Europe. Historians have long traced patterns of European development far back time. Berend (2013, p. 324) writes that 'the long survival of the *acien regime* [broadly conceived] was accompanied by a lack of education, and mass illiteracy.' Berend gives some figures on southern Italy and Spain, inheritors of the Spanish Habsburg institutions, showing that illiteracy was generally high in Italy (75 to 80 per cent in 1890), but 'much higher in the south of the country.' In Spain, Berend (2013, p. 324) writes, 'secular and scientific thinking' was 'suppressed' - 'an oppressive legacy of the Inquisition since the sixteenth century.' For all its flaws, an instrumental variable strategy specifies the long historical roots of underdevelopment more clearly. This, as opposed to coal, appears to be a more promising line of research, especially constructing new measures of subnational institutions.

Chapter 5

An Alternative Explanation: Market Potential

5.1 Introduction

Can market potential explain Europe's late-ninteenth regional per capita income structure? In *Chapter 3*, I outlined Europe's spatial income structure, which by 1910 concentrated in the northwest, relegating the rest of the Continent to the economic periphery. In *Chapter 4*, I showed that, for most of Europe, the cost of accessing coal deposits was not correlated with higher regional per capita income levels, on account of the dramatic reductions in transport costs. In contrast, my measure of regional institutions is significantly correlated with per capita income levels, indicating that in this period of state-formation regional *de facto* institutions played a role in the location of economic activity. Still, these institutions cannot fully account for Europe's spatial income structure or, more specifically, the northwestern clustering in regional per capita income.

In this chapter, I provide an alternative explanation; one that can account for the spatial and temporal distribution of regional income and fits with the broader historical context of falling barriers to trade. Since the pioneering work of Krugman (1991), market potential has been used to empirically explain county-level wages in the late-twentieth century United States, regional income in late-twentieth century Europe, and industrial location in interwar Poland, for example (Hanson, 2005; Wolf, 2007). The basic idea is that the level of economic activity (which can be measured by per capita income) in a region is conditioned by that region's access to markets for its goods. Economies of scale and trade costs - transport and tariff costs - created demand linkages between regions which contribute to agglomeration. Producers are drawn to economically active regions by the prospect of serving their large markets at low trade costs. Congestion costs, which come with higher property prices and labour costs, act to limit the degree of geographical concentration.

A corollary of this argument is that when trade costs are very high or very low, regional income levels will be dispersed. At very low costs, production would not need to concentrate in a particular region: this would create congestion costs that are greater than serving the regional market from a distance. At very high costs, markets would be dispersed and production would focus on serving distinct regional markets.

A further corollary is that even with perfect institutions everywhere, the integration of regional markets may lead to economic divergence. This is an important point in light of the *Chapter 4*'s finding that regional institutions have a significant effect on per capita incomes.

The market potential idea is based on Harris (1954), who argued that the demand for goods produced in a region is the sum of purchasing power in surrounding regions, weighted by trade costs to those regions. Since Harris (1954), geographers have used the ad hoc formulation to successfully explain urban patterns and income structures (Keeble et al., 1982). More recently, economists have derived estimates of market potential from formal models of bilateral trade

The rest of this chapter is organised as follows. First, I examine the neo-mercantilist arguments put forward in the historiography. I then describe my measures of market potential and the sources used, giving further details in *Appendix C*. I then set out my empirical strategy, which is geared towards uncovering baseline effects of market potential on income, differences between foreign and domestic market potential, and whether the relationship is uniform across regions. The final section concludes with a discussion of the issues raised.

5.2 The Age of Neo-Mercantilism?

Pollard (1981, p. 252) described the 1870 to 1914 period as one of growing neo-mercantilist tendencies among European states. He writes that the openness that allowed for Europe's industrialisation running up to 1870 began to decline thereafter, as political authority became progressively more powerful. In his words,

> *...governmental actions increasingly came to disrupt the relatively easy and free intercourse of commodities as well as factors of production between nations on which the successful industrialisation of Europe had so largely depended.*

In particular, Pollard (1981) bases his characterisation on three different trends. Firstly, there was growing nationalism across the Continent. As a force of action, this was actually a positive trend: feudalism was destroyed; new institutions like the Napoleonic Code introduced; and mass literacy, required for urban living and factory work, was promoted by states. The reaction, however, was negative. Russia, for example, deliberately held back Polish efforts at industrialisation, lest they led to Polish independence. Secondly, as economies and the civil services managing them grew, so did the opportunities for taxation. Greater tax revenues were required to fund previously non-existent public services like sanitary and safety interventions in working class areas and on sea vessels, education, and a police force. Thirdly, industrialisation was concentrated in a small proportion of northwestern regions, as we saw in *Chapter 3*. These industrial regions were surrounded and outnumbered by agrarian regions which, given the growing franchise and democratisation of political power, were able to push for protectionism, particularly in agricultural goods and as a response to the influx of cheaper grain from the US and Russia (O'Rourke, 1997).

Pollard (1981, p. 258) illustrated the practical effect of these trends through import tariffs. German rates were initially moderate, around 10 to 15 per cent of value on industrial goods and five to seven on agricultural goods, but these rate rose in 1885 and then again in 1887, by when the corn tariff increased five-fold. In France, agricultural protection increased in 1885 and again in 1887, in synch with Germany. The Meline Tariff of 1892 made clear France's protectionism. Austria-Hungary implemented similar tariff hikes starting 1878; while Italy, which was initially free-trade, became clearly projectionist especially in manufactures by 1894.[1]

A salutary lesson in considering variables in isolation

While the historical argument for neo-mercantilism is appealing, and potentially devastating to my own own market potential thesis where declining barriers to trade is central, it relies on selective evidence. First, by looking at specific products or industries, the general trend, which is after all what Pollard (1981) seeks to explain, is obscured. We cannot describe this period as neo-mercantilist by looking at tariffs on industrial or agricultural goods in isolation. Issues arise with the grouping of products and industries, and how well they are reflected in trade volumes and values. According to Pollard (1981, p. 259), the average *ad valorem* tariffs on industrial goods in circa 1914 were 13 per cent in Germany, 18 per cent in Italy and Austria-Hungary, and 20 per cent in France. The average *total ad valorem* tariff levels, calculated as total customs revenue over total import value from Mitchell (2003), were 7.6 per cent in Germany, six per cent in Austria-Hungary, 7.7 per cent in Italy, and 8.2 per cent in France. The general tariff level is less than half the specific industrial good levels reported in Pollard (1981), and certainly lower the agricultural levels reported in O'Rourke (1997).

Secondly, rising tariffs would only matter if they affected consumers' and producers' decisions. The assumption in Pollard (1981) is that they did but, as with the cost of transporting coal discussed in the previous chapter, actors consider costs *in toto*. Even if we do accept that general tariff levels were rising or that it is a country's specific tariff structure that matters, as argued by Lehmann and O'Rourke (2011), it does not necessarily imply that the overall costs of trading were also rising. The historiography also makes clear that there was a transport revolution underway during this same period (Berend, 2013; O'Rourke and Williamson, 1999; Pollard, 1981). We can see the extensive growth of Europe's transport network in figure 4.1 in the previous chapter. Indeed, O'Rourke and Williamson (2000, p. 17) write that in the late nineteenth century 'rising tariffs were mainly a defensive response to the competitive winds of market integration

[1] O'Rourke (1997) provides a more detailed analysis of the agricultural tariff hikes and their effects.

as transport costs declined.' An example using my data is illustrative.

To take two random regions, the cost of transporting a ton of coal-grain (averaged) by rail, road, waterway and ocean from Aveyron (France) to Abruzzi (Italy) in 1870 was \$182.9.[1] The 1870 Italian tariff level was 8.3 per cent. This tariff added \$18.98 to the transport cost, making the total trade cost \$201.87.[2] In 1900, the same journey - without tariffs - cost \$102.7; a 44 per cent decrease. Italian tariffs, however, were hiked by 5.2 percentage points, making for an additional cost of \$17.33 and a total trade cost of \$120 in 1900. The first point is that tariffs accounted for very little of total trade costs: 9.4 per cent in 1870 and 14.4 per cent in 1900. Second, despite the 5.2 percentage point increase in tariffs from 1870 to 1900, total trade costs were dropping drastically due to improvements in shipping and rail transport. The journey cost dropped by 44 per cent excluding tariffs, and by 41 per cent with tariffs. Indeed, if a 5.2 percentage point increase in tariffs shaved three percentage points off the total trade cost decrease, then it would have taken an additional *71.1 percentage point* increase ($= 41 \times 5.2/3$) to keep journey costs constant. Despite the worst efforts of trade policy, Europe remained open for business. This can be seen clearly in the calculation of European regional market potential.

5.3 Measuring market potential

Constructing market potential using data on transport and tariff costs and transport networks is the approach taken in the economic history literature and when subnational units are used (Crafts, 2005a; Schulze, 2007b). This is because market potential *estimation* requires data on bilateral trade flows which are scarce the further back in time we go, and are especially scarce for subnational units. Still, there are ways around this, and I implement both approaches to ensure robustness.

[1]This average cost adds in grain, as a representative good, in contrast to the use of coal only in the previous chapter. More details on this in the following section.

[2]Tariff to transport cost conversion is done using the technique in Estevadeordal et al. (2002), which estimates a gravity model for trade where distance has an elasticity of -0.8 and a tariff elasticity of -1.0.

5.3.1 Constructing market potential

To construct market potential, I use the Harris (1954) type function, which is more popular in the economic history literature (Crafts, 2005b; Schulze, 2007b). In this formulation, purchasing power is weighted inversely to distance, so that the MP of region i at year t is

$$MP_{it} = \sum_{j} \frac{1}{D_{ijt}} P_{jt} \tag{5.1}$$

where D is the distance from i to j, and P_j is the purchasing power at j. Following the literature, I take P to be the GDP data underlying figure 3.2. D is the transport cost, where I took the average of coal and grain as representative freight goods, as well as tariffs (customs revenue over import value, from Mitchell (2003) along with additions from Schulze (2007b, Appendix)) for international connections, between nodes in regions i and j. I converted the tariffs into transport equivalents using the technique in Estevadeordal et al. (2002), which estimates a gravity model for trade which has a distance elasticity of -0.8 and a tariff elasticity of -1.0. This is also the technique used by Crafts (2005a) and Schulze (2007b). As in Keeble et al. (1982), I control for own-regional distance: I add onto the denominator distance a third of the radius of the square root of a circle the size of the given region.

To put the regions in the wider global economy, I included important out-of-sample economies, drawing GDP from Bolt and Van Zanden (2013): Argentina (Buenos Aires), Australia (Sydney), Belgium (Brussels), China (Shanghai), Denmark (Copenhagen), India (Bombay), Japan (Tokyo), the Netherlands (Amsterdam), Norway (Oslo), Portugal (Lisbon), Russia (Moscow), Switzerland (Zurich), Turkey (Istanbul), and the United States (New York). These economies were not necessarily major trading partners of the European regions in my sample, but the point is to capture *potential* rather than *actual* market access. It is a different story altogether whether regions failed to capitalise on (trade with) markets that were there. Where country-specific railway costs were unavailable, I used the "Europe" series from Schulze (2007b, Appendix). In practice, this rarely applied since overland out-of-sample country connections only existed with Denmark, Portugal, the Netherlands, and Norway, and most of them could be reached at a much lower rate by ocean transport. As with all other countries, *ad valorem* tariffs for these countries came from Mitchell (2003).

The transport network I used here is the same GIS as that from section 4.2.1 of the previous chapter, save for two changes. First, to allow the aforementioned out-of-sample connections, I supplemented the RRG (2012) shipping lines map by digitising the global steamship lines from the 'World Trade Routes, 1912' map in Rodrigue (2013) and digitising the international ports from V. Alexander & Co.'s 'Seaports of the World by Country' dataset (V. Alexander and Co., 2013). Second, I took coal *and* grain as representative goods, using the same sources for ocean and railway freight, which also provide grain freight rates (Cain, 1980; Kaukiainen, 2006; Noyes, 1905; Schulze, 2007b; U.S. Bureau of Railway Economics, 1915). Owing to data scarcity and a lack of research, for roads and waterways, I had no option other than assuming the ratio of coal to grain transport costs for railways was the same as it was for roads and waterways (Moulton, 1914; Van Vleck, 1997).

The railway transport costs I are in table 5.1. While terminal costs were equally high - except for Italy - across countries, variable costs varied by country. France, for example, had variable railway costs three to four times lower than Spain's. These relatively high costs exacerbated Spain's backwardness. Another notable feature in table 5.1 is that some countries lowered their railway costs dramatically over the period, as Austria-Hungary did, while others remained unchanged, as with Italy. Italian railway costs, in fact, stand out as persistently low. It is the only country were terminal costs were, at \$8.35 to \$6.22, were lower than \$20. This is the result of generous state subsidies to railway companies. Schram (1997, p. 46-9) writes that the new Italian state was so enamoured by railways and their perceived power to unify the country that, in the 1865 Railway Act, provided a guarantee on gross revenue, ensuring that it would never fall below a certain level, and allowing railway companies to keep fares low. In short, this table shows the importance of using *costs* rather than simple *distances* along a transport network, as in A'Hearn and Venables (2013).

Table 5.2 shows the tranport costs for road, waterway, and sea transport. As discussed, road and waterway transport costs are the same across the sample, given data constraints. Road transport costs are high both in terminal and vari-

Table 5.1: Railway transport costs

	1870	1900	1910
Austria-Hungary	20.73	12.25	11.78
	(0.51)	(0.31)	(0.29)
Germany	24.97	24.50	22.15
	(0.21)	(0.20)	(0.18)
Spain	28.74	25.45	23.10
	(0.41)	(0.36)	(0.33)
Europe	28.74	25.45	23.09
	(0.21)	(0.19)	(0.17)
France	29.22	28.74	25.92
	(0.12)	(0.12)	(0.11)
Britain	8.95	9.42	8.48
	(0.32)	(0.33)	(0.30)
Italy	8.35	6.85	6.22
	(0.25)	(0.20)	(0.18)
Sweden	28.74	25.45	23.09
	(0.12)	(0.16)	(0.14)

Notes: Costs are in \$. Variable costs, costs per ton-mile, are in brackets. Non-brackted figures are terminal costs. See text for details on sources. AH is Austria-Hungary; DE is Germany; ES is Spain; FR is France; GB is Britain; IT is Italy; SE is Sweden; and EU is Europe.

able components. This reflects the arguments in, among others, Pollard (1974, p. 38), who argues that roads were unable to support commerce beyond local transactions: 'the farthest possible distance for transporting timber or grain was about twelve miles: beyond it the cost of freight began to exceed the value of goods...' Waterways were substantially cheaper than this - cheaper, even, than railway transport - but were naturally limited in their geographical coverage. As Moulton (1914) comprehensively argued, this limiting factor meant that very few waterways earned a profit, and quickly lost traffic to railways (see upcoming discussion around figure 5.1). The costs that stand out here are variable shipping (sea) costs. The terminal component starts out in 1870 at an unusually high level, but quickly drops to a level similar to the cheapest railway terminal costs seen in table 5.1. The variable costs, however, are persistently low: at every year, cheaper than all other modes of transport. The advantage of coastal regions versus landlocked regions, in promoting trade, market integration, and economic development, is widely discussed in the literature (Easterly, 2003; Odell, 1989; Sachs et al., 1999). Here, as the empirical analysis makes clear, it manifests itself in higher market potential values for coastal regions.

Table 5.2: Road, waterway, and sea transport costs

	1870	1900	1910
Road	30.55	30.55	30.55
	(0.43)	(0.43)	(0.43)
Waterways	12.32	12.32	12.32
	(0.17)	(0.17)	(0.17)
Sea	46.41	23.06	20.49
	(0.02)	(0.01)	(0.01)

Notes: Costs are in $. Variable costs, costs per ton-mile, are in brackets. Non-brackted figures are terminal costs. See text for details on sources.

By way of practical example, it cost $201.87 to transport one ton of coal-grain from Aveyron (France) to Abruzzi (Italy) in 1870. Of this cost, as we have already seen, 9.4 per cent can be attributed to tariffs; 19.2 per cent to rail transport; 33.2 per cent to road transport; 8.9 per cent to waterway transport; and 29.3 per cent to ocean transport. Figure 5.1 shows the trade cost breakdown by year for this pair of regions. It makes clear some important features that characterised trade during this period. First, it shows the well-known dramatic decline in trade costs, especially between 1870 and 1900 (Berend, 2013; O'Rourke and Williamson, 1999). Second, it highlights the minimal role of tariffs in those trade costs - by 1910, 10 per cent ($9.10) of the total cost. Third, it shows how road and waterway haulage declined in importance once Europe's railway network was built by 1900 (Marti-Henneberg, 2013; Moulton, 1914).

To arrive at D for Aveyron, I followed the same procedure by year for every one of the other nodes along with Abruzzi; summing them all up yields Aveyron's D. Applying D to the GDP data as in equation 5.1, yields the MP variable, which is summarised in table 5.3.

Table 5.3: Constructed market potential summary statistics.

Year	Mean	Maximum	Minimum	Standard Deviation
1870	4,795	9,177	1,586	1,010
1900	13,848	20,942	4,031	3,760
1910	17,202	26,523	5,265	4,670
Sample	11,900	26,500	1,590	6,310

Notes: Regional constructed market potential in millions of 1990 Geary-Khamis dollars.

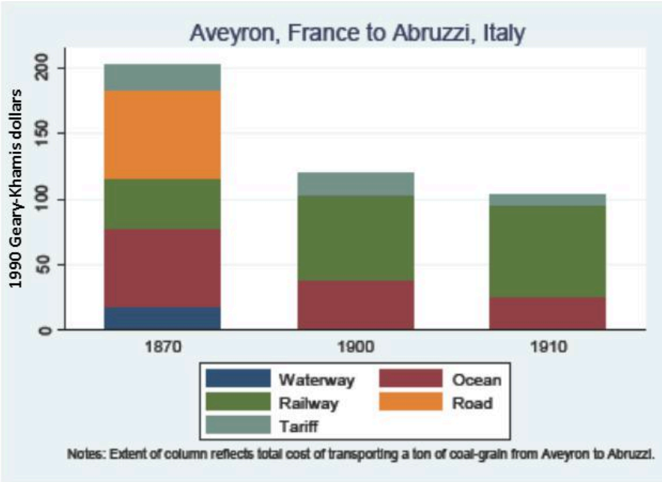

Figure 5.1: Trade cost ($) breakdown: Aveyron and Abruzzi

Notes: All costs are in 1990 Geary-Khamis $. Extent of column is the total trade cost of transport a ton of coal-grain in that year. See text for sources.

5.3.2 Estimating market potential

While trade data do exist for some European regions (see, for example, Wolf (2007)), most countries did not record regional trade flows during this period. For this reason, I follow the procedure in **?** and use national-level bilateral trade data to arrive at estimates for regional market potential. The trade data are from Jacks et al. (2011), which covers global bilateral trade flows between 1870 and 2000. Trade values and GDP levels are all measured in 1990 dollars, making the data consistent with the GDP data I use here.[1]

The strategy is to use information contained in international trade flows to get estimates for price indices and bilateral trade costs, and apply these estimates to regions.[2] The assumption is that interregional trade flows follow the same patterns as international ones. This assumption is supported by studies that are able to exploit interregional trade data (Combes et al., 2005). **?** proposes a number of adjustments to make this assumption more reasonable. First, I restrict the data to exports within my sample of countries, and from my sample countries to the rest of the world. This captures the notion that trade flows

[1]Full details on sources and adjustments are in *Appendix C*.

[2]For a more detailed exposition of the theory, readers should refer to **?**, or the broader literature on gravity trade models and market access as in Redding and Venables (2004).

(and so market potential) may operate differently in different parts of the world, especially when trading areal units are at different levels of development. Second, I also control for factors other than bilateral distance. As **?** points out, this is particularly important in a regional implementation, as trade between regions of the same country is usually a multiple of trade between regions with similar bilateral features, but in different countries (McCallum, 1995). To capture this, I include a set of dummies that indicate whether countries share a border, and whether countries share an official language.

Despite these controls, some problems of course persist. There is, in particular, high variability in regional output structures, which is not borne out in national aggregate data. Given these empirical constraints, I view this estimation as a cross-check on the previously constructed market potential measure rather than a central or stand-alone variable.

More formally, I assume that bilateral trade costs between any two countries i and j are given by:

$$T_{ij} = dist_{ij}^{\beta_0} \times (\exp(\text{border}_j i)^{\beta_1}) \times \exp(\text{language}_{ij})^{\beta_2} \tag{5.2}$$

In this expression, border and language are the dummies discussed earlier, and β_0 and β_2 are the elasticities of trade cost with respect to its different components. Inserting a time dimension yields the following econometric implementation:

$$\begin{aligned} \ln\left(\frac{x_{ijt}}{E_{it}E_{jt}}\right) &= \alpha + \gamma_{1t}\ln(dist_{ij}) + \sum_i \gamma_{2it} border_{ij} + \gamma_{3t} language_{ij} \\ &+ \delta_{1it} exporter_{it} + \delta_{2jt} importer_{jt} + \varepsilon_{ijt} \end{aligned} \tag{5.3}$$

Where x_{ijt} is the value of exports from i to j at year t, and E_{it} and E_{jt} are the trading partners' GDPs. The coefficients on exporter and importer dummies, $\delta_{1it}$and δ_{2jt}, are used to obtain estimates for price indices, since relative prices affect trade flows, but are unobservable. To arrive at trade costs for each benchmark year in my sample, I estimate model 5.3 on 14-year windows of theJacks et al. (2011) data between 1870 and 1920.[1]

As is standard in the literature, market potential is the trade cost and price index weighted sum of GDPs of all surrounding regions and countries, that is, regions in the same country (cty), and regions in the rest of the sample (ROS). I use the results from the gravity equation to calculate market potential for each region i at each year t for all countries j as follows:

$$MP_{it} = \sum_{j\varepsilon ctyi} \left(e^{\widehat{\lambda 3t}} dist_{ij}^{\widehat{\lambda 1t}} \right) \delta_{jt}^{-1} E_{jt} \sum_{j\varepsilon ROS} \left(e^{\widehat{\lambda 2it}} e^{\widehat{\lambda 3 language eij}} dist_{ij}^{\widehat{\lambda 1t}} \right) \delta_{jt}^{-1} E_{jt} \tag{5.4}$$

where δ_{jt}^{-1} and the parameters $\widehat{\lambda_{1t}}$, $\widehat{\lambda_{2it}}$, and $\widehat{\lambda_{3t}}$ were estimated in the gravity trade equation 5.3, and E_{jt} is again proxied by a region or country's GDP in year t. The results are summarised in table 5.4.

Table 5.4: Estimated market potential summary statistics.

Year	Mean	Maximum	Minimum	Standard Deviation
1870	13,144	386,501	22	48,971
1900	8,868	170,875	29	18,856
1910	33,016	499,064	387	55,621
Sample	18,343	499,064	22	45,315

Notes: Regional estimated market potential in millions of 1990 Geary-Khamis dollars.

5.3.3 Comparing two measures of market potential

The main purpose of estimating market potential is to provide a cross-check, given its centrality to my research, of the constructed market potential variable. I include both in my empirical implementation, but it is worth looking at the relationship between the two first.

Looking at the summary statistics in tables 5.4 and 5.3, the estimation procedure produces a market potential variable with much greater variation, as can be seen in the standard deviations of $6,310mn for constructed market potential

[1] The results of these estimations are in *Appendix C*. The length of the window is, roughly, the length of the entire period, divided by three. I experimented with different windows, finding more or less the same results.

(MP^C) and \$45,315mn for estimated market potential (MP^E). While the minimum value of MP^E (\$22mn) is much smaller than that of MP^C (\$1,590mn), the estimation generally produces larger values, with a mean value that is 54 per cent larger. Differences are to be expected, given the imperfection of both methods used to caluclate market potential, but what matters in the ultimate empirical analysis is whether the variables are monotonically different. The Spearman rank correlation between MP^E and MP^C, at 0.584 and significant at the one per cent level, shows that this is - by and large - not the case.

As a more rigorous paramteric test, I standardised both variables, so that they both have a mean of zero and standard deviation of one, and regressed MP^C on MP^E. Standardising the variables removes the issue of different levels, which we already know exists from the summary statistics in tables 5.4 and 5.3. This OLS correlation can be seen in figure 5.2. The coefficient implies that for every standard deviation increase in MP^E, MP^C increases by 0.612 standard deviations. This is not a perfect correlation, but with a T-statistic of 18.89, it is a highly statistically significant one. The R^2 is a respectable 0.375.

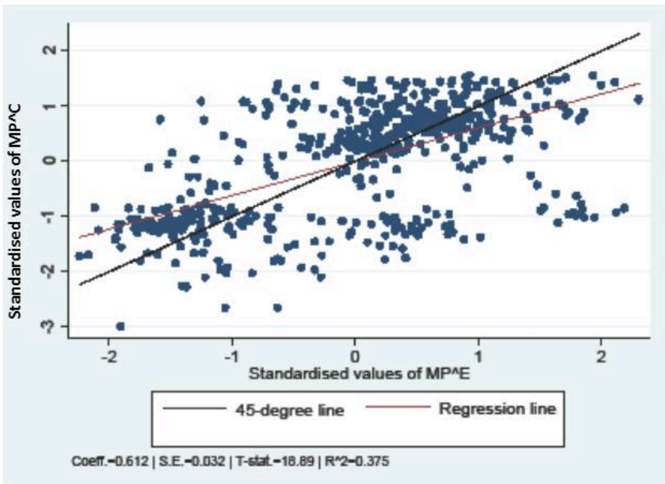

Figure 5.2: Constructed versus estimated market potential

Notes: Values on y-axis represent the standardised values of MP^C, the constructed market potential measure, and the x-axis represents the standardised values of MP^E, the estimated market potential measure. The variables are standardised so that they have a mean of zero and standard deviation of one.

The most reassuring similarity between the two standardised variables is their (log-normal) distribution. The kernel density plot in figure 5.3 shows that despite their differences both variables capture the sample's bi-modal distribution of market potential. This distribution, incidentally, is a necessary condition for an explanation of regional per capita income, which is also bi-modal.

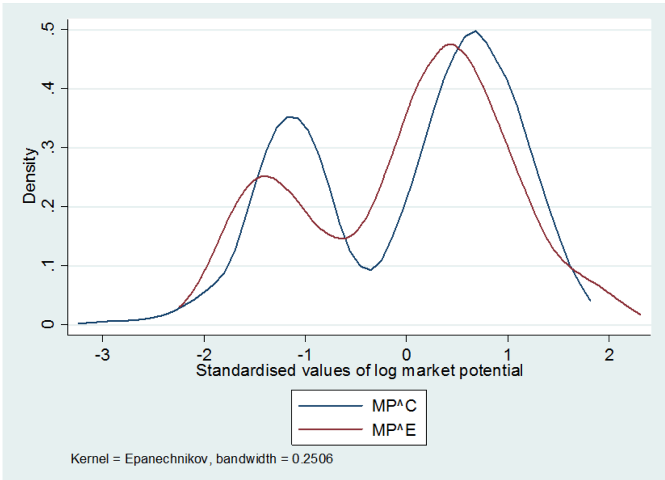

Figure 5.3: Log-normal distribution of constructed and estimated market potential

Notes: Estimated using Epanechnikov kernels. Values on the x-axis represent the standardised values of MP^E*, the estimated market potential measure, and* MP^C*, the constructed market potential measure. The variables are standardised so that they have a mean of zero and standard deviation of one.*

All in all, both variables capture the "true" variation in market potential in the sample. Indeed, the first principal component of the two variables accounts for 81 per cent of their total variation, and both variables are correlated with the component at 0.90, significant at one per cent.[1] Again, the purpose of estimating market potential was to provide a cross-check on the main variable of interest, MP^C. The checks in this section show us that the differences in measurement, an important but un-tested concern in Hanson (2005), are overstated. This cross-check also adds weight to market potential work in the economic history literature that has been unable to work with anything other than Harris-type forumaltions (Crafts, 2005a; Schulze, 2007b).

[1] I extracted this component using principal components analysis. Its eigenvalue is 1.612.

5.4 Empirical strategy

The goal here is to estimate equation 5.5 where, for the sake of brevity, MP refers to both the constructed and estimated measure of market potential:

$$\begin{aligned} Y_{it} = \alpha + \pi \ln Coal\_Transport_{it} + \rho INST_{it} + \ln MP_{it} \\ + \gamma_c + \theta_t + \mu_{it}. \end{aligned} \tag{5.5}$$

The issues with $INST$'s endogeneity remain, however, and there is now the threat of market potential-related endogeneity that I must deal with. Given these concerns, I implement following system of equations, where in equation 5.7 I first instrument $INST$ using the same procedure in the preceding chapter and then in equation 5.8 instrument market potential. Since MP includes own-region demand, increases in income lead to increases in MP. I then use the predicted values $\widehat{MP}$ and $\widehat{INST}$ to estimate the second-stage equation 5.6:

$$\begin{aligned} Y_{it} = \alpha + \pi \ln Coal\_Transport_{it} + \rho \widehat{INST}_{it} + \ln \widehat{MP}_{it} \\ + \gamma_c + \theta_t + \mu_{it} \end{aligned} \tag{5.6}$$

$$\begin{aligned} INST_{it} = \alpha + \pi \ln Coal\_Transport_{it} + \varnothing SH_i + \chi \ln Dist\_London_{it} \\ + \gamma_c + \theta_t + \varepsilon_{it} \end{aligned} \tag{5.7}$$

$$\begin{aligned} MP_{it} = \alpha + \pi \ln Coal\_Transport_{it} + \varnothing SH_i + \chi \ln Dist\_London_{it} \\ + \gamma_c + \theta_t + \varepsilon_{it}. \end{aligned} \tag{5.8}$$

Identifying market potential

The instrument for MP is $Cost\_London$: the distance from each regional node to London, measured in kilometers, multiplied by the cost of getting there in dollars using the same GIS as in section 5.3.1. Distance-based instruments for market potential are the convention in the literature (Head and Mayer, 2006; Redding and Schott, 2003; Redding and Venables, 2004; ?). I have extended it here with a time-varying measure so as to make it compatible with regional fixed effects. This technique has also been used in studies of market potential, with Holl (2011) interacting historical road placement with national railway network growth rates.

Both components of this instrument capture the same idea. They provide exogenous geographical variation that captures the market potential advantage of locations close to the economic centre, London. The choice of centre depends on the sample, so **?** uses Brussels since his sample is contemporary Europe while Redding and Venables (2004) use London, New York, and Tokyo since they work with a contemporary global sample. London makes most sense here because it was by far the most important economic node in my sample across all benchmark years.

Head and Mayer (2006) point out that restricting a sample to European regions implicity determines a centre to begin with, and the location and relative prosperity of the continent was itself the outcome of an endogenous process. To get at this concern they construct an instrument, 'Global Centrality', which equals the distance from each region to the centre of every inhabited one-by-one degree cell in the world population grid. Using this instrument, the elasticity on market potential they get is 0.0877, compared to 0.0996 when using 'distance to Brussels' and 0.0790 when using distance to the centre of their sample. In short, there is in effect no empirical difference between these instruments. I have experimented with using as an IV distance to my sample's geographical centre, the sum of regions' inverse distances (suggested in Head and Mayer (2006)), distance to New York (as an alternative to London), and distance to my sample's economic centre (calculated using the technique in Mathys and Grether (2010)). Of all these, $Cost\_London$ shows the strongest correlation and as all of them get at essentially the same idea, making the choice an empirical question, I stuck with distance to London.[1]

The Pearson correlations between $Cost\_London$ and MP^C and MP^E are respectively -0.385 and -0.378, both significant at the one per cent level. Applying

[1] The correlations between $\ln MP^C$ and log distance to the sample geographical centre, log distance to New York, log distance to the sample economic centre, and log sum of inverse distances are, respectively, -0.024, -0.007, -0.108, 0.024. For $\ln MP^E$, the correlations in the same order are: 0.012, -0.015, -0.083, 0.032. Only distance to the sample economic centre is statistically significant, but at a lower level than $Dist\_London$.

some more rigour, table 5.5 shows the results from regressions of $\ln MP^C$ and $\ln MP^E$ on $\ln Cost\_London$. The R^2s are low, at around 14 per cent (column 4), but both the coefficients (column 1) are significant at the one per cent level: MP^C with a T-statistic of -5.15 (column 3), and MP^E with a T-statistic of -7.17.The standardised-βs (column 5) show that $Cost\_London$ has a very similar effect on either market potential measure, which is reassuring. Column (6) shows the T-statistics from a robust regression estimation, desgined to control for outliers, which are present in $Cost\_London$ in the form of, for example, the Canary Islands (Spain). The estimation first runs the normal OLS estimation, calculates the $Cook'sD$ for each observation, and then drops any observation where $Cook'sD$ 1.[1] The T-statistics maintain the same levels of statistical significance for both variables, indicating that even when dropping outliers the picture does not change. Column (7) gets at the same issue by running a jackknife estimation: the T-statistics again maintain their levels of significance. In short, $Cost\_London$ is a relevant predictor of market potential.

Table 5.5: OLS correlations between $Cost\_London$ and market potential

	(1) β	(2) R.S.E.	(3) T-stat.	(4) R^2	(5) Std.-β	(6) Rob.	(7) Jack.
$\ln MP^C$	-0.039***	0.007	-5.15	0.148	-0.385	-23.51	-4.98
$\ln MP^E$	-0.136***	0.019	-7.17	0.143	-0.378	-11.2	-6.97

*Notes: The independent variable is $\ln Cost\_London$. MP^C (MP^E) is the constructed (estimated) market potential measure. R.S.E. is the robust standard error, clustered by region, where *** indicates statistical significance at the one per cent level. Std.-β is the standardised regression coefficient. Rob. is the T-statistic from the robust regression. Jack. is the T-statistic from the jackknifed estimation.*

Following Clogg et al. (1995), I test for the equality of the coefficients reported in column (2) of table 5.5 using a Z-test, subtracting the coefficient of MP^C from MP^E, and dividing the result by the square root of the sum of their standard errors. The resulting Z-score is -0.60, implying no statistically significant difference.

The exclusion restriction is that $Cost\_London$ does not have any effect on

[1] $Cook'sD$ is the squared difference between the predicted values of a dependent variable from a normal OLS and the prediction of that same variable in which an observation has been omitted, divided by the number of fitted parameters in the model times the mean square error of the model.

real per capita income levels other than through its effect on market potential. The presence of an unobserved variable that is correlated with $Cost\_London$ would break this restriction, since the effect of $Cost\_London$ would be working through a channel other than market potential. While we can never test the exclusion restriction directly, we can at least test an auxiliary hypothesis that lends credence to it. Since $Cost\_London$ is spatial in nature, the competing hypothesis must also be spatial and as I discussed in the previous chapter, spatial explanations for Europe's regional income are not forthcoming.

Once such hypothesis is that climate, which is spatially correlated with $Cost\_London$'s gradient, also affects economic activity, providing an alternative channel to market potential. The climate argument has been made by among others Sachs et al. (1999) and Easterly (2003), and is often proxied using latitude and longitude coordinates (Easterly, 2003; Gallup and Sachs, 2001; Rodrik et al., 2004; Sachs et al., 1999). In table 5.6, I regress market potential on $Cost\_London$ and control for latitude and longitude. The results for $\ln MP^E$ show that both latitude and longitude are statistically insignificant, and that $Cost\_London$ retains the same magnitude and significance as in table 5.5. The results for $\ln MP^C$ show the same picture. While a fuller test on the usefulness of $Cost\_London$ as an IV will have to wait until the IV estimation in the following section, it is reassuring that popular geographical controls used in the literature have no effect on market potential, when $Cost\_London$ is controlled for, which goes some way in supporting the exclusion restriction here.

Another hypothesis is that distance to London is correlated with the diffusion of technology from Britain to Continental Europe. This is difficult to test empirically here, but a long line of literature has debunked the idea of a technological diffusion gradient across Europe. As I wrote in **?**, in this "leader-follower" scheme, Britain industrialised first, which raised its productivity in manufacturing above European levels. Then, in varying degrees, European countries are argued to have adopted Britain's productive technology. The pace at which they did this, the scheme goes on, determines the extent of their income lag behind Britain. Rostow (1960) provided the GDP figures and one of the scheme's keywords: 'discontinuity', which described the movement from one of his stages of growth to another. Gerschenkron (1962), focusing on industrial production instead, wrote

Table 5.6: Latitude and longitude versus *Cost_London*

	$\ln MP^C$	$\ln MP^E$
Longitude	0.008	-0.018
	(0.007)	(0.026)
Latitude	-0.001	0.006
	(0.005)	(0.020)
$\ln Cost\_London$	-0.039***	-0.137***
	(0.007)	(0.019)
Constant	23.784***	26.798***
	(0.439)	(1.387)
R^2	0.150	0.143
N	597	597

Notes: Latitude and longitude are measured in decimal degrees, taken from a WGS84 projection. MP^C *(*MP^E*) is the constructed (estimated) market potential measure. Robust standard errors clustered by region are reported in brackets. *** indicates statistical significance at the one per cent level.*

of economic 'backwardness' in peripheral Europe, giving the scheme its next keyword. Landes (1969) and Pollard (1973) wrote of the 'diffusion' of industrialisation across Europe (Pollard, 1973, 1981). Following this scheme, explanations of income differentials would be the determinants of the speed of adoption of the leader's technology. The authors assumed that new technology is adopted only if more capital is available, and so European economic history was one of differential rates of capital accumulation. Gerschenkron (1962), for example, spent a lot of time on the role of banks and states in overcoming underinvestment.

This scheme was undermined by the production of quantitative evidence - ironically, given this was Rostow's aim - which was occurring in parallel. Rostow's "discontinuities" and "take-offs" in European economic development proved to be hard to pin down. Gerschenkron's higher shares of production-goods in the output of backward economies, and the importance of investment banks, did not show up in the data underlying the comprehensive book of Milward and Saul (1973). Nor did the theory hold up well to newer economic ideas (Sylla and Toniolo, 1992). The jump in capital accumulation was not necessitated by manufacturing, but railways. In Britain, railways accounted for a much higher proportion of the economy's capital stock until the interwar years (Feinstein and Pollard, 1988).

5.4.1 Baseline market potential effect

Before getting into the instrumented results of model 5.6, it is worth comparing the effects of MP^C and MP^E on regional per capita income, Y, in a OLS setting with region and year fixed effects. The results in table 5.7 show that both measures have a highly statistically significant effect on Y. The magnitude of MP^C, with an elasticity of 0.083, is larger than that of MP^E, with an elasticity of 0.008. Both are statistically significant at the five per cent level. While the differences in magnitude appear large, their t-statistics, at 2.14 and 2.26 respectively, are similar, indicating less precission for MP^C. This is supported by its larger $RMSE$ and slightly smaller $\widehat{R}^2$. Following Clogg et al. (1995), I test for the equality of these coefficients using a Z-test, subtracting the coefficient of MP^E from MP^C, and dividing the result by the square root of the sum of their standard errors. The resulting Z-score is 0.35, implying no statistically significant difference.

Table 5.7: OLS market potential and income

	$\ln Y$	$\ln Y$
$\ln MP^C$		0.081**
		(0.038)
$\ln MP^E$	0.008**	
	(0.004)	
Constant	7.268***	5.645***
	(0.07)	(1.331)
Region F.E.	Yes	Yes
Region F.E.	Yes	Yes
$\widehat{R}^2$	0.955	0.857
$RMSE$	0.096	0.905
N	597	597

*Notes: Dependent variable is log of regional per capita income, Y. MP^C (MP^E) is the constructed (estimated) market potential measure. RMSE is root mean-square error. Region FE is a regional fixed effects term. Year FE is a year fixed effects term. Robust standard errors clustered by region are reported in brackets. ** indicates statistical significance at the five per cent level; *** at the one per cent level.*

Table 5.7 shows that, despite their different conceptual and empirical bases, the coefficients on MP^C and MP^E are similar in magnitude and statistical significance, but what does the effect mean in practice? The MP^C effect implies a 0.135 standard deviation increase in income for every standard deviation increase in market potential. This would, according to the sample's income standard deviation of \$842, add \$114 onto a region's per capita income level. This is the equivalent to five per cent of the sample mean per capita income level, \$2,350. As another comparator, it is also the equivalent to 19 per cent of the sample minimum per capita income level, \$606 (Galicia, Spain in 1870).

Does instrumenting market potential and the introduction of controls change the picture? Table 5.8 contains the results, for both measures of market potential, of model 5.6. Starting with the first stage for market potential, we can see that $Cost\_London$ is a strong predictor of market potential. For both MP^C and MP^E, the coefficient is correctly signed and highly statistically significant. Both also have large F-statistics, backing up the tests on $Cost\_London$ in tables 5.5 and 5.6. The first stage for $INST$, which is of course the same for both measures of market potential, shows that SH continues to be a strong predictor of $INST$: the coefficient is correctly signed and highly significant, and the F-statistic on its exclusion is also large in both cases.

Taking the predicted values for $INST$ and MP, we see strong results in the second stage. It is unsurprising given the first stage results that the second stage estimations both pass a Cragg-Donald Wald joint F-test for weak identification and an Anderson underidentification test. They also both pass the Angrist-Pischke F-test on their joint exclusion. The elasticity at which Y responds to MP^C is a significant 0.403. The coefficient on $INST$ is insignificant. The elasticity on MP^E is much smaller, at 0.0.67, but still highly statistically significant. The result for $INST$ is larger here, but is still insignificant. $\ln Coal\_Transport$ continues to be insignificant in all estimations.

Are the market potentials comparable? While their coefficients a very different, their t-ratios are similar at 5.10 for MP^E and 4.81 for MP^C. Further, conducting a z-test for the equality of coefficients following Clogg et al. (1995) produces a z-score of 1.08, implying no statistically significant difference between them. There is simply less precision surrounding the constructed measure of market potential, as discussed in the previous section.

The insignificance of regional institutions, after controlling for market potential, is perhaps unsurprising when we consider that the explanation cannot account for spatial patterns. There are no conceptual priors that tell us efficient institutions, especially on the sub-national level, should be co-located.The empir-

Table 5.8: IV market potential and income

	Constructed MP		
	IV-1: MP	IV-1: $INST$	IV-2: Y
$\ln \widehat{MP}$			0.403**
			(0.084)
$\widehat{INST}$			0.522
			(0.505)
$\ln Coal\_Transport$	0.084	-0.011	0.021
	(0.089)	(0.044)	(0.046)
$\ln Cost\_London$	-0.361***	-0.047***	
	(0.008)	(0.005)	
SH	0.002*	-0.002**	
	(0.001)	(0.000)	
$Year F.E.$	Yes	Yes	Yes
$Region F.E.$	Yes	Yes	Yes
$\widehat{R}^2$	0.909	0.457	0.837
N	597	597	597
Angrist-Pischke F-test	9.46	15.96	
Cragg-Donald Wald joint F-statistic			15.83
Anderson canon. corr. LM-statistic			13.04
	Estimated MP		
	IV-1: MP	IV-1: $INST$	IV-2: Y
$\ln \widehat{MP}$			0.066***
			(0.013)
$\widehat{INST}$			1.031
			(1.268)
$\ln Coal\_Transport$	1.989	-0.011	-0.058
	(1.745)	(0.044)	(0.085)
$\ln Cost\_London$	-1.084***	0.047	
	(0.067)	(0.005)	
SH	-0.068**	-0.0002**	
	(0.008)	(0.000)	
$Year F.E.$	Yes	Yes	Yes
$Region F.E.$	Yes	Yes	Yes
$\widehat{R}^2$	0.526	0.457	0.489
N	597	597	597
Angrist-Pischke multivariate F-test	13.26	59.5	
Cragg-Donald Wald joint F-statistic			17.82***
Anderson canon. corr. LM-statistic			24.75***

*Notes: *** denotes significance at one per cent and ** at five per cent. Robust standard errors clustered on countries, and reported in brackets. Angrist-Pischke F-test is on the excluded IV. Cragg-Donald Wald joint F-statistic for weak identification is conducted jointly on the excluded IVs. Anderson canon. corr. LM statistic is for under-identification. Y is GDP per capita; $INST$ is a measure of regional institutional efficiency; $Coal\_Transport$ is transport cost to nearest coal deposit; SH is the Spanish Habsburg dummy; MP is market potential; and $Cost\_London$ is cost to London. IV-1 (IV-2) is a first (second) stage. Mark Schaffer at Heriot-Watt University supplied the Stata code for this estimation.*

ical results shown here find much in common with Redding and Venables (2004) who, in their analysis of international per capita income, find that institutions - risk of expropriation and length of Soviet rule - are insignificant in their full specifications that include total, as opposed to foreign-only, market potential.

The MP^C coefficient in table 5.8 is similar to those found in the literature. In their IV regressions of international-level log real GDP per capita on log market potential (derived using a trade equation), Redding and Venables (2004, p. 69) estimate a coefficient of 0.256 compared to my 0.403. The IV regressions of European regional log GVA per capita on market potential in **?**, p. 609 show coefficients on market potential ranging from 0.275 to 0.309.

The IV results in table 5.8 imply a one standard deviation in market potential results in a 0.65 standard deviation increase in per capita income. That is, an increase of \$547 on the sample mean income of \$2,350 or almost equal the sample minimum income of \$606 (Galicia, Spain in 1870).

5.4.2 Domestic versus foreign market potential

Were domestic or foreign markets the income drivers behind the baseline results in table 5.8? This question is related to a large literature on global integration and trade liberalisation versus inward-looking policies like import-substitution industrialisation in economic development. Dollar (1992, p. 523), summarising the World Bank's position on the matter, wrote that accessing foreign markets 'generally results in more rapid growth of exports, and there may be externalities associated with exporting that cause open economies to grow more rapidly over long periods of time.' The work of Sachs and Warner (1995) shows differential levels of foreign market potential can explain the divergence and inequality I found in *Chapter 3*. They write that 'open economies tend to converge, but closed economies do not. The lack of convergence in recent decades results from the fact that the poorer countries have been closed to the world' (Sachs and Warner, 1995, p. 3). In their empirical analysis, Redding and Venables (2004, p. 65) find that the per capita income effect of foreign market potential is three-times greater than that of domestic market potential.

In this subsection, I disaggregate MP^C into foreign and domestic regional market components, and use them in the same IV setting as model 5.6. MP^E has served its purpose as a cross-check on MP^C in the preceding section. The domestic market potential term - DMP - is composed using the formulation 5.1, but restricting the sample to same-country regions. This means a region's market potential is only affected by its neighbours in its own country, and so tariffs drop off, leaving transport costs as the weighting parameter. The foreign market potential term - FMP is constructed by restricting the sample to out-of-country regions, so that a region's market potential is affected only by foreign markets. A region's MP is thus the sum of its foreign and domestic components, as in $MP_{it} = DMP_{it} + FMP_{it}$. The results are displayed in table 5.9.

The first stage results for $INST$ are the same as those reported in table 5.9, and so I have excluded them here. I have reproduced the estimations for total MP^C, for the sake of comparison. The *Cost_London* instrument continues to work well across domestic and foreign market potential. It comfortably passes the F-test on its exclusion, and under- and weak-indentification tests.

In the second stage results, we that foreign market potential yields a coefficient as large and significant as total market potential, but that domestic market potential is insignificant. This implies that most of the gains of market potential were working through increasing *foreign* not domestic market access. This fits with the historical literature's emphasis on the pull of expanding core markets in Europe, and overseas (Bairoch, 1997; **?**). The effect of foreign market potential implies a \$488 increase in income for its every standard deviation increase.

Redding and Venables (2004, p. 65) derive similar results. First, the size of their foreign potential coefficient, at 0.311, is close to the one reported here. Second, their estimations consistently show foreign markets are more important than domestic ones. Given that foreign market potential is always larger than domestic - by construction of my sample and that in Redding and Venables (2004) there are more regions outside a country than it - a one per cent increase in foreign market potential in reality corresponds to a much larger income effect. The result, however, also has theoretical support. According to NEG theory, a reduction in international trade costs, which is what we see in this historical context, makes domestic markets less attractive relative to foreign markets. These results are also in line with work done outside NEG. Both Dollar (1992) and Sachs and

Table 5.9: IV domestic versus foreign market potential estimates.

	(1)	(2)
	IV-2	IV-2
	$\ln Y \ln Y$	
$\ln \widehat{MP^C}$	0.403**	
	(0.084)	
$\ln \widehat{DMP^C}$		0.016
		(0.208)
$\ln \widehat{FMP^C}$		0.365**
		(0.153)
$\ln Coal\_Transport$	0.021	-0.0005
	(0.046)	(0.088)
$\widehat{INST}$	0.522	0.323
	(0.505)	(0.286)
Year F.E.	Yes	Yes
Region F.E.	Yes	Yes
$\widehat{R}^2$	0.837	0.836
N	597	597
Angrist-Pischke multivariate F-test	5.48	18.87
Cragg-Donald Wald joint F-statistic	10.19	13.76
Anderson canon. corr. LM-statistic	13.04	19.77
	(4)	(5)
(6)		
	IV-1	IV-1
IV-1		
	$\ln DMP^C$	$\ln FMP^C$
$\ln Coal\_Transport$	-0.230	0.161
	(0.179)	(0.196)
SH	-0.001	-0.006***
	(0.001)	(0.001)
$\ln Cosst\_London$	-0.250***	-0.338***
	(0.012)	(0.012)
heightYear F.E.	Yes	Yes
Yes		
Region F.E.	Yes	Yes
Yes		
$\widehat{R}^2$	0.847	0.916
N	597	597
597		

*Notes: *** denotes significance at one per cent and ** at five per cent. Robust standard errors clustered on countries, and reported in brackets. Angrist-Pischke F-test is on the excluded IV. Cragg-Donald Wald joint F-statistic for weak identification is conducted jointly on the excluded IVs. Anderson canon. corr. LM statistic is for under-identification. Y is GDP per capita; INST is a measure of regional institutional efficiency; Coal_Transport is transport cost to nearest coal deposit; SH is the Spanish Habsburg dummy; MP is market potential; FMP is foreign market potential; DMP is domestic market potential; and Cost_London is cost to London. IV-1 (IV-2) is a first (second) stage. Mark Schaffer at Heriot-Watt University supplied the Stata code for this estimation.*

Warner (1995) found that openness to foreign markets resulted in faster export growth and income convergence among developing countries. While lower access to foreign markets may not have been a conscious policy choice in this context and was, rather, the effect of poor transport infrastructure, the effect is the same: domestic markets were unable to support per capita income development as much as foreign markets were.[1]

5.4.3 Falling behind

In the previous subsection, I showed that foreign markets had a bigger effect on per capita incomes. To the extent that regions experienced varying relative levels of foreign and domestic market potential growth, this can explain some of the income divergence and inequality I found in *Chapter 3*. There is, however, a more immediate cause behind Europe's periphery falling behind. NEG holds that as trade costs drop to an intermediate level, economic activity locates near large markets to save on trade costs. The relocation of this activity further increases the size of the market, drawing in more activity at the expense of peripheral regions. This is the "unequalising spiral" that Krugman and Venables (1995) described theoretically. As Crafts and Wolf (2013a) note, the 'first globalisation' of the nineteenth century - driven by improvements in transport technology - saw the simultaneous processes of industrialisation in Europe and de-industrialisation in Asia. On a less grand scale, this section shows the same processes were at work within Europe.

Following Clingingsmith and Williamson (2008) who work with employment data rather than income, a region experiences *strong* de-industrialisation if its absolute per capita income level falls. It experiences *weak* de-industrialisation if its percentage share of sample income ($GDP_{it}^{share} = 100 \times GDP_{it}/\sum_{n=1}^{j} GDP_t$) falls. These are my measures of de-industrialisation. To measure the extent to which regions' market potentials are falling behind, I create two variables. The variable MP_{it}^{max} is the ratio of the sample maximum regional MP^C to region i's MP^C at year t ($MP_{it}^{max} = max(MP_t^C)/MP_{it}^C$). The variable MP_{it}^{med}

[1] For an econometric analysis on the debilitating effects of poor infrastructure on trade, see Limao and Venables (2001).

is the ratio of the sample median regional MP^C to region i's MP^C in year t ($MP_{it}^{med} = median(MP_t^C)/MP_{it}^C$). Table 5.10 displays a regression analysis of these variables, where I instrument MP^{med} and MP^{max} with $Dist\_London$.

As the Krugman and Venables (1995) argument is about the effects of increasing market potential *on peripheral countries*, I restrict the sample here to regions that fall within the bottom two quartiles of distance from London. That is, the regions farthest away from the sample's economic core.

Table 5.10: De-industrialisation and market potential

	IV-1 MP^{med}	IV-1 MP^{max}	IV-2 $\ln Y$	IV-2 $\ln Y$	IV-2 GDP^{share}	IV-2 GDP^{share}
$\ln Cost\_London$	0.020***	0.047***				
	(0.003)	(0.010)				
MP^{med}			-0.699**		-0.017**	
			(0.334)		(0.006)	
MP^{max}				-0.290**		-0.009**
				(0.147)		(0.466)
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Region F.E.	Yes	Yes	Yes	Yes	Yes	Yes
$\hat{R}^2$	0.477	0.435	0.072	0.327	0.554	0.539
N	597	597	597	597	597	597
Angrist-Pischke	44.33	21.30				
Anderson			14.24**	37.58	24.17	24.01

*Notes: Sample restricted to bottom quartiles of distance to London (those farther away). *** denotes significance at one per cent and ** at five per cent. Robust standard errors clustered on countries, and reported in brackets. Angrist-Pischke is the F-test is on the excluded instrument, Cost_London (cost to London). Anderson is the canon. corr. LM statistic is for under-identification. Y is GDP per capita; GDP^{share} is regional GDP percentage share in sample GDP; MP^{med} is sample median MP^C relative to regional MP^C; and MP^{max} is sample maximum MP^C relative to regional MP^C. Mark Schaffer at Heriot-Watt University supplied the Stata code for this estimation.*

The results in table 5.10 show that $Dist\_London$ is a good predictor of the two relative market potential measures, MP^{max} and MP^{med}, as indicated by its statistical significance and the large Angrist-Pischke F-statistics. In this case, the coefficients on $Cost\_London$ are positive since increasing distance to London is correlated with a greater ratio of maximum (or median) market potential to regional market potential. Both MP^{max} and MP^{med} have a large negative effect on regional per capita income levels, implying strong deindustrialisation. For the former, a one unit increase in the ratio of median market potential to regional market potential decreases income by 0.70 per cent. For the latter, the coefficient is a smaller, but still highly significant, -0.30 per cent. In reality such large effects are unlikely to have been realised, and there is likely to be imprecision around the estimate, but these results are still consistent with the Krugman and Venables (1995) hypothesis that as a region's market potential decreases relative to the "global" level, its output

This is backed up by the results on GDP^{share}. Both MP^{max} and MP^{med} have a large negative effect on regional GDP shares. The coefficients indicate that for every unit increase in the ratios, a region's share in the sample output will drop by, respectively, 0.02 and 0.01 percentage points. Again, while I am only describing patterns in historical data here and there is bound to be imprecision around point estimates, these results are at least clear in direction: a decline in *relative* market potential, caused by declining trade costs, resulted in absolute losses in per capita income and losses in sample output shares. This is consistent with the concentration of income in the northwest of Europe and the decline of the periphery, analysed in *Chapter 3*.

5.4.4 Why isn't everything on the coast?

Figure 5.4 maps the values of MP^C in standard deviations at each year. Two general patterns are immediately visible. First, there was indeed a concentration of market potential in the northwestern core. This was already visible by 1870, but was much more so in 1910. This trend was accompanied by a relative decline of market potential in the periphery. The interior of France - the Massif Central - was "hollowed" out by 1900. Sweden became increasingly peripheral over the period so, that by 1910, only the southern most tip of the country, Stockholm, and the mineral-rich north had middling levels of market potential. The second pattern, which can be seen at each year, is that coastal regions - even if they were poor - had high market potentials. By 1910, this is so clear that the map looks as though dark perimeter surrounds the Continent, running from East Prussia, along the the Bay of Biscay, and then along the western Mediterranean coast.

There are, of course, pockets of high market potential where we would expect them - the industrial heartlands of eastern Germany and northern England, and the London-agglomeration. Still, the high-market potential coast calls out for an explanation. Not in the sense that is unusual - it is widely acknowledged that shipping freight was so much cheaper than any form of overland transport that coastal regions are at a locational advantage - but that not all of these regions were developed (Crafts, 2005a; Schulze, 2007b). That is, if the coastal regions had such high market potential, why did economic activity not follow instead of locating inland?

To be sure, many of the most developed regions in my sample are either coastal - Catalonia (Spain), Stockholm (Sweden), the Basque Country (Spain), Bouches-du-Rhone (France), the Northeast and Northwest (Britain), for example - or very close to the coast or on a river - London, the Rhineland and Hamburg (Germany), for example. The high market potential of poor Sicily (Italy) or France's western coast, however, shows the correlation between income and market potential, while strong, is not uniform.

This can be seen in figure 5.5, which maps the residuals from a regression of log per capita income on log MP^C, instrumented with log $Dist\_London$, controlling for region and year fixed effects. Robust errors from this regression are clustered by region. Figure 5.5 shows, in cross-sectional standard deviations, the size of regions' over or under-performance in per capita income relative to their market potentials. A large positive standard deviation indicates that a region is performing better than expected given its market potential, and vice-versa.

Looking across the maps in figure 5.5, a clear "capital city advantage" can be seen. All major capitals stand out as over-performers, particularly the inland ones: Madrid, Paris, Brandenburg, and Vienna. This fits with NEG. Early on, Krugman (1993) wrote how Chicago, the central city of the American heartland, developed in the nineteenth century without any distinctive natural advantages. The city first stood on a plain, near a barely-nvigable river, and on inadequate, silty lakeside harbour. It was connected to the great Mississippi watershed by a canal, but that canal only had a few years of economic importance before the railways came along. Still, that was long enough for Chicago to become a central market - a focal point for transport and commerce that fed on itself through increasing returns. In Europe, a number of cities followed this same pattern Madrid, for example, arbitrarily became Spain's capital when Philip II moved his court there from Toledo in 1561, making Madrid the political centre of the monarchy. The area is dry, at the time poorly inhabited, and far from the coast and other cities in Spain, but it persisted despite this low market potential. Why? 'Cities are costly to build. Since fixed costs are a source of increasing returns, "sunk investments" ("history" in the Krugamanian sense) may account for path dependence' (Jedwab et al., 2014, p. 2). The same can be argued for Brandenburg (home to Berlin), which was a "sandy plain" before the reign of Friedrich Wilhelm I (reigned 1713 to 1740). Determined to build a great military power, he promoted immigration of Protestants from across Germany and France and Switzerland. These ideas have found strong empirical support in Redding et al. (2011), who found that the explanation for Frankfurt airport's dominance in passenger flows is its status as a hub - not its locational fundamentals or its proximity to other cities. Establishing its "hub" status was the event that set increasing returns in motion.

Another interesting feature of figure 5.5 is the over-performance of Austrian and northeastern Hungarian regions. These are all very low-market potential regions, and so have higher per capita income levels than the model predicts. This suggests that the literature's focus on peripherality is correct, but that some countervailing influence existed (Schulze, 2007b).

Most regions are quite close to the yearly mean, but a few under-performers stand out. In 1870, East Prussia is a clear under-performer, falling into the negative 2.5 standard deviation interval. East Prussia had a lot going for it in market potential terms: coastal, and part of a fast-growing country. Its economic structure, being predominantly agricultural, however, meant that the region had little value-added to export. For an increase in its income based on market potential, it would have to draw in footloose industry. To some extent, this was achieved. Its per capita income level rose, as its share of industry relative to agriculture grew. By 1910, its performance relative to market potential was in mean territory. We see the same story with other coastal under-performers like Dalmatia (Austria-Hungary), Galicia (Spain), and Finistere (France). All are coastal regions, and while the first two were part of low-income countries, all had potential access to large markets beyond their shores.

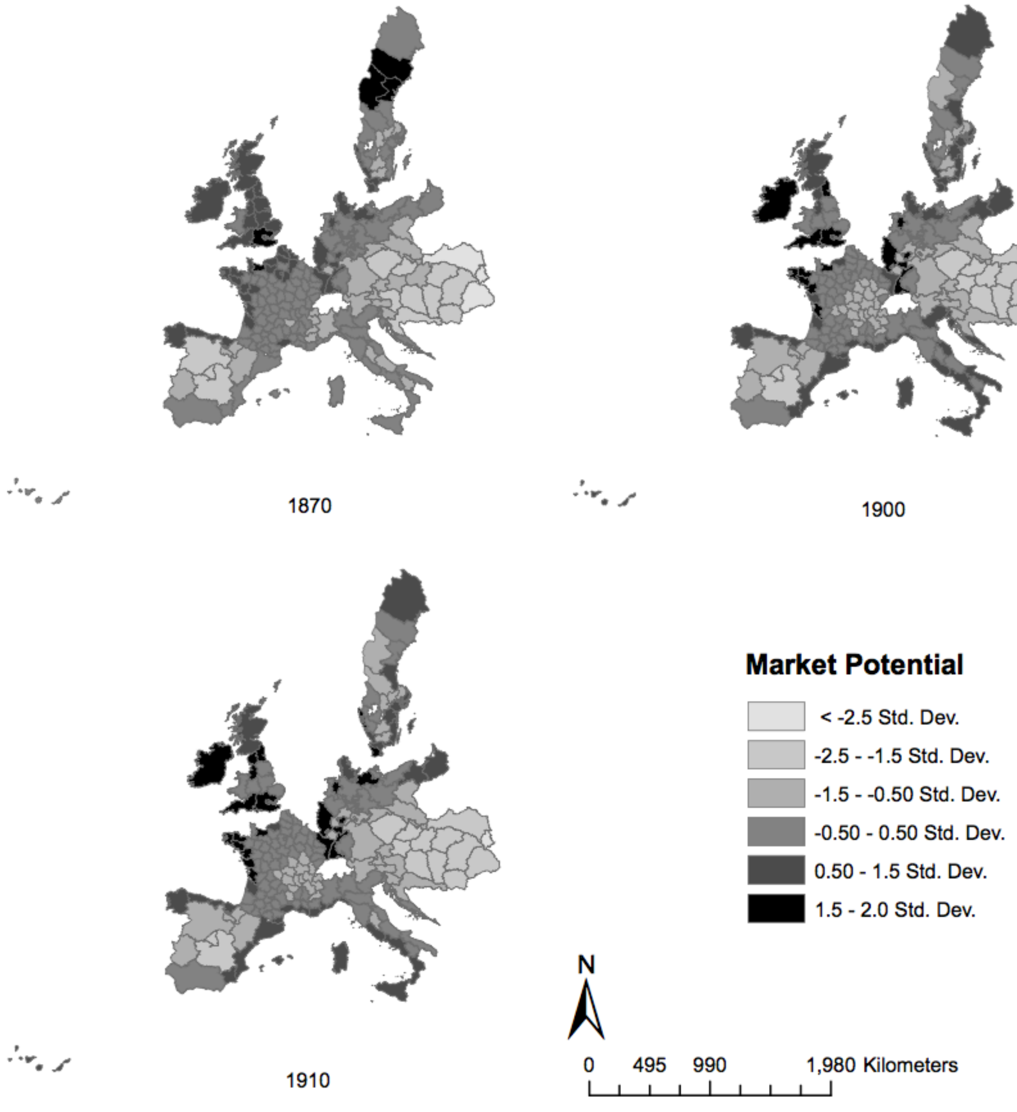

Figure 5.4: Market potential map

Notes: Data are constructed market potential, MP^C*, and originally in \$. The standard deviations are calculated at each year.*

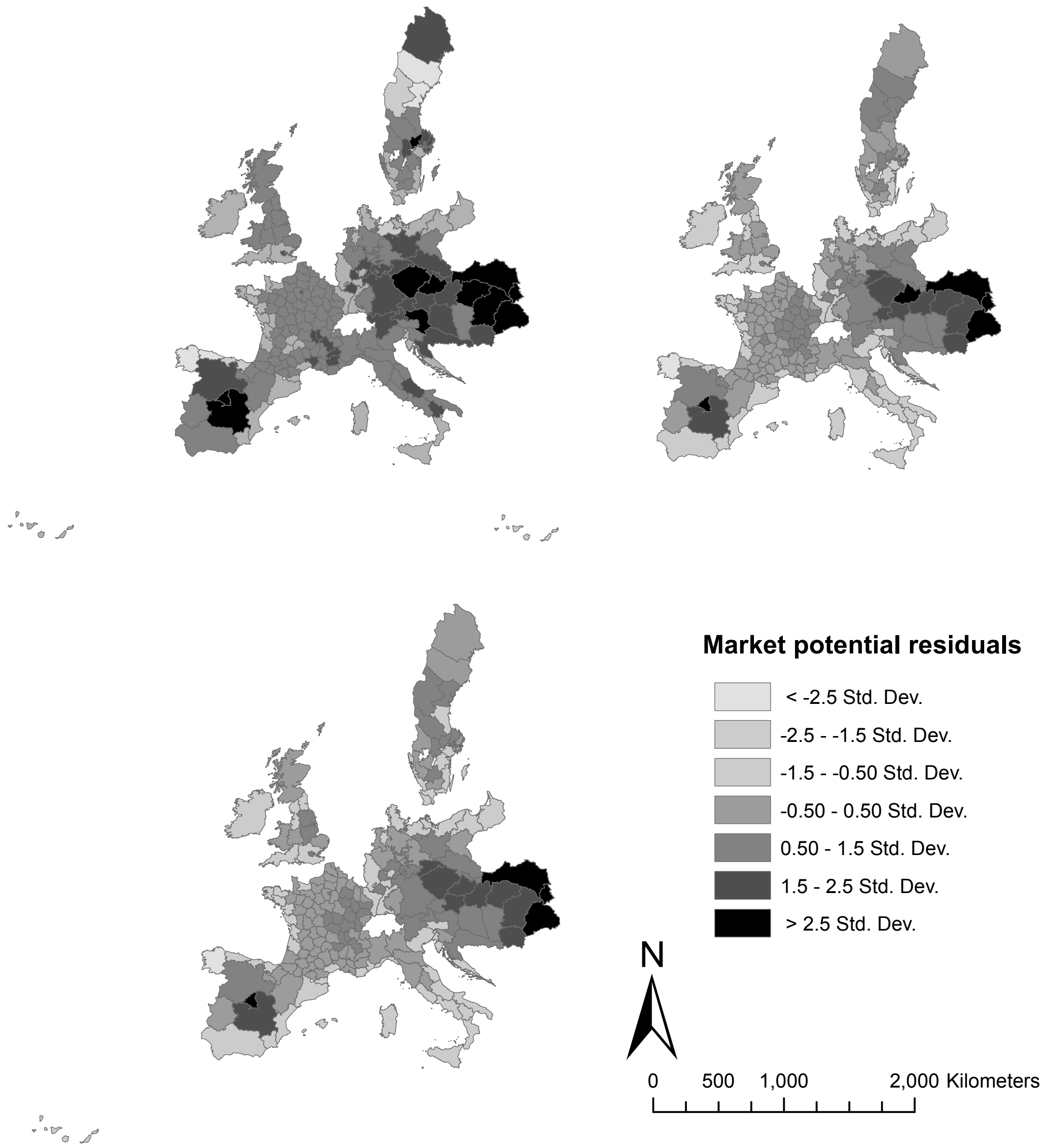

Figure 5.5: Market potential residuals map

Notes: Residuals from a regression of log per capita income on log MP^C, instrumented with log Dist_London, controlling for region and year fixed effects. Robust errors from this regression are clustered by region. The standard deviations are calculated at each year.

5.5 What about Belgium and the Netherlands?

Two small, rich countries are noticeably missing from my sample of regions: Belgium and the Netherlands. I was not able to include becase there were no estimates of regional GDP available at the time of writing. Both countries are included in my calculations of market potential, as foreign markets. Their exclusion may matter because (1) they are both home to coal deposits, (2) they enjoyed the kind of institutional change that is argued to improve growth prospects, (3) they are located close to large markets, and (4) they are both rich by European standards.

In other words, their exclusion might be exerting a bias towards zero for the institutional, cost to coal, and market potential variables. While this would affect my empirical point estimates, the biases are all in the same direction. This means that the hierarchy of coefficients, or the relative importance of the coal, institutions, and market potential explanations, are less likely to be affected.

This is, unfortunately, one other reason besides measurement error that makes the estimation of point estimates difficult in historical work. My strategy here has been more modest: uncovering which explanations mattered a lot or a little for regional income differentials.

Still, we can get an idea of what might happen with the inclusion of Belgium and the Netherlands by looking at a recent paper on market potential and American productivity by Liu and Mesissner (2014). The authors measure market potential for 27 countries including Belgium and the Netherlands in 1900 and 1910. They find that larger market potential is positively associated with GDP per capita the sample, but find heterogeneous effects depending on countries' sizes and borders. They single out Belgium and the Netherlands as two small countries that have high foreign market potential. In a counterfactual exercise where they eliminate international borders, they find welfare gains (in real income) of 25 per cent for the two countries, among the largest gains across all countries. They control for factor endowments, human capital, and geographical positions of countries.

The results of Liu and Mesissner (2014), then, suggest that had Belgium and the Netherlands been included in my sample, we might have seen a yet greater effect of market potential on income. That is, a larger point estimate or more reason for market potential to be ranked above institutional and coal-based explanations.

5.6 Discussion

The general findings of this chapter are as follows: (1) whether constructed or estimated, market potential has an economically meaningful and statistically significant effect on per capita income levels; (2) foreign markets have a larger effect on per capita income levels than domestic markets; (3) increasing core market potential relative to peripheral market potential results in an absolute decline in peripheral per capita income levels and a decline in peripheral share's of total GDP; and (4) the relationship between market potential and per capita income is not uniform. Residuals suggest the influence of increasing returns, and smaller role for factor endowments and locational fundamentals.

Given recent work, the baseline finding should not be surprising. Roses (2003), looing at Spain between 1797 and 1910, shows that as transport costs declined and internal barriers to domestic trade were eliminated, Spanish industry concentrated in Catalonia and the Basque Country. Wolf (2007) shows that after Polish unification in 1918, when internal tariffs were removed and transport infrastructure improved, Polish industry relocated based on skilled labour endowments as well as access to markets. Crafts (2005a), looking at Britain between 1871 and 1931, shows that the North, Scotland, and Wales were much more economically peripheral before WWI than after. This, he argues, was due to the rise of road haulage at the expense of coastal shipping. The novelty here, if I can claim any, is using the tried-and-tested concept of market potential to explain a specific historical episode: the crystalisation of Europe's core-periphery spatial income structure, *in aggregate*. A number of explanations have been put forward for it. In the previous chapter, I discussed the two dominant ones - coal endowments and institutions. While the latter explanation held up well against the coal explanation, it loses all significance when controlling for market potential.

The value in analysing Europe in aggregate is that economic activity was not simply responding to markets within its own country, but to markets across the Continent. If this were not the case, then one important feature in the data would not be there - the consolidation of Europe's core-periphery pattern - neither would the two following findings - the importance of foreign versus domestic market potential, and the de-industrialisation effect of relative declines in market potential.

The few analyses of foreign versus market potential that we have show the assert the importance of foreign markets (Redding and Venables, 2004). This is uncontroversial: a reduction in international trade costs, an important feature of this period, makes foreign markets more attractive since, in aggregating all markets in the sample, they are much larger than domestic markets. In other words, even if market sizes (the numerator) were held constant and international trade costs (the denominator) dropped, foreign market potential would grow faster than domestic market potential. This is in effect the same argument found in the development literature, which argues that openness to foreign markets induces export growth and income convergence, while the opposite is true for countries that are close to international markets (Dollar, 1992; Sachs and Warner, 1995).

In some senses, the historical literature has implicitly dealt with these ideas. Kenwood and Lougheed (1982), for example, write about Britain's pace of industrialisation accelerating relative to Europe on the back of its cotton-goods exports to the United States. While we do observe these aggregate effects, in reality the pull of markets only applied to a few select areas *within* countries, which is why see so much regional diversity in Europe's income map. Pollard (1981, p. 8-9) gives the example of the way in which the British linen industry dramatically relocated from its main centre of the West Country and East Anglia (south east of Britain) to the West Riding of Yorkshire (north west of Britain), as demand markets shifted from Europe and the Mediterranean to the United States. To use the current terminology, Yorkshire's market potential was greater than East Anglia's, for two reasons. The size of the American market was expanding faster than Continental Europe's, and the cost of transport from Yorkshire to the United States was lower than that from East Anglia, which is better placed to serve the Continent. The result was production moving to Yorkshire. Scandinavia's rising income after 1870 was export-led rather than based on domestic markets - a consequence, as Pollard (1981, p. 201, 236) put it, of its high degree of integration with 'Inner Europe.' Sweden, in particular, still had a small population, agriculture was stagnant, and it was often easier to transport goods from abroad than from another region in the country - a fact that can be seen in the previous chapter's transport map, figure 4.1. It cost less to ship goods from Stockholm to Continental Europe than from Stockholm up to Norbotten in the north of Sweden, which was not much of a market in any case. Swedish production thus concentrated in the southern tip of the Swedish peninsula. Moving to Central and Eastern Europe, we see that the Czech lands, then part of Austria-Hungary, came close to the industrial regions of northwestern Europe. The thriving textile industries there were certainly not supported by the poor domestic market of peripheral Austrian and Hungarian regions, but by the booming export markets of neighbouring German Saxony and Silesia. This is why the income gradient of Austria-Hungary declined as one moved east, which brings us to de-industrialisation.

When markets exert their pull, there will be winners and losers. The winners are regions that have successfully attracted economic activity; the losers, the source of that activity. NEG holds that firms locate near large markets when trade costs drop to an intermediate level, further epxanding the size of those markets - or setting in increasing returns. This was theorised to occur at an international level by Krugman and Venables (1995), and has received some attention in the historiography (Crafts and Wolf, 2013a). This explanation is consistent with Europe's core-periphery structure at large, but is also visible within countries. As industrialisation progressed through Britain, as an early example, the cotton weavers in rural counties around Manchester were replaced by power looms in the city. The weaving sheds did not revert back to agriculture, but were abandoned by their labourers, who moved into industry in the Lancashire industrial complex (Crafts and Wolf, 2013a; Pollard, 1981). The effect on those rural Manchester regions was an absolute loss of income. Production also moved to regions much farther afield, in spite of the fact that the mobility of production factors was lower between than within countries - the relative size of foreign markets, as we have seen, compensated for this. For example, as markets grew in northwestern Europe, the Languedoc wool industry lost its demand markets across Europe to

Yorkshire exporters, who found the benefit of already-established markets with all their distribution facilities and infrastructure (Pollard, 1981, pp. 29, 123). This not only made the Languedoc area poorer than Yorkshire, but poorer in absolute terms. By the start of the twentieth century, it had completed its transformation from an industrial area successful in woolens to an agricultural area, producing (good) wine (Johnson, 1995). While these examples show the pull of markets and their losers, some regions held onto high (low) levels of income despite their low (high) market potentials, as seen in the map of residuals from the income-market potential relationship (figure 5.5.

In the discussion around figure 5.5, I related the large positive residuals for inland cities to increasing returns. For whatever reason "historical accident" as Krugman (1991) calls it, these cities were established as locations of activity. These fixed costs then became sources of increasing returns, creating income levels that are at odds with their market potentials, as exemplified by Madrid. To be sure, locational fundamentals did matter - to the extent that they also affected market potential. Keeping to Spain, the other centres of economic activity were coastal: Catalonia and the Basque Country profited from cheap ocean transport. At the other end of the spectrum we have Dalmatia (Austria-Hungary), which in spite of its coastal access and so high market potential, was consistently among the poorest regions in my sample.

The implication of this is that locational fundamentals are often less important than increasing returns (and the historic event that set them in motion) in the distribution of economic activity (Davis and Weinstein, 2002; Krugman, 1991). Market potential should in theory draw in economic activity to these regions but, for example, Dalmatia was poor long before 1870 and, despite its coastal access, remained poor since. The operative word in the term is *potential*. That Dalmatia, despite its potential access to foreign markets failed to draw in activity is an important point. It shows that even dramatic changes in trade costs, and hence market potential, cannot always undo historic events (Davis and Weinstein, 2002).

There are, of course, other determinants that may be at play. Wolf (2007), for one, found that market potential mattered in attracting industry *in conjunction* with other factor endowments, particularly skilled labour. The relationship, however, is not necessarily so clear-cut. The location of skilled labour, or human capital more generally, is itself determined by market potential. As Redding and Schott (2003) argue, besides the direct trade costs savings that matter to central regions, market potential also provides longer-run incentives for human capital accumulation by increasing the demand for skilled labour. As the development of skills requires investment, the decision to invest is based on the expected returns of that investment. Labourers would only acquire skills if they foresaw improvements in market potential that would, given the established relationship, lead to improvements in production. A negative residual may thus indicate a failure, for whatever reason, to invest in other determinants of production despite high market potential levels. That is, market potential may have been attracting production, but lagging in its ability to incentivise investment in factor endowments.

Chapter 6

Addendum: Market Potential and Human Capital

6.1 Introduction

Did market potential have a direct effect on other determinants of income, namely human capital? In the preceding chapter I showed that regions with high market potential generally enjoyed higher per capita income levels. I also showed that while market potential is a strong determinant of income, there is - as always - a residual. Authors have emphasised the joint role of market potential and factor endowments. Wolf (2007) nested both skilled labour and market potential in his model of Polish inter-war industrial location, and Crafts and Wolf (2013b) consider both physical geography, like coal deposits and climate, as well as market potential in their explanation of the location of nineteenth-century British cotton textile industires.

Redding and Schott (2003) and **?**, however, show us that market potential can have causal effects on those very factor endowments. Their work shows that market potential provides incentives for 'long-run human capital accumulation by increasing the premium for skilled labour' (**?**, p. 610). We would expect people to respond to the potential gains from an increasingly larger market by investing in their human capital levels. That is, increasing market potential increases the return to skill. The authors build on the standard two-sector agriculture and manufacturing NEG model from Fujita et al. (1999b), to let unskilled workers endogenously decide whether to invest in education. The basic logic being that an increase in remoteness, measured by market potential, inflicts higher trade costs on firms selling their products, which has the same effect as a reduction in the price of a manufactured good relative to an agricultural one. Therefore, as manufactured goods are more skill-intensive than agricultural ones, firms had less value-added to pay and retain their skilled workers, and workers had less of an incentive to acquire skills. Redding and Schott (2003) empirically show that countries located far away from centres of world economic activity tend to have lower levels of educational attainment. Lopez-Rodriguez et al. (2007) and **?** used the same framework to motivate their empirical analyses, which both found that increases in market potential had significant positive effects on education levels in contemporary Europe.

My hypothesis here is that, generally, regions far from the northwestern European core had low stocks of human capital, as measured by literacy (more on this later), but as economic distance, as measured by market potential, between the core and periphery dropped over the late-nineteenth century, the periphery's incentives to invest in human capital increased. This produced an *international* convergence in literacy (Cipolla, 1969; Crafts, 1997). National law that made primary enrolment compulsory is an alternative explanation for this convergence but, as I show in this chapter, regional variation in literacy was higher within than between countries. Further, regional literacy rates were spatially correlated. These features in the data are consistent with my hypothesis: market potential, itself spatially correlated, varied by region causing the incentives for human capital investment to also vary by region. The emphasis here is on market potential providing an incentive on which agents can decide to act: it reflects the demand for, rather than supply of (as with national law), for human capital. A negative residual in the market potential relationship, as highlighted in the conclusion of *Chapter 5*, may be interpreted as agents failing to invest in their human capital despite their regional level of market potential.

In the previous two chapters, I argued through my measure of regional institutions that one of the functions of institutions whether de facto or de jure is to provide goods like human capital; in this case, literacy. Here I am arguing that growing industry created incentives for people to educate themselves. How do the two arguments tie in together? As the previous chapters showed, high regional literacy differentials existed. My regional institutions variable measures the extent to which these differentials were a function of de facto political choices. A look back to the frontier model results of this exercise in table 4.2 shows that the formulation does not explain *all* the variation in literacy. As I go on to argue in this chapter, governments made serious, almost concerted attempts at legislating for the public provision of education, and many of these attempts were successful, but some were unsuccessful. A portion of the within- and between-country literacy differentials, I argue here, can be attributed to varying incentives for accumulating human capital. In short, there is room for both the incentives and institutional explanation.

The rest of this chapter is organised as follows. In the following section, I review the relevant historiography, which also serves to introduce this chapter's conceptual basis. I then discuss my data on regional literacy rates in more depth than in *Appendix B*. The empirical strategy follows, with a discussion of estimation issues and then results. The final section concludes.

6.2 Demand verus supply of human capital

While the paths taken varied from country to country, most European states institutionalised mass education - primary enrolment - before WWI. Table 6.1 shows differences in the timing of compulsory schooling among countries and, as a measure of the latter's success, their primary enrolment ratios in 1870. Prussia was unusually early in enacting compulsory primary enrolment, thanks to the enlightened Frederick II, who set up a national education system to 'save the souls' of his subjects (Soysal and Strang, 1989, p. 278). The system failed to meet its goals of universal schooling. Still, enrolment ratios grew rapidly as part of Prussia's attempt to nation-build after its 1806 defeat to Napoleon.

Conflicts between secular and religious authorities delayed the enactment

Table 6.1: Compuslory schooling introduction and enrolment in 1870.

Country	Introduction of Compulsory Schooling Laws	Primary Enrolment Ratio (%) in 1870
Prussia	1763	67
Spain	1838	42
Sweden	1842	71
Austria	1869	40
Italy	1877	29
Britain	1880	49
France	1882	75
Ireland	1892	38
Portugal	1844	13
Greece	1834	20

Notes: taken from Soysal and Strang (1989, p. 278). Primary enrolment is for both sexes, aged between five and 14. The years refer to the introduction of compulsory school attendance - not the establishment of public schools or important, generic education reforms.

of laws in other countries. The French state, for example, struggled with the Catholic Church for control over schooling over most of the nineteenth century. The Church and landed gentry wanted to ensure the perpetuation of conservative sympathies, while the state was after building a republican France. The passage of the "Jules Ferry Laws", which established free education in 1881 and then mandatory and laic education in 1882, are in fact interpreted as part of a broader anti-clerical campaign in France (Barnard, 1969). There were earlier attempts: laws enacted in 1791 and 1833 to establish a state education system were repealed by Napoleon and Louis Napoleon. Similarly, in Britain a number of powerful interest groups prevented the state from controlling education until the Foster Education Act of 1870, which provided a framework for national schooling, and the 1880 Act which made schooling compulsory. Anglicans, Non-conformists, and Dissenters tried to expand their congregations through schooling. The middle class was intent on maintaining its sons' advantages over the working class (Soysal and Strang, 1989). These groups and their aims blocked the supply of a unified national system until a late date, by European standards.

Crucially, table 6.1 also shows us that there is a weak relation between the timing of compulsory schooling laws and primary enrolment ratios. France achieved an exceptionally high ratio by 1870, when its laws were enacted in 1882. Spain, by contrast, enacted compulsory schooling lawas unusually early, in 1838, but only managed a ratio of 42 per cent by 1870. Portugal and Greece had compulsory schooling laws for at least 25 years before registering some of the lowest enrolment ratios in 1870.

While counter examples like Prussia and Sweden exist, these numbers indicate a problematic relationship between national legislation and outcomes. That is, the increases in literacy, numeracy, and enrolment ratios observed by historians like Cipolla (1969) and in the lieracy data under study here cannot be solely a *supply*-side story: laws sometimes came after enrolment improved, most notably in France. Neither can it solely be a supply-side story in terms of resources: Britain, with a developed economy, large taxable base, and stable institutions, had persistently low literacy rates until the turn of the nineteenth century (Mitch, 1992). Further, even after the enactment of compulsory schooling laws, there was considerable regional variation in literacy and enrolment rates *within* countries. A Theil Index shows that within-country inequality in literacy rates was twice as high as inequalities between countries: 0.062 for within country inequality, and 0.036 for the between-country component. This either implies that the *de facto* enforcement of schooling laws varied within countries, much like the ideas underlying my regional institutions ($INST$) measure, or that enforcement was consistently poor wtihin countries, and variation in educational outcomes came from variation in *demand*-drivers.

What could those demand-drivers be? Redding and Schott (2003) developed a model that allows unskilled workers to endogenously decide whether to invest in education based on the economic (export) prospects of their location. The main explanatory variable is market potential. The intuition is an increase in market potential lowers trade costs for firms exporting their products, increasing the price of manufactures relative to agricultural goods. As manufactured goods are more skill-intensive, firms would have less value-added to pay and retain skilled workers, and so workers had less of an incentive to acquire skills. This prediction has found empirical support in contemporary settings at international and subnational levels (Lopez-Rodriguez et al., 2007; Redding and Schott, 2003; **?**). In the context under study here, however, we first need establish some preliminary foundation that is usually taken for granted in contemporary settings: in particular, we need to ask whether the demands for human capital were coming from industry - rather than agriculture or services - in the “Second Industrial

Revolution."

Industrial demands for literacy

Becker et al. (2011b) provide us with evidence of this at the subnational level in nineteenth century Prussia. The authors show that for Prussia, which was a technology follower rathern than leader like Britain, formal education mattered. They show that primary enrolment, across 344 counties, in 1816 had a strong positive effect on the presence of metal manufacturers in 1849 and 1882. That is, education is significantly correlated wtih industrialisation in the early and later phases of the Industrial Revolution. As Becker et al. (2011b, p. 120) conclude,

> *Industrial development...depended on the availability of an educated population that was earlier aware of the producive potential of new technologies and more capable of adjusting to change, Some regions lacked these skills necessary to adopt the new industrial technologies from and catch up to Britain.*

Using some of my own data on French and German regional sectoral employment (see *Chapter 3*) and on regional literacy rates, we can find further support for this. Literacy, as I discuss in section 6.3, is a noisy but useful measure of human capital. Table 6.2 shows the correlations between regional literacy rates (L) and agricultural and industrial employment shares in the regional labour force ($\frac{A}{LF}$ and $\frac{I}{LF}$, respectively). I control for year fixed effects, since literacy and industrial employment trend up over time while agricultural employment trends down, and for country fixed effects, since there a differences in the meausurement of literacy by country[1] and because they introduced compulsory schooling laws at very different times (see table 6.1). Put together, French and German regions make up 53 per cent of my sample of regions. The results in columns 1 and 2 show that the share of agricultural employment in total employment is negatively associated with literacy rates, and that the opposite is true for industrial employment; both coefficients are highly significant. Introducing lags in columns 3 and 4, we see that higher agricultural employment shares in previous years (the data cover 1870, 1900, and 1910) predict lower literacy rates; the opposite being true for

[1] I expand on this issue in section 6.3.

industrial employment shares. Finally, in column 5 we see that as the ratio of industrial to agricultural employment shares ($\frac{I}{LF}/\frac{A}{LF}$) increases, literacy rates also increase.

Table 6.2: OLS French and German literacy rates versus sectoral employment

	(1)	(2)	(3)	(4)	(5)
$\frac{A}{LF}$	-14.952***				
	(0.157)				
$\frac{I}{LF}$		15.157**			
		(5.177)			
$\frac{A}{LF}_{t-1}$			-3.180**		
			(1.431)		
$\frac{I}{LF}_{t-1}$				5.755**	
				(2.173)	
$\frac{I}{LF}/\frac{A}{LF}$					0.457***
					(0.084)
Country F.E.	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes
N	324	324	216	216	324
$\hat{R}^2$	0.663	0.671	0.222	0.100	0.658

Notes: Dependent variable is L, regional literacy rates measured in percentages. $\frac{A}{LF}$ *is the share, from zero to one, of a region's labour force in agriculture;* $\frac{I}{LF}$ *is the same for industry.* $\frac{I}{LF}/\frac{A}{LF}$ *is the ratio of industrial to agricultural employment shares. Estimated using OLS, with robust standard errors clustered on regions reported in brackets. The results do not change when using a Tobit estimation procedure, with an upper censoring limit of 100 and lower of 0. Panel consists of 23 German and 85 French regions over 1870, 1900 and 1910. Employment data same as used in Chapter 3.*

This is not to say that human capital, or literacy as measured here, was useless to agricultural workers. As Lockheed et al. (1980) show in their rigorous survey, there are positive and significant returns to education in the form of increased agricultural productivity and higher wages. The point of the model is to say that the returns to human capital, proxied by literacy, were much higher in the high value-added manufacturing sector relative to the low value-added agricultural sector (the empirical results in *Chapter 3*, where I estimate regional income using employment shares, makes this point on the relative productivity differences between the two sectors).[1] Indeed, even when it comes to agriculture,

[1] Parnam (2012) finds substantial returns to education for farmers in the American Midwest during the early-twentieth century. This is all well and good, but his study does not provide estimates of these returns relative to non-agricultural sectors, which is crucial to the conceptual

Lockheed et al. (1980) stress that returns are highest in modernising economies, where there are new technologies (for example, chemical fertilisers) to adopt.

The rate of technological progress was also higher in the manufacturing sector than in any other. As Mokyr (2002a, p. 67) discusses, industrialisation was about the development of and commitment to technical knowledge, which depended on literacy. In Britain this happened earliest. By the early-nineteenth century, the commitment to 'technical knowledge', as Mokyr (2002a) calls it trickled down from the elite to the middle classes, to the extent that one observer noted, 'In every town, nay almost in every village, there are learned person running to and from with electrical machines, galvanic troughs, retorts, crucibles, and geologist hammers' (cited by Mokyr (2002a, p. 73)). These ideas were not new to West (1985, p. 231), who argued that as Lancashire's indigenous population moved into 'key growth areas like manufacturing', where he says 'literacy was of more consequence', Irish immigrants moved into the low-skilled jobs they left behind. West (1985, p. 231) makes the additional point that as the literacy of the Irish immigrants improved, compared to normal levels in their place of origin, 'we can still speak of this as a growth in education, and one that was associated with industrialization.'

Why is manufacturing rather than *services* central to the explanation? Following on from the discussion above, returns on education were higher in the manufacturing sector because economically useful knowledge was increasing fastest there. For example, London, already then a services centre, was as Van der Beek (2010) writes less affected by technological progress because of its economic structure. Secondly, market potential is more relevant to manufacturers since, unlike most services, they export their production and are, unlike agriculture, footloose. The cost of distance has a more immediate effect on manufacturers than on services where, during this period, output was mainly non-traded.[1] Traded goods, also contain embedded knowledge, which influences human capital levels (Grossman and Helpman, 1991). More fundamentally, manufacturing is at the centre of the model because this was an industrial age and we are trying to understand,

model I use here. We cannot know whether education was better rewarded in agriculture if we do not know how well rewarded it was in manufacturing or commerce.

[1]Arguably, the traded proprotion of the London services sector was high enough for market potential to have an effect. We cannot test this due to data unavailability.

like Becker et al. (2011b), whether regions adapted to the growing and spreading industrial economy; in this case, in terms of their human capital levels.

The foundation we have to lay down concerns the explanatory variable: why should we expect the Redding and Schott (2003) market potential explanation to work in this context? First, table 6.1 has already shown us that the argument cannot be simply an institutional one. Laws that unified and established national compuslory schooling were enacted some times after primary enrolment improved, as in France, and some times long before, with no improvement in enrolment afterwards, as in Spain. Secondly, the Theil Index of regional literacy rates brought up earlier shows us that the within-country component of inequality accounts for two-thirds of the total inequality in literacy rates (the rest accounted for by the between-country component). Not only was regional literacy rate variation high, but like per capita income and market potential, it was spatially autocorrelated, as can be seen in figure 6.1.

We can be more systematic about this and calculate Moran's I statistic of spatial autocorrelation. The statistic is calculated as:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(L_i - \bar{X})(L_j - \bar{L})}{\sum_i (L_i - \bar{L}})^2 \tag{6.1}$$

where N is the number of regions indexed by i and j; L is regional literacy rate; $\bar{L}$ is the cross-sectional mean of L; and w_{ij} is an element of the spatial weights matrix used in *Chapter 3*. It is a squared inverse distance spatial matrix, where Euclidean distance is measured between all regional nodes, but only geographically close regions are substantially associated with one another. That is, closer regions get bigger weights. Values for I range between -1 (perfect dispersion) to 1 (perfect spatial correlation), and its expected value under the null hypothesis of no spatial autocorrelation $E(I)$ equals $-1/N - 1$. The results in table 6.3 show that at each year regional literacy rates were spatially autocorrelated, that is, high (low) regional literacy rates were located near other high (low) regional literacy rates. I is not particularly large, but it is positive, and highly statistically significant.

High within-country variation and evidence of spatial autocorrelation in literacy rates imply that changes on the supply-side, in states enacting laws that

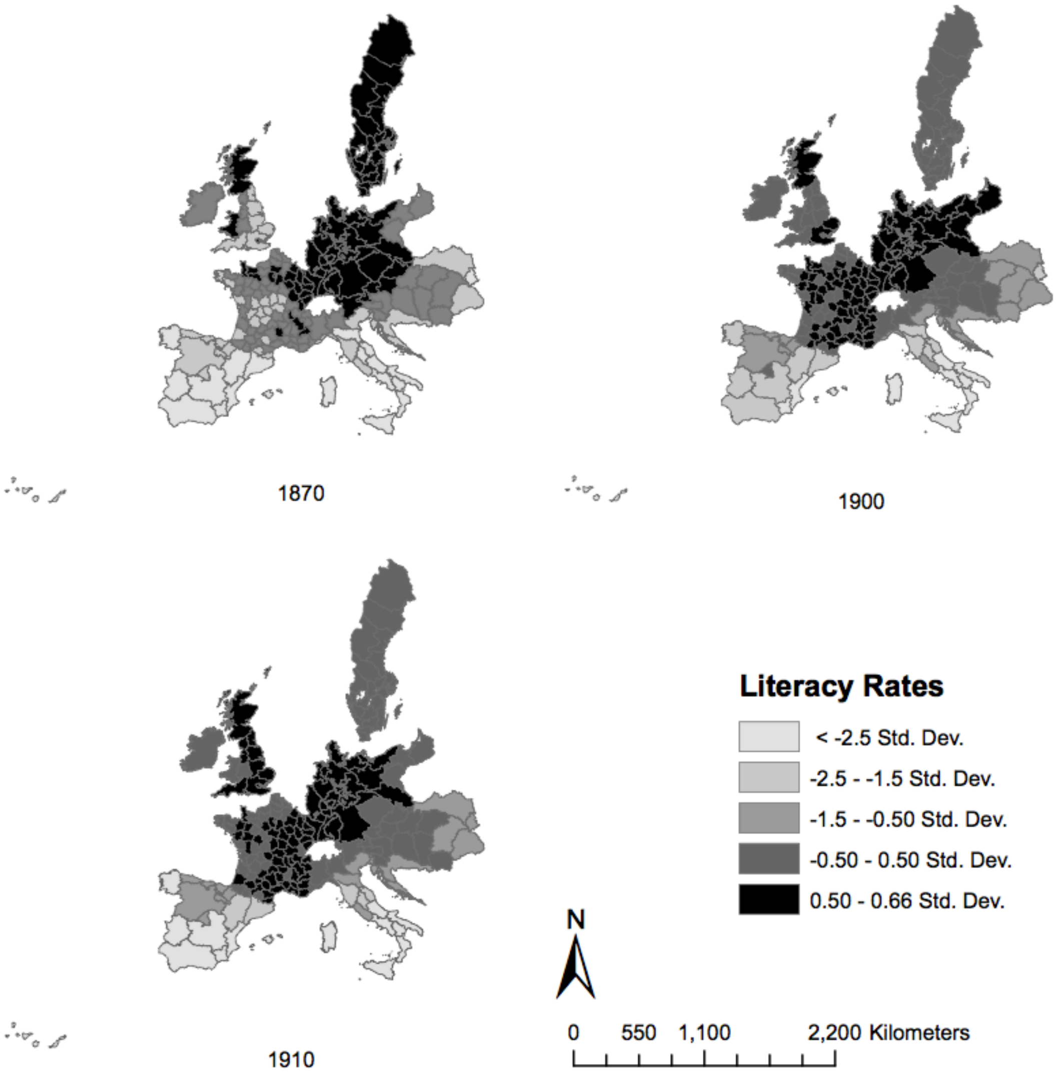

Figure 6.1: Map of regional literacy rates

Notes: Underlying data are in percentages, and are expressed as cross-sectional standard deviations on the map.

Table 6.3: Spatial autocorrelation in regional literacy rates

	1870	1900	1910	Sample
I	0.131	0.137	0.125	0.183
$E(I)$	-0.005	-0.005	-0.005	-0.002

Notes: I is the Moran's I statistic; $E(I)$ is its expected value. All values of I are statistically significant at the one per cent level, as calculated with a one-tail test.

made primary education compulsory and in building schools, do not really work as an explanation. As Murtin and Viarengo (2011) discuss, enforcement of these laws was weak due to budget constraints, making primary schooling more of an option than a rule. So, while the option of education may have been enabled by these laws, regional literacy variation shows that the option was valued differently across space. This calls for an explanation of literacy rates that is spatial and sub-national in nature. Market potential can, in theory, account for both these observations in the literacy data: as the preceding chapter has shown, it predicts clustering and operates at subnational levels.

Indeed, **?**, p. 593 finds that access to markets of demand play an important role in Europe's contemporary regional income structure, but that the benefits of market access, he argues, come from 'increased incentives for physical and human capital accumulation and not from direct trade cost savings.' This is an expected finding for a European economy with no tariffs and low transport costs that is mainly based on tradeable services, but in the 1870 and 1910 period, when economies were industrialising and there was no customs or monetary union to speak of, the human capital channel is unlikely to be as substantial as the direct trade cost savings channel. Still, the intention here is not to set up a horse race between the two channels; it is only to show that both mattered, that is, "geography" mattered in more ways than one, and can in fact integrate explanations (as in, human capital versus geography) that are often pitted against each other (Hanson, 2005; Midelfart et al., 2000). Before getting into the empirical analysis, it is important to examine whether literacy rates proxy human capital levels well enough for my purposes.

6.3 Literacy Rates as human capital

Woessmann (2003) discuss a range of human capital proxies. Many of them, like "education-augmented labour input," are difficult to use in historical work due to data constraints. Even if the data do exist, however, the time and labour needed to "augment" them are beyond what I can do here (Cvrcek and Zajicek, 2013; Hippe, 2012). As such, (Hippe, 2012, p. 1530) writes that 'the most important proxies for human capital in Europe are literacy (the ability to read and write, and to sign a register) and, more recently, numeracy.' Other economic historians have used school enrolment ratios and book production, but even these variables are difficult to find in most cases (Baten and Van Zanden, 2008; Becker et al., 2011b). Here I use the regional literacy data used in *Chapter 4*, and detailed in *Appendix B*, but given its more central role in this chapter, it is worth reviewing the use of literacy in historical work and its associated issues.

Literacy has long been a popular measure of human capital. It is a noisy "output" (or stock) measure of human capital, as opposed to "input" (or flow) measures like school enrolment. As Vincent (2004, p. 24) put it, literacy is a 'tool for enabling individuals and social groups to extend theur understanding of themselves and their world.' Graff (1991) categorises work done on historical literacy into three generations. First, authors like Schofield (1968), Cipolla (1969), and Stone (1969) laid the groundwork for future research in the field by emphasising the importance of literacy and its changes - in the case of Cipolla (1969), on a global scale. It also provided us with extensive sources of data, Building on their work, the second generation exploited more detailed quantitative information on literacy with more advanced techniques (Mitch, 1992; Schofield, 1981). Lastly, the third generation has been combining the study and collection of literacy data with other measures, such as numeracy (Crayen and Baten, 2010; Hippe, 2012; Reis, 2005). Throughout all these generations the main aim has been to measure human capital through literacy, and to check this measure against other possible ones.

What are the issues with using literacy? First, observations may be of people who can read but are not necessarily capable of understanding the content of what they are reading. Second, we might also be observng people whose only writing skill is that of signing a contract or marriage register - like Vincent (2000), I use marriage registers to measure nineteenth century British literacy. Third, literacy may be measured differently across countries. For Germany, my literacy data is drawn from adult military recruits listed in yearbooks, while for France, it is "*jeunes gens*" listed in census books. Fourth, literacy may be a relevant proxy of human capital, but is more relevant in certain periods than in others. That is, literacy may only contain useful information on human capital in periods where economies and societies were literacy-scarce. Further Woessmann (2003, p. 243) makes the point that while literacy rates 'certainly reflect a component of the relevant stock of human capital...they miss out most of the investments made in human capital because they only reflect the very first part of these investments.' That is, literacy rates exclude the acquisition of more "advanced" human capital, like logical or analytical reasoning. I deal with these points below.

6.3.1 Literacy in the Second Industrial Revolution

Perhaps it is best to start with the criticism that, during the Second Industrial Revolution, literacy was not a relevant form of human capital. That is, during this period of industrialisation acquiring education was often done in the form of on-the-job learning, apprenticeships, and general work-experience, rather than learning how to read and write at school. A corollary of this point is that the mass of (factory) workers during this period did not require, and so did not achieve, literacy in their working lives.[1]

The regression results in table 6.2 have already shown us that as a the share of a region's labour force in industry increased, and increased relative to agricultural shares, its literacy rate also increased. If literacy were irrelevant to industrial employment during this period, then we would expect a negative or insignificant correlation. Even if literacy preceded industrialisation, as it did in Sweden or Britain, industry still placed demands for *more* literacy: I am not after explaining the initial spurt in industry or literacy (Sandberg, 1979). As Becker et al. (2011b) argue, education may not have been important for the first industrialiser, but was important in the Continent's catch-up industrialisation. The corollary of this is that regions starting from a low-GDP base would invest more in education in an effort to catch up, as evidenced by the German experience (Becker et al., 2011b). Further, even if we accept that low-level factory work may not have *required* literacy, industrialisation more generally 'created more jobs, both in skilled artisanal and supervisory capacities...for which literacy was, if not an absolute requirement, at any rate a great advantage.' (Laquer, 1974, p. 102).

[1]These pessimistic views of the effects of industrialisation are were summarised quite early by West (1978). It is surprising that the tradition persists.

Moreover, it is a misconception that this phase of industrialisation bred a mass of low-level, illiterate factory operatives. Even in towns were industry was the largest employer, factory operatives formed only a minority of the labour force. In, for example, nineteenth century Manchester and Salford, 'only 25 per cent of the working classes - which themselves were said to constitute only 66 per cent of the community - worked in any capacity in textile factories' (Laquer, 1974, p. 102). In Bury, another town dominated by the cotton industry, only 40 per cent were so employed. In mid-nineteenth century industrial Preston, only 26 per cent of the male population over 20 were ccategorised as in factory employment (Laquer, 1974, p. 103). Industrialisation did not create, as Laquer (1974, p. 103) put it,

> *...an urban economy and society which was dominated by the lower factory operative. Not only did the economy require supervisory personnel for whom literacy was necessary, or very nearly so, but it created a whole mass of ancillary jobs - in engineering, transport, trade, retailing, finance, and the older artisanal trades.*

Indeed, West (1978, p. 381) identifies 1867 - the approximate start of the "Second Industrial Revolution" - as a truning point in the decline of illiteracy in Britian. That is three years *before* the Forster Act; 13 years before national compulsion was enacted; and 15 years before the state provided "free" schooling.

Finally, even if education was acquired through on-the-job learning or apprenticeships, this would not exclude the potential for workers to become literate. First, it is important to note that apprenticeships were more a system of the early-modern economy rather than the Second Industrial Revolution (Leunig et al., 2012). Second, Humphries (2010, p. 342) shows that, in Britain, the mean years of schooling for the 1851-1878 cohort was 10 per cent higher for those in mining and metals than it was for those in agriculture. Third, Laquer (1974, p. 106-7) gives evidence that shows factory workers would attend Sunday schools - set up for the express purpose of schooling factory workers - to improve their literacy and numeracy skills, and move into other occupations. Indeed, Becker et al. (2011b) show that primary schooling was the most important source of literacy acquisition.[1]

6.3.2 Reading and writing quality

Another objection to using literacy is that it is too basic: that people may have been able to read, but not understand the content of what they are reading, and that, in the case of the British data, could only sign their name, but not write. One way of getting at these concerns is to compare my literacy data with other data on human capital, wherever it is available. Hippe (2012) shows, for a number of European regions over the nineteenth century, strong positive correlations between adult (ages 23 to 72) literacy rates and numeracy (as measured by age-heaping). This encouragingly indicates that those recorded as literate were also likely to be numerate.[2] In this vein, I draw out a number of comparisons between my literacy data and the available school enrolment data. While literacy is a stock and enrolment a flow measure of human capital, flows of human capital depend on contemporaneous levels of human capital. The return, and so acquisition, of human capital depends on its stock. As Mokyr (2002b, p. 2) counter-intuitively argues, skilled workers are much more productive when they are among other skilled workers - *not* when they are the scarce resource. If there is a strong positive correlation between the literacy data and enrolment data, where more than a basic reading and writing ability is required, then it tells us two things: literacy does not occur in isolation of other forms of human capital (it is unlikely to have a literate, and unschooled population), and it is correlated with the "true" variation in human capital.

Starting with some straightforward national enrolment numbers, table 6.4 shows the Spearman rank correlations between my literacy data, averaged by country-year, and the national primary and secondary enrolment rates in Lindert (2004), for Austria, Britain, France, Italy, Spain, Sweden, and Fernandes and Schulze (2009) for Hungary. The ranks show strong consistency between the two measures, adding weight to the validity of the literacy data, and the notion that

[1] Primary schooling data cannot be used here because it is unavaible for most regions and years under study. See Hippe (2012) for human capital data constraints and alternatives.

[2] Hippe (2012) was, unfortunately, unable to share any of his data.

literacy is correlated with other measures of human capital.

Table 6.4: Literacy and primary and secondary enrolment rank correlations

	1870	1900	1910
1870	0.77***		
1900		0.83***	
1910			0.84***

*Notes: Literacy rates are means by country-year, and enrolment rates are from Lindert (2004). Figures are Spearman rank correlation coefficients, where *** indicates statistical significance at one per cent. Countries: Austria, Britain, France, Italy, Spain, Sweden, and Hungary. Values for Hungary from Fernandes and Schulze (2009). The cross-sections in Lindert (2004) only match with my 1870 (1880 in Lindert (2004)), 1900, and 1910 data.*

Moving onto regions, we can draw comparisons between enrolment and literacy for Austria, Britain, and Italy. The data for Austrian regions are from Cvrcek and Zajicek (2013), and record the school enrolment for both boys and girls between the ages of six and 12, which was made compulsory in 1869. The authors provide data on enrolment by generation for the 1830 to 1850 period. The data for British regions are from the 1851 Education Census, and record "elementary school enrolment," which is the the number of pupils in attendence in day schools (not evening or Sunday schools) on census day in 1851, divided by population aged between 5 and 19 inclusive. I also collected additional literacy data from the *Annual Register* books for the 1841 to 1861 period to draw further comparisons in this exercise.[1] The data for Italian regions are from Felice (2012), and are unusually detailed. He provides, for 1871, 1891, and 1911, regional data on primary and secondary enrolment (ages six to 14); tertiary enrolment (ages 14 to 19); and on total enrolment (all ages).

In table 6.5 I show the OLS correlations between all the available enrolment data and my literacy dataset. The average R^2 of all regressions is 68 per cent; a strong correlation, given the varying leads and lags used, and the small number of regions available per regression. It also tells us that the literacy data are a fair reflection of what was going on, in terms of human capital, in a region. Of all 44 regressions, the only statistically insignificant coefficient is that for the correlation between British literacy in 1910 and school enrolment in 1851, with a P-value of 0.33. Including this figure, the average probability value is 0.01,

[1]See *Appendix 2* for more details on sources of literacy data.

Table 6.5: OLS correlations between literacy and other human capital measures

Country Group	Literacy Rate in:	Ind. Variable	Coeff.	T-stat.	P-value	R^2	N
Britain	1841	Enr., 1851	2.117	3.540	0.008	0.296	10
Britain	1851	Enr., 1851	1.353	10.340	0.000	0.520	11
Britain	1861	Enr., 1851	1.048	6.790	0.000	0.503	11
Britain	1870	Enr., 1851	0.831	4.300	0.002	0.523	11
Britain	1881	Enr., 1851	0.587	4.270	0.002	0.520	11
Britain	1891	Enr., 1851	0.271	3.440	0.007	0.540	11
Britain	1900	Enr., 1851	0.069	2.070	0.068	0.466	11
Britain	1910	Enr., 1851	0.029	1.030	0.331	0.147	11
Austria	1870	Enr., 1830 generation	46.481	3.910	0.002	0.707	13
Austria	1870	Enr., 1840 generation	50.811	4.550	0.001	0.778	13
Austria	1870	Enr., 1850 generation	55.849	5.020	0.000	0.835	13
Austria	1900	Enr., 1830 generation	27.504	3.020	0.012	0.609	13
Austria	1900	Enr., 1840 generation	30.355	3.350	0.007	0.681	13
Austria	1900	Enr., 1850 generation	30.355	3.350	0.007	0.681	13
Austria	1910	Enr., 1830 generation	12.073	4.580	0.001	0.755	13
Austria	1910	Enr., 1840 generation	13.163	5.760	0.000	0.827	13
Austria	1910	Enr., 1850 generation	14.491	7.260	0.000	0.889	13
Italy	1870	Total Enr., 1871	5.313	15.280	0.000	0.879	19
Italy	1870	Total Enr., 1891	5.243	17.210	0.000	0.928	19
Italy	1870	Total Enr., 1911	5.099	8.590	0.000	0.842	19
Italy	1870	P. + S. Enr., 1871	1.121	34.290	0.000	0.963	20
Italy	1870	P. + S. Enr., 1891	1.244	7.400	0.000	0.898	20
Italy	1870	P. + S. Enr., 1911	1.425	7.890	0.000	0.851	20
Italy	1870	Tertiary Enr., 1871	49.965	2.950	0.009	0.375	19
Italy	1870	Tertiary Enr., 1891	24.136	3.420	0.003	0.436	19
Italy	1870	Tertiary Enr., 1911	14.215	3.020	0.008	0.325	19
Italy	1900	Total Enr., 1871	6.518	11.820	0.000	0.827	19
Italy	1900	Total Enr., 1891	6.746	32.240	0.000	0.961	19
Italy	1900	Total Enr., 1911	6.731	13.100	0.000	0.918	19
Italy	1900	P. + S. Enr., 1871	1.144	9.600	0.000	0.898	20
Italy	1900	P. + S. Enr., 1891	1.362	23.760	0.000	0.965	20
Italy	1900	P. + S. Enr., 1911	1.541	14.640	0.000	0.892	20
Italy	1900	Tertiary Enr., 1871	64.878	2.950	0.009	0.396	19
Italy	1900	Tertiary Enr., 1891	31.366	3.730	0.002	0.461	19
Italy	1910	Tertiary Enr., 1911	18.846	3.270	0.005	0.358	19
Italy	1910	Total Enr., 1871	6.308	9.330	0.000	0.772	19
Italy	1910	Total Enr., 1891	6.734	24.060	0.000	0.954	19
Italy	1910	Total Enr., 1911	6.880	19.250	0.000	0.955	19
Italy	1910	P. + S. Enr., 1871	0.989	5.530	0.000	0.793	20
Italy	1910	P. + S. Enr., 1891	1.239	16.680	0.000	0.942	20
Italy	1910	P. + S. Enr., 1911	1.411	8.390	0.000	0.882	20
Italy	1910	Tertiary Enr., 1871	66.176	3.040	0.007	0.410	19
Italy	1910	Tertiary Enr., 1891	31.804	3.860	0.001	0.472	19
Italy	1910	Tertiary Enr., 1911	19.457	3.350	0.004	0.379	19

Notes: See text for sources of data. Underlying data are regions, grouped by countries, where "N" is the number of regions. "Ind. Variable" is the independent variable used in the regression. "Coeff." is the regression coefficient on that variable. 'Enr.' stands for enrolment; 'P+S' for primary plus secondary enrolment. N=number of regions. Estimated using OLS, with robust standard errors clustered on regions. Rates are in percentage form. The OLS results do not change in a Tobit estimation.

implying statistical significance at the five per cent level. The coefficients reflect the difference in timining of the variables. In Austria, the largest coefficient is that for the correlation between enrolment of the 1850s generation and literacy in 1870, by when that generation would have made it into the adult literacy sampling. Still, the coefficient on the 1830s generation is only 17 per cent smaller, so we cannot be sure that schooling was the only source of literacy. In Italy, we see similar effects. There is, for example, a particularly strong correlation (R^2 of 96 per cent) between literacy in 1910 and total school enrolment in 1911. While it is normal in the literature to think of school enrolment as an input and literacy as an output, these regression show us that the relationship can be interpreted in much broader terms. There were, according to this estimates, other sources of literacy, and the contemporaneous effects (of, say literacy and enrolment in the same year being highly correlated) indicate that *literate societies* were ones where schooling and literacy were both important. Quantitatively speaking, the measurement error in literacy rates (as a measure of human capital) is not great enough (or random) to distort its correlation with other, more formal measures of human capital, so it is not unreasonable to assume that literacy is correlated with the "true" and unobservable value of human capital. We can extend this idea in another check.

Taking Felice (2012) data series on regional literacy, primary and secondary enrolment, and tertiary enrolment, I extracted the first principal component of all the regional cross-sections put together to approximate something a "true" human capital component. This first component has an eigenvalue of 2.62 and explains 87 per cent of the variation across all three series. Its correlation coefficient with tertiary enrolment is 0.88; 0.95 with primary and secondary enrolment; and 0.97 with literacy, where all coefficients are significant at the one per cent level. While its strongest correlation is with literacy, we should not put too much weight on this exercise; it is here to support the results in table 6.5 and their interpretation.

6.3.3 Cross-country measurement of literacy

One final issue is the whether regional literacy rates are comparable between countries, given their different samples (for example, military recuits in Germany or youth in France) and different measurement (for example, signing a register in Britain or comprehension tests in Germany).

Given the results in table 6.2 establish strong correlations between literacy rates and industrial employment for a sample of French and German regions, we take this relationship as a baseline correlation and see how country effects differ. If differences in the way literacy was measured between the two countries produced systematically different variation in their regional literacy series, then we would expect those series to have different effects on the same target variable: industrial employment. If differences in measurement produced no systematic difference in variation, then they should have similar effects on industrial employment.

In this estimation, I replace country fixed effects with regional ones, so we are looking at variation within regions rather than countries; I standardise the data (0 mean; 1 standard deviation) and run regressions for both countries together and then separately. The results are in table 6.6.

Table 6.6: Cross-country comparison of literacy

	(1)	(2)	(3)
	L	L	L
$\frac{I}{LF}$	0.979***	0.767**	1.025***
	(0.240)	(0.224)	(0.290)
Country FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
R^2	0.101	0.331	0.094
N	324	255	69
Sample	FR+DE	FR	DE

*Notes: *** denotes statistical significance at one per cent; and ** at five per cent. Robust standard errors are clustered on regions, and are reported in brackets.* $\frac{I}{LF}$ *is a region's share of industrial employment* I *in its labour force* LF*; and* L *is literacy rate. FR denotes French regions, and DE German ones. The OLS results do not change in a Tobit estimation.*

Column (1) shows a standardised coefficient of 0.979, significant at one per cent, when the sample includes both countries. This is the baseline effect. Column (2) shows a coefficient that is 0.21 standard deviations smaller. The T-ratios, at 4.079 for column (1) and 3.424 for column (2) are also very similar. Column (3), the French regions sample, produces a coefficient of 1.025, which is 0.26 standard deviations larger than the German sample one. This may appear to be a more meaningful difference, but with a T-ratio of 3.534, the estimate is actually almost identical to that of the France sample. In practice, these differences are not large enough to produce any substantive differences between the two countries.

This admittedly brief check shows us that while there were differences between countries in the way literacy was measured, they do not produce any systematic biases. A possible explanation is that adult literacy is likely to have been correlated with youth literacy, and the literacy of military recruits is likely to have been correlated with that of the general adult population. Further, when it comes to the econometrics, it is possible to look at variation within countries - the level at which measurement methods varied. It should also be noted that when it comes to analysis of my full panel of data, a systematic bias can only occur if literacy measured in one country is correlated with the panel model's error term, and there is no reason that I am aware of to suspect this.

6.4 Empirical Strategy

In their empirical analyses, Redding and Schott (2003) and **?** regress contemporaneous levels of human capital on levels of market potential. They find a significant and robust positive relationship. Regressing my literacy rates in percentage form on my log market potential (MP^C), I get an R^2 of 0.226 and a coefficient, significant at one per cent, of 2.87 per cent ($= \ln[17.72]$). With region and year fixed effects, the R^2 rises to 0.241 and the coefficient, significant at five per cent, falls slightly to 2.10 per cent ($= \ln[7.85]$). There is clearly a strong contemporaneous correlation, given the correlation between stocks and flows of human capital. Market potential would affect, contemporaneously, flows, but I can only measure stocks. Redding and Schott (2003) and **?** do not address this issue, but in my OLS estimations I experiment with lags, showing the same effect.

A more important issue is the endogeneity between literacy and market potential. Data on GDP forms a large part of MP^C, and a number of studies have established a causal relationship running from human capital to income (Barro, 1991, 1997). To alleviate the risk of endogeneity, I first-difference the variables in a simple OLS setting, and then instrument market potential, as in *Chapter 5*, with distance to London ($Dist\_London$).

The OLS implementation takes the following form:

$$L_{it} - L_{it-1} = \alpha + \pi[\ln(MP^C_{it}) - \ln(MP^C_{it-1})] + \eta\mathbf{X} + \gamma_c + \theta_t + \mu_{it} \tag{6.2}$$

where L is the literacy rate in percentage form of region i at year t; MP^C is the constructed market potential measure; $\mathbf{X}$ is matrix of control variables that may affect changes in regional literacy rates; θ is a year fixed effects term to control for common shocks and upward trends in literacy growth rates; and γ is a fixed effects term for regions, c.

Under $\mathbf{X}$, I control for two other variables that proxy for supply-side determinants of literacy. First, regions' shares of sample GDP by year, GDP^{share}_{it} $(= 100 \times GDP_{it} / \sum_{n=1}^{j} GDP_t)$. Second, population density, PD. The rationale for the latter is that services like education are more efficiently supplied among dense populations (Olsson and Hansson, 2011). The rationale for the former is based on historical accounts of education as a way of improving or maintaining an economy's standing. For example, according to Sanderson (1999, p. 15-6), Baron Lyon Playfair, a Scottish chemist and Liberal MP for the Universities of Edinburgh and St. Andrew's, later Postmaster General under Gladstone, took the view that

> *...if Britain were to remain without tariffs, as he [Playfair] advocated, then it would be self defeating unless British industry were made as efficient as that of countries from which we imported. Education for economic efficiency was the only true protection.*

This view was motivated by the rise of industrial Germany, which had come to dominate the world chemicals and dye market by 1913. It was echoed by Sir Philip Magnus, an educational reformer, MP for the London University constituency, and director of the City and Guilds, who argued for '...state support for science and education to create industries from chemistry and electrical physics based on creatable and educable expertise...' Sanderson (1999, p. 15-6). Along with PD, this variable is intended to control for the supply-side story; or at least, the *motivations* for the state supply of human capital. If a region's share of output fell from one year to the next, or even if it were at a low level, policymakers would be incentivised to invest in the provision of education. We see this in the quotes of Baron Lyon Playfair and Sir Philip Magnus, but also in official reports like that of Mann (1846) or Stowe (1834), who toured Europe, examining national public education systems in the service of American states. Their reports make clear the internationally competitive nature of public education provision, even in the early nineteenth century. As Stowe (1834, p. 8) writes,

> *In short the world seems to be awake and combining in one simultaneous effort for the spread of education; and sad indeed will be the condition of that community which lags behind in this universal march.*

The quantitative analysis in Benavot and Riddle (1989) of a global sample of countries between 1870 and 1913 shows that this synchronised move towards public education was motivated by international economic and political competition. Were these forces feeding through to the regional level? I can do no better than quote from Cousin's report on Prussian public education for the Count of Montalivet, French Minister of the Interior. Full of praise for public education in Prussia, Cousin (1834, p. 21) writes that,

> *In short, the end of the entire organization of public instruction in Prussia is, to leave details to the local powers, and to reserve to the minister and his council the direction and general impulse given to the whole.*

We have seen in the preious chapter that *relative* market potential affects regional GDP shares, which may present multicolinearity here. The Pearson (Spearman) correlation between log market potential in its *absolute* form and GDP shares, however, at 0.030 (-0.005) is small and insignificant. The variable GDP_{share} is interesting because both a negative and positive signed coefficient would be consistent with its conceptual basis. If fears of falling *further* behind dominate public policy (the "follower" effect), then a negative coefficient should emerge. If fears of losing an initially high standing dominate (the "leader" effect), then a positive coefficient should emerge. It is difficult to form expectations based on the historiography, which provides evidence for both, so this is a purely empirical question.

Growth in literacy can arguably affect growth in GDP (and so market potential). One way of dealing with this source of potential bias is to use an IV strategy. I instrument market potential in the same way as the previous chapter, by using distance to London, $Dist\_London$, and control for the same variables in $\mathbf{X}$. The IV implementation is as follows:

$$L_{it} = \alpha + \pi \ln \widehat{MP_{it}^{C}} + \eta \mathbf{X} + \gamma_c + \theta_t + \mu_{it} \tag{6.3}$$

$$\ln MP_{it}^{C} = \alpha + \chi \ln Dist\_London_{it} + \eta \mathbf{X} + \gamma_c + \theta_t + \mu_{it} \tag{6.4}$$

where equation 6.4 is the first stage regression, in which market potential is regressed on distance to Lonon and the other independent variables, and equation 6.3 is the second stage, in which the predicted values of market potential are used to explain literacy rates. While an IV estimation gets us closer to a causal explanation, there a number of things that can be improved; more detailed data on human capital (wider coverage of school enrolment or age-specific literacy), for example. Still, a significant and positive coefficient on market potential here would lend further empirical support to the Redding and Schott (2003) theory, and provide an explanation that is consistent with regionally clustered and unequal literacy rates, as seen in figure 6.1 and table 6.3.

6.4.1 Results

Table 6.7 shows the OLS results. Column (1) starts off the estimations, where all variables are specified as contemporaneous changes, controlling for region and year fixed effects. The market potential effect is large and highly significant. The coefficient implies a three per cent (= ln[21.886]) increase in regional literacy rates for every one per cent increase in market potential. In standardised form, this a 0.685 standard deviation effect on literacy - or 14.72 percentage points of literacy. In 1870, three Italian regions and one Spanish region had lower literacy rates than this: Lucania, Calabria, Sardinia, and the Canary Islands. GDP^{share} (in percentage form) and $\ln PD$ are insignificant, supporting the demand-side explanation for literacy. Given its statistical insignificance, we should not place too much emphasis on GDP^{share} , but that it is positive supports the “leader” effect. That is, regions that grew their share of total output also improved their literacy rates.

In column (2), I use a lag of market potential ($\ln MP^{C}_{t-1}$), which is significant at the five per cent level. The coefficient implies a much smaller 5.7 per cent effect.This reduction in magnitude is likely to be a result of the large gaps between each benchmark year. The other independent variables remain insignificant.

Table 6.7: OLS literacy and market potential

	(1)	(2)	(3)	(4)	(5)
$\Delta \ln MP^C$	21.886***		21.886***	1.893*	7.415**
	(1.501)		(1.501)	(1.044)	(3.717)
ΔGDP^{share}	1042.5	-468.4**	1042.6	-106.59	821.09
	(710.26)	(351.184)	(810.26)	(765.183)	(664.67)
$\Delta \ln PD$	-10.943	3.339	-10.943	-2.559	-1.243
	(9.645)	(2.883)	(9.645)	(2.076)	(11.437)
$\Delta \ln MP_{t-1}$		5.699**			
		(1.641)			
Year F.E.	Yes	Yes	Yes	Yes	Yes
Region F.E.	Yes	No	Yes	Yes	Yes
λ			0.737***		
$\widehat{R}^2$	0.562	0.084	0.542	0.081	0.083
N	398	199	398	143	56

*Notes: Dependent variable is the percentage point change in literacy rates, ΔL. *** denotes statistical significance at one per cent; ** at five per cent; and * at 10 per cent. Robust standard errors are clustered on countries, and are reported in brackets. MP^C is constructed market potential; PD is population density; GDP_{share} is a region's cross-sectional GDP share in sample GDP in percentage form; and . λ is the coefficient on the spatially-weighted error term in the spatial error model. The dependent variables in columns 4 and 5 are, respectively, the top two and bottom two quartiles of regional literacy levels. The OLS results do not change in a Tobit estimation.*

In the remaining columns, I deal with three potential sources of error. In column (3), I use a spatial error model to correct for spatial autocorrelation in the model's error term, given the high spatial correlation of literacy seen in table 6.3. This might bias the standard errors reported in column (1). A spatial error model applies a spatial weights matrix to the model's error term. The matrix I use is the same as that in Chapter 3 and underlying table 6.3. It is a squared inverse distance spatial matrix, where Euclidean distance is measured between all regional nodes, but only geographically close regions are meaningfully associated with one another. The coefficient on this saptially-weighted error term, λ, is positive and significant, indicating the presence of spatial autocorrelation between the errors. As the spatial error model treats spatial effects as a spatial nuissance to be corrected, this is only of interest if it affects the model's other coefficients. This it does not do: the magnitude and significance of all coefficients are the same as previous estimations.

In columns (4) and (5), I run the same specification as in column (1), only dividing the samples into the top two and bottom two quartiles of literacy levels. Column (4) is the top two, and column (5) is the bottom two quartile sub-sample. Most regions enjoyed high and increasing literacy rates throughout the period, so the distribution of literacy is skewed to the right. This skew may result in coefficients that are biased to high literacy regions. It is also interesting to see how the market potential effect varied according to literacy levels. The coefficients are significant across both subsamples, but less so for the high literacy regions, as we would expect.

Table 6.8 contains the results of the IV implementation 6.4 and 6.4. Column (1) shows the first-stage estimation, that is, a regression of market potential on the cost to London instrument ($\ln Cost\_London$), population density and GDP^{share}, while controlling for region and year fixed effects. We see that the instrument continues to have in this specification a highly significant negative effect on market potential, as in previous chapters. This results is backed up by the Angrist-Pischke F-statistic on excluding the instrument, which is far above the 12-points rule of thumb mark. The Anderson LM-statistic for weak identification, at 67.09, is also re-assuring.

The second stage results in column (2) stand alongside the OLS equivalent in column (3). The coefficients on market potential are both highly significant. The IV estimate (a 2.9 per cent effect) is similar to the OLS estimate (a 3.11 per cent effect). The IV results also support the OLS results in table 6.7. They show us that the spatial distribution of human capital was not random, and that it was driven by demand-side variables, namely market potential. A market potential explanation deals with the spatial structure of the distribution, and the paradox of

Table 6.8: OLS literacy and market potential

	(1) IV-1	(2) IV-2	(3) OLS	
$\ln Cost\_London$	-0.091***			
	(0.022)			
$\Delta \ln MP^C$		19.346**	22.553***	
		(1.152)	(1.091)	
ΔGDP^{share}	-125.82	-522.52	-59.21	
	(123.46)	(328.96)	(327.47)	
$\Delta \ln PD$	2.655***	-11.983**		
	(0.250)	(4.254)	(9.645)	(4.282)
Year F.E.	Yes	Yes	Yes	
Region F.E.	Yes	Yes	Yes	
$\widehat{R}^2$	0.552	0.706	0.718	
N	597	597	597	

*Notes: *** denotes statistical significance at one per cent; ** at five per cent; and * at 10 per cent. Robust standard errors are clustered on countries, and are reported in brackets. $IV-2$ is the second stage IV estimation; and $IV-1$ is the first stage. L is literacy rate in percentage form; MP^C is constructed market potential; PD is population density; GDP_{share} is a region's cross-sectional GDP share in sample GDP in percentage form; and $\ln Cost\_London$ is distance to London. Angrist-Pischke multivariate F-test is on the excluded IVs for MP first stage. Cragg-Donald Wald F-statistic for weak identification is conducted jointly on the excluded IVs. Anderson canon. corr. LM statistic is for underidentification. The OLS results in column 3 do not change in a Tobit estimation. Mark Schaffer at Heriot-Watt University supplied the Stata code for this estimation* national laws enforcing compulsory education along with the presence of high regional inequalities in literacy rates. To the extent that industrialisation depended on human capital, and authors have shown this to be the case, these results have implications for the Continent's catch-up industrialisation (Becker et al., 2011b; Leunig et al., 2011). According to this line of argument, market potential did not only affect industrialisation through the trade cost savings channel, but also through its effects on human capital.

6.5 Discussion

In this chapter, I have shown that changes in human capital, measured by literacy rates, were caused by changes in market potential, as people sought to capitalise on higher returns to education. To the extent that technological and industrial progress depended on a region's human capital, the implications of these results are that market potential shaped a region's ability to industrialise. Leunig et al. (2011, p. 434) conclude their analysis of London apprenticeships and human capital formation by writing, 'where barriers exist that prevent people exploiting their talents, societies as well as individuals will suffer.' This much we know, but the debate centres on *what* these barriers were. This chapter's results show that trade costs were an important barrier: they not only determined the location of economic activity directly, but also indirectly, through its demands for human capital.

For Pollard (1981, p. 249), 'adverse social milieu' (typified, partly, by low literacy rates) were defining characteristics - rather than a cause or effect - of underdeveloped regions. In his words, 'illiteracy figures of countries like Russia or Italy were extremely high even at the end of the [nineteenth] century, and there was '*Gefaelle*' [gradient] of literacy across Europe as there was of national incomes.' This gradient can be explained by the cultural characteristics Pollard alludes to,[1] but such a cultural explanation should not be the first port of call: growth in literacy was largely a function of, quite simply, expanding markets.

[1]'Italian entrepreneurs were unwilling to take risks...Spanish society was too rigid and there was social resistance to change...' (Pollard, 1981, p. 249).

Similarly, Berend (2013, p. 348) described the core's incentives as 'the enticing spectacle of Western industrial transformation,' but did not explicitly connect this to education or literacy. Like Pollard (1981), he blames illiteracy on social characteristics and government inaction, but does not explain *why* governments might have been inactive or that it might have been people who were unresponsive. The gradient of illiteracy can be explained through rational human capital investment decisions in the periphery. It is not 'market consciousness and capitalist spirit' that people had to wait for, but market potential (Pollard, 1981, p, 249). Investments in education were costly, especially in countries with inefficient public service provision, and it would make no sense in making them unless the expected returns were high.

Market potential thus shaped Europe's regional income structure through the trade cost and market potential channel, as well as through the incentives it provided for the accumulation of human capital. This second effect also goes some way in explaining why some regions, despite having high levels of market potential, failed to exploit it. That failure may have been a result of blockages in the market potential-human capital channel; perhaps changes in market potential were not high enough to warrant human capital investments. Low levels of human capital, in turn, hindered regions' abilities to develop. More broadly, this chapter gives us another way of thinking about the relationship between factor endowments and market potential in industrial location, in contrast to pitting the variables against each other.

Chapter 7

Conclusion

The literature on late-nineteenth century European regional development has been empirically and theoretically incomplete. Starting with Pollard (1973), historians picked Europe apart into regions, and explained industrialisation across the Continent in terms of "diffusion", as though the only barrier to industrial location was time. Many years later, economists formalised old geographical ideas like market potential, providing us with a theoretical and empirical framework to understand regional development: NEG (Harris, 1954; Krugman, 1991). Some years later, economic historians applied these ideas to individual European countries, finding that per capita incomes were higher in regions with better market potential and that as trade costs dropped, production concentrated spatially (Combes et al., 2011; Crafts, 2005a; Roses, 2003). Building on this progress, I extended the same theoretical framework over Europe during a period of fast-falling trade costs, in search of an explanation of its spatial income pattern.

At the start of the period, trade costs were high, and so economic activity - long concentrated in Britain - was spread out more or less evenly across Europe. By the end of the period, when trade costs dropped, economic activity concentrated in the northwest of Europe, being better placed to serve the larger markets there and also able to serve more distant markets at lower trade costs. The point at which trade costs drop so much that economic activity is free to locate anywhere - as it would be equally cheap to serve any market - had not been reached by this point. The resulting pattern was the consolidated "Golden Triangle." This pattern persists until today. This, however, is only the broader picture. Foreign markets, for example, mattered more than domestic ones, and increasing market potential in the core relative to the periphery resulted in an absolute and relative decline in peripheral income levels, as predicted by NEG. The explanation can thus explain the de-industrialisation of Europe's periphery, like Languedoc going from a textile powerhouse to a glorified vineyard.

Four analytical chapters, on the distribution of regional income, the traditional explanations of this distribution, on the role of market potential in explaining it, and the role of market potential in other determinants of income, all point towards accepting the hypothesis that NEG was an important reason for Europe's regional income structure. More specifically, I find that: (1) per capita income inequalities were higher within than between countries, a fact that does not sit well with traditional institutional or Solowian explanations, and that regional income clustered with a northwestern European bias; (2) traditional explanations based on coal deposits work only when there are no controls for regional institutions, and that regional institutions work only when there are no controls for market potential; (3) market potential has, across a number of different specifications and calculations, a significant positive effect on per capita income levels but relative declines in market potential have a negative effect on peripheral regions' income; and (4) market potential also has a positive effect on regional human capital levels. Together, these relationships provide reasons as to why some regions developed while others did not, and more importantly, why regional income formed clear spatial clusters.

This dissertation modestly revises the way in which we look at the Continent's economic development in the 'Second Industrial Revolution' (Mokyr, 1990). Most histories of the topic were written either as disjointed national narratives, as with Milward and Saul (1973), or as pan-European histories that glide over the spatial patterns that I stress, taking them as self-evident (Berend, 2013; Pollard, 1981). Even when I consider regional *de facto* institutions, I find the institutional explanations washes out after the inclusion of more "basic" ideas, like market potential. This is perhaps unsurprising, given institutional explanations are a-spatial in nature and there were clear spatial income clusters, but still contrasts with the influential work of, for example, Acemoglu et al. (2001) and North and Weingast (1989). Likewise, the coal-based explanations, persistent from Pounds and Parker (1957) and Parker (1961) to Wrigley (2010), do not to hold up well. The growing economic history literature that applies NEG methods is also, again, limited to national cases (Crafts, 2005a; Henning et al., 2011; Roses, 2003; Schulze, 2007b). It provides us with useful historical evidence, and shows how economic geographical patterns play out within the border of a nation, but misses the patterns that are only visible when looking at a broader map of Europe's regions. When this is done, we find that European regions interacted with one another to a greater extent than we thought, where in particular foreign markets are shown to be more important to domestic ones, which contrasts to the 'Age of Neo-Mercantilist' thinking that has framed the debate for some time (Pollard, 1981).

This dissertation also provides some novelty in its empirics. Regional institutions, for example, are often ignored, as they cannot be directly observed and are controversial to measure. While the regional institutions I propose here are a distance from the ideal, they do give us some new ways of thinking about and approaching the issue. Ignoring the existence of *de facto* institutions, as in Redding and Strum (2008), is not a promising route. There is also some novelty in the estimation of regional income for France and Germany, where I provide a method that can be applied to other countries (Sweden and Austria-Hungary at least, as in the robustness checks) when data are scarce.

Indeed, empirical researchers working on European economic history face considerable obstacles in their work in terms of data availability and especially comparability.[1] These problems mean that testing relationships drawn from economics is difficult, and so we find a historiography of Europe's economy that is broad in scope and sweep, but perhaps not as analytical and hypothesis-driven as an economic historian might like it to be (Berend, 2013). This dissertation shows that it is possible to do *European* economic history covering a number of regions from different countries over different years, with a single hypothesis-driven empirical strategy, and with little pomp and little funding.

A causal explanation of Europe's regional income differentials required a variety of different techniques. Drawing inspiration from empirical economics and quantitative economic history, the search for appropriate instrumental variables,

[1] Perhaps not as many obstacles as those working on African economic history (Jerven, 2013)

for example, took the dissertation deeper into Europe's history. The Habsburg Division of 1521, that I used to instrument regional institutions, threw new light on late-nineteenth century institutional differences that economic historians often accept as a product of their own time or less-credibley attribute to legacies like the Roman Empire (Tortella, 1994). The emphasis on spatiality required the construction of a GIS of historical roads, railways, waterways, ports, shipping lines, and ton-mile and tariff costs. It adds to a growing literature in economic history that exploits the benefits of attributing spatial information to historical data (Atack, 2013). As always, there is a feedback between the collection and analysis of quantitative data, and the historical record. Both are used to inform one another, and neither can work alone.

An essay in prediction

There is one final task of the economic historian: the essay in prediction. Advice based on historical evidence for those in the present is normative by definition. A dissertation should not be normative, and so no concrete prescriptions are offered here. Still, it is hoped that the analysis of the role of market potential in late-nineteenth century European regional income differentials can provide a historical antidote to presentist accounts of European regional inequalities as a particularly twenty-first century issue. Solow (1970, p. 103) famously wrote that most explanations of relative economic performance end in 'a blaze of amateur sociology.' With this in mind, what follows is more indulgence than policy advice; a speculative end to pages of academic restraint.

Starting from the observation that Europe's map of regional income has not changed much since 1870, I explained its late-nineteenth century map through a simple NEG framework. These NEG frameworks set up a deep structure to economies, shaped by geography, initial conditions, and trade costs. Multiple equilibria can found in that differences and random events in initial conditions - historical accident - can lead to various results. I could not go back further then 1870, but we saw that Europe's core-periphery pattern really consolidated by 1910; the spatial distribution of income was relatively much more even in 1870.

Is "geography" destiny? If initial conditions are set long ago, countries are locked onto a path, incapable of breaking free until another shock resets the conditions. Irwin (2006, p. 42), on thinking about this, is reminded of the Englishman driving through rural Ireland, who 'stopped by the side of the road to ask a farmer, "What's the best way to get to Dublin?" To which the farmer replied, "Well, if I were you, I wouldn't start from here."

While some regions may find themselves at the periphery of an industrial boom, there are things that can be done to overcome their "bad geography" - or unfortunate initial conditions. Australia is a good example of an economy overcoming its peripherality through good transport, trade policies, and communication systems. Limao and Venables (2001) show that landlocked African countries have not simply been penalised by exogenous distance from markets, but by poor domestic transport infrastructure. Donaldson (2012) showed that large-scale railway construction in colonial India decreased trade costs and interregional price gaps, increased interregional and international trade, and increased real income levels. That the constraints of geography can be alleviated through transport and trade policy is clear. Indeed, the 2012 to 2015 World Bank Group budget allocated 43 per cent of total assistance to transport infrastructure - it is now 'the single largest business of the group' making the Group 'the largest multilateral development financier in infrastructure' (World Bank Group, 2014). Similarly, the OECD (2013a), an organisation of brain but not brawn, advised that 'enhancing the capacity of infrastructure primarily transport systems is a priority in some member countries.'

Crafts and Wolf (2013a) show us that there are other policy options. Looking at nineteenth-century Lancashire, they ask how this small region of a small country produced most of the world's cotton textiles, when the technology was universally known, and it was a high-wage, adamantly free-trade economy, in a world of fast-dropping trade costs and cheap Asian competition. Besides the large private investments in the region's transport system, which were encouraged by parliamentary legislation, there were two other reasons for Lancashire's success. First, the region's urban and industrial development was not constrained by land-use planning regulations. Second, while the provision of public goods like healthcare increased the supply price of labour, they reduced not only the health risks of working in textile towns, making them more attractive places to settle and work. In short, Crafts and Wolf (2013a) conclude that a policy stance that is more supportive to agglomeration is the key to success.

Is this what European policymakers are implementing? The ongoing Eurozone crisis is increasingly being seen as one of fundamental differences in national competitiveness, although the debate rages on (Draghi, 2012; Trichet, 2011; Wyplosz, 2013). Greece's troubles, for example, are framed as result of its high unit labour costs, built up during the heady capital inflows of the pre-crisis boom years. Since Eurozone economies cannot devalue their currency, they are advised to internally devalue 'through wrenching labour-market adjustments' (Blanchard et al., 2013). While these reforms have had an effect on the Eurozone's recovery, progress on the fronts highlighted by Crafts and Wolf (2013a) - land-use regulation and public goods provision - has been mixed.

The OECD (2013a) report on economic policy reform notes that land reform, particularly housing policies, can boost labour productivity and utilisation. Strict regulation blocks investment and supply of residential and commercial units, limits labour mobility, preventing the re-allocation of labour across sectors and regions. Unlike nineteenth-century Lancashire, a number of European countries are bound up in strict regulation. Since 2010, the OECD (2010) has been advising a revision of land-use regulation in Denmark, Israel, Luxembourg, the Netherlands, Poland, Slovak Republic, Sweden and the United Kingdom. They maintain these revisions for beyond 2013 (OECD, 2013a). While the OECD (2013a) report notes that major healthcare reforms are needed in some member states, like Russia, European countries have been, in a climate of fiscal contraction, making headway. For example, Greece introduced a new output-based hospital funding system; Italy is reforming its uncompetitive pharmaceutical distribution market, as is Spain; and Portugal is investing in performance management systems for hospitals (OECD, 2013b). These reforms are promising, and are in line with the public-good provision policy implications of Crafts and Wolf (2013a). There is no reason why Europe's periphery should remain the periphery for another 140 years. Greece was the centre of the ancient world, Italy of the medieval, and Spain of the early-modern. Where to next?

Appendix A

Regional income

A.1 Comparison of French and German Estimates to New Estimates

In on-going research with Joan Roses, Nikolaus Wolf estimated the NUTS-2-level GDPs of France in 1990 borders for decadal benchmark years from 1900 to 2000, using the Geary and Stark (2002) method. These estimates were unavailable to me at the initial stages of my research. The provided estimates are expressed as regional shares of national nominal GDP. I use them here as a cross-check on my own estimates derived in *Chapter 3*. I aggregated all 85 French *departments* into their NUTS-2 groupings, expressing their nominal GDPs as shares of the national GDP for 1900 and 1910, which are the only overlapping years between the datasets.

The summary statistics are in table **??**. They show reassuring similarity. The means are the same, and a t-test accepts the null hypothesis of equal means between the two series with a t-statistic of 0.018. The standard deviations, as well as the minimum and maximum values, are also similar.

The Pearson correlation between my estimates and the Wolf-Roses estimates is 0.93, and the Spearman rank correlation is 0.80. Both correlation coefficients

Table A.1: France NUTS-2: Own Estimates vs Wolf-Roses.

	N	Mean	Std. Dev.	Min.	Max.
Own	40	0.05	0.03	0.01	0.16
$Wolf - Roses$	40	0.05	0.04	0.02	0.22

are significant at the one per cent level.

As a parametric check, I regressed my estimates on the Wolf-Roses estimates. For a sample of 40 observations, the regression yields an R^2 of 0.87, an F-statistic of 249.08, and a root mean-square error of 0.011. The coefficient of 0.653 is significant at one per cent, and the constant term of 0.017 is also significant at one per cent.

Roses-Wolf also provided estimates of Germany's Regierungsbezirk-level GDPs from 1900 to 2000, estimated using the Geary and Stark (2002) method. It is not possible to run the same kind of comparison as for France, as their German estimates are for 1990 borders, and are again expressed in shares.

However, they also provided an estimate for Alsace-Lorraine's regional GDP level in dollars for 1907. This is helpful as a look back to *Chapter 3* shows I had no income estimates for this region, and so resorted to using an average instead.

My estimate for the region's 1907 GDP is 62 per cent of the Wolf-Roses estimate. This is a large discrepancy, but does it affect the estimation procedure? This is what matters in the end, as I am scaling to known national GDP numbers. In table A.2 I compare the cross-section regression results with my estimate of Alsace-Lorraine and that of Wolf-Roses.

While the size of the F-statistic has dropped to 246.3 from 672.6, it remains large and highly significant. The other noticeable change is that the $\frac{S}{LF}$ coefficient is significant at the 10 per cent level rather than the five per cent level. The sizes of the coefficients have changed slightly, but none to a statistically significant degree. Z-tests for the equality of coefficients accept the null of no significant differences for all coefficients, with a z-score of 0.02 for P; -0.02 for $\frac{I}{LF}$; $\frac{S}{LF}$ for -0.06; and -0.103 for the $Constant$.

Given that the replacement of Alsace-Lorraine did not change the regression results, it is unsurprising that the predicted GDP estimates (after a linear transformation) are highly correlated. The Pearson correlation between the two

Table A.2: Regional GDP estimation model: with alternative Alsace-Lorraine

	Own	Wolf-Roses
P	1.038***	1.034***
	(0.026)	(0.044)
$\frac{I}{LF}$	0.527***	0.534***
	((0.057)	(0.094)
$\frac{S}{LF}$	0.156**	0.179*
	(0.067)	(0.101)
Constant	6.662***	6.766***
	(0.387)	(0.638)
$\hat{R}^2$	0.98	0.97
F-Stat.	672.6	246.3
N	23	23

*Notes: Own column contains the original estimates; Wolf-Roses column contains the estimates with the new Alsace-Lorraine number. Statistical significance: *** 1 per cent, ** 5 per cent. Standard errors in brackets.*

predicted GDP series is 0.99, and the Spearman correlation is also 0.99. Regressing the Wolf-Roses predicted series on the original predicted series, with a sample size of 92, I get an R^2 of 0.99, a coefficient of 0.996 that is significant at one per cent, a constant of 0.108 that is also significant at one per cent, and a root mean-square error of 0.008.

As far as these tests go, the income estimation method presented here produces similar results to the Geary and Stark (2002) method.

A.2 Re-Aggregation of French Regions

I aggregated all 85 French *departments* into their 19 NUTS-2 groupings, and re-ran the empirical analysis in *Chapter 3* to see whether this re-aggregation affects results.

First, as can be seen in A.3, while the Theil index for the new aggregation shows lower absolute levels of within inequality in GDP per capita (T_w), which is what we would expect from broader aggregations, the gap between the within and between components (T_b), and the trend in the gap, remains comparable from the new to the old sample.

Second, as can be seen in table A.6, the aggregation has not changed Frances

Table A.3: France NUTS-2: Theil Indices.

Sample	1870	1900	1910
New			
T_w	0.03148	0.0338	0.03346
T_b	0.02662	0.02802	0.03102
Old			
T_w	0.0375	0.0406	0.0412
T_b	0.0330	0.0278	0.0298

Notes: All per capita income data are in 1990 Geary-Khamis dollars. T_w refers to inequality in regional GDP per capita within countries. T_b refers to inequalities between countries.

country-year coefficients of variation in a meaningful way. In the unweighted results, the new sample shows lower absolute levels of dispersion, but similar trend growth from one cross-section to the next (Pearson correlation of 0.96). The same goes for the population-weighted results, which are yet more comparable (Pearson correlation of 0.99).

Table A.4: France NUTS-2: Coefficients of Variation.

Sample	1870	1900	1910
Unweighted			
New	0.191	0.233	0.265
Old	0.302	0.314	0.344
Weighted			
New	0.480	0.601	0.636
Old	0.961	1.243	1.304

Notes: Panel A contains the unweighted data; Panel B contains the population-weighted data. All per capita income data are in 1990 Geary-Khamis dollars. Coefficient of variation is the standard deviation over the mean for each country at each year. A value of zero indicates perfect equality.

Third, as can be seen in figure A.2, the new-sample boxplot shows no change beyond London appearing as an outlier in 1870 and 1910. This is the result of Frances richer regions forming part of larger, poorer regions, smoothing out the distribution, and making London more exceptional than it originally was.

Fourth, as can be seen in table **??**, the β-convergence regression results are close to identical for the new and old population-weighted samples, showing divergence in both cases, and showing highly significant convergence for the unweighted samples. Here again the point estimates are similar.

Fifth, as can be seen figure **??**, the Kernel density distribution of log GDP

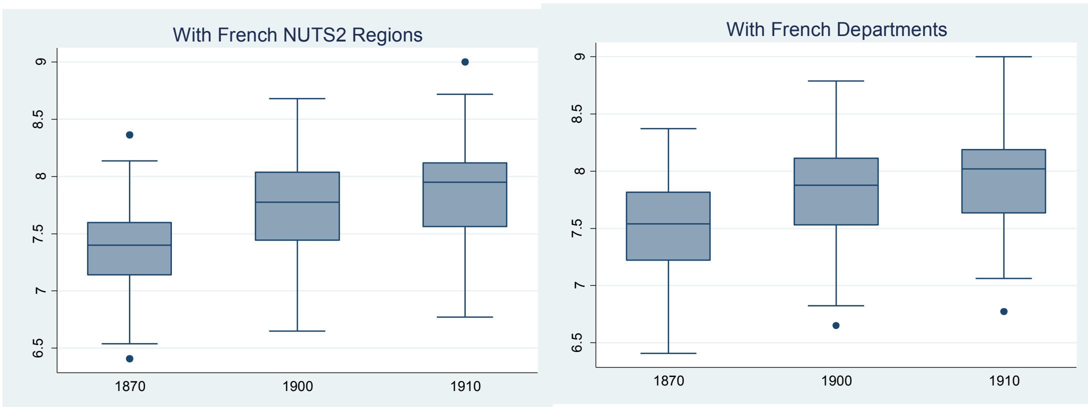

Figure A.1: Box plot of regional income with French NUTS2 regions

Notes: Per capita income in 1990 Geary-Khamis dollars: 1870 (left), 1900 (middle), and 1910 (right). See text for underlying sources.

Table A.5: France NUTS-2: β-convergence.

	Weighted		Unweighted	
	New	Old	New	Old
β	0.232	0.216	-0.012	-0.014
T-stat.	11.01	6.66	-4.13	-16.55
R^2	0.349	0.371	0.701	0.582
N	228	398	228	398

Notes: Dependent variable is first difference in GDP per capita. Independent variable is initial level of GDP per capita. Weights are population shares. All per capita income data are in 1990 Geary-Khamis dollars.

per capita levels follows the same pattern over the period: a normal distribution in 1870, and the emergence of a bi-modal distribution by 1910. This is reflected in the similarity of the Kernel bandwidths for the new and old samples: 0.1145 versus 0.1224.

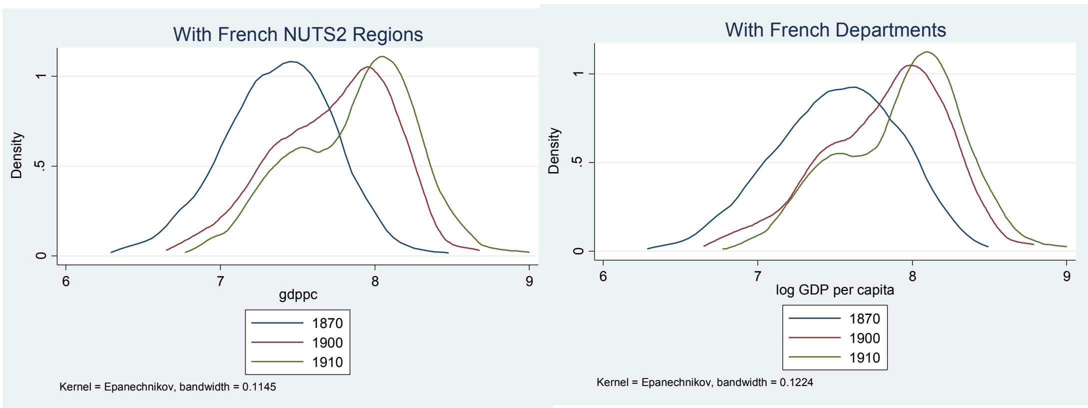

Figure A.2: France NUTS-2: Kernel denstiy distributions.

Notes: Per capita income in 1990 Geary-Khamis dollars.See text for underlying data. Estimated using Epanechnikov kernels.

Sixth, as can be seen in table **??**, I have re-calculated Morans I statistics for the new and old samples. These statistics were used to plot the original map, so any difference in the numbers presented here would result in a different map. As can be seen, however, the Morans I statistics are very similar in size and statistical significance. This implies that the re-aggregation of French regions does not change the broad clustering of income: a look back at figure 3.6 shows that rich cluster in the northwest of France maps perfectly onto Nord-Pas de Calais.

Table A.6: France NUTS-2: Moran's I statistic.

	Statistic	Z-score	p-value
New	0.125	35.01	0.000
Old	0.112	43.86	0.000

Notes: All per capita income data are in 1990 Geary-Khamis dollars.

Appendix B

Literacy and institutions

B.1 Distance to coal deposit centroid versus coalfield border

Chapter four showed a strong correlation between the distance to coal deposit centroid and distance to coalfield border measures. Here I test whether any of the "coal variables" has an effect on per capita income, which is the point of the exercise.

The results in table B.1 show the OLS correlations between log GDP per capita and the various coal measures, controlling for year fixed effects and regional institutions. I cannot control for regional fixed effects as the coal distance measures are time-invariant. Backing up the results in *Chapter 4*, none of the coal measures are correlated with per capita income even in these basic specifications. As argued in *Chapter 4*, the improvements in transport technology that this period saw rendered distance to coal irrelevant.

Table B.1: Three measures of coal and income OLS estimates

	(1)	(2)	(3)
	$\ln Y$	$\ln Y$	$\ln Y$
$\ln Coal\_Transport$	-0.029		
	(0.028)		
$INST$	0.203***	0.524***	0.427***
	(0.046)	(0.080)	(0.082)
$\ln Coal\_Distance$		-0.017	
		(0.017)	
$\ln Coal\_Field$			-0.041
			(0.032)
Year F.E.	Yes	Yes	Yes
Region F.E.	No	No	No
$\hat{R}^2$	0.868	0.327	0.350
N	597	597	597

*Notes: *** denotes significance at one per cent. Robust standard errors clustered on regions, and reported in brackets. Y is GDP per capita; $INST$ is a measure of regional institutional efficiency; $Coal\_Transport$ is transport cost to nearest coal deposit; $Coal\_Distance$ is the distance to coal deposit centroid measure; and $Coal\_Field$ is the distance to coalfield border measure.*

B.2 Testing for Ruggedness

Another first-nature geographical variable that may affect regional income differentials is terrain ruggedness. Research has shown more rugged areas are less productive, due higher transport costs are less suitable land for production (Combes, 2010).

The baseline data to calculate this indicator was the Digital Elevation Model of all Europe, obtained from GISCO (European Commission/Eurostat). It covers 39 member and cooperating countries, with a 3D raster dataset with elevation.[1] The dataset used was made available as tiles (5x5 degree), and georeferenced in ETRS89 coordinates system. Using this source, I calculated the Riley et al. (1999) ruggedness index to quantify topographic heterogeneity.

More specifically, I implemented a GIS model to do the following operations. First, extract the elevation information of every raster cell, obtaining a grid of points that covers all the studied countries. Second, create meshes of 3x3 elements.Third, calculate the elevation's difference between neighbour nodes follow-

[1]Available at: http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/digital-elevation-model

ing the index formulation in Riley et al. (1999). Finally, calculate the average value between all values within each region.

Regressing log GDP per capita on this index, as can be seen in the table B.2, yields a significant negative coefficient, but this effect washes out with the cost to coal deposit and regional institution variables, even before introducing region or/and year fixed effects.

Table B.2: Ruggedness, Coal and Institutions OLS estimates

	(1) $\ln Y$	(2) $\ln Y$	(3) $\ln Y$
Rugg.	-0.0003**	-0.0002	
	(0.000)	(0.000)	
$\ln Coal\_Transport$		-0.187***	-0.189
		(0.034)	(0.034)
$INST$		0.818***	0.817***
		(0.053)	(0.053)
Constant	7.733***	7.785***	7.769**
	(0.027)	(0.157)	(0.154)
Year F.E.	No	No	No
Region F.E.	No	No	No
$\hat{R}^2$	0.014	0.392	0.393
N	597	597	597

*Notes: *** denotes significance at one per cent and ** at five per cent. Robust standard errors clustered on regions, and reported in brackets. Y is GDP per capita; $INST$ is a measure of regional institutional efficiency; $Coal\_Transport$ is transport cost to nearest coal deposit; and Rugg. is the ruggedness index.*

B.3 Literacy data sources

Literacy data come from a number of primary and secondary sources. It is well-known that historical literacy data are difficult to work with. What I have done here is use primary official sources wherever possible; use general (rather than, say, military) rates when they are available; and use standard, very straightforward and transparent extrapolation for missing regions or years. The goal was not to arrive at a precise figure, but at an approximate level, which allows us to understand the way resources are distributed. There is no reason to think that a persistent bias, that for example literacy measured in one particular country is correlated with the empirical model's error term, exists in this dataset. *Chapter 6* provides a series of robustness checks of these literacy data. In biref, I find that these data are correlated with school enrolment rates from the literature, and with alternative primary sources of literacy data.

Austria-Hungary. Regional indices (ratios relative to national rate) are from Good (1984, p. 156). I converted these into real rates, using the national rates in Cipolla (1969, pp. 127, 118).

Britain. Rates for England and Wales are from volumes 33, 63, and 73 of Registrar-General of Births, Deaths, and Marriages in England and Wales (1910). These figures refer to the rate of people who are unable to sign their own marriage register. Rates for Ireland are from Irish Census (1911) and Irish Census (1910). Scottish rates come from Cipolla (1969, p. 127).The Scottish figure for 1910 is an average for Wales, England, and Ireland.

France. Rates are from the census books. For 1870, Statistique de la France (1872). For 1886, Statistique de la France (1886). For 1901, Statistique de la France (1901). For 1911, Statistique de la France (1911).

Germany. Rates for Prussian regions are from Cipolla (1969, p. 91). I then took the rates of illiterate military recruits from the yearbooks Deutsch Statistischen Bureau (1882) and Deutsch Statistischen Bureau (1910), and used these values to linearly extrapolate the values for 1900 and 1910, as well as the non-Prussian regions in 1870. It is worthwhile pointing out that Cipolla (1969) also used illiterate military recuits as his measure.

Italy. Rates come from Felice (2012). The rates for 1910 are from Cipolla (1969, p. 19).

Spain. Nunez (1990), provides provincial literacy rates, split by gender. I took the average of this split to indicate overall literacy. Some provinces are missing from my list due to differences in aggregation. Missing regional rates are proxied with those of neighbours.

Sweden. Regional rates for 1930 are from Statistiska Centralbyrån (1935) These are the earliest we know of. To extrapolate back in time, I used the annual growth rate of 0.25 per cent presented in Sandberry and Steckel (1997).

Table B.3 shows country's mean level and standard deviation of literacy by year. The statistics fit expectations of high German and Swedish literacy, as

Table B.3: Summary of literacy rates

	1870	1900	1910
Austria-Hungary	63.36	75.16	82.55
	(16.685)	(10.544)	(4.563)
Germany	89.77	99.26	99.39
	(9.175)	(1.129)	(1.261)
Spain	28.94	49.09	56.18
	(10.641)	(16.333)	(17.657)
France	65.29	95.79	97.31
	(15.815)	(3.511)	(2.016)
Britain	57.21	93.58	97.12
	(15.133)	(2.966)	(2.045)
Italy	26.81	47.37	56.83
	(14.757)	(19.859)	(19.093)
Sweden	85.17	92.35	94.74
	(0.037)	(0.040)	(0.041)

Notes: Literacy rates are in percentage form. Standard deviations are in brackets. See this appendix for details on sources. AH is Austria-Hungary; DE is Germany; ES is Spain; FR is France; GB is Britain; IT is Italy; and SE is Sweden.

examined in Becker and Woessmann (2009) and Sandberg (1979) and Ljunberg and Nilsson (2009), and of low initial levels of British literacy, but fast catch-up growth, as discussed in Vincent (2000). The table also shows that regional variation in literacy, as measured by the standard deviation, was highest were literacy levels were lowest: Italy, Spain, and Austria-Hungary. Regional variation in literacy is analysed in *Chapter 6*.

B.4 Iron ore and regional income

In *Chapter 4*, in the discussion around figure 4.9 I highlighted the correlation between regional per capita income levels and the distance to iron ore deposits. This correlation, in simple OLS form, is stronger than the coal and income correlation, raising the possibility that using iron instead of coal would affect my empirical results. Bearing in mind that here I am using distance to iron ore deposits, rather than the cost of transport, I run the same IV exercise as in *Chapter 4*'s table 4.7. The results are in table B.4.

In the reduced form estimaiton in column (1), we can see that the farther regions were from iron ore deposits, the lower their income levels. The effect is

Table B.4: Iron and institutions IV estimates

	(3) OLS $\ln Y$	(4) IV-2 $\ln Y$
$INST$	0.643***	0.618***
	(0.085)	(0.086)
$\ln Iron\_Distance$	-0.142*	-0.081
	(0.071)	(0.052)
$Constant$	7.855***	7.041**
	(0.129)	(0.493)
Year F.E.	Yes	Yes
Country F.E.	Yes	Yes
$\hat{R}^2$	0.650	0.604
N	597	597
Angrist-Pischke multivariate F-test		16.78***
Anderson canon. corr. LM statistic		16.62***
	(1) RF $\ln Y$	(2) IV-1 $INST$
SH	-0.274***	-0.145***
	(0.038)	(0.012)
$\ln Iron\_Distance$	-0.189***	-0.064**
	(0.022)	(0.022)
$Constant$	8.297***	0.972***
	(0.097)	(0.061)
Year F.E.	Yes	Yes
Country F.E.	Yes	Yes
$\hat{R}^2$	0.465	0.534
N	597	597

*Notes: *** denotes significance at one per cent and ** at five per cent. Robust standard errors clustered on countries, and reported in brackets. Angrist-Pischke multivariate F-test is on the excluded instrument, SH (Spanish Habsburg Regions). Anderson canon. corr. LM statistic is for under-identification. Y is GDP per capita; INST is a measure of regional institutional efficiency; and and Iron_Distance is distance to nearest iron deposit. Column (1) is the reduced form estimation; (2) is the first stage; (3) is the OLS estimation; and([4) is the second stage. Mark Schaffer at Heriot-Watt University supplied the Stata code for this estimation.*

large and significant. It does not come at the cost of a change in magnitude or significance of the Spanish Habsburg IV (SH) for regional institutional efficiency ($INST$). The same holds for the first stage estimation in column (2), where SH retains is strength. In the second stage in column (4), however, while the coefficient on distance to iron ore deposits ($Iron\_Distance$) grows in magnitude from -0.064 to -0.81, its error also increased, making it statistically insignificant. This contrasts to its significance in the OLS estimation in column (3), and to the relationship I highlighted in 4.9. Instrumented $INST$, on the other hand, has retained its magnitude and significance, showing similar results to those reported in table 4.7. That $Iron\_Distance$ is negatively and significantly correlated with $INST$ in column (2), the first stage, implies that the simple iron-income correlation was perhaps spurious, and was working through the variation in $INST$. A full econometric investigation is beyond the scope of what I am trying to do here, but the point remains that when controlling for institutional efficiency in an appropriate IV framework, access to either coal or iron do not exert a significant influence on regional per capita income levels.

Appendix C

Market potential

C.1 Estimating market potential

Bilateral trade data are from Jacks et al. (2011), which provides flows for all countries starting in 1870. The section I take is 1870 to 1920. There are no data prior to 1870, which is why the final window must extend slightly beyond my precise period. The length of the window is, roughly, the length of the entire period, divided by three. I experimented with different windows, finding the same results.

My GDP data are the same as those underlying *Chapter 3*, and are detailed in the appendix. Great circle distances from regional nodes (regional or provincial capitals) to one another were calculated using a geographical information system; and language and border data were taken and cross-checked using the 1912 *Cambridge Modern History Atlas* (William Ward et al., 1912). Following **?**, I adjusted for internal (within-regional) distances as $dist_{ii} = 0.66 \times \left(\frac{area_i}{\pi}\right)^{0.5}$ where $area_i$ is region i's area in square-kilometres. This formula, or variants of it (for example, Crafts (2005b); Schulze (2007b)) is often used in the literature, and gives the average distance in a circular location under the assumption that economic activity occurs in the centre and consumers are spread evenly across space. As in **?**, equation 5.4 assumes that price indices are identical across regions within the same country. There is no way around this since the trade model yields only one estimate δ_{jt} per year per country. It is an inevitable part of using national-level trade data to derive regional level market potential values.

Table C.1: Results of bilateral trade model.

	1870-1883	1900-1906	1907-20
dist	-0.845	-0.726	-0.51
	(0.386)**	(0.444)*	(0.302)*
border	0.541	0.629	0.782
	(0.426)	(0.331)*	(0.254)**
language	3.202	2.268	2.334
	(0.527)***	(0.471)***	(0.389)***
Importer F.E.	Yes	Yes	Yes
Exporter F.E.	Yes	Yes	Yes
$\widehat{R}^2$	0.74	0.65	0.67
N	1284	1295	1079

*Notes: robust standard errors clustered at trading-pair level reported in brackets. *** denotes statistical significance at one per cent; ** at five per cent; and * at 10 per cent. Sample restricted to exports within my sample of countries, and from my sample countries to the rest of the world.*

The econometric results are displayed in table C.1. Each period provides a sizeable number of observations, and the adjusted-R^2 values lie between 0.74 and 0.65. This simple specification explains a considerable proportion of variation in bilateral trade flows. Look more specifically at the variables, sharing a common language has a strong positive effect on bilateral trade flows, as in **?**. Except for the initial period, sharing a national border also has a strong positive effect on trade flows. These results find much in common with the Schulze and Wolf (2012) findings that political borders and ethno-linguistic networks matter for economic integration. Unlike Schulze and Wolf (2012), however, I am not interested in uncovering the precise underlying mechanisms. The literature on this is vibrant enough, and I am both ill equipped to enter it and content with using these implied effects to estimate market potential as in **?**. As in all gravity models, distance is highly significant and negative - in all periods. This captures the high trade costs (mainly transport) that come with increasing distance. It is interesting to see that the size of the coefficient on distance is declining over time, in line with the late-nineteenth century's transport revolution. Jacks et al. (2011) see the same decline in the size of distance coefficient over the 1870 to 1939 period for similar reasons.

References

Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *Journal of Economic History*, 46(2):385–406. 24

Abreu, M., de Groot, H., and Florax, R. (2005). A meta-analysis of beta-convergence: the legendary two-percent. *Journal of Economic Surveys*, 19:389–420. 62

Acemoglu, D., Bautista, M., Querubin, P., and Robinson, J. (2008). Economic and Political Inequality in Development: The Case of Cundinamarca, Colombia. In Helpman, E., editor, *Institutions and Economic Performance*, number 5. Harvard University Press, Cambridge, M. 116

Acemoglu, D., Cantoni, D., Johnson, S., and Robinson, J. (2011). The Consequences of Radical Reform: The French Revolution. *American Economic Review*, 101(7):3286–3307. 94, 95, 101

Acemoglu, D. and Dell, M. (2010). Productivity Differences between and within Countries. *American Economic Journal: Macroeconomics*, 2(1):169–188. 24, 70, 74, 87, 88, 116

Acemoglu, D. and Johnson, S. (2005). Unbundling institutions. *Journal of Political Economy*, 113(5):949–995. 25, 74

Acemoglu, D., Johnson, S., and Robinson, J. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *The American Economic Review*, 91(5):1369–1401. 25, 69, 73, 74, 95, 96, 97, 98, 184

Acemoglu, D., Johnson, S., and Robinson, J. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *The Quarterly Journal of Economics*, 117(4):1231–1294. 69

Acemoglu, D., Johnson, S., and Robinson, J. (2005). Institutions as a Fundamental Cause of Long-Run Growth. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1.A., pages 385–414. Elsevier B.V., Amsterdam. 13, 17, 22, 26, 59, 87, 88, 90

Aguilar-Retureta, J. (2014). The GDP per capita of the Mexican regions (1895-1930): new estimates. *Asociacion Espanola de Historia Economica, Documentos de Trabajo*, (1415):1–44. 34

A'Hearn, B. and Venables, A. J. (2013). Internal Geography and External Trade: Regional Disparities in Italy, 1861-2011. In Toniolo, G., editor, *The Oxford Handbook of the Italian Economy since Unification*, number 12. Oxford University Press, New York. 123

Aigner, D., Knox Lovell, C., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production functions. *Journal of Econometrics*, 6:21–37. 89

Aitken, N. (1973). The Effect of the EEC and EFTA on European Trade: A Temporal Cross-Section Analysis. *American Economic Review*, 63(5):881–92. 2

Aldcroft, D., editor (1994). *The European Economy, 1750-1914: A Thematic Approach*. Manchester University Press, Manchester. 26

Allen, R. (2001). The Great Divergence in European Wages and Prices from the Middle Ages to the First World War. *Explorations in Economic History*, 38:411–447. 113

Allen, R. (2006). The British Industrial Revolution in Global Perspective: How Commerce Created The Industrial Revolution and Modern Economic Growth. http://www.nuff.ox.ac.uk/users/allen/unpublished/econinvent-3.pdf. 114

Allen, R. (2009). *The British Industrial Revolution in Global Perspective*. Cambridge University Press, Cambridge. 69, 71

American Society for Testing and Materials, Subcommittee: D05.18 (1998). Standard Classification of Coals by Rank. Legally Binding Document, Office of the Federal Register ASTM D388, US Office of the Federal Register, Washington, D.C. 80

Anselin, L. (1995). Local indicators of spatial association - LISA. *Geographical Analysis*, 27:93–115. 65

Armstrong, J. (1987). The Role of Coastal Shipping in UK Transport: an Estimate of Comparative Traffic Movements in 1910. *Journal of Transport History*, 8:164–178. 79

Atack, J. (2013). On the Use of Geographic Information Systems in Economic History: The American Transportation Revolution Revisited. *Journal of Economic History*, 73(2):313–338. 15, 186

Badia-Miro, M., Guilera, J., and Lains, P. (2012). Reconstruction of the Regional GDP of Portugal, 1890-1980. *Documents de Treball de la Facultat d'Economia i Empressa, Universitat de Barcelona*, (E12/280). 34

Baines, D. (1994). European labour markets, emigration, and internal migration, 1850–1913. In Hatton, T. and Williamson, J., editors, *Migration and the International Labour Market, 1850–1939*, pages 35–54. Routledge, London. 59

Bairoch, P. (1997). *Victoires et deboires. Histoire economique et sociale du monde du XVIe. siecle a nos jours*. Editions Gallimard, Paris. 11, 140

Balassa, B. (1961). *The Theory of Economic Integration*. R.D. Irwin, New York. 2

Banfield, E. (1958). *The Moral Basis of a Backward Society*. The Free Press, Glencoe, IL. 87, 88

Barnard, H. (1969). *Education and the French Revolution*. Cambridge University Press, Cambridge. 158

Barro, R. (1991). Economic growth in a cross section of countries. *The Quarterly Journal of Economics*, 106(2):407–443. 24, 174

Barro, R. (1997). *The Determinants of Economic Growth: A Cross-Country Empirical Study*. MIT Press, Cambridge, M. 174

Barro, R. and Sala-i Martin, X. (2004). *Economic Growth*. MIT Press, Cambridge, M. 2, 59

Bartholomew, J. (1907). Europe and Near East - General Commercial Chart. Suez Canal. In Bartholomew, J., editor, *Atlas of the World's Commerce, A New Series of Maps with Descriptive Text and Diagrams showing Products, Imports, Exports, Commercial Conditions and Economic Statistics of the Countries of the World, Compiled from the Latest Official returns at the Edinburgh Geographical Institute*. George Newnes, Ltd., London. 77

Baten, J. and Van Zanden, J. (2008). Book production and the onset of modern economic growth. *Journal of Economic Growth*, 13:217–235. 166

Baumol, W. (1986). Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show. *American Economic Review*, 76(5):1072–85. 24

Becker, S., Boeckh, K., Hainz, C., and Woessmann, L. (2011a). The Empire is Dead, Long Live the Empire! Long-run Persistence of Trust and Corruption in the Bureaucracy. *CEPR Discussion Papers*, (8288). 98, 100, 103

Becker, S., Hornung, E., and Woessmann, L. (2011b). Education and Catch-up in the Industrial Revolution. *American Economic Journal: Macroeconomics*, 3(3):92–126. 160, 163, 166, 167, 168, 181

Becker, S. and Woessmann, L. (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *The Quarterly Journal of Economics*, 124(2):531–596. 199

Benavot, A. and Riddle, P. (1989). The Expansion of Primary Education, 1870-1940: Trends and Issues. *Sociology of Education*, 61(3):191–210. 176

Berend, I. (2013). *An Economic History of Nineteenth Century Europe: Diversity and Industrialisation*. Cambridge University Press, Cambridge. 28, 97, 98, 104, 115, 116, 120, 125, 182, 184, 185

Berenger, J. (1994). *A History of the Habsburg Empire, 1273-1700*. Longman, London. 99, 100, 103

Blanchard, O., Jaumotte, F., and Loungani, P. (2013). Voxeu: Unemployment, labour-market flexibility and IMF advice: Moving beyond mantras. http://www.voxeu.org/article/unemployment-labour-market-flexibility-and-imf-advice. 188

Bolt, J. and Van Zanden, J. (2013). The First Update of the Maddison Project; Re-Estimating Growth Before 1820. *Maddison Project Working Paper*, (4). 14, 34, 53, 54, 122

Bordo, M. (1981). The Classical Gold Standard: Some Lessons for Today. *Federal Reserve Bank of St. Louis Review*, pages 2–17. 20, 24

Bourguignon, F. and Morrisson, C. (2002). Inequality Among World Citizens: 1820-1992. *American Economic Review*, 92(4):727–744. 58

Breinlich, H. (2006). The spatial income structure in the European Union - what role for Economic Geography? *Journal of Economic Geography*, 6(5):593–617. 10

Broadberry, S., Federico, G., and Klein, A. (2010). An Economic History of Modern Europe: Sectoral Developments, 1870-1914. In Broadberry, S. and O'Rourke, K., editors, *Unifying the European Experience: An Economic History of Modern Europe*. Cambridge University Press, London. 13, 33, 35, 38

Buck, T. (2006). Internal market thrives within golden triangle. *The Financial Times*, (13). 13

Buyst, E. (2009). Reversal of Fortune in a Small, Open Economy: Regional GDP in Belgium, 1896-2000. *VIVES Research Centre for Regional Economics*, (8). 34, 55

Cain, P. (1980). Private enterprise or public utility? Output, pricing and investment on English and Welsh railways, 1870-1914. *Journal of Transport History*, 1(1):9–28. 77, 123

Cameron, R. (1985). A New View of European Industrialization. *Economic History Review*, 38(1):1–23. 22, 73

Caron, F. (1979). *An Economic History of Modern France*. Methuen, London. 23

Caruana Galizia, P. (2013). Indian Regional Income Inequality: Estimates of Provincial GDP, 1875-1911. *Economic History of Developing Regions*, 28(1):1–27. 14, 34

Caruana Galizia, P. and Ma, Y. (2014). Chinese Regions in the Great DIvergence: Provincial GDP per capita, 1873-1918. Under review at journal. 14, 34

Chandler, A. (1990). *Scale and Scope: The Dynamics of Industrial Capitalism*. Belknap Press of Harvard University Press, Cambridge, M. 22, 23

Chang, H.-J. (2010). Institutions and economic development: theory, policy and history. *Journal of Institutional Economics*, 7(4):473–498. 15, 25, 96

Chenery, H. and Syrquin, M. (1975). *Patterns of Development, 1950-1970*. Oxford University Press, Oxford. 35, 42

Cipolla, C. (1969). *Literacy and Development in the West*. Pelican, Baltimore. 17, 156, 159, 166, 198

Cipolla, C., editor (1976). *The Fontana Economic History of Europe: The Industrial Revolution, 1700-1914*, volume 3. Barnes and Noble, New York. 25

Clingingsmith, D. and Williamson, J. (2008). Deindustrialization in 18th and 19th century India: Mughal decline, climate shocks and British industrial ascent. *Explorations in Economic History*, 45(3):209–34. 142

Clogg, C., Petkova, E., and Haritou, A. (1995). Statistical Methods for Comparing Regression Coefficients Between Models. *American Journal of Sociology*, 100(5):1261–1293. 133, 136, 137

Combes, P., Lafourcade, M., Thisse, J., and Toutain, J. (2011). The rise and fall of spatial inequalities in France: A long-run perspective. *Explorations in Economic History*, 48(2):243–71. 16, 29, 30, 35, 37, 39, 40, 41, 42, 56, 183

Combes, P.-P. (2010). *Agglomeration Economics*, chapter Estimating Agglomeration Economies with History, Geology, and Worker Effects, pages 15–66. University of Chicago Press, Chicago, IL. 196

Combes, P.-P., Lafourcade, M., and Mayer, T. (2005). The trade creating effects of business and social networks: evidence from France. *Journal of Institutional Economics*, 66:1–29. 126

Combes, P.-P., Mayer, T., and Thisse, J. (2008). *Economic Geography: The Integration of Regions and Nations*. Princeton University Press, Princeton, N.J. 9, 10

Cousin, M. (1834). *Report on the State of Public Instruction in Prussia: Addressed to the Count de Montalivet, translated by Sarah Austin*. Effinghma Wilson, London. 176

Crafts, N. (1984). Patterns of Development in Nineteenth Century Europe. *Oxford Economic Papers*, 36(3):438–58. 22, 35, 36, 74

Crafts, N. (1997). The Human Development Index and changes in standards of living: Some historical comparisons. *European Review of Economic History*, 1(3):299–322. 17, 156

Crafts, N. (2005a). Market potential in British regions, 1871-1931. *Regional Studies*, 39(9):1159–1166. 6, 29, 121, 122, 130, 145, 150, 183, 185

Crafts, N. (2005b). Regional GDP In Britain, 1871-1911: Some Estimates. *Scottish Journal of Political Economy*, 52(1):54–64. 14, 34, 47, 54, 122, 202

Crafts, N. and Mulatu, A. (2006). How did the location of industry respond to falling transport costs in Britain before World War I? *Journal of Economic History*, 66(3):575–607. 29

Crafts, N. and Wolf, N. (2013a). The Location of the UK Cotton Textiles Industry in 1838: a Quantitative Analysis. *CEPR Discussion Paper*, (DP9626). 142, 152, 187, 188

Crafts, N. and Wolf, N. (2013b). The Location of the UK Cotton Textiles Industry in 1838: A Quantitative Analysis. *CEPR Discussion Paper*, (9626). 155

Craig, L. and Fisher, D. (1997). *The Integration of the European Economy, 1850-1913*. Macmillan Press, London. 100, 102

Crayen, D. and Baten, J. (2010). Global trends in numeracy 1820-1949 and its implications for long-run growth. *Explorations in Economic History*, 47:82–99. 166

Cvrcek, T. and Zajicek, M. (2013). School, what is it good for? The politics and economics of public education in the 19th Century Habsburg Empire. 165, 170

Daudin, G., Morys, M., and O'Rourke, K. (2010). Globalization, 1870-1914. In Broadberry, S. and O'Rourke, K., editors, *The Cambridge Economic History of Modern Europe: Volume 2, 1870 to the Present*, number 1. Cambridge University Press, Cambridge. 12, 19, 20, 24

Davis, D. and Weinstein, D. (2002). Bones, Bombs and Break Points: The Geography of Economic Activity. *American Economic Review*, 92(5):1269–89. 153

de Grauwe, P. (2010). Voxeu: Fighting the wrong enemy. http://www.voxeu.org/article/europe-s-private-versus-public-debt-problem-fighting-wrong-enemy. 103

de Grauwe, P. (2012). The Governance of a Fragile Eurozone. *Australian Economic Review*, 45(3):255–68. 64

Deane, P. (1965). *The First Industrial Revolution*. Cambridge University Press, Cambridge. 73, 75

Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica*, 78(6):1863–1903. 115, 116

DeLong, B. (1989). Productivity Growth, Convergence, and Welfare, a Comment. *American Economic Review*, 79:1138–59. 24

Deutsch Statistischen Bureau (1871). *Statistik des Deutschen Reichs 1871*, volume A.F. 14/2. Mitteilungen des statistischen Bureaus in Berlin, Berlin. 37

Deutsch Statistischen Bureau (1882). *Statistik Des Deutschen Reichs 1882*, volume N.F.111. Mitteilungen des statistischen Bureaus in Berlin, Berlin. 37, 198

Deutsch Statistischen Bureau (1895). *Statistik Des Deutschen Reichs 1895*, volume N.F.111. Mitteilungen des statistischen Bureaus in Berlin, Berlin. 37

Deutsch Statistischen Bureau (1910). *Deutschen Reichs - Statistisches Jahrbuch - 1910*, volume 31. Mitteilungen des statistischen Bureaus in Berlin, Berlin. 37, 198

Deutsch Statistischen Bureau (1912). *Deutschen Reichs - Statistisches Jahrbuch - 1912*, volume 33. Mitteilungen des statistischen Bureaus in Berlin, Berlin. 37

Diamond, J. and Robinson, J. (2010). *Natural Experiments of History*. Harvard University Press, Cambridge, M. 15

Dimico, A., Isopi, A., and Olsson, O. (2012). Origins of the sicilian mafia: The market for lemons. *University of Gothenburg, Working Papers in Economics*, (532). 87

Dollar, D. (1992). Outward-Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976-1985. *Economic Development and Cultural Change*, 40(3):523–44. 139, 140, 151

Donaldson, D. (2012). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. 187

Draghi, M. (2012). Competitiveness of the euro area and within the euro area. Speech at at the colloquium "Les défis de la compétitivité", organised by Le Monde and l'Association Française des Entreprises Privées (AFEP), Paris. 188

Drelichman, M. and Voth, H.-J. (2011). Lending to the Borrower from Hell: Debt and Default in the Age of Philip II. *The Economic Journal*, 121(557):1205–1227. 103

Easterly, W. (2003). Tropics, germs, and crops: how endowments influence economic development. *Journal of Monetary Economics*, 50:3–39. 124, 134

Elliott, J. (1961). The decline of Spain. *Past and Present*, 20:52–75. 103

Elliott, J. (2009). *Spain, Europe, and the Wider World, 1500-1800*. Yale University Press, New Haven, CT. 103

Enflo, K., Henning, M., and Schon, L. (2010). Swedish regional GDP 1855-2000: Estimations and general trends in the Swedish regional system. *Universidad Carlos III de Madrid Working Papers in Economic History*. x, 14, 34, 47, 48, 49, 55

Engerman, S. and Sokoloff, S. (2000). History Lessons: Institutions Factor Endowments, and Paths of Development in the New World. *Journal of Economic Perspectives*, 14(217-232). 91

Epstein, P., Howlett, P., and Schulze, M.-S. (2003). Distribution dynamics: stratification, polarization, and convergence among OECD economies, 1870-1992. *Explorations in Economic History*, 40:78–97. 12, 60, 62, 63, 64, 70

Estevadeordal, A., Frantz, B., and Taylor, A. (2002). The rise and fall of world trade, 1870-1913. *National Bureau of Economic Research, Working Paper*, (9318). 121, 122

Esteves, R. (2011). The Belle Epoque of International Finance: French Capital Exports, 1880-1914. *Oxford Department of Economics Discussion Paper Series*, (534). 20

European Commission (2011). Eurostat - Regional Statistics Database. epp.eurostat.ec.europa.eu/portal/page/portal/region$_c$ities/regional$_s$tatistics/data/database.2

Fabricant, S. (1942). *Employment in Manufacturing, 1899-1939*. National Bureau of Economic Research, New York. 35

Farrel, M. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society*, 120:253–281. 89

Feinstein, C. and Pollard, S. (1988). *Studies in Capital Formation in the UK, 1750-1920*. Clarendon Press, Oxford. 21, 135

Felice, E. (2009). Estimating regional GDP in Italy (1871-2001): sources, methodology, and results. *Universidad Carlos III de Madrid, Working Papers in Economic History*, (09-07). 14, 34, 47, 52, 53

Felice, E. (2012). Regional convergence in Italy, 1891–2001: testing human and social capital. *Cliometrica*, 6(3):267–306. 170, 172, 198

Ferguson, N. (2012). Ninth Semi-Annual Munk Debate - Has the European Experiment Failed? The European project is a total failure. http://www.munkdebates.com/debates/europe. 3

Fernandes, F. and Schulze, M.-S. (2009). Human Capital Formation in Austria-Hungary and Germany: Time Series Estimates of Educational Attainment, 1860-1910. In Bela, B., editor, *A Felhalmozás Míve: Tönténeti Tanulmányok Kövér György Tiszteletére*. Szászadvég Kiadó, Budapest. 169, 170

Fernihough, A. and O'Rourke, K. (2014). Coal and the European Industrial Revolution. *University of Oxford, Discussion Papers in Economic and Social History*, (124). 14, 73, 75, 76, 80, 113

Foreman-Peck, J. (1995). A Model of Later Nineteenth Century European Economic Development. *Revista de Historia Económica / Journal of Iberian and Latin American Economic History*, 13(3):441–71. 21, 22, 23

Foreman-Peck, J. (1999). The Zenith of European Power, 1870–1918. In Aldcroft, D. and Sutcliffe, A., editors, *Europe in the International Economy: 1500-2000*, chapter 3. Edward Elgar, London. 20

Frank, H. (1994). *Regionale Entwicklungsdisparitäten im deutschen Industrialisierungsprozess 1849-1939: eine empirisch-analytische Untersuchung*. LIT Verlag, Hamburg. x, 35, 37, 38, 44

Fredriksen, K. (2012). Income Inequality in the European Union. *OECD Economics Department Working Papers*, (952):1–26. 1

Fujita, M., Krugman, P., and Mori, T. (1999a). On the evolution of hierarchical urban systems. *European Economic Review*, 43(2):209–251. 8

Fujita, M., Krugman, P., and Venables, A. J. (1999b). *The Spatial Economy: Cities, Regions, and International Trade*. MIT Press, London. 27, 156

Fukayama, F. (2001). Social capital, civil society and development. *Third World Quarterly*, 22(1):7–20. 88

Gallup, J. and Sachs, J. (2001). The Economic Burden of Malaria. *The American Journal of Tropical Medicine and Hygiene*, 64(1,2). 134

Geary, F. and Stark, T. (2002). Examining ireland's post-famine economic growth performance. *The Economic Journal*, 112(482):919–935. 35, 45, 47, 51, 54, 55, 189, 190, 191

Gerschenkron, A. (1962). *Economic backwardness in historical perspective, a book of essays*. Belknap Press of Harvard University Press, Cambridge, M. 21, 26, 69, 134, 135

Giusti, U. (1914). *Annuario statistico delle città Italiane, 1913-14*. Alfani e Venturi, Florence. 52

Global Financial Data (2014). Exchange Rates to 1383. 83, 113, 114

Goethe, J.W. von (1961). *Faust*. Doubleday, New York. 98

Good, D. (1984). *The Economic Rise of the Habsburg Empire: 1750-1914*. University of California Press, California. 198

Good, D. and Ma, T. (1998). *Von der Theorie zur Wirtschaftspolitik - ein österreichischer Weg*, chapter New estimates of income levels in central and eastern Europe, 1870-1910. Lucius und Lucius, Stuttgart. 35

Goubert, P. (2002). *The Course of French History*. Routledge, London. 101

Grafe, R. and Irigoin, M. (2006). The Spanish Empire and Its Legacy: Fiscal Re-distribution and Political Conflict in Colonial and Post-Colonial Spanish America. Working Papers of the Global Economic History Network (GEHN) 23/06, London School of Economics, London. 102

Graff, H. (1991). *The Literacy myth. Cultural Integration and Social Structure in the Nineteenth Century*. Transaction Publishers, New Brunswick. 166

Grossman, G. and Helpman, E. (1991). Trade, knowledge spillovers, and growth. *European Economic Review*, 35:517–526. 162

Hadfield, C. (1968). *The Canal Age*. David and Charles, London. 78

Hall, P. (1993). Forces Shaping Urban Europe. *Urban Studies*, 30(6):883–98. 13

Hamilton, E. (1938). The decline of Spain. *Economic History Review*, 8(2):168–79. 103

Hanson, G. Market potential, increasing returns and geographic concentration. 6

Hanson, G. (2005). Market potential, increasing returns and geographic concentration. *Journal of International Economics*, 67(1):1–24. 6, 8, 28, 118, 130, 165

Harris, C. (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. 6, 118, 122, 130, 183

Head, K. and Mayer, T. (2006). Regional wage and employment responses to market potential in the EU. *Regional Science and Urban Economics*, 36:573–594. 131, 132

Head, K. and Meyer, T. (2004). Market potential and the location of japanese investment in the european union. *Review of Economics and Statistics*, 86(4):959–972. 6

Helpman, E. (1998). *Topics in Public Economics: Theoretical and Applied Analysis*, chapter The Size of Regions. Cambridge University Press, Cambridge. 7, 8

Henning, M., Enflo, K., and Andersson, F. (2011). Trends and cycles in regional economic growth: How spatial differences shaped the Swedish growth experience from 1860–2009. *Explorations in Economic History*, 48(4):538–555. 28, 60, 185

Heywood, C. (1995). *The development of the French Economy, 1750-1914*. Cambridge University Press, Cambridge. 67, 73

Hippe, R. (2012). How to measure human capital? The Relationship between Numeracy and Literacy. *Economies et sociétés*, 45:1527–1554. 165, 166, 169

Hobsbawm, E. (1987). *The Age of Empire: 1875-1914*. Weidenfeld and Nicolson, London. 19, 20

Hobson, J. (1902). *Imperialism: A Study*. Cosimo, London. 19, 20

Hoffman, W. (1965). *Wachstum der Deutschen Wirtschaft seit der Mitte des 19. Jahrhunderts*. Springer, Berlin. 37, 38, 44, 45, 46, 54

Holl, A. (2011). Market potential and firm-level productivity in Spain. *Journal of Economic Geography*, 12(6):1191–1215. 132

Humphries, J. (2010). *Childhood and Child Labour in the British Industrial Revolution*. Cambridge University Press, Cambridge. 168

INSCR (2012). Integrated Network for Societal Conflict Research - Polity IV Annual Time-Series 1800-2012. 95

Irish Census (1910). *General Report and Ireland Census 1901, Part II*. H.M. Stationery Office, London. 198

Irish Census (1911). *Census of Ireland: General Report*. H.M. Stationery Office, London. 198

Irwin, D. (2006). Commentary: Shifts in Economic Geography and Their Causes. In *Proceedings-Economic Policy Symposium-Jackson Hole*, pages 41–8, Kansas City, MO. Federal Reserve Bank of Kansas City, Federal Reserve Bank of Kansas City. 187

Jacks, D. (2005). Intra- and international commodity market integration in the atlantic economy, 1800-1913. *Explorations in Economic History*, 42(3):381–413. 53

Jacks, D., Mesissner, C., and Novy, D. (2011). Trade booms, trade busts, and trade costs. *Journal of International Economics*, 83(2):185–201. 126, 127, 202, 203

Jacobs, A. and Richter, H. (1935). *Die Großhandelspreise in Deutschland von 1792 bis 1934*. Number 37 in Sonderhefte des Instituts für Konjunkturforschung. Hanseatische Verlagsanstalt Hamburg, Berlin. 113, 114

Jedwab, R., Kerby, E., and Moradi, A. (2014). History, Path Dependence and Development: Evidence from Colonial Railroads, Settlers and Cities in Kenya. 146

Jerven, M. (2013). *Poor Numbers: How We Are Misled by African Development Statistics and What to Do about It*. Cornell University Press, New York. 185

Jevons, H. S. (1915). *The British Coal Trade*. Kegan Paul, Trench, Trubner and Co. Ltd., London. 110, 111, 112, 113, 114

Jevons, W. S. (1865). *The Coal Question; An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our Coal Mines*. Macmillan and Co., London. 111

Johnson, C. (1995). *The Life and Death of Industrial Languedoc, 1700-1920*. Oxford University Press, New York. 153

Jones, E. (1981). *The European Miracle: Environments, Economies and Geopolitics in the History of Europe and Asia*. Cambridge University Press, Cambridge. 22

Jorberg, L. (1973). *A History of Prices in Sweden, 1732-1914, Volume I: Sources, Methods, Tables*. CWK Gleerup, Lund. 53

Kaplan, M. (1957). *System and Process in International Politics*. European Consortium for Political Research Press, Glasgow. 19, 20

Kaukiainen, Y. (2006). Journey Costs, Terminal Costs and Ocean Tramp Freights: How the Price of Distance Declined from the 1870s to 2000. *International Journal of Maritime History*, 18(2):17–64. 5, 77, 123

Keeble, D., Owens, P., and Thompson, P. (1982). Regional accessibility and economic potential in the European community. *Regional Studies*, 16(6):419–32. 118, 122

Kennedy, W. (1986). *Industrial Structure, Capital Markets, and the Origins of British Economic Decline*. Cambridge University Press, Cambridge. 22

Kenny, C. (2005). Why Are We Worried About Income? Nearly Everything that Matters is Converging. *World Development*, 33(1):1–19. 59

Kenwood, A. and Lougheed, A. (1982). *Technological diffusion and industrialisation before 1914*. Croom Helm, London. 74, 76, 151

Klemann, H. and Schenk, J. (2013). Competition in the Rhine delta: waterways, railways and ports, 1870–1913. *Economic History Review*, 66(3):826–47. 78

Krugman, P. (1991). *Geography and Trade*. MIT Press, Cambridge, M. 3, 4, 6, 7, 9, 11, 27, 67, 117, 153, 183

Krugman, P. (1993). First nature, second nature, and metropolitan location. *Journal of Regional Science*, 32(2):129–144. 27, 145

Krugman, P. and Venables, A. J. (1990). Integration and the Competitiveness of Peripheral Industry. In Braga de Macedo, J. and Bills, C., editors, *Unity with Diversity in the European Economy: The Community's Southern Frontier*, number 3. Centre for Economic Policy Research, London. 3

Krugman, P. and Venables, A. J. (1995). Globalisation and the inequality of nations. *The Quarterly Journal of Economics*, 110(4):857–880. 27, 142, 143, 144, 152

Landes, D. (1965). Technological Change and Development in Western Europe, 1750-1914. In Habakkuk, H. and Postan, M., editors, *The Cambridge Economic*

History of Europe, volume VI, pages 274–585. Cambridge University Press, Cambridge. 25, 67, 110

Landes, D. (1969). *The Unbound Prometheus: Technological Change and Industrial Revolution in Western Europe from 1750 to the Present*. Cambridge University Press, Cambridge. 5, 21, 22, 25, 135

Laquer, T. (1974). Literacy and Social Mobility in the Industrial Revolution. *Past and Present*, 64:96–107. 167, 168

Lehmann, S. and O'Rourke, K. (2011). The structure of protection and growth in the late nineteenth century. *Review of Economics and Statistics*, 93(2):606–616. 120

Leunig, T., Minns, C., and Wallis, P. (2011). Networks in the Premodern Economy: The Market for London Apprenticeships, 1600-1749. *Journal of Economic History*, 71(2):413–443. 181

Leunig, T., Minns, C., and Wallis, P. (2012). Networks in the Premodern Economy: The Market for London Apprenticeships, 1600-1749. *Journal of Economic History*, 71(2):413–443. 168

Levy-Leboyer, M. and Bourguignon, F. (1990). *The French Economy in the Nineteenth Century: An Essay in Econometric Analysis*. Cambridge University Press, Cambridge. 23

Lieberman, S. (2013). *The Contemporary Spanish Economy: A Historical Perspective*. Routledge, London. 115

Limao, N. and Venables, A. J. (2001). Infrastructure, Geographical Disadvantage, Transport Costs and Trade. *World Bank Economic Review*, 15(3):451–79. 142, 187

Lindert, P. (2004). *Growing Public: Social Spending and Economic Growth since the Eighteenth Century*, volume 1. Cambridge University Press, Cambridge. 169, 170

Liu, D. and Mesissner, C. (2014). Market Potential and the Rise of US Productivity Leadership. *Journal of International Economics*, In Press(xx):xx. 149

Ljunberg, J. and Nilsson, A. (2009). Human Capital and Economic Growth: Sweden, 1870-2000. *Cliometrica*, 3(1):71–95. 199

Lockheed, M., Jamison, T., and Lau, L. (1980). Farmer Education and Farm Efficiency: A Survey. *Economic Development and Cultural Change*, 29(1):37–76. 161, 162

Lopez-Rodriguez, J., Andres Faina, J., and Lopez-Rodriguez, J. (2007). Human capital accumulation and geography: Empirical evidence from the European Union. *Regional Studies*, 41(2):217–234. 156, 159

Lösch, A. (1939). *The Economics of Location*. Yale University Press, New Haven, CT. 64

Luebke, D. (2014). Map of Europe in 1700. University of Oregon. 99

Maddison, A. (1995). *Monitoring the World Economy, 1820-1992*. OECD, Paris. 32, 54

Maddison, A. (2001). *The world economy: a millennial perspective*. OECD, Paris. 54

Maddison, A. (2007). *Contours of the World Economy 1-2030 AD: Essays in Macro-Economic History*. Oxford University Press, Oxford. 34, 38, 53, 55, 83

Mann, H. (1846). *Report of an Educational Tour in Germany, and Parts of Great Britain and Ireland: Being Part of the Seventh Annual Report of Horace Mann*. Simpkin, Marshall, and Co., London. 176

Marchand, O. and Thelot, C. (1991). *Deux Siecles de Travail en France: Population active et structure sociale duree et productivite du travial*. Institut national de la statistique et des etudes economiques, Paris. 41, 46

Marshall, A. (1920). *Principles of Economics*. Macmillan Press, London. 6, 27, 28

Marti-Henneberg, J. (2013). European integration and national models for railway networks (1840–2010). *Journal of Transport Geography*, 26:126–138. 125

Martínez-Galarraga, J. (2014). Market potential estimates in history: a survey of methods and an application to Spain, 1867-1930. *EHES Working Paper*, (51). 6

Mathias, P. (1983). *The First Industrial Nation: An Economic History of Britain 1700–1914*. Methuen, London. 75

Mathys, N. and Grether, J.-M. (2010). Is the World's Economic Center of Gravity Already in Asia? *Area*, 42(1):47–50. 132

McCallum, J. (1995). National borders matter. *American Economic Review*, 85:615–623. 127

McGillivray, M. and White, H. (1993). Measuring Development? The UNDP's Human Development Index. *Journal of International Development*, 5(2):183–192. 92, 93

Midelfart, K., Overman, H., Redding, S., and Venables, A. J. (2000). The location of European Industry. Technical Report 142, European Commission. 165

Milanovic, B. (2007). Globalization and Inequality. In Held, D. and Kaya, A., editors, *Global Inequality: Patterns and Explanations*, number 2. Polity Press, Cambridge. 59

Milward, A. and Saul, S. (1973). *The Economic Development of Continental Europe 1780-1870*. George Allen and Unwin, London. 21, 135, 184

Ministero di Agricoltura, I. e. C. (1887). *Annuario Statistico Italiano, 1886*. Ministero di Agricoltura, Industria e Commercio, Rome. 52

Mitch, D. (1992). *The Rise of Popular Literacy in Victorian England: The Influence of Private Choice and Public Policy*. University of Pennsylvania Press, Philadelphia. 159, 166

Mitchell, B. (2003). *International Historical Statistics: Europe, 1750-2000*. Basingstoke. 120, 122, 123

Mokyr, J. (1983). *Why Ireland Starved: A Quantitative and Analytical History of the Irish Economy, 1800-1850*. Allen and Unwin, London. 113

Mokyr, J. (1990). *The Lever of Riches. Technological Creativity and Economic Progress*. Oxford University Press, Oxford. 184

Mokyr, J. (1998). The Second Industrial Revolution, 1870-1914. Northwestern University, Evanston IL. 20, 30

Mokyr, J. (2002a). *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton University Press, Princeton, N.J. 162

Mokyr, J. (2002b). Thinking about Technology and Institutions. Macalaster International College Roundtable. 169

Montinola, G., Qian, Y., and Weingast, B. (1995). Federalism, Chinese Style: The Political Basis for Economic Success in China. *World Politics*, 48(1):50–81. 88

Morris, C. and Adelman, I. (1988). *Comparative Patterns of Economic Development, 1850-1914*. Johns Hopkins Press, Baltimore. 22

Moulton, H. (1914). *Waterways Versus Railroads*. Riverside Press, New York. 78, 123, 124, 125

Murtin, F. and Viarengo, M. (2011). The Expansion and Convergence of Compulsory Schooling in Western Europe, 1950–2000. *Economica*, 78(311):501–522. 165

Naritomi, J., Soares, R., and Assuncao, J. (2012). Institutional Development and Colonial Heritage within Brazil. *Journal of Economic History*, 72(2):393–422. 116

North, D. (1981). *Structure and Change in Economic History*. W.W. Norton, New York. 22, 101

North, D. (1989). Comparative Patterns of Economic Development, 1850-1914. by Cynthia Taft Morris; Irma Adelman Review by: Douglass C. North. *Journal of Economic Literature*, 27(1):89–90. 22

North, D. and Thomas, R. (1973). *The Rise of the Western World: A New Economic History*. Cambridge University Press, Cambridge. 24, 73, 74, 87

North, D. and Weingast, B. (1989). Constitution and commitment: The evolution of institutional governing public choice in seventeenth-century england. *Journal of Economic History*, 49(4):803–832. 69, 184

Noyes, W. (1905). *American Railroad Rates*. Boston. 77, 123

Nunez, C. (1990). Literacy and Economic Growth in Spain, 1860-1977. In Tortella, G., editor, *Education and Economic Development since the Industrial Revolution*. Generalitat Valenciana, Valencia. 198

Nunn, N. (2014). Historical Development. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 2, chapter 7. Elsevier B.V., North Holland. 23

O'Brien, P. and Keyder, C. (1978). *Economic Growth in Britain and France: 1780-1914*. George Allen and Unwin, London. 22, 23, 35, 56, 69

Odell, K. (1989). The Integration of Regional and Interregional Capital Markets: Evidence from the Pacific Coast, 1883–1913. *Journal of Economic History*, 49(2):297–310. 124

OECD (2010). *Economic Policy Reforms 2010: Going for Growth*. OECD Publishing, Paris. 188

OECD (2013a). *Economic Policy Reforms 2013: Going for Growth*. OECD Publishing, Paris. 187, 188

OECD (2013b). *Health at a Glance 2013: OECD Indicators*. OECD Publishing, Paris. 188

Olson, M. (1982). *The Rise and Decline of Nations: Economic Growth Stagflation and Social Rigidities*. Yale University Press, New Haven, CT. 22, 23

Olsson, O. and Hansson, G. (2011). Country size and the rule of law: Resuscitating montesquieu. *European Economic Review*, 55:613–629. 175

Opensahw, S. (1984). *The Modifiable Areal Unit Problem*. GeoBooks, Norwich. 25

O'Rourke, K. (1997). The European Grain Invasion, 1870-1913. *Journal of Economic History*, 57(4):775–801. 30, 119, 120

O'Rourke, K. and Williamson, J. (1997). Around the European Periphery, 1870-1913: Gloablisation, Schooling and Growth. *European Review of Economic History*, 1:153–190. 24

O'Rourke, K. and Williamson, J. (1999). *Globalization and History: The Evolution of a Nineteenth-Century Atlantic Economy*. MIT Press, Massachusetts. 12, 19, 20, 24, 70, 112, 120, 125

O'Rourke, K. and Williamson, J. (2000). When did Globalization begin? *National Bureau of Economic Research, Working Paper*, (7632). 120

Orsagh, T. (1968). The Probable Geographical Distribution of German Income, 1882-1963. *Journal of Institutional and Theoretical Economics*, 124(2):280–311. x, 43, 44

Palmer, R., Colton, J., and Kramer, L. (2002). The Transformation of Eastern Europe. In Palmer, R., Colton, J., and Kramer, L., editors, *A History of the Modern World, 9th edition*, number 23. McGraw Hill, New York. 99

Parker, W. (1961). Comment. In Spengler, J., editor, *Natural Resources and Economic Growth*, page 160. Resources for the Future, Washington, D.C. 73, 185

Parnam, J. (2012). Good schools make good neighbors: Human capital spillovers in early 20th century agriculture. *Explorations in Economic History*, 49(3):316–334. 161

Platt, D. (1984). *Foreign Finance in Continental Europe and the USA, 1815-1870: Quantities, Origins, Functions, and Distribution*. George Allen and Unwin, London. 23

Pollard, S. (1973). Industrialization and the European Economy. *Economic History Review*, 26(4):636–48. 1, 5, 21, 26, 68, 69, 88, 135, 183

Pollard, S. (1974). *European Economic Integration, 1815-1970*. Thames and Hudson, London. 124

Pollard, S. (1981). *Peaceful conquest: The industrialization of Europe, 1760-1970*. Oxford University Press, Oxford. 5, 14, 15, 17, 19, 20, 21, 26, 27, 28, 29, 30, 67, 68, 71, 72, 73, 75, 76, 119, 120, 135, 151, 152, 153, 181, 182, 184, 185

Pounds, N. and Parker, W. (1957). *Coal and Steel in Western Europe; the Influence of Resources and Techniques on Production*. Indiana University Press, Indiana. 17, 73, 184

Prados de la Escosura, L. (2000). International comparisons of real product, 1820-1990: An alternative data set. *Explorations in Economic History*, 37(1):1–41. 53

Quah, D. (1996). Twin Peaks: Growth and Convergence in Models of Distribution Dynamics. *Economic Journal*, 102:1045–1055. 63, 71

Ralph, S., Tarr, B., and McMurry, F. (1910). *New Geographies - Second Book*. The MacMillan Company, New York. 81

Redding, S. and Schott, P. (2003). Distance, skill deepening, and development: will peripheral countries ever get rich? *Journal of Development Economics*, 72:515–541. 131, 154, 155, 156, 159, 163, 174, 177

Redding, S. and Strum, D. (2008). The costs of remoteness: Evidence from german division and reunification. *The American Economic Review*, 98(5):1766–1797. 25, 88, 185

Redding, S., Sturm, D., and Wolf, N. (2011). History and Industry Location: Evidence from German Airports. *The Review of Economics and Statistics*, 93(3):814–32. 146

Redding, S. and Venables, A. J. (2004). Economic geography and international inequality. *Journal of International Economics*, 62:53–82. 10, 28, 126, 131, 132, 139, 140, 151

Registrar-General of Births, Deaths, and Marriages in England and Wales (1870-1910). *Annual Report of the Registrar-General of Births, Deaths, and Marriages in England and Wales*. H.M. Stationery Office, London. 198

Reis, J. (2005). Economic Growth, Human Capital Formation and Comsumption in Western Europe before 1880. In Allen, R., Bengtsson, T., and Dribe, M., editors, *Living Standards in the Past*, pages 195–225. Oxford University Press, Oxford. 166

Rhode, P. (2008). William N. Parker, interviewed by Paul Rhode. In Lyons, J., Cain, L., and Williamson, S., editors, *Reflections on the Cliometrics Revolution: Conversations with Economic Historians*, pages 177–193. Routledge, London. 16, 25

Riley, S., DeGloria, S., and Elliot, R. (1999). A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences*, 5(2-4):23–27. 196, 197

Rodrigue, J. (2013). *The Geography of Transport Systems*. Routledge, New York. 123

Rodrik, D. (2004). Institutions and Economic Performance - Getting Institutions Right. *CESifo DICE Report*, 2(2):10–15. 97

Rodrik, D., Subramanian, A., and Trebbi, F. (2004). Institutions rule: The primacy of institutions over geography and integration in economic development. *Journal of Economic Growth*, 9(2):131–165. 87, 134

Rosenberg, N. and Birdzell, L. (1987). *How The West Grew Rich: The Economic Transformation Of The Industrial World*. Basic Books, New York. 22

Roses, J. (2003). Why isn't the whole of Spain industrialized? new economic geography and early industrialization, 1797-1910. *Journal of Economic History*, 63(4):995–1022. 28, 29, 30, 150, 183, 185

Roses, J., Martinez Galarrga, J., and Tirado, D. (2010). The upswing of regional income inequality in Spain (1860-1930). *Explorations in Economic History*, 47:244–257. 14, 15, 34

Rostow, W. (1960). *The Stages of Economic Growth: A Non-Communist Manifesto*. Cambridge University Press, Cambridge. 21, 134, 135

RRG (2012). Rrg spatial planning and geoinformation. 77, 80, 123

Sachs, J., Gallup, J., and Mellinger, A. (1999). Geography and economic development. *International Regional Science Review*, 22(2):179–232. 34, 124, 134

Sachs, J. and Warner, A. (1995). Economic reform and the process of global integration. *Brookings Papers on Economic Activity*, 26(1):1–118. 139, 140, 151

Sala-i-Martin, X. (1997). I Just Ran Two Million Regressions. *The American Economic Review*, 87(2):178–183. 24

Sandberg, L. (1979). The Case of the Impoverished Sophisticate: Human Capital and Swedish Economic Growth before World War I. *Journal of Economic History*, 39(1):225–241. 22, 102, 167, 199

Sandberry, L. and Steckel, R. (1997). Was Industrialisation Hazardous to your Health? Not in Sweden! In Steckel, R. and Floud, R., editors, *Health and Welfare during Industrialisation*. University of Chicago Press, Chicago, IL. 198

Sanderson, M. (1999). *Education and Economic Decline in Britain, 1870 to the 1990s*. Cambridge University Press, Cambridge. 175

Schofield, R. (1968). The measurement of literacy in pre-industrial England. In Goody, J., editor, *Literacy in Traditional Societies*, number 10, pages 310–326. Cambridge University Press. 166

Schofield, R. (1981). The Dimensions of illiteracy in England, 1750-1850. In Graff, H., editor, *Literacy and social development in the West: a reader*, number 10, pages 201–213. Cambridge University Press, Cambridge. 166

Schram, A. (1997). *Railways and the Formation of the Italian State in the Nineteenth Century*. Cambridge University Press, Cambridge. 123

Schulze, M.-S. (2007a). Origins of catch-up failure: Comparative productivity growth in the Habsburg Empire, 1870-1910. *European Review of Economic History*, 11:189–218. 39

Schulze, M.-S. (2007b). Regional income dispersion and market potential in the late nineteenth century Hapsburg Empire. *LSE Economic History Working Papers*, (106/07). x, 5, 14, 29, 34, 47, 49, 50, 51, 52, 55, 56, 77, 121, 122, 123, 130, 145, 146, 185, 202

Schulze, M.-S. (2011). Revised Estimates of Regional GDP in Imperial Hungary, 1870-1910. Mimeo. 34, 47

Schulze, M.-S. and Wolf, N. (2012). Economic nationalism and economic integration: the Austro-Hungarian Empire in the late nineteenth century. *Economic History Review*, 65(2):652–673. 70, 97, 203

Simpson, J. (1997). Economic development in Spain, 1850-1936. *Economic History Review*, 50(2):348–59. 110, 114

Smits, J., Woltjer, P., and Ma, D. (2009). A dataset on comparative historical national accounts, ca. 1870-1950: A time-series perspective. *Groningen Growth and Development Centre Research Memorandum GD-107*. 53, 54

Solow, R. (1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics*, 70(1):65–94. 21, 24

Solow, R. (1970). Science and Ideology in Economics. *The Public Interest*, (21):94–107. 186

Soysal, Y. and Strang, D. (1989). Construction of the First Mass Education Systems in Nineteenth Century Europe. *Sociology of Education*, 62(4):277–288. 157, 158

Statistique de la France (1872). *Statistique de la France - Tome XXI - Population*. Imprimerie Nationale, Paris. 37, 40, 198

Statistique de la France (1886). *France Dénombrement I - Résultats Statistique - Premiere Partie*. Berger-Levrault et Cie, Paris. 37, 40, 198

Statistique de la France (1901). *France Récensement - Tome IV - Résults Généraux*. Imprimerie Nationale, Paris. 37, 40, 198

Statistique de la France (1911). *France Récensement - Tome I - Premiere Partie - Population Légale ou De Résidence Habituelle*. Imprimerie Nationale, Paris. 37, 40, 198

Statistiska Centralbyrån (1935). *Folkråkningen den 31 december 1930*. Kung. boktryckeriet, P.A. Norstedt and Söner, Stockholm. 198

Stone, L. (1969). Literacy and Education in England, 1640-1900. *Past and Present*, 2:69–169. 166

Stowe, C. (1834). *Report on Elementary Public Instruction in Europe: Made to the Thirty-sixth General Assembly of the State of Ohio, December 19, 1837*. Dutton and Wentworth, State Printers, Columbus, OH. 176

Strassman, W. (1959). Creative Destruction and Partial Obsolesence in American Economic Development. *The Journal of Economic History*, 19(3):335–49. 110

Swan, T. (1956). Economic Growth and Capital Accumulation. *Economic Record*, 32(2):334–61. 21

Sylla, R. and Toniolo, G. (1992). *Patterns of European Industrialization: the nineteenth century*. Routledge, London. 21, 69, 135

Tabellini, G. (2010). Culture and institutions: Economic development in the regions of europe. *Journal of the European Economic Association*, 8(4):677–716. 98, 104

Taylor, A. (1948). *The Habsburg Monarchy 1809-1918: A History of the Austrian Empire and Austria-Hungary*. Penguin, London. 100

The Economist (2011). Regional inequality: internal affairs. Print Edition. 1

Thompson, I. and Yun, B. (1994). *The Castilian Crisis of the Seventeenth Century: New Perspectives on the Economic and Social History of Seventeenth-Century Spain*. Cambridge University Press, Cambridge. 103

Tortella, G. (1994). Economic retardation and recovery in south-western europe in the nineteenth and twentieth centuries. *Economic History Review*, 47(1):1–21. 186

Toutain, J. (1981). The Uneven Growth of Regional Incomes in France from 1840 to 1970. In Bairoch, P. and Levy-Leboyer, M., editors, *Disparities in Economic Development since the Industrial Revolution*, pages 302–315. Macmillan Press, London. 39

Toutain, J. (1987). Le Produit Interieur Brut de la France de 1789 a 1982. *Economies et Societes*, 15:1–237. 39, 40, 45, 46, 53, 54

Trichet, J. (2011). Competitiveness and the smooth functioning of EMU. Lecture at the University of Liège, Liège. 188

University of Lleida (2012). HGISE: Historical GIS of Europe. 77, 78

U.S. Bureau of Railway Economics (1915). *Comparison of Railway Freight Rates in the United States (and) The Principal Countries of Europe, South Australia, and South Africa*. U.S. Bureau of Railway Economics, Washington, D.C. 77, 123

USGS (2013). Mineral Facilities of Europe - Open-File Report 2010-1257. 80, 86, 110

V. Alexander and Co. (2013). Seaports of the World by Country. http://www.valexander.com/wtr/wtr$_site/world_ports.asp.$ 123

Van der Beek, K. (2010). Technology-Skill Complementarity on the Eve of the Industrial Revolution: New Evidence from England (1710-1772). Ben-Gurion University Working Paper. 162

Van Vleck, V. (1997). Delivering Coal by Road and Rail in Britain: The Efficiency of the "Silly Little Bobtailed" Coal Wagons. *The Journal of Economic History*, 57(1):139–60. 79, 123

Vincent, D. (2000). *The Rise of Mass Literacy: Reading and Writing in Modern Europe*. Polity Press, Cambridge. 166, 199

Vincent, D. (2004). *Literacy and Popular Culture: England, 1750-1914*. Cambridge University Press, Cambridge. 166

Voigt, S. (2013). How (Not) to Measure Institutions. *Journal of Institutional Economics*, 9(1):1–26. 87

Wang, Z. and Winters, L. (1992). The Trading Potential of Eastern Europe. *Journal of Economic Integration*, 7(2):113–36. 2

Ward, M. and Devereux, J. (2003). Measuring British Decline: Direct Versus Long-Span Income Measures. *Journal of Economic History*, 63(3):826–851. 53

Weinstein, W. and Adam, J. (2008). *Guesstimation: Solving the World's Problems on the Back of a Cocktail Napkin*. Princeton University Press, New Jersey. 83

West, E. (1978). Literacy and the Industrial Revolution. *Economic History Review*, 31(3):369–383. 167, 168

West, E. (1985). Literacy and the Industrial Revolution. In Mokyr, J., editor, *The Economics of the Industrial Revolution*. Rowman and Littlefield, Lanham. 162

Wilcox, W. and Fercenzi, I. (1929a). *International Migrations, Volume I: Statistics*, chapter Part II: Statistics of Migrations, International Tables, pages 255–352. National Bureau of Economic Research, New York. 11

Wilcox, W. and Fercenzi, I. (1929b). *International Migrations, Volume I: Statistics*, chapter Intercontinental Emigration According to National Statistics: Europe, pages 85–136. National Bureau of Economic Research, New York. 11

William Ward, A., Prothero, G., Mordaunt Leathes, S., and Benians, E., editors (1912). *The Cambridge Modern History Atlas*. Cambridge University Press, London. 99, 202

Williamson, J. (1995). The Evolution of Global Labor Markets since 1830: Background Evidence and Hypotheses. *Explorations in Economic History*, 32(2):141–196. 59

Williamson, J. (1996). Globalisation, convergence, and history. *Journal of Economic History*, 56:277–306. 32, 59, 60, 70, 71

Williamson, J. (1998). Globalization, Labor Markets and Policy Backlash in the Past. *Journal of Economic Perspectives*, 12(4):51–72. 19, 20, 42

Woessmann, L. (2003). Specifying Human Capital. *Journal of Economic Surveys*, 17(3):239–270. 165, 167

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. 28, 29, 34, 118, 126, 150, 153, 155

Wolf, N. (2010). Regional GDP Across Germany: Some first estimates at the level of NUTS-1, 1895-2000. Mimeo. 35

World Bank Group (2014). Portfolio: World Bank Group Infrastructure Commitment: FY00 through FY12. 187

Wright, G. (1990). The origins of American industrial success, 1879–1940. *American Economic Review*, 80(4):651–68. 76, 110

Wrigley, E. (1961). *Industrial growth and population change; a regional study of the coalfield areas of north-west Europe in the later nineteenth century*. Cambridge University Press, Cambridge. 5, 75

Wrigley, E. (1988). *Continuity, Chance and Change: The Character of the Industrial Revolution in England*. Cambridge University Press, Cambridge. 75

Wrigley, E. (2010). *Energy and the English Industrial Revolution*. Cambridge University Press, Cambridge. 14, 17, 29, 73, 75, 185

Wyplosz, C. (2013). Eurozone Crisis: It's About Demand, not Competitiveness. www.tcd.ie/Economics/assets/pdf/Not$_c$*ompetitiveness.pdf*. 188

Yan, M. (2014). Dijkstra's algorithm. Presentation, MIT, http://math.mit.edu/ rothvoss/18.304.3PM/Presentations/1-Melissa.pdf. 13, 84

Young, A., Higgins, M., and Levy, D. (2008). Sigma Convergence versus Beta Convergence: Evidence from U.S. County-Level Data. *Journal of Money, Credit and Banking*, 40(5):1083–1093. 61

Zamagni, V. (1993). *The Economic History of Italy, 1860-1990*. Oxford University Press, Oxford. 115