European Regional Income, Institutions, and Market Access, 1870-1910[1]

Paul Caruana-Galizia[2]

25 March 2013

Abstract
This paper documents large regional income differentials for a number of late-nineteenth century European countries. Such differentials do not fit well with standard explanations from growth theory that emphasise the importance of physical capital, which is freely mobile within countries. Using new data on regional institutions and market access, it provides empirical support for institutional explanations of growth and, to a greater extent, market access – or second-nature geographical – explanations.

[1] Prepared for the All-UC-Huntington-Caltech conference 'Economic and Political Inequality: Social Mobility, Economic Growth, and Inequality,' 5 April 2013.This paper draws on my doctoral research, supervised by Max-Stephan Schulze and co-supervised by Nicholas Crafts, funded by the UK Economic and Social Research Council and an EU-Malta Government Steps Scholarship.
[2] Ph.D. candidate, Department of Economic History, London School of Economics, p.caruana-galizia@lse.ac.uk.

1. Introduction

Regional per capita income levels in Europe have long displayed high inequality and a persistent spatial structure. Studer (2009), for example, traces the rise of north-western Europe relative to central and southern Europe back to the commodity market integration it enjoyed in the 17th century. Using real wage series, Allen (2001) shows the rise of this same region, particularly London, at an equally early date. This pattern of income distribution has persisted to the present day, with wealth concentrated in the north-west and low income levels around the eastern and southern geographical periphery.

This paper is an attempt at understanding these patterns. I systematically document the levels of within country per capita income inequality for late-nineteenth century Europe, when industrialisation was spreading across the continent and widening income differentials. Focusing on within-country developments sheds light on the extent to which the causes of income differentials are national, regional or idiosyncratic. This has important implications for the theories and concepts we use to explain those income levels. Within-country differences in physical capital, for example, are unlikely to be the main cause, as capital mobility within national borders is free. There are, therefore, differences in the efficiency of production across subnational regions.

As Acemoglu and Dell (2010) write, the dominant empirical approach for this type of work is the neoclassical (Solowian) growth model. In this model, growth and output are functions of human capital, physical capital, and technology. Technology is exogenous, so emphasis is placed on the capital stock dynamics. This paper, like Acemoglu and Dell (2010), argues that, given the mobility of physical capital within countries, the neoclassical model offers limited insight into the production efficiencies and hence income levels of regions within countries. The authors put forward institutions – '…defined as the rules determining how collective decisions are made…' - as an explanation of within country differences. These "local" institutions, they argue, determine the provision of local public goods and the security of local property rights, which are the promoters of economic growth (Acemoglu and Dell 2010: 170).

This focus on within-country institutions is unusual. Most of the economics and economic history literature, either explicitly or implicitly, argues that while institutions are useful in explaining cross-country income differentials, they are not useful in explaining within-country differences, as regions share the same national institutions. Redding and Sturm (2008), for example, exploit the division of Germany following WWII, and the reunification of Germany in 1990, as a natural experiment to provide evidence for the importance of market access in economic development. The units of analysis are cities, which showed varying rates of population change, a proxy for economic performance. Western cities near the division border went from being at the heart of an integrated Germany pre-division, to being at the periphery of Europe post-division. They find that cities in the Western half closest to border suffered population decline whereas cities farther away from the border, still in the West, enjoyed population growth. As the authors conclude after their empirical analysis,

> *It would be difficult to explain the observed pattern of estimates with the other…leading explanations for differences in economic performance, namely differences in institutions or natural endowments… As both border and non-border cities are part of the same country during all years of our sample, there are no obvious differences in institutions between our treatment and control cities that could be responsible for the decline of the border cities* (Redding and Sturm 2008: 1782-3).

The authors do, however, show the empirical importance of market access in explaning within-country differentials – something Acemoglu and Dell (2010) ignore. Indeed, in recent empirical study on European regions in the late-twentieth century, Breinlich (2006: 608) finds that a one per cent increase in a region's market access is associated with up to a 0.25 per cent increase in its per capita income level. This backs up the importance of market access on a European regional level.

In this paper, I contribute to the debate, and our understanding of Europe's spatial income structure, in three ways. First, I use a direct measure of regional institutions, rather than Acemoglu and Dell's (2010) proxy of paved roads. This variable also measures the provision of goods and distribution of resources, being drawn from a production model of human capital (output) and income (input). Second, I assess the relative importance of this institutional measure against a newly estimated regional market access variable. The latter is estimated using a gravity equation of national bilateral trade, applied to the regional level. Third, I use an instrumental variable framework to deal with the endogeneity issues that plague these analyses. As always, instruments do not solve everything, but, along with the new market access and institution data, they do get us closer to an explanation of within-country income differentials.

Summing up these issues, my directional research questions are: What are the relative contributions of market access and institutions to European regional per capita income? Do the empirical results fit with existing (national-level) concepts and theories?

2. Theories of income differentials

The neoclassical growth model is the starting point for most explanations of growth rates and income levels between countries, as well as between regions within countries. The model has no theory of technology differences, and has little to say of differences in human capital. Its central focus is the dynamics of human capital. Mankiw et al. (1992) seminally argued that the model explains a good deal of cross-country differences in income per capita, without significant technology differences. Another leading paper, by Barro and Sala-i-Martin (1995), finds convergence across OECD countries, across US regions, and argues that cross-country growth dynamics are best explained through the closed-economy neoclassical growth model.

While influential, this approach poses a number of problems. First, to account for the large income differentials, technology differentials would also need to be large and this is often not the case (Hall and Jones 1999). Second, as is the thrust of Acemoglu and Dell (2010), the *closed-economy* neoclassical growth model is unsuitable for the explanation of within country differences, given the absence of internal barriers to physical capital mobility. There is also a more generic criticism of this growth model. It is unable to provide a *fundamental* explanation of growth and income levels. The factors used to explain growth, '...are not causes of growth; they *are* growth' (North and Thomas 1973: 2). This is what leads onto the explanation of institutions – broadly speaking, the constructs that shape human behaviour, through property rights and politico-legal systems.

A fundamental approach is what is needed here because, as mentioned at the outset, the spatial structure of European regional income levels has been persistent over time, indicating the work of fundamental causes. That the pattern is intrinsically spatial, also calls for some sort of spatial explanation. The most successful type has been the market access approach, sadly ignored by Acemoglu and Dell (2010). Various empirical studies have shown that market access, the accessibility of markets for demand (consumers) and supply (production inputs), is a powerful explanatory tool for spatial wage structures (Redding and Venables 2004; Breinlich 2006).

My approach therefore emphasises the following potential determinants of per capita income levels in within-country settings:

- Institutional efficiency varies on the subnational level, and determines subnational per capita income levels. (Acemoglu and Dell 2010; Acemoglu et al. 2005).
- Market access varies on the subnational level, determines the location of production, and hence per capita income levels (Breinlich 2006). That it is an essentially spatial concept makes it appropriate for the analysis of spatial income structures, as found in late-nineteenth century Europe.

3. European regional income differentials

Caruana-Galizia and Marti-Henneberg (2013) provide a standardised data set of European regional GDP per capita in 1990 Geary-Khamis dollars. The data cover 199 regions from seven countries for benchmark years between 1870 and 1910, and are displayed in figure1.

FIGURE 1 AROUND HERE

The geographical spread of these data is even. It covers northern Europe (Britain, 12 regions), the eastern periphery (Austria-Hungary, 22 regions), Scandinavia (Sweden, 24 regions), the south-west (Italy, 16 regions and Spain, 17), along with central (Germany, 23 regions), and the Western Europe (France, 85 regions). This coverage provides much exploitable heterogeneity both not just within, but also between countries. Secondly, determinants of income are often related to an economy's location (say, climate). Focusing on one particular area might exclude confounding variables, but perhaps more importantly would introduce a sample selection bias.

Regions are generally at the NUTS-2 level, but there is some variation.[3] Sizes of regions vary according to historical accident. Germany, once a collection of some 300 principalities and other units, is a federation of very different historical entities. Hamburg (technically, a "free city," originally part of the Hanseatic League) and Bavaria, for example, maintain enormous differences in size to this day. In Sweden, northern units are much larger than those in the south because they are more sparsely populated. Does variation in regional size affect my empirical results? Not substantively, which is unsurprising given the findings in Briant et al. (2010) that areal shapes and sizes are of at best secondary importance to specification issues.[4] Ultimately, I analysing the effects of institutions and market access on per capita GDP, so regional income levels are already normalised by regional population sizes.

Table 1 shows the summary statistics for the GDP per capita data by benchmark year. The mean shows generally increasing per capita income over the period: at a compound annual rate of 1.05 per cent. The difference between the richest and poorest regions is striking, and increases over time. The highest income region is richer than the lowest by a factor of seven in 1870; by a factor of eight in 1900; and by a factor of nine in 1910. While the standard deviation also increases over the period, the coefficient of variation, which can be taken as a measure of income inequality, remains at the same level; declining only slightly between 1870 and 1900.

TABLE 1 AROUND HERE

When calculated in this way, however, the coefficient of variation is obscuring much of the income variation of interest. As we saw in figure 1, within-country variation is high. Calculating Theil indices allows us to see how much of inequality stems from differences between countries and within countries. The Theil index's strength is that it can be decomposed into nested geographical levels, that is, an index of inequality *between* and *within* spatial units. In this, setting inter-*regional* inequality can be decomposed into

3 NUTS (Nomenclature of Units for Territorial Statistics) is a European Union level of aggregation largely based on historical administrative boundaries. The levels range from NUTS-1 (for example, groups of states); to NUTS-2 (for example, states or provinces); and NUTS-3 (for example, counties).

4 Dropping France altogether from my OLS baseline estimation, model (10), does not change coefficient magnitudes or significance, except for the coefficient on *Dtcoal* (distance to coal), which turns out to be insignificant in all other standard estimations. A jackknife estimation by region of the OLS baseline produces identical standard errors.

inequality within and between *countries*. Formally, for a given level of income distributed across the number D of regions, the Theil index is defined as follows:

(1) $$T = \sum_{d=1}^{D} \frac{Y_d}{Y} \ln \frac{Y_d}{Y/_D}$$

where Y_d is the actual level of income (GDP per capita) within region d, while $R = \sum_{d=1}^{D} Y_d$ denotes the total level of income. The index is equal to zero when the level of income is uniformly distributed across regions: $Y_d=Y/D$ for all d. Conversely, if the whole level of income were to be concentrated in only one region, the index takes on the positive value lnD=2.298 (for D=199). The greater the value of the Theil index, the higher the level of inequality.

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

(2) $$T= T_w+T_b.$$

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

(3) $$T_w = \sum_{r=1}^{R} \frac{Y_r}{Y} T_r$$

where R is the number of countries, and $Y_r = \sum_{d=1}^{D_r} Y_d$ 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:

(4) $$T_r = \sum_{d=1}^{D_r} \frac{Y_d}{Y_r} \ln \frac{Y_d}{Y_r/D_r}.$$

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

(5) $$T_b = \sum_{r=1}^{R} \frac{Y_r}{Y} \ln \frac{Y_r/D_r}{Y/D}.$$

Table 2 shows the index values obtained for GDP per capita. The last row shows the percentage difference between the within- and between-country values.

TABLE 2 AROUND HERE

The indices show that GDP per capita inequalities *within* countries are much higher than inequalities *between* countries. Indeed, Bourguignon and Morrison (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, inequalities within countries grew. These figures are interesting because they show that, despite within-country differentials being larger, most theoretical and empirical work is directed at understanding cross-country differentials.

4. Measuring regional institutions

Institutions have been successful in explaining cross-country income differentials, but it is usually assumed that they cannot account for subnational differentials since they are national structures (Redding and Sturm 2008). Acemoglu and Dell (2010) show that both de jure (federal states) and de facto (Mafia control of southern Italy) institutions vary subnationally. This much is easy to agree with, and it has long been recognised by historians (Banfield 1958; Pollard 1973; Fukayama 2001). My conceptual approach differs, however, in explaining subnational institutions themselves.

According to Acemoglu et al. (2005: 389-90) institutions are exogenous: '[Institutions]...influence not only the size of the aggregate pie, but how this pie is divided among different groups and individuals in society....(the distribution of wealth, physical or human capital).' But as Clark (2007) argues, while institutions may determine the efficiency of resource allocation, they are determined by economic forces, and vary across time and place because of differences in relative prices, technology, and consumption patterns. That is, institutional efficiency matters, but is driven by economic fundamentals, and will in the long run move toward efficiency as all actors gain from efficient institutions. I am not as optimistic as Clark (2007), given we have numerous historical examples of the persistence of inefficient institutions (Spanish Inquisition lasting for almost four centuries), but do agree that it is institutional efficiency that matters – at least in the short or medium run, by Clark's (2007) argument - and that it is shaped by underlying economic forces. Taking this into account, I use a production frontier model to estimate institutional efficiency as the rate at which economic production (GDP per capita) is converted into "human development" (literacy):[5]

$$\ln L_i = \alpha + \sigma \ln Y_i + \varepsilon_i \qquad (6)$$

where L is the literacy rate of region i (sources detailed in the appendix), and Y is GDP per capita. In a frontier model, the error term ε can be decomposed into two elements: v and u. The former random disturbance, and u is asymmetric disturbance, or technical efficiency, calculated as the expectation of the term u conditional on the composed error term ε. It is the ratio of observed output in i relative to potential (frontier) output, and is $0 \leq 1$. I run cross-sectional models, producing efficiency estimates for each region relative to the benchmark year frontier.

[5] Frontier model rather than OLS, since we are not interested in the means, but the relative regional input-output ratios.

While this approach is admittedly not ideal, it has some advantages. First, it is data-driven and systematic, avoiding intractable institutional explanations as in Greif (2006). Second, it can be applied to different contexts unlike the case study specific proxies as in Acemoglu and Dell (2010) with paved roads. Third, it can provide data for empirical analysis at any for which we have relevant input and output data. In this case, the relationship has a sound historical basis. In the late-nineteenth century, governments began public education plans in earnest, with specific goals to improve literacy; GDP proxies the taxable base available to governments to fund those plans.[6]

The results in table 3 show an expectedly persistently significant and positive relationship between literacy rates and per capita income. $\ln(\sigma_u^2)$ is the parametised asymmetric error component u; and $\ln(\sigma_v^2)$ is the parametised random error component. The two variances of the two error components indicate that the inefficiency component u is much more statistically significant than the random component v. This implies that inefficiency u makes a more important contribution to the variability of the total error in the cross-sectional frontier model, and that inefficiency is highly significant across regions and years. Exponentialising $\ln(\sigma_u^2)$ parameters, we see a similar declining effect of efficiency on literacy rates, as with its relationship with GDP per capita: a one percentage point increase in efficiency resulted in a 50 per cent increase in literacy rates in 1870; 15 per cent in 1900; and 9 per cent in 1900.

TABLE 5 AROUND HERE

Table 6 summarises the regional (institutional) efficiency (*INST*) results, calculated as $\exp[-E(u_i|\varepsilon_i)]$. The results indicate stark differences between countries. Italy and Spain persistently stand out as mean low-efficiency countries, with high standard deviations of within-country efficiency. That is, their mean scores are low and there is high inequality of efficiency. Germany along with Sweden represent the frontier throughout the estimations. Sample mean efficiency rose over time, while its standard deviation declined. Looking across the sample, the main implication of these results is that the average region could have reduced inputs (GDP per capita) by around 20 per cent without reducing their output (literacy rate), simply by improving their institutional efficiency.

TABLE 6 AROUND HERE

This approach uses per capita income as an input to capture the idea that institutions are not completely independent of the economy. Since I am

[6] Engerman and Sokoloff (2000: 227) on Latin America during this period: '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.'

ultimately regressing per capita income on *INST*, there is a risk of redundancy. In their analysis of the Human Development Index (HDI) and its redundancy versus GDP per capita, McGillivray and White (1993) propose two criteria of 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 criterion, since the only component or input is GDP per capita, I can test the first. The Pearson (Spearman) correlation coefficients are: 0.34 (0.26) in 1870; 0.64 (0.57) in 1900; and 0.57 (0.48) in 1910. While the correlation coefficients are statistically significant, as we would expect them to be, they pass the 'Level 1' and 'Level 2' redundancy criteria. That is, *INST* contains useful information beyond its GDP per capita input.

4.2. Measuring regional market access

Market access is based on the circular logic that producers locate in sites with good access to (potential) consumers. Sites with good access are ones that already have producers, which draw in consumers. What came first – producers or consumers – is historical accident. The point is that once this circular logic is set in motion, it is self-reinforcing, as producers enjoy increasing returns and economies of scale (Krugman 1991). While market access can be measured in various ways, it is most often in the literature estimated using some form of a gravity equation of bilateral trade. While trade do exist for some European regions (Wolf 2007), most countries did not record regional trade flows. For this reason, I follow the procedure in Breinlich (2006) and use national-level bilateral trade data to arrive at estimates for regional market access. The trade data are from Jacks et al. (2005), 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.

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.[7] 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). Breinlich (2006) 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 (and so market access) 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

[7] For a more detailed exposition of the theory, readers should refer to Breinlich (2006), or the broader literature on gravity trade models and market access as in Redding and Venables (2004).

distance. As Breinlich (2006) 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.

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

$$(7) \qquad T_{ij} = dist_{ij}^{\beta_0} \times \left(\prod i \in i \ldots j \exp(border_i)^{\beta_1}\right) \times \exp(language_{ij})^{\beta_2}.$$

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:

(8)
$$\ln\left(\frac{X_{ijt}}{E_{it}E_{jt}}\right) = \alpha_t + \gamma_{1t}\ln\left(dist_{ij}\right) + \sum_i \gamma_{2it}\, border_i + \gamma_{3t} language_{ij} + \delta_{1it} exporter_{it} + \delta_{2jt} importer_{jt} + \varepsilon_{ijt}$$

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, δ_{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 (2) on 14-year windows of Jacks et al.'s (2005) data between 1870 and 1920.[8] There are no data prior to 1870, which is why the final window must extend slightly beyond my precise period. The econometric results are displayed in table 3. 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 Breinlich (2006). 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 Schulze and Wolf's (2009) findings that political borders and ethno-linguistic networks matter for economic integration. Unlike Schulze and Wolf (2009), 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 access as in Breinlich (2006). As in all gravity models, distance is highly significant and negative – in all periods. This captures

[8] 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.

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, indicating some sort of improvement in transport technologies. Jacks et al. (2005) see the same decline in the size of distance coefficient over the 1870 to 1939 period for similar reasons.

TABLE 7 AROUND HERE

As is standard in the literature, market access 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 access for each region *i* at each year *t* for all countries *j* as follows:

$$(9) \qquad MA_{it} = \sum_{j\varepsilon ctyi}\left(e^{\hat{\gamma}_{3t}}dist_{ij}^{\hat{\gamma}_{1t}}\right)\hat{\delta}_{jt}^{-1}E_{jt} + \sum_{j\varepsilon ROS}\left(e^{\hat{\gamma}_{2it}}e^{\hat{\gamma}_{3}language_{ij}}dist_{ij}^{\hat{\gamma}_{1t}}\right)\hat{\delta}_{jt}^{-1}E_{jt}$$

where $\hat{\delta}_{jt}^{-1}$ and the parameters $\hat{\gamma}_{1t}$, $\hat{\gamma}_{2it}$, and $\hat{\gamma}_{3t}$ were estimated in the gravity trade equation (7), and E_{jt} is again proxied by a region or country's GDP in year t. My GDP data has already been described, and is 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 a variety of atlases. Following Breinlich (2006), 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 km². This formula, or variants of it (e.g. Schulze 2007; Crafts 2005), 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 Breinlich (2006), equation (8) 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 $\hat{\delta}_{jt}$ per year per country. It is an inevitable part of using national-level trade data to derive regional level market access values.

It is worth comparing the market access values I estimated here, using this method, to those constructed in Caruana-Galizia (2012), which were derived for the sample, using a historical geographical information system of observed transport networks and costs. The two are not strictly comparable since the latter measure is ad hoc, and not used when enough data for a trade equation are available (most often in regional settings). Caruana-Galizia's (2012) measure follows the Harris (1954) formulation, which is the sum of incomes divided by distance. The GDP data are the same as here, and distances were calculated using a network that covers railways and shipping lines, using freight rates taken from a variety of sources. While we should expect some correspondence between the two measures, there is also likely to be a gap from measurement error. Breinlich (2006, p.605), for example, regresses regional gross value-added on both the estimated and Harris-type measure, finding that the coefficient on the Harris-type measure is 44 per cent larger. Both are highly significant. Using the same specification on my data, I also get a coefficient on the Harris-type measure that is 26 per cent larger than the one estimated here. Both coefficients are statistically significant, and the root mean square errors of both regressions are essentially the same: a difference of seven per cent. The constant terms are also very similar, showing a difference of only 14 per cent. Table 4 shows some more summary statistics for the two measures.

TABLE 8 AROUND HERE

Both variables have means and medians that are effectively the same, differing by only one per cent in both cases. The maximum values show some divergence (15 per cent), but then the minimum values show a difference of less than one per cent. The biggest differences between the two variables are to be found in their variance and standard deviations. The constructed measure of market access shows much less variance. Most of its values are concentrated on the right tail of its distribution, with a skewness value of -2.376 compared to -0.288 for the estimated variable, which shows a much more varied distribution. The Pearson correlation between the two is 0.21, significant at the one per cent level. In sum, while there are differences between the two measures, given they are similar in many ways and correlated; both capture or correlate with the "true" value of market access. Hanson (2005) argues that the estimated version of market access is a more useful measure than the Harris-type construction, as the estimation procedure nests purchasing power with other variables and lends structural interpretations to the estimated coefficients. These structural parameters reflect the magnitude of economies of scale and transport or trade costs.

9. Empirical strategy

The aim is to estimate the following model:

$$(10) \quad \ln(Y_{it}) = \alpha + \gamma \ln(MA_{it}) + \beta \ln(INST_{it}) + \theta' X_{it} + \varepsilon_{it}$$

where Y GDP per capita in region i at year t, *MA* is the estimated market access term, *INST* is the measure of institutions discussed earlier, and α and ε are a constant and an error term with standard properties. 'X is a matrix of potential determinants of income levels. Natural resources figure prominently in the debate on European industrialisation. Allen's (2009) argument of early British industrialisation rests on Britain's availability of cheap coal. This argument has not gone uncontested. Clark and Jacks (2007) find that English possession of coal reserves made a negligible contribution to Industrial Revolution incomes. I calculate the log great circle distance from each regional node to the nearest coal deposit for a measure of *X*. I expect an inverse relationship between distance to coal deposits and income levels, but how strong or significant this relationship would be is less clear. Throughout the paper, I will be examining the size, sign, and significance of the three coefficients γ, β, and θ.

It is worth looking at the simple correlations between GDP per capita and each one of the variables. Figure 2 shows scatter plots of GDP per capita against each variable. All plots show a clear relationship with GDP per capita, perhaps less for the distance to coal measure. Given there is debate in the literature about the importance of coal access during industrialisation, this weaker (negative) correlation is unsurprising. The other two plots show a more unambiguous positive relationship. It appears that any one of these variables has the potential to explain GDP per capita levels. The signs of institutions (positive), market access (positive) and distance to coal (negative) are all as expected. The *INST* and *MA* coefficients are statistically significant at the one per cent level; *Dtcoal* at five per cent. Regions with stronger institutions, better market access, and ones that are closer to coal deposits are more likely to have higher levels of income.

FIGURE 2 AROUND HERE

Things are not so straightforward, unfortunately. Reverse causality, omitted variable bias, and measurement error plague these relationships. Higher income, for example, leads to higher market access levels, since *MA* includes own-region demand (income). By construction of *INST*, there is also reverse causality the institutions-income relationship. Further, there is likely to be measurement error in *INST*, as institutional efficiency is difficult to quantify. While no method is perfect in dealing with these issues, I employ a two-stage least squares estimation strategy. The identification strategy is to use instrumental variables (IVs) to isolate plausibly exogenous variation in market access and institutions – distance to coal is exogenous – and use that variation to explain regional income levels.

9.1. Identification

To instrument market access, I use the log great circle distance to London (for each regional node), *DTL*. Both Redding and Venables (2004) and Breinlich (2006) use to distance to Luxembourg and Brussels, which provides exogenous geographical variation that captures the market access advantage of locations close to the economic centre of Europe. Distance to London makes more sense here, as it was by far the most important *economic* node in my sample, across all benchmark years.

As an instrument for institutions, I exploit a historical critical juncture. In 1556, the Habsburg family officially split its rule over Europe – in reality the split started in the 1520s. The Spanish Habsburgs took control of Spain, and Southern Italy and Milan, while the Austrian branch took control of Austria-

Hungary and indirect control of the Holy Roman Empire (what is mostly Germany today). Becker et al. (2011), in a particular relevant paper on the Austrian Habsburgs, empirically show the current effects of long-gone formal institutions, which operate through the channels of current cultural norms, values, beliefs and formal institutions. Unlike Clark's (2007) argument, their research shows that historical institutions do leave a legacy – even in the long run. This is what made the Habsburg division matter.

The Austrian Habsburgs, led by Ferdinand I, established royal authority over fragmented principalities. Ferdinand was quick to harmonise legal and financial systems across Austria-Hungary. As early as 1526, before the official split, he issued a decree which formalised a central government, composed of an Aulic Council (judicial), Privy Council, court chancellery, and a chamber of accountants. This was a stark difference to the preceding fragmented and often conflicting authorities of Bohemia, Moravia, and all other "sovereign" provinces. As Berenger (1990: 162) put it, wrote, 'jurisdiction extended over the whole monarchy, without distinction between countries and particular privileges.' In short, it was the start of institutional consolidation that continued, as Becker et al. (2011) show, into the nineteenth century. The authors give the example of Czech public offices still opening nowadays at 7am, as Emperor Franz Josef (who died in 1916) required. A.J.P. Taylor (1948: 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.'

We see this same consolidation in the other countries in my sample – except the Spanish Habsburg lands. Britain, as every economic historian knows, went through the Glorious Revolution in 1688; France enjoyed a period of state consolidation even earlier, in the early sixteenth century. Its eighteenth century revolutions, while bloody in the short term, paid off soon after (Acemoglu et al. 2011). While not prosperous, Sweden was stable and peaceful. Two years before the Glorious Revolution, it passed the Church Law of 1680, obliging priests to make sure that every young person in their parish could read (the Bible). This was the same year as Sweden's "Great Reduction:" when Charles XI recaptured lands granted to the nobility, and greatly curbed the power of aristocratic families to make the state solvent and able to pay its debts. Its unintended consequence was the development of a strong, clear organisation of Sweden's finances and government, more generally (Roberts 1968). These changes were not taking place across Spanish Habsburg territory.

Charles V shared little in common with his reformer brother, Ferdinand I. He had an 'aristocratic and medieval conception of patrimony…matters of government and matters of family for him were closely connected' (Berenger 1990: 140-4). It is with this approach that he governed his territory. Berenger (1990: 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 V was financing his wars with France, as he spent money on little else, through credit from German and Antwerp bankers. The uncontrolled accumulation debt throughout Habsburg Spain is a clear indicator of weak institutions We can see more institutional decay quite dramatically in the Spanish Inquisition, which Philip II greatly expanded in the mid to late sixteenth, making church orthodoxy a goal of public policy. The Inquisition was ended in 1834. Following from this, we see the expulsion of the industrious Moriscos, contrasting greatly with the societal trust between citizens and their public institutions Becker et al. (2011) found in the Austrian dominions. Indeed, attempts at reforming Spain's bloated and inefficient bureaucracy running through to the mid seventeenth century were met with staunch resistance (Elliott 2009).

The Spanish Habsburgs also ruled over the Milan region (equivalent to Lombardy), Sardinia, Sicily, and much of southern Italy. We normally think of Lombardy as a developed region not just of Italy, but of Europe. Indeed, while under Spanish rule, Lombardy managed a mean $INST$ score of 0.738: respectable, but still below the sample mean in table 6. As Tabellini (2010) documents, even though Charles V 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. The Lombard Senate had, in the words of Tabellini (2010: 142) '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: 145) describes simply as 'absolutist and autocratic,' giving them the lowest rank in his institutional scoring system. Berend (2013: 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 striking historical institutional differences between Spanish Habsburg regions and all others, and given the persistence of institutions, I define my IV $SH = \{1\ if\ under\ Spanish\ Habsburg\ rule\ by\ 1700; 0\ otherwise\}$, where 1700 marks the year of the Spanish War of Succession, when the Austrians won back most of the Spanish territories. SH=1 for 12 per cent of my regions.

9.2. Implementation and results

In the first stage models (12) and (13), $INST$ and MA are regressed on the exogenous determinant $Dtcoal$ and the IVs, DTL and SH. The predicted values, along with $Dtcoal$, are then used in the second stage GDP per capita model (11):

$$(11) \quad \ln(Y_{it}) = \alpha + \gamma\ln\left(\widehat{MA}_{it}\right) + \beta\ln\left(\widehat{INST}_{it}\right) + \theta Dtcoal_i + \delta_t + \varepsilon_{it}$$

$$(12) \quad \ln(INST_{it}) = \alpha + \pi\ln(DTL_i) + \varphi\ln(SH_i) + \theta Dtcoal_i + \delta_t + \varepsilon_{it}$$

(13) $\ln(MA_{it}) = \alpha + \pi\ln(DTL_i) + \varphi\ln(SH_i) + \theta Dtcoal_i + \delta_t + \varepsilon_{it}$

I included country-by-year fixed effects - δ_t - to control for national-level policies. This is particularly relevant, since during this period governments were expanding public education programmes, and because it will absorb national-level de jure institutional effects, leaving the de facto institutional effects we are interested in with *INST*.

Table 9 shows the results of these estimations. The column OLS shows the baseline OLS implementation. Backing up figure 2, the coefficients are all highly significant, and show the expected signs. The hierarchy seems clear, with *INST* at the top, followed by *MA*, but to compare the coefficients more accurately, we can standardised them as $S\beta_x=\beta_x\times(S_x/S_y)$, where the standardised coefficient ($S\beta$) on some independent variable x is the normal coefficient, multiplied by the ratio of its standard deviation to the dependent variable's (y) standard deviation (S). This scales the coefficient to reflect the difference between the "spread" of the two variables. The satndardised coefficient on *INST* is 0.213; 0.210 on *MA*; and -0.124 on *Dtcoal*. The first two coefficients, when standardised, are very similar, indicating that a one standard deviation rise in either variables results in a 0.21 standard deviation rise in per capita income. This is a meaningful effect. The sample mean standard deviation of per capita income is $1,065, so the effect equals an additional $224 – or 39 per cent of the sample minimum per capita income. The effect of increasing a region's distance to coal by one standard deviation (161 kilometres) is a reduction on per capita income of $132. Do the IV results tell a different story?

The first stage estimations show the IVs perform well.[9] *SH* has, as expected, a very strong significant negative effect on institutional efficiency. Is F-statistic is also large and significant. According to these results, there is something to be said for institutional legacies. The inefficiency of Spanish Habsburg institutions outlived the Spanish Habsburgs themselves. *DTL* is also a strong predictor, showing that regions far from London have significantly lower market access. The benefits of proximity to what was Europe's richest region are captured through *MA*. The F-statistic here is admittedly small, but still significant at the five per cent level.

TABLE 9 AROUND HERE

Moving onto the second stage, *Dtcoal* remains negative, but is now insignificant. Instrumenting for institutional efficiency and second-nature geography (market access), access to coal was not that important. This would contradict some, like Pomeranz (2000) where a large part of the "great

[9] The reduced form is also encouraging. Controlling for year-by-country fixed effects, and clustering robust standard errors on regions, the effect of *SD* (*DTL*) on GDP per capita is -0.397 (-0.186), significant at the one (one) per cent level.

divergence" narrative is based on coal access, but fits with more recent empirical research that shows the limited contribution of coal to industrialising economies (Clark and Jacks 2007). This result does, however, highlight the importance of correctly specifying "geography." Acemoglu et al. (2002) take natural endowments as their measure of geography and Rodrik et al. (2004) use a host of first-nature (physical) geographical controls. Using these measures, both find that "geography does not matter." They fail to consider relative, or second-nature, geography. That is, the spatial linkages between countries and regions that may affect income levels and economic activity more generally. This is the point of market access: it captures the benefits of being economically central. It is particularly relevant when income patterns show a clear spatial structure, as in figure 1.

In fact, standardising the *MA* and *INST* coefficients, we see that market access has a 1.95 standard deviation effect on per capita income, compared to a 0.763 effect for institutional efficiency. This simple second-nature geographical variable has a stronger effect on per capita income levels than what is the dominant – at least at the national level – explanation at the moment. The market access effect is a substantive revision upwards from the OLS estimate. It equals an additional $2,076 on per capita income; rather than the $224 reported earlier. It is likely that the application of national-level trading costs overstate costs when applied to the regional-level, which would bias the OLS estimate downwards. *INST*'s IV estimates may be larger because the exploited assignment of Spanish Habsburg rule picks up returns on institutional efficiency for a group for whom it is particularly large. Given these considerations, it is expected that IV estimates will be larger in magnitude.

By way of illustration, if Sicily improved its institutional efficiency from 29 per cent in 1900 to the mean level of 85 per cent; its resulting improvement in per capita income would be an additional $3,284, which is already higher than the sample mean per capita income.[10] Unfortunately, Sicily's institutional efficiency only made it to 42 per cent by 1910, but the result does indicate that Sicily was economically over-performing relative to its (low) institutional efficiency. While it was dragged down by low institutional efficiency, the island enjoyed a market access level that was in 1900 4.3-times greater than the sample mean level; a function of its central Mediterranean location and relative proximity to markets like those of Lombardy and Catalonia.

The IV results tell us that while de facto subnational institutions matter for development, they are not as important – in terms of magnitude - as market access; coal is not important at all. That is, the potential gains from improving market access were greater than those from improving institutional efficiency. This raises a number of broader issues with the literature on relative economic performance.

10. Some Implications

[10] The increase in institutional efficieny is equivalent to 2.7 standard deviations, so 2.7×0.763

The regional GDP per capita inequalities I document and analyse in this paper cannot be explained – at least entirely, as neoclassical theory would suggest – by differences in physical capital intensity of production. This is because within countries there are no formal barriers to capital mobility. Inequalities within countries, as I have shown, are determined by region-specific characteristics.

More specifically, regions' have de facto institutions, and do not simply operate under their national de jure institutions. These de facto institutions have effects on regional incomes, in the same way that cross-country income differentials are explained by institutional differences. This is not just a matter of different geographical scales or increasing granularity. Again, neoclassical theory depends on factor immobility *between* countries, but not within them. Recent work is beginning to deal with this concept of varying de facto institutions on different geographical scales (Naritomi et al. 2012; Acemoglu and Dell 2010), but it still relies on proxy variables to measure institutions. A lack of evidence of regional institutions has been an issue in empirical work. Indeed, Redding and Sturm's (2008) conclusion that institutions do not matter because their observations are part of the same country is to say that absence of evidence is evidence of absence. Before concluding that within-country institutions do not matter or do not exist, we should first look for them. I have used a direct measure of regional institutions, finding it to be a substantive determinant of per capita incomes. It will be the task of future research to subject this measure to scrutiny, but at least for now we have direct preliminary evidence that systematic institutional differences exist within countries, that they affect income differentials within countries, and that this means neoclassical theories of growth are not applicable to inequalities within countries.

This is of course not to say that other within-country determinants of income are irrelevant. Redding and Sturm (2008) emphasise market access, as does Breinlich (2006), but it is ignored by Acemoglu and Dell (2010). My analysis has shown that market access is in fact a more substantive determinant of within-country inequalities than regional institutions. This finding highlights the importance of geography in the perhaps still unconventional sense. A number of authors (Acemoglu et al. 2001 or Rodrik et al. 2004) have dismissed "geography" altogether, but they only looked at first-nature (or physical) geographical variables. My distance to coal measure, a first-nature geographical variable, comes up as insignificant in my main IV regressions. This fits with the dim view of first-nature geography, but there are clearly other ways in which geography affects income. Market access is one of them. It is particularly useful because it is a spatial concept, and as shown in figure 1, Europe's regional per capita structure had clear spatial qualities. Rich regions were concentrated in the northwest while the geographical periphery (extreme west, east and south) was the *economic* periphery. These spatial patterns are ignored by standard growth theories.

The implications this analysis raises are both empirical and conceptual. Future research would do well to test the variables I have constructed and used here, and to construct or estimate new measures of these variables and apply them in different contexts. Acemoglu and Dell (2010), for example document considerable within-country disparities in Latin America, and account for them using a proxy for local institutions (proximity to paved roads). It would be interesting to see how far market access can go in explaining these differentials, seeing as there is a clear spatial income structure in their data.[11] There is also much progress to be made in empirically measuring within-country institutions. There are a number of ideas out there, but they are based on proxies or very specific historic events, and so have limited applicability. A more systematic method of quantifying subnational institutional characteristics, which I have tried to provide here, would be more useful. Conceptually, the empirics of my paper have highlighted some issues with the applicability of neoclassical growth theory. While we have empirically sound ideas as to what causes within-country income differentials, we are less able to fit these ideas into a broader conceptual framework. How, for example, can we integrate institutions and market access in a more theoretically rigorous way? Is it that market access and institutions act as de facto barriers to within-country physical capital mobility? The challenge of future work is to integrate empirical insights into a coherent conceptual framework.[12]

[11] A map is presented in figure A2 of the *NBER* working paper version, available at www.nber.org/papers/w15155.

[12] Acemoglu and Dell (2010) attempt to provide such an account, but do not consider the spatial structure of their labour income data.

11. References

Acemoglu, D. and Dell, M. (2010). Productivity Differences Between and Within Countries. *American Economic Journal: Macroeconomics* **2**(1), 169-188.

Acemoglu, D., Johnson, S. and Robinson, J.A. (2005). 'Institutions as a fundamental cause of long-run growth' in P. Aghion and S.N. Durlauf (Eds.) *Handbook of Economic Growth, Vol. 1A*, pp. 386-414.

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

Allen, R.C. (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.

Allen, R.C. (2009). *The British Industrial Revolution in Global Perspective*. Cambridge: Cambridge University Press.

Banfield, E. (1958). *The Moral Basis of a Backward Society*. New York: Free Press.

Barro, R.J. and Sala-i-Martin, X. (1995). *Economic Growth*. Highstown, NJ: McGraw-Hill.

Berenger, J. (1990). *A History of the Habsburg Empire, 1273-1700*. London: Longman.

Becker, S., Boeckh, K., Hainz, C. and Woessman, L. (2011). The Empire is Dead, Long Live the Empire! Long-run Persistence of Trust and Corruption in the Bureaucracy. *CESifo Working Paper Series* **3392**.

Brienlich, H. (2006). The spatial income structure in the European Union – what role for Economic Geography? *Journal of Economic Geography* **5**, 593-617.

Buyst, E. (2010). Reversal of Fortune in a Small, Open Economy: Regional GDP in Belgium, 1896-2000. *VIVES Discussion Paper* **8**.

Caruana-Galizia, P. (2012). Market Access and Human Capital: European Regions, 1870-1910. *Under review at journal.*

Caruana-Galizia, P. (2013). Estimating French Regional Income: Departmental Per Capita Gross Value Added, 1872–1911. *Research in Economic History* **29, 71-95.**

Caruana-Galizia, P. and Marti-Henneberg, J. (2013). European Regional Railroads and Real Income, 1870-1910: A Preliminary Report. *Scandinavian Economic History Review* **61**(2).

Clark, G. (2007). *A Farewell to Alms: A Brief Economic History of the World.* Princeton, N.J.: Princeton University Press.

Crafts, N.F.R. (2005). Regional GDP in Britain, 1871-1911: Some Estimates. *Scottish Journal of Political Economy*, **52**(1), 54-64.

Engerman, S. and Sokoloff, K. (2000). History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World. *Journal of Economic Perspectives* **14**, 217-232.

Fukayama, F. (2001). Social capital, civil society and development. *Third World Quarterly* **22**(1), 7-20.

Hall, R.E. and Jones, C.I. (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics* **114**(1), 83-116.

Harris, C.D. (1954). The market as a factor in the localization of industry in the United States. *Annals of the Association of American Geographers* **44**(4), 315-331.

Krugman, P. (1991). *Geography and Trade*. Cambridge, MA: MIT Press.

Mankiw, N.G., Romer, D., and Weil, D.N. (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics* **107**(2), 407-437.

North, D.C. and Thomas, R.P. (1973). *The Rise of the Western World: A New Economic History*. Cambridge: Cambridge University Press.

Pollard, S. (1973). Industrialisation and the European Economy. *The Economic History Review* **26**(4), 636-648.

Redding, S.J. and Sturm, D. (2008). The Costs of Remoteness: Evidence from German Division and Reunification. *American Economic Review* **98**(5), 1766-1197.

Redding, S.J. and Sturm, D.M. (2008). The Costs of Remoteness: Evidence from German Division and Reunification. *American Economic Review* **98**, 1766-11797.

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

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.

Roberts, M. (1968). *The Early Vasas: A History of Sweden, 1523-1611*. Cambridge: Cambridge University Press.

Studer, R. (2009). Does trade explain Europe's rise? Geography, market size and economic development. *LSE Economic History Working Papers* **129/09**.

Schulze, M.-S. (2007). Regional income dispersion and market potential in the late nineteenth century Hapsburg Empire. *LSE Economic History Working Paper* **106/07**.

Schulze, M.-S. and Caruana-Galizia, P. (2013). Empires Diverging: A Spatial Analysis of Austro-Hungarian and German Regional GDP, 1870-1910. *LSE Economic History, mimeo*.

Tabellini, G. (2011). Culture and Institutions: Economic Development in the Regions of Europe. *Journal of the European Economic Association* **8**(4), 677-716.

Taylor, A.J.P. (1948). *The Habsburg Monarchy, 1809-1918: A History of the Austrian Empire and Austria-Hungary*. Chicago: University of Chicago Press.

12. Appendix

12.2. Regional GDP data

Table 12.1 below summarises the existing regional GDP data from Caruana-Galizia and Marti-Henneberg (2013), showing the years for which data are available, number of regions, method of estimation, and currency of estimates. As mentioned earlier, most of the papers come from an international research project that aims to estimate European regional GDP in the long run, co-ordinated by Joan Roses and Nikolaus Wolf. Two sources, providing the regional GDP data for France and Germany, are independent of this project, but share its aims (Schulze and Caruana-Galizia 2013; Caruana-Galizia 2013).

TABLE 12.1 AROUND HERE

The benchmark years do not always match, but are never too far off to make comparisons across time unreliable. What we are after here is not so much precise figures of GDP at precise points in time, but approximations of relative GDP levels at intervals. Taking this approach, rather than using some method of interpolation or extrapolation, is more transparent and leads to fewer assumptions of the data. In this paper, I use only the benchmark years 1870, 1900, and 1910, as these years are the closest-common years across all regions, and because other data (literacy, mainly) were only available for these intervals.

There are two-methods of estimation used in the construction of this regional GDP dataset. Top-down approaches consist of structural equations, where regional GDP levels are specified as functions of sectoral employment, wages, and value-added, as first set out by Geary and Stark (2002). Bottom-down approaches involve the painstaking application of national accounting methods to regional-level data. Does the use of different regional GDP estimation methods introduce a bias? This would matter if, for example, top-down methods consistently under- or over-estimated income levels. This is unlikely. While short-cut methods are less accurate, Geary and Stark (2002), Enflo et al. (2010), and Buyst's (2009) robustness checks against official estimates show that the margin of error is tolerable, according to national accounting standards, and that short-cut methods do not produce results that veer far off from officially bottom-up-constructed series. Their checks show no persistent directional bias. Furthermore, in our sample, it is only Schulze's estimation of Austria that was done using a bottom-up approach. Given the above points, in the unlikely event that these estimates do stand at odds with the rest of the sample, Austria only comprises 14 of the 200 regions, that is, 7 per cent of the total sample. Any bias would thus be contained.

More interesting, is the variation in currency. For our analysis, it was necessary to standardise and deflate all the estimates to make comparison across both space and time possible. We converted them into 1990 Geary-Khamis dollars ($GK), which is a standard unit in much of the economics and economic history literature. Conveniently, Schulze's (2007) estimates for Austria-Hungary come in $GK. The estimates for Italy and Germany were already deflated; the first come in 2000 euros and second in 1913 marks, and so their conversion was straightforward. The remainders – France, Britain, Spain, and Sweden –required both deflation and conversion.

The approach we took is straightforward. We deflated nominal regional GDP estimates using a national GDP deflator, and then converted those estimates using the exchange rates implicit in Maddison's (2003) widely used data, which are in $GK. According to Prados de La Escosura (2000), these data are the best of their kind available. For regional GDP estimates that were already in constant terms (Italy and Germany), we converted directly from Maddison. For consistency, we always derived the exchange rate as the period starting year Maddison national GDP per capita divided by the starting year national GDP per capita in our sample. For example, for Italy, which had a starting year of 1871, $ER_{Italy}=GDPpc_{Maddison,1871}/GDPpc_{Felice,1871}$. We then used this same exchange rate to convert all deflated regional GDP estimates. Deflators came from Smits et al. (2009), which provides datasets of nominal and real GDP, as well as GDP deflators for, among other countries, Spain (1958 pesetas) and Britain (1913 pounds). France's deflator came from Toutain (1987). Enflo et al. (2010) provide both real and nominal figures in their paper.

Regional price variation is the main issue when deflating and standardising regional GDP figures. Ideally, regional price indices should be used to account for differences in prices across countries. This is rarely ever applied for three reasons. First, it is very data intensive. Very often regional prices for a broad basket of goods defined by region just do not exist. Indeed, Wolf writes of the difficulty in finding such data for Germany, a usually well-documented country (Wolf 2010). This alone makes using national price indices the only way forward. Second, there is a more basic methodological concern. Apart from Schulze's (2007) data, which does not need any work anyway, the regional nominal GDP estimates are derived using the Geary-Stark method. Researchers who have used this method often proxied regional wages using national wages and sometimes proxied regional wages using neighbouring regions' wages. Furthermore, these estimates are ultimately scaled to a given national GDP figure. Altogether, this makes using regional deflators inconsistent. Regional wage data often do not match and such deflators would invalidate the scaling procedure. These finer points aside, Felice (2009) argues that it is unlikely that, at least in the case of Italy, regional prices were so different that they caused differences in regional income levels. His view is backed up by some of the sporadic data we have collected on regional staple goods prices. As measured by the coefficient of variation, the average regional wheat price variation in Sweden between 1870 and 1914 is just 7.62 per cent (Jorberg 1972). In line with Felice's (2009) comments, the data underlying Jacks' (2005) work shows that at least between 1870 and 1877, variation of wheat prices between the Italian regions of Brescia, Padua, and Rome was an average of 5.95 percent. Ward and Devereux's (2003) flour price data covering 12 British cities in 1872 show that variation was only 6.09 per cent. These low levels of regional price variation are not high enough to re-order rankings of relative regional GDP per capita levels, which is the potential fundamental issue.

12.2. Regional literacy data

This is the same data set in Caruana-Galizia (2012). There is heterogeneity in the sample and types of sources, but the point was not to derive a precise figure of literacy. The point is to approximate regional human capital levels, as an output for institutional efficiency, in a way that would not bias the empirical results. There is no reason to think that a persistent bias, that for example literacy measured in one country is correlated with the empirical model's error term, exists in this dataset.

- *Austria-Hungary*. Regional indices (ratios relative to national rate) are from page 156 of Good, D.F. (1984) *The Economic Rise of the Habsburg Empire, 1750-1914,* Berkeley: UC Press. I converted these into real rates, using the national rates on pages 127 and 118 in Cipolla, C. (1969), *Literacy and Western Development*, London: Penguin.
- *Britain*. Rates for England and Wales are from volumes 33, 63, and 73 of the *Annual Report of the Registrar-General of Births, Deaths, and Marriages in England and Wales,* London: H.M. Stationery Office. These figures refer to the rate of people who are unable to sign their own marriage register. Rates for Ireland are from *Ireland Census 1911, General Report* and *Ireland Census 1901, General Report, Part 2*, London: H.M. Stationery Office. Scottish rates come from page 127 in Cipolla, C. (1969). *Literacy and Western Development*, London: Penguin. 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 - Tome XXI – Population*, Paris: Imprimerie Nationale. For 1901, France Récensement 1901 - Tome IV - Résults Généraux. Paris: Imprimerie Nationale. For 1901, *France Récensement 1901 - Tome IV - Résults Généraux*, Paris: Imprimerie Nationale. For 1911, *France Récensement 1911 - Tome I - Premiere Partie - Population Légale ou De Résidence Habituelle*, Paris: Imprimerie Nationale.
- *Germany*. Rates for Prussian regions are from page 91 in Cipolla, C. (1969). *Literacy and Western Development*, London: Penguin. I then took the rates of illiterate military recruits from the 1890 and 1880 *Statistisches Jahrbuch für das Deutsche Reich*, 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 recruits as his measure.
- *Italy*. Rates come from Felice, E. (Forthcoming). Regional Convergence in Italy, 1891-2001: testing human and social capital, *Cliometrica* (February 2012). The rates for 1910 are from page 19 in Cipolla, C. (1969). *Literacy and Western Development*, London: Penguin.
- *Spain*. Nunez, C-E. (1990) 'Literacy and Economic Growth in Spain, 1860-1977', in G. Tortella (ed.), *Education and Economic Development since the Industrial Revolution*, Valencia: Generlitat Valencian, pp. 125-151, 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 *Sverige Folkrakningen 1930*. 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 Sandbery, L. and Steckel, R.H. (1997) 'Was Industrialisation Hazardous to your Health? Not in Sweden!', in R.H. Steckel and R. Floud (eds.) *Health and Welfare during Industrialisation*, Chicago: University of Chicago Press, pp. 127-160.

13. Tables and figures

Table 1: Regional GDP per capita summary statistics.

	1870	1900	1910
Mean	1965	2681	2986
Std. Dev.	746	995	1146
Max.	4317	6552	8109
Min.	606	771	874
C.V.	0.38	0.37	0.38

Notes: all GDP per capita data are in 1990 Geary-Khamis dollars. These statistics summarise regional values across all countries.

Table 2: Theil indices for GDP per capita.

	1870	1900	1910
(1) Within Countries	0.0375	0.0406	0.0412
(2) Between Countries	0.0330	0.0278	0.0298
% Difference (1) and (2)	14	46	38

Notes: all GDP per capita data are in 1990 Geary-Khamis dollars. Within refers to inequality in regional GDP per capita within countries. Between refers to inequalities between countries.

Table 5: Production frontier model (6) results.

	1870	1900	1910
$\boldsymbol{\alpha}$	6.546	4.498	4.554
	(0.000)***	(0.000)***	(0.000)***
ln*(Y)*	0.015	0.014	0.006
	(0.000)***	(0.000)***	(0.000)***
$\ln(\sigma_u^2)$	-0.700	-1.928	-2.458
	(0.000)***	(0.000)***	(0.000)***
$\ln(\sigma_v^2)$	-36.413	-37.103	-37.626
	(270.941)	(270.812)	(270.360)
Log-likelihood	-75.132	47.723	100.676
N	199	199	199

*Estimated using maximum likelihood. Standard errors are in parentheses. *** denotes statistical significance at the one per cent level. Dependent variable is log of regional literacy rates.* $ln(\sigma_u^2)$ *is the parametised asymmetric error component u; and* $ln(\sigma_v^2)$ *is the parametised random error component. Exponentialising these parameters gives their per cent effect on literacy rates.*

Table 6: Mean regional *INST* scores by country and year.

	1870	1900	1910
Austria-Hungary	0.66	0.75	0.83
	[0.10]	[0.10]	[0.04]
Germany	0.93	0.99	0.99
	[0.09]	[0.01]	[0.01]
Spain	0.30	0.49	0.56
	[0.11]	[0.16]	[0.18]
France	0.67	0.95	0.97
	[0.16]	[0.03]	[0.02]
Britain	0.61	0.93	0.97
	[0.13]	[0.03]	[0.02]
Italy	0.28	0.46	0.57
	[0.15]	[0.20]	[0.19]
Sweden	0.88	0.92	0.95
	[0.00]	[0.00]	[0.00]
Sample	0.66	0.85	0.89
	[0.24]	[0.20]	[0.17]

Mean regional efficiency scores and their standard deviations (square brackets) by country and by year. A value of 1 indicates maximum efficiency (that is, the frontier) for the year. Parameters derived from model (6); efficiency calculated as $exp[-E(u_i|\varepsilon_i)]$.

Table 7: Gravity trade model results.

	1870-1883	1893-1906	1907-1920
dist	-0.845	-0.726	-0.551
	[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 FE	Yes	Yes	Yes
Exporter FE	Yes	Yes	Yes
R^2	0.74	0.65	0.67
N	1284	1295	1079

Notes: Robust standard errors clustered at trading-pair level reported in parentheses. *** denotes statistical significance at 1%; ** at 5%; and * at 10%. *Sample restricted to exports within my sample of countries, and from my sample countries to the rest of the world.*

Table 8: Comparison of estimated and constructed market access.

	ln Estimated M.A.	ln Constructed M.A.
Mean	21.853	21.652
Median	22.292	21.998
Standard Deviation	2.204	1.081
Maximum	26.935	23.486
Minimum	16.913	16.802

Notes: Figures are in logs. Estimated M.A. refers to the market access data derived in this paper. Constructed M.A. refers to the market access variable constructed following the Harris-formulation in Caruana-Galizia (2012).

Table 9: OLS and IV estimation results.

	OLS	IV-2
INST	0.229	0.822
	[0.039]***	[0.219]***
MA	0.041	0.380
	[0.006]***	[0.141]**
Dtcoal	-0.054	-0.087
	[0.019]**	[0.056]
Country-Year F.E.	Yes	Yes
	IV-INST	**IV-MA**
SH	-0.973	1.081
	[0.057]***	[0.435]**
DTL	0.012	-0.505
	[0.023]	[0.179]**
Dtcoal	-0.008	0.196
	[0.018]	[0.136]
Country-Year F.E.	Yes	Yes
F-Statistic	12.07***	4.22**
N	597	597

*INST is institutional efficiency in logs; MA is market access in logs; Dtcoal is distance to coal from regional node in logs; SD is the Spanish Habsburg excluded IV; DTL is the distance to London from regional node in logs excluded IV. F-Statistics are on excluded IVs. Robust standard errors clustered on regions reported in square brackets. All estimations include country-by-year fixed effects. Column OLS is the baseline OLS estimation; IV-2 is the second-stage least squares estimation; IV-INST is the first-stage for INST; and IV-MA is the first-stage for MA. *** denotes statistical significance at one per cent; ** at five per cent.*

Table 12.1: Summary of regional GDP data.

Country	Years	Authors	Method	Currency	Regions
Spain	1860;1900;1910	Roses et al. (2011)	Top-down	Cur. Pesetas	17
Britain	1871;1881;1891;1901;1911	Crafts (2005)	Top-down	Cur. Pounds	12
Italy	1871;1881;1891;1901;1911	Felice (2009)	Top-down	Con. Euros	16
A-H	1870;1880;1890;1900;1910	Schulze (2007)	Bottom-up	Con. Dollars	22
Sweden	1870;1880;1890;1900;1910	Enflo et al. (2010)	Top-down	Con. Kronor	24
France	1872;1886;1901;1911	Caruana-Galizia (2013)	Top-down	Con. Francs	85
Germany	1871;1882;1895;1907	Schulze and Caruana-Galizia (2013)	Top-down	Con. Marks	23

Figure 1: Regional GDP per capita in 1990 Geary-Khamis dollars, expressed as standard deviations at each year. Moving from left to right, maps are for 1870, 1900, and 1910.

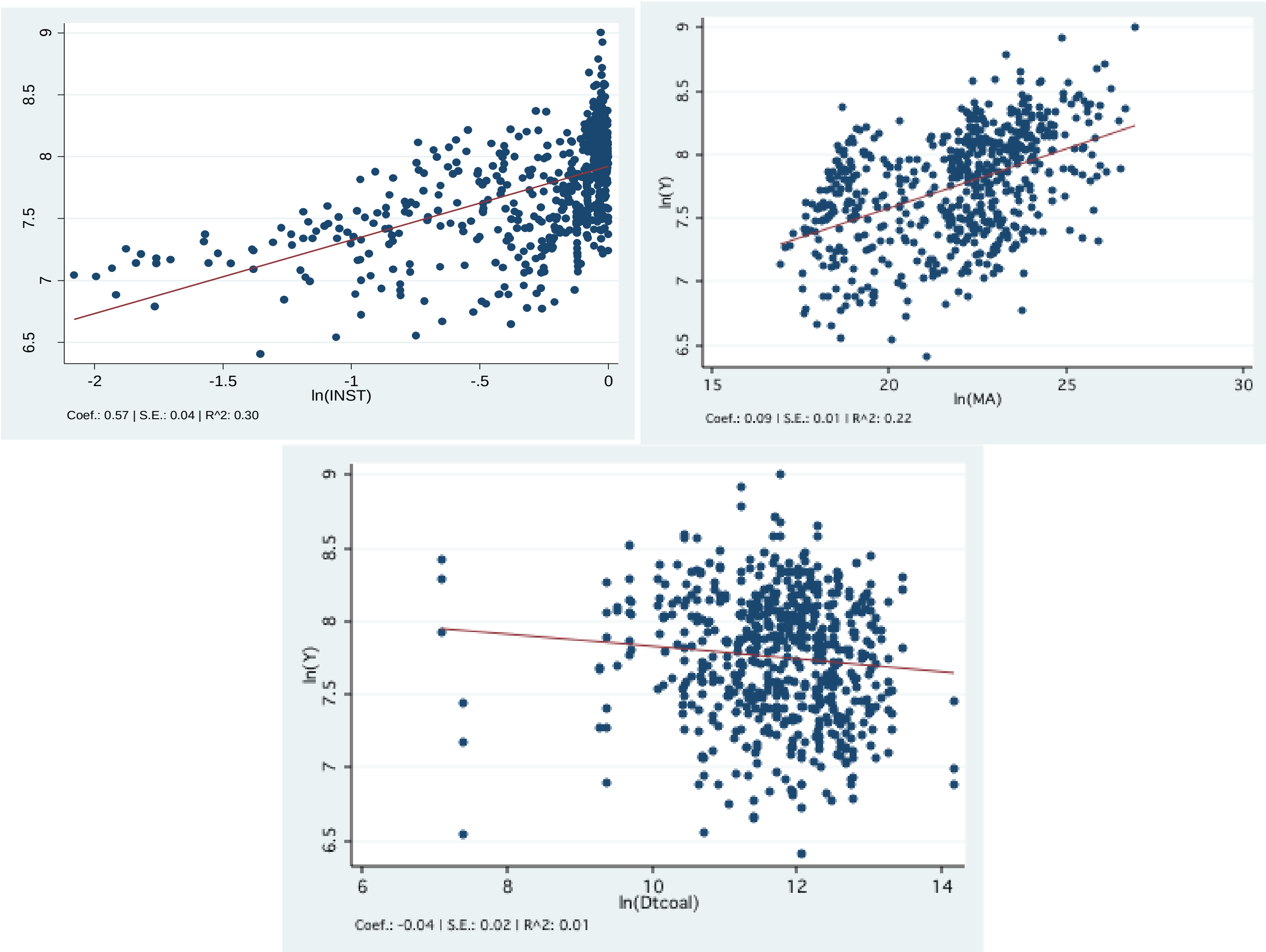

Figure 2: Pooled OLS correlations. Left is *INST*; middle is *MA*; right *Dtcoal*.