The Location of the UK Cotton Textiles Industry in 1838: a Quantitative Analysis

Nicholas Crafts
University of Warwick

and

Nikolaus Wolf
Humboldt University Berlin

Abstract

We examine the geography of cotton textiles in Britain in 1838 to test claims about why the industry came to be so heavily concentrated in Lancashire. Our analysis considers both first and second nature aspects of geography including the availability of water power, humidity, coal prices, market access and sunk costs. We show that some of these characteristics have substantial explanatory power. Moreover, we exploit the change from water to steam power to show that the persistent effect of first nature characteristics on industry location can be explained by a combination of sunk costs and agglomeration effects.

Keywords: agglomeration; cotton textiles; geography; industry location

JEL Classification: N63; N93; R12

Introduction

Cotton was a key sector in 19th-century industrialization. During the Industrial Revolution, the industry demonstrated remarkable technological progress and productivity advance. Both the markets for its raw material and its finished products were global. At the same time, the industry remained very spatially concentrated. In 1800, the UK had 95 percent of the world's spindles, in 1850 it had 69 percent and in 1900 still 58 percent (Farnie and Jeremy 2004). The cotton industry accounted for 22 percent of British industrial value added and 50 percent of British merchandise exports in 1831; during the years 1780-1860 it sustained an average rate of TFP growth of 2.6 percent per year while contributing about a quarter of the economy's total TFP growth (Crafts 1985).

Even within the UK, the industry was highly localized and increasingly so. In 1850, Lancashire accounted for 66 percent of spindles and in 1903 that share had risen to 79 percent (BPP 1850; BPP 1903). Given that the cotton industry witnessed substantial change in terms of power technology, mechanization and processing techniques, this locational persistence is quite remarkable. It is also well known and is a staple of accounts of the development of the industry. For example, Douglas Farnie (1979) devotes a lengthy chapter to this phenomenon and its extensive historiography. He identifies a number of plausible hypotheses but recognizes the absence of a proper explanation. In fact, we lack any quantitative analysis that tackles an interrelated set of questions: Why was so much of the UK industry in Lancashire? Why did it stay there? Why did other areas of the UK, which had at least some of the characteristics believed to be attractive to the industry, fail to develop cotton textiles?

Given the role of cotton textiles in 19th century industrialization, explaining its location is clearly an important task for economic history. Recent historiographical trends add value to such an investigation. To begin the new economic geography, including models featuring the possibility of "geographical lock-in" such as Paul Krugman (1991) or more recently Stephen Redding et al. (2011) has sparked an interest in location decisions. This has produced work like Joan R. Roses' (2003) study of industrialization in Catalonia. Second, economists are now giving much attention to 'directed technological change' (Acemoglu 1998) in which both natural endowments and market demand conditions are seen as driving decisions to develop new technologies. Here the key economic history reference is Robert Allen (2009). He emphasizes the joint role of cheap energy and high wages in directing the development of textiles technology during the British industrial revolution. In this context, understanding the interaction between factor prices, other locational characteristics and choice of locations assumes added significance.

In this article, we try to account for the location of the UK cotton textiles industry in 1838 using a detailed statistical report that resulted from the implementation of the Factory Act (BPP, 1839). The report was the most detailed ever published by the Factory Inspectors and gave information on the distribution of employment and the use of power both at the county level for all parts of the UK (including Ireland) and for individual towns. It has already been used for a number of statistical analyses about the nature of the industry within Lancashire (for example, Gattrell 1977; Rodgers 1960). However, no one has used it to shed light on the spatial distribution of the cotton industry across Britain.

Literature Survey

A good starting point is the classic study by Farnie (1979, ch. 2) who stated that: "The original advantages of Lancashire comprised its poverty, its climate, its water supply, its textile tradition and its mechanical inventions. The acquired advantages included its supply of coal, machinery and labour, its access to the markets of Liverpool and Manchester, its low transport cost, and its auxiliary industries" (1979, p. 46).

In Farnie's elaboration of these original advantages, 'poverty' stood for the county's poor soils; 'climate' involved the damp air that increased the pliability of cotton fibers and reduced the risk of breakages in the yarn; 'water supply' denoted the abundance of fast-running streams that can provide water power. Overall, this represents an argument that original natural advantages were later augmented by investment and culminated in agglomeration advantages.

The same factors have been highlighted by other writers. The importance of Lancashire's proto-industrial traditions is emphasized by Mary B. Rose (1996). She does caution that the emergence of rural manufacturing in impoverished upland areas with poor-quality agricultural land was not an automatic precursor to industrial cotton manufactures. Nevertheless, she points out that there was an almost perfect geographic coincidence of 18th century fustian manufacture with the location of the modern cotton industry (1996, p. 15). Rose's account resonates both with Farnie's list of Lancashire's original advantages and with his acquired advantages in sustaining the location of the industry.

Many people have remarked upon the eventual tendency of the cotton industry within Lancashire to locate on the coalfield in an era when coal was very expensive to transport over land. Rodgers (1960) represents a very strong version of the argument that the location of the industry was basically determined by the price of coal, an important intermediate input, at least once steam power took over from water power in cotton mills. Nick von Tunzelmann (1986) cited estimates that the transport cost for coal was about 2.5 pence per mile in the early 1840s which would have doubled the pithead price at a location 10 miles away.

Theo Balderston (2010) suggested that agglomeration advantages were the real secret of Lancashire's success and provides a long list of what these comprised including tacit knowledge, dense labor markets, specialist suppliers who inter alia lowered machinery costs, and the unique density of the Liverpool raw cotton market. This claim resonates with ideas from Alfred Marshall (1919) and the new economic geography. One obvious feature of successful agglomerations is that they tend to bid up the price of labor but keep unit labor costs competitive through external economies of scale.

The ultimate importance of agglomeration advantages for Lancashire's success is generally agreed, as is the observation that, as steam power superseded water power, the industry became more spatially concentrated even within Lancashire itself. This process of further spatial concentration and agglomeration entailed a shrinking of the industry's geographical breadth rather than a move to completely new locations, although there were other areas with lower wages, cheap coal, access to ports, high humidity etc. Nevertheless, older plants stayed in business into the second half of the 19th century by adopting auxiliary steam engines and by substituting steam engines for water wheels (Taylor 1949, see also section 3 below).

The nature of the original advantages resulting from climate and natural resources has proved contentious. In claiming that humidity was an important asset, Farnie was reasserting a traditional

belief that had been strongly disputed and somewhat discarded following a critical analysis by H.W. Ogden (1927). Ogden himself, along with Rollin Sallisbury Atwood (1928), argued that the real advantage was in the ready availability of soft water which was helpful for the finishing industries but also for the boilers of steam engines. In turn, this claim was vigorously attacked as a myth by H.B. Rodgers (1960), who stressed that pollution had negated any possible advantage from water quality.

Overall, no-one is proposing that there is a single factor explanation for the dominance of Lancashire in cotton textiles. If anything, as with Farnie (1979), we are presented with a 'laundry list'. Only occasionally is there a sense of how much of an advantage or handicap a given attribute offered. For example, Farnie (1979, p. 49) suggested that Lancashire's natural humidity could be seen as the equivalent of a 10 percent tariff. Rodgers argued that a mill located far enough from a coal mine that it paid twice the pithead price faced a cost penalty equivalent to a 20% wage increase (1960, p. 140) but that the annual saving from clean soft water for spinners and weavers was only equivalent to 5 percent of the cost of fuel (1960, p. 142). We do not, however, have an analysis which carefully examines the trade-offs between different locations. This is unfortunate because, as we shall see, Lancashire does not appear to have been uniquely blessed with any one cost advantage other than, eventually, the 'acquired' benefits of agglomeration.

Empirical Strategy and Theoretical Framework

Our aim is to put these claims from historical narratives into a theoretical framework and confront them with statistical evidence. To this end, we first discuss more generally estimation methods before we provide a theoretical framework that specifies how the explanatory variables interact. Following the literature on location choice starting with Daniel McFadden (1978) and Dennis Carlton (1979), we consider the location choice of a profit-maximizing mill owner in 19th century Britain. The fact that we use here only one cross-section of data implies we must focus on the long-run determinants of location choice, because our data includes newly established mills as well as mills founded decades ago in the mid-18th century. Hence, our results will reflect the effect of different factors on a mill-owner's decision to produce at a particular location – by starting a new mill or by deciding to continue production in an existing mill. Because we have only one cross section, some of our results will suffer from the effect of omitted variables not reported at that time. In order to address this issue, our estimation allows for region-specific variation in constants (region fixed effects) and in the distribution of errors (clustered and heteroskedasticity-robust standard errors).

In modeling location choice, we assume first that a (current or prospective) mill owner can locate his plant in one of J possible locations, where we treat the location choices as independent of one another. In particular we use a Poisson model, which is, as shown by Paulo Guimaraes et al. (2003) observationally equivalent to conditional logit models, but much more tractable. The profit to a firm i (=1,.., 1823) of locating in j (=1, …148) can be described with an unobservable latent variable π* as

$$\pi_{ij}^* = \beta x_j + \varepsilon_{ij}, \tag{1}$$

where the first is a deterministic part and the second a stochastic part. In equation (1), the letter x denotes a vector of location-specific variables at location j. (One might also want to include a vector of possible industry or firm-level characteristics but we do not have those). The decision maker now selects one out of the J mutually exclusive alternatives and we assume that the locations we observe in our data are those that maximize the profit π*. With this, the probability that we observe a cotton mill i at location j is

$$P_{ij} = P\left(\pi^*_{ij} > \pi^*_{ik} \forall k = 1, \dots J : k \neq j\right) = P\left(\varepsilon_{ik} - \varepsilon_{ij} \leq V_{ij} - V_{ik} \forall k = 1, \dots J : k \neq j\right). \qquad (2)$$

Hence, the probability of observing location choice j for mill i depends on the distribution of the stochastic part ε, given the (observed) deterministic part V. The Poisson estimator assumes that the expected number of mills at j, E(nj) is independently Poisson distributed with region-specific mean as

$$E(n_j) = e^{xj\beta} \qquad (3)$$

Hence, we assume here that the decisions on plant location can be treated as independent across mill owners. We will estimate the vector of coefficients β in equation (2) using maximum likelihood. To ease the interpretation of estimation results, Kurt Schmidheiny and Marius Brülhart (2011) note that for the Poisson-estimator the elasticity of the expected number of firms in location j, E(nj), with respect to a change in the k-th locational characteristic x_k of region j is simply given by its coefficient in the profit function:

$$\varphi_{jj} = \frac{\partial logE(n_j)}{\partial x_{jk}} = \beta_k. \qquad (4)$$

We will make use of this property in our empirical analysis below.

The next step is to formulate hypotheses about locational characteristics that would matter for location choice. In terms of equations (1) and (2), we therefore need to distinguish between various types of location characteristics, x_{jk}. Some characteristics were already highlighted in the historical literature on the cotton industry, including access to water power or humidity. The recent theoretical and empirical literature on location choice would call these "first nature" characteristics, because they can be treated as fixed. In addition, the literature highlights "second nature" factors, which are directly affected by (endogenous to) location choices and arise from the interaction between locations.

To understand how "first nature" and "second nature" characteristics play together, consider the simple framework proposed by Pierre-Phillipe Combes et al. (2008, chapters 4 and 12). They show that in a standard model with monopolistic competition, transportation costs and several inputs, total profits for a firm located in j can be formulated as a function of three groups of factors: variable costs, fixed costs and benefits from market access. To keep things simple, we can assume that differences in variable costs mainly reflect differences in "first nature" characteristics between locations, while variations in fixed costs and market access are endogenous to earlier location choices. Hence they reflect "second nature" characteristics. Optimal location choice therefore involves searching for the place that maximizes profits with respect to variable and fixed costs and benefits of market access:

$$Max(Profit_j) = Max[cm_j^{-(\sigma-1)}MP_j - F_j], j \in \{1, \dots, J\}, \qquad (5)$$

where m stands for the marginal costs of some input factor, c and σ are parameters (both >0), F stands for fixed costs and MP gives the benefits from access to markets that will increase in own market size and good access to neighboring markets.[1] In this case earlier location choices affect

---

[1] Head and Mayer (2004) for example estimate this model under the assumption that fixed costs are equal across locations and therefore can be neglected. They provide empirical evidence that both differences in variable costs and access to markets matter for location choice, which implies that there is scope for positive feedback effects or geographical lock-in as argued in Krugman (1991).

future choices because they change market sizes and the size of neighboring markets. Therefore, even if a location starts to become less attractive in terms of variable costs, the benefits from a large market (that comes from having a large installed capacity) can still outweigh this disadvantage.

For our analysis of the location of the cotton industry, we need to modify this theoretical framework to account for the effect of technological change. This change should have led to a relocation of cotton mills, as some "first nature" characteristics became obsolete (e.g. water power), while others (e.g. coal) started to become essential. However, fixed costs can delay or even prevent such relocation. With their assumption that fixed costs are equal across locations, Head and Mayer (2004) neglect that a considerable part of the fixed costs are typically sunk costs in the sense that they have been incurred in the past (such as costs for a building), and they are location-specific so that they cannot be recovered upon exit (Motta and Thisse 1994). In the presence of sunk costs some fraction of fixed costs will now matter for a firm that considers relocating, even if fixed costs are perfectly equal across locations. More generally, this introduces an extra cost of relocation: the existence of sunk costs will *dampen* the impact of current or prospective changes in the differences in variable costs as well as market access between locations and will *increase* the impact of past differences (see appendix for a theoretical discussion).

In our context, consider the investment of a cotton mill owner into power equipment. A waterwheel installed in the past is tailor-made to the nearby river or stream and will have little or no resale value. However, a part of the old construction and machinery can be reused with the arrival of the new steam power technology or combined with it, saving the costs of relocation. For example, at the Quarry Bank Mill in Styal near Manchester, the first steam engines were installed in 1810 to provide auxiliary power to the existing water wheels and as a precaution against water shortages. After several decades the mill was still operational at its old location but nearly entirely steam powered.[2] In addition, the waterway infrastructure (canals, weirs, water reservoirs) put in place for the old water-power technology remained useful for the new steam power technology as well for two reasons: first, because steam engines needed access to water for feeding boilers and condensing steam (Holden 1999), second because waterways reduced the cost of transporting raw materials and finished products (Maw et al. 2012).

To summarize, we see that there are three main types of locational characteristics that affect location choice: variable costs (or "first nature" characteristics) as well as fixed costs and benefits from market access (or "second nature characteristics"), which interact over time. In general, contemporaneous variation in all three types of factors should matter for location choice. However, past variations in operating profits can modify the impact of contemporaneous variation in variable costs due to two types of path dependency: first, there can be positive feedback effects due to market access as highlighted in new economic geography models in the wake of Krugman (1991). Second, sunk costs can introduce another form of hysteresis in location choice that can delay relocation (similar to the mechanism discussed in Redding et al. (2011)). Without direct evidence on the extent of sunk costs, such an effect will always be implicit in the estimated effect of variable costs (or "first nature" characteristics) and market access. In particular, omitting sunk costs will bias these coefficients upwards. In the next section we now turn to empirical evidence on the effect of specific types of costs and market access effects.

Data on the Location of the British Cotton Industry and Its Potential Determinants

---

[2] This was not unusual; other instances are discussed in Ashmore (1969) and Williams (1992).

We seek to explain the location decisions which underlie the observed spatial distribution of mills and employment in the UK cotton textile industry in 1838. We consider employment in each county but also divide Lancashire into sub-divisions based on Poor Law Unions. Our dependent variable number of mills or employment at a given location is derived from the Factory Inspectors' report for 1838 (BPP 1839). Table 1 provides descriptive statistics for the location of cotton mills and their employment, maps 1 and 2 show the substantial geographical variation in concentration.

By 1838 much of the industry was in the North West, and a very high share in Lancashire alone, it was also highly concentrated at the county level. Many counties had no cotton textile employment at all. In fact 92 of the 148 locations in our data set have no cotton industry and even within Lancashire 10 of the 31 Poor Law Unions had no cotton mills or employment. On maps 1 and 2 the green no cotton industry area covers most of Britain.

To measure local characteristics, we begin with aspects of what Farnie (1979) called "original advantages": poverty, climate, water flow, mechanical inventions, and water quality. Farnie suggests 'poverty' as an attractor of cotton textiles. In a comparative advantage view of pre-industrial England, textiles were most likely to be found in areas with low agricultural incomes and low land values. As a proxy we used the land tax which was imposed in the 1690s for England and Wales and in the 1710s for Scotland and for which the valuations did not change thereafter (Ginter 1992). Valuations are reported for the individual parish as well as at the county level in BPP (1844b). These have been converted to a tax rate per acre using the areas reported in John George Bartholomew (1904) and BPP (1862). As shown in map 3, there is a clear gradient rising from north to south, with show relatively low land-values for Lancashire compared with southern England.

Given that we lack comparable data for Ireland, we provide another measure for the suitability of land for agriculture: ruggedness of the terrain. The data is provided by Nathan Nunn and Diego Puga on http://diegopuga.org/data/rugged/, which we recalculated to match historical counties and Poor Law Unions. There is a tight correlation (-0.66) between the log of early 18th century land tax rates and average terrain ruggedness. Similarly, if we exclude Ireland, the correlation between late 18th century wages in agriculture and average terrain ruggedness is also high (-0.56). Therefore, we use ruggedness for which we have data for all of the British Isles rather than land tax rates as our proxy for the suitability of land for agriculture to capture Farnie's notion of poverty as an "original advantage".

To measure humidity, which Farnie (1979) saw as the key aspect of climate, we use data on vapor pressure and on mean daily temperature averaged over 1901-2006, which permits calculation of an estimate of relative humidity.[3] The data is shown on map 4. It is informative to see that the lowest estimated value of relative humidity is 83.5% - well above the 50-54% considered by Cecil H. Lander (1914) to be the minimum for cotton spinning. While levels of humidity do vary over the year and within a day, they typically stay well above this critical level at any time. In fact all locations in the UK

---

[3] The University of East Anglia Climate Research Unit (2008), *A Gridded Time-series Dataset, CRU TS 3.0*. To calculate the relative humidity utilized in this paper, we first calculate the dew point using the following equation: $T_D = \left(243.5 * ln\left(\frac{v}{6.112}\right)\right) \Big/ \left(17.67 - ln\left(\frac{v}{6.112}\right)\right)$ where v is the actual water vapor pressure, in units of millibar. Because the dew point (T_d) decreases by about 1°C for every 5% decrease in relative humidity (RH), we can approximate the RH using the following equation from Lawrence (2005): $RH \approx 100 - 5(T - T_d)$, where T (temperature) and T_d are measured in degrees Celsius.

met the required level. Lancashire observations are all very similar and lie just below the average for all of the British Islands. Moreover Lancashire has substantially *lower* levels of relative humidity than other parts on the Atlantic coast, notably western Ireland and the Southwest of England.

The availability of suitable water flows is another aspect of nature which we capture through modern data. We use data from the UK Hydrometric Register (2008), which provides basic statistical information on the average extent and volatility of water flows, based on 1,454 "gauging stations" across the UK in modern boundaries (including Northern Ireland but not the Republic of Ireland). Matching flow data for Ireland were obtained from Ireland's Environmental Protection Agency (EPA, at http://hydronet.epa.ie/) for 125 gauging sites. Most of the data is averaged over several decades, typically over periods of 30 to 50 years. For our purposes these data include two useful variables, namely the "mean flow" and the "q95". The former gives the average water flow of rivers in cubic meters per second and the latter gives the flow which was equaled or exceeded for 95% of the time. The mean flow data ranges from values of below 1 cubic meter per second up to more than 89 in England and even 165 if we include Scotland, with an average of 4 and 6 cubic meters per second for England and the UK, respectively. Data on "q95" range from near zero to more than 42 cubic meters per second in the UK, with an average of 1.1 meters. Mean flow and q95 values for rivers in Lancashire, including the Irwell, Darwen, Mersey or Ribble in Lancashire along which many of the early cotton mills were located, are typically above the English average. As shown in map 5, Lancashire has good access to rivers with high mean flow levels, but it was not the only county so endowed.

To assess Ogden's (1927) argument, we also collected data on water quality, specifically on water hardness (the concentration of calcium carbonate ($CaCO_3$)). Data for average total water hardness for about 2,500 sites across the UK was kindly provided by the Environmental Change Research Centre, UCL.[4] We complement this data with information from the Irish Water Framework Directive (WFD) River Basin District Management Systems, Working Group on Groundwater Guidance (2005). Again, we find considerable variation over space. Lancashire has moderately soft water.

To capture the idea of an earlier textile tradition as well as a location's technical capabilities to produce mechanical inventions, we use James Dowey's novel data set on the location of patent holders, weighted by their impact.[5] The variable (Patent Citations) is a location-specific index based on the number of patents that people resident at that location hold, weighted by the number of times this patent was cited in the technology literature through 1862. The data set starts in the early 17th century and includes all patents relevant to the textiles industry in a broad sense. Dowey shows that it is important to de-trend this data to take into account that older patents would have a higher probability of being cited independent of their technical importance. We use only patents awarded up to 1815 to limit endogeneity issues. Many regions of Britain produced innovators. Before 1815, Lancashire was not a particularly innovative area, but it was near to regions with very high innovative activity.

---

[4] This centre is funded by the Department for Environment, Food and Rural Affairs. We are grateful to Dr. Pete Smyntek, Queen Mary University London for giving us access to this data.
[5] The original patent data are reported in Woodcroft (1854) (1862). We are very grateful to James Dowey for supplying us with his dataset. This will become publicly available after Dowey completes his Ph. D. dissertation. Woodcroft's Reference Index is described in detail in Nuvolari and Tartari (2011) who used it as the basis of a measure of patent quality and show that it is suitable for this purpose.

Next, let us consider several factors, which Farnie (1979) termed the 'acquired advantages' relevant to decisions about where to locate the cotton industry. This category must include coal because substantial effort is necessary to find and exploit a location's coal reserves and because the price of coal in a location reflects both how far it is from natural coal resources and the transport infrastructure. Cheap coal in Lancashire emerged as a supply response to industrial growth (Langton 1979). Information exists for years close to 1842 and 1843 on coal prices paid by Poor Law Unions in England and Wales (BPP 1843) and in Scotland (BPP 1844a) and from Robert Kane (1845) for Irish regions. These data, displayed in map 6, show a very wide range of prices. Coal was most expensive at Bradfield in Berkshire (40 shillings per ton). In rural southern England it was not uncommon to pay 30 shillings or more. The lowest price was 4 shillings and 2d at Leigh in Lancashire but prices only marginally higher were recorded at Poor Law unions in Durham and Northumberland. Map 6 again shows the data at the county level and for Poor Law Unions within Lancashire. But how much of the variation in coal prices is due to variation in first nature? We cannot tell exactly, because the first comprehensive set of data on British coal deposits was collected by the Coal Commission for all known coalfields around 1870 (BPP 1871). From this source we estimated the amount of coal measured in statute tons in each of our 148 locations. Using this measure we find that about 20% of the variation in 1840s coal prices can be explained by local coal deposits as known in 1871. For example, the region with the highest estimated coal deposits (Carmarthenshire in Wales) shows around 1840 coal prices similar to the sample average, probably due to the fact that large-scale coal mining in the region started only in the 1850s. In addition, a measure of access to coal deposits taking into account transport costs based on von Tunzelmann (1986) can explain over 50% of coal price variation. We take this as evidence that coal prices, which are the relevant variable facing a mill-owner, reflect geology to some extent but also indicate that cheap coal was an 'acquired advantage'.

A simple way to capture foreign market access is to consider average distance to major ports weighted by the tonnage that they handled, based on the list of 12 principal ports in BPP (1854). An alternative is to restrict the measure to distance to Liverpool or London, the 2 most important ports when cotton took off. The literature on the location of the industry in Lancashire might even hypothesize that proximity to Liverpool would suffice given its ultimate dominance. It is obvious that measured in this way Lancashire is a county with excellent access to markets, but so are several counties close to London.[6]

Next, we explore the role of benefits from agglomeration in terms of a proxy variable, domestic market access. These advantages which accrue from tacit knowledge, deep labor pools, dense networks of suppliers etc. are generally thought to have been important in the 19th century cotton industry (Hanlon 2013; Leunig 2003; Broadberry and Marrison 2002). As a first approximation, the extent of external economies of scale might be captured by distance-weighted employment for which we could use the Factory Inspectors' Report (BPP 1839). However, such a measure is directly related to the variable that we are trying to explain, namely the number of cotton mills and their employment in 1838. To limit this problem of endogeneity, we note that a location's access to overall population around 1800 is highly correlated (with a correlation coefficient of about 0.77) to agglomeration in terms of cotton textiles employment in 1838 but can be regarded as exogenous to the number of mills and employment a generation later (see the scatterplot in figure 1). Map 7

---

[6] Foreign and domestic demand were equally important in the 1830s when the share of exports in gross output was close to 50% (Deane and Cole 1962).

shows the resulting measure of market potential approximated by population in 1801. South East Lancashire clearly scores very highly in terms of market potential.

Lancashire may have benefited from agglomeration but its corollary was high wages. The only source that has a national scope for our period is the data on agricultural wages compiled by Artur Lyon Bowley (1900). Lancashire was clearly a relatively high wage location in the late 1830s, certainly compared with many agricultural areas of the United Kingdom. High wages were no doubt an endogenous result of the concentration of the industry and a cost to be accepted because it was compensated by high productivity rather than an advantage per se.[7]

Table 2 summarizes this discussion of the data. We see that the Lancashire locations (Manchester and Preston) were well placed compared with the average of other locations in most respects. On the other hand, few Lancashire observations were best-placed on any single characteristic. Turning to Lancashire as a whole, the county has the location with the best score in terms of market access and coal prices but not for anything else.

Empirical Results

a) Benchmark Results

In this section we fit a Poisson model for the location of the cotton textiles industry across the British Islands according to the Factory Returns for 1838, as discussed in section 3. We first estimate the expected number of cotton mills for each location as a function of locational characteristics as described above (Table 3, columns 1 and 2). Next, we estimate the expected number of employed persons in the cotton industry for each location (columns 3 and 4). Finally we estimate the expected number of employed persons per mill or average mill size, where in this last case the number of persons per mill is set to zero whenever there is no industry at a location (column 5). In each specification we add the (log of) the area of various locations in km^2 as a control to account for the fact that, if location choices were entirely random, smaller locations should always have fewer cotton mills than larger locations, ceteris paribus (Ellison and Glaeser, 1997). We also add a set of region dummies and allow for regional variation in the distribution of errors (heteroskedasticity) in all specifications.

Table 3 shows that a larger area seems to increase the likelihood of finding a mill and cotton industry employment but the effect is not always statistically significant, depending on the specification. A location's terrain as measured by its average ruggedness tends to have a large and significant effect. It is interesting to note that this variable only matters for the location of mills. If we interpret ruggedness as a measure of the unsuitability of land for agriculture, this lends strong support to Farnie's notion of poverty as an "original advantage"[8] Our next finding is quite revealing: higher levels of average relative humidity appear to reduce a location's attractiveness for the cotton

---

[7] Prior to the industrial revolution, Bowley's data show that in the 1770s Lancashire was a county with relatively low agricultural wages.

[8] This interpretation is also suggested by the fact that ruggedness has an effect similar to that of land taxes, but with opposite sign. In an alternative specification (not shown) we re-estimated the model for all those locations, for which we have data on land taxes (all except Ireland) with similar results compared to the full sample. We next estimated the model replacing ruggedness by log(land tax), which again produced very similar results, except that the coefficient on land tax is strongly negative, as expected. Moreover, we used the data from Bowley (1900) on agricultural wages around the 1770s for the counties to re-estimate the model at the county level (consider table 4a). This leaves our results nearly unchanged.

industry, not increase it. Although the effect is large it is never significantly different from zero in any estimation). It seems that humidity was not a determining factor. Similarly, the quality of water (hardness), does not matter much: the estimated coefficients are far from significant. This might be due to measurement problems because the quality of water varies substantially between neighboring locations, which we cannot capture perfectly. Therefore, we also provide results after dropping the humidity and hardness variables, which remain largely unchanged (compare columns 1 with 2 and 3 with 4).

Next, we find that the availability of water power, measured here as proximity to rivers and streams with fast and steady water flow (water power: q95) exerted a strong influence. In all specifications, we find that good access to water power increases a location's probability of having a cotton industry, and also that it increases the expected size of that industry. Finally, we turn to the last aspect of Farnie's list of original advantages: a location's mechanical inventions. Our proxy, based on weighted patents issued before 1815, comes out with a strong positive coefficient, which is about as robust as our finding on access to water power. The elasticity of the expected number of cotton mills as well as the expected employment in any location with respect to this measure is between 70% and 80%. Obviously, there may still be an issue of endogeneity, if the incentive to innovate came from the existence of an industry, but we show at least that an earlier track record of innovation up to 1815 had lasting effects into the 1830s.[9]

Turning to Farnie's "acquired advantages", we first have a surprising result: a location's coal price has a negative but typically insignificant effect on the expected number of cotton mills, that is on the probability of finding a cotton mill at all. This is interesting because in 1838, the share of steam power in total installed power (measured in horse power) was already about 54 percent, and within Lancashire it was 73 percent. Therefore, we might have expected to find a much stronger effect of the coal price. This might be affected by the way we measure market access, because the price of coal and the size of an agglomeration in the 1830s might be positively correlated (Balderston, 2010).[10] What we find is that the probability of finding a cotton mill at all seems to be much more affected by access to waterpower than access to coal. In contrast, the results in columns 3, 4, and 5 show that coal prices had a strong impact on industry size, measured in terms of total employment and even more so employment per mill: while some part of the industry could survive with above average coal prices, dear coal apparently put a clear limit on the growth of the industry.[11] The average cotton employment at locations with above average coal prices is 83 (median is zero), while the average employment at locations with below average coal prices is 3737 (median is 243).

Distance to major ports did matter for the geography of the industry. This makes sense for an industry that was so heavily dependent on imported raw materials, and increasingly relied on foreign demand for its products. However, this effect is less robust than others, possibly because the location of large ports is again correlated with other factors such as market size. Interestingly, we see that the location of larger mills is much more strongly affected by distance to ports than mill location per se. This suggests that mills engaged in foreign trade (via both import and export

---

[9] We also generated an alternative variable based on patents up to 1780 only and repeated all estimations. This reduces the estimated coefficient a bit, but does not change anything of substance.

[10] However, we do not find any positive but rather a (weak) negative correlation between market access (measured by our variable mp01) and coal prices.

[11] If we use our estimate of 1870 coal deposits instead of coal prices this leaves our results on the remaining variables nearly unchanged. Only the coefficient on market potential is slightly reduced but stays highly significant.

markets) tended to be larger, which is in line with the recent literature on heterogeneous firms in international trade (Melitz 2003; Redding and Melitz 2012). Finally, we explore the effect of the size of agglomerations on the expected number of cotton mills and industry size. Using the sum of distance-weighted population as of 1801 across locations as a proxy, we find that agglomerations make it more likely to find a mill and correspondingly that such locations have higher total employment in this industry (columns 1-4). However, we also find that agglomerations are associated with a smaller average mill size, which is in line with earlier findings by Broadberry and Marrison (2002) and Leunig (2003). A notable fact about Lancashire's cotton industry was its network of a large number of small but highly specialized and interconnected firms.

b) Robustness Checks and Extensions

As a robustness check, we rerun our Poisson model with two alternative samples. First, we pool all 31 poor law unions in Lancashire into one observation, the county of Lancashire (Table 4a). Second, we rerun the model for the 31 poor law unions of Lancashire only (Table 4b).[12]

Our results are reasonably robust to changes in regional aggregation, with some interesting exceptions. If we pool the Lancashire poor law unions (Table 4a), we find that ruggedness has no significant effect on the location of cotton mills, a negative effect on employment, but again no detectable effect on the average size of mills. Variation within Lancashire (Table 4b), which is substantial, produces an opposite result. The coefficient on access to water power still has the expected positive sign but is no longer significant. Hence, the finding in Table 3 seems to be largely driven by variation across counties, similar to the finding on distance to ports. However, note that in Table 4b we have only 31 observations, which will reduce levels of statistical significance. The positive coefficient on water hardness within Lancashire is very likely to be spurious, as our data show very little variation across Poor Law Unions.[13] In comparison, our findings on the negative effect of coal prices and positive effects of innovative activity before 1815 and market access are rather robust to regional aggregation. Higher coal prices make it less likely that a location develops a cotton industry and notably limit the size of mills, while a strong tradition in innovation has the opposite effect. Good market access is a strong predictor for the development of an industry, but seems rather to increase the number of mills, while limiting their size, in line with our findings for the full sample.

Our findings cannot be interpreted without considering several important issues. So far we have neglected contemporary wages as a determinant of industry location. This may seem odd, given that wages accounted for a much higher share in total costs than either energy or cotton (see for example Kane 1845, p. 64). A first intuition would suggest that locations with lower wages should be particularly attractive to the industry and this may indeed have been one of Lancashire's original advantages.

For a mill owner in 1838, however, low wage labor could only be accessed in areas where the cotton industry was small or non-existent, in other words, only by foregoing agglomeration benefits. Table 3 shows that agglomeration exerted a strong positive effect on the location of cotton textiles. Theory predicts that successful agglomerations will pay higher wages and in modern data areas with high market potential pay more (Head and Mayer, 2006). Data from the 1906 *Earnings Enquiry* show

---

[12] We also re-estimated our model with logit, probit and conditional logit with qualitatively very similar results.

[13] To be precise, the finding is entirely driven by the Poor Law Union of Clitheroe.

that higher wages were paid to operatives in districts with larger cotton employment (Broadberry and Marrison, 2002). A simple OLS regression suggests that the pattern of wage increases across locations (proxied by agricultural wages due to limits to data availability) is strongly positively affected by the size of the cotton textile industry.[14] So, we can reasonably conclude that wages were endogenous to the size of a location's industry and that the estimated impact of market potential in Table 3 is less than the pure productivity effect of agglomeration because it is dampened by higher wages.

The reader should also consider that cotton was a very dynamic industry, characterized by substantial technological change and probably by high rates of firm and plant entry and exit. Although a panel would be desirable we only have one (albeit very detailed) cross-section of data, 1838. In particular, we argued in section 3 that market access and sunk costs can give rise to a geographical lock-in or path-dependency where previous cost advantages (for example, good access to water power) continue to exert an influence, even though technological change seemingly had rendered them irrelevant. One way to go further is to focus on one very particular finding from Table 3: we find that access to water power exerts a strong influence on the location and size of the cotton industry in 1838, while we know that at this time already more than half of all power (measured in hp) in the cotton industry was generated by steam engines. To make the point more explicit, we use the same set of variables as in table 3 above to estimate in Table 5 the expected number of steam engines across locations.

We find that the effect of access to water power on the probability of finding steam engines in the cotton industry is comparable to the effect of this factor on the probability of finding a cotton mill even after we control for all the other variables discussed in section 4 (Table 5, column 1). Clearly, this cannot reflect any direct effect of water, but suggests that other, indirect channels matter. In our theoretical section 3, we argued that three types of factors should affect location choices: variable costs, fixed costs (due to the effect of sunk costs), and benefits from market access, which interact over time. The benefits from having good access to water power in 1838 should have become more limited, while the benefits from having good access to coal should have become much more important. Moreover, we would expect to see this effect most strongly for the location of steam engines in the industry.

What then explains the strong effect of water on steam engines? We argued that past variations in local characteristics (such as better access to water decades ago) can modify the impact of contemporaneous variation in variable costs (better access to coal in 1838) due to two types of path dependency. First, there can be positive feedback effects because of market access, as highlighted in new economic geography models since Krugman (1991). Second, sunk costs can introduce another, additional form of hysteresis in location choice that might delay relocation. In this context, it is noteworthy that our proxy for agglomeration (population in 1801 in neighboring regions) does not eliminate the significant coefficient on water power (compare Table 5, column 1). To investigate further, we can introduce a more direct measure using access to the employment in the cotton industry in 1838 in neighboring regions (ignoring endogeneity problems for the moment). Doing so reduces the estimated coefficient on water power somewhat but it stays positive and highly significant (Table 5, column 2). Given the endogeneity bias, the continuing effect of water power is

---

[14] For an OLS regression we find: ln(wage1830/wage1770) =0.384 (0.019) +0.067 (0.011)lncottonmills; n = 148, $R^2 = 0.14$.

probably understated so we can conclude that path dependence due simply to agglomeration does not really explain our finding.

Finally we use the number of waterwheels in the cotton industry as a simple measure of the level of sunk costs in power equipment. Water power should have exerted an indirect effect on the location of steam engines because it had promoted the installation of waterwheels and auxiliary investment such as buildings and of course spinning or weaving machinery using power equipment. Insofar as these could be re-vamped using steam engines (Taylor 1949), their presence might dampen the incentive to relocate away from a location in response to contemporaneous differences in variable costs, like coal prices. As shown in Table 5, column 3, there is indeed strong evidence that such sunk costs mattered. Controlling for both access to markets and the number of waterwheels, which should capture the essence of the mechanism, the coefficient on access to water power for the number of steam engines across locations stays positive but it becomes statistically insignificant. We conclude that our results on the role of "original" and "acquired" advantages using cross-sectional data reflect much more than the contemporaneous variation across locations in 1838, including also some relevant interactions over time.

c) Discussion

Clearly, Lancashire's dominance of the cotton textiles industry was not based on having the best position with regard to any of Farnie's 'original advantages' and, in turn, none of these was a sufficient condition for cotton textiles to be present in 1838 (see our maps and Table 2). For example, much of Scotland was better placed with regard to 'poverty' and water power, most of Ireland in terms of humidity. Softer water was to be found in parts of Wales. Yet, most of these locations had no cotton mills. Often, however, they were remote, being quite far away from big population centers and large ports.

If there was a necessary condition for having a cotton textiles mill in 1838, a cheap source of power is the nearest candidate. Only 20 mills out of 1823 (7 of which were in Ireland) were in locations where both the price of coal was above average and also water flow was below average. Proximity to population and ports did not prevent a complete absence of cotton mills as was the case in Bedfordshire, Berkshire and Buckinghamshire each of which had high coal prices and weak water flows.

It seems that the cotton industry's preference from Lancashire stemmed from a combination of good first and second nature geography. In terms of what Farnie called 'acquired advantages,' Lancashire benefited from cheap coal, proximity to ports, and excellent market access, which made it much more attractive to mill owners than the Celtic Fringes of the British Isles. Table 6 uses our regression results to explore these points and to provide some quantification of Lancashire's advantages. Using model 4 of Table 3, we investigate counterfactual changes in cotton textile employment to the statistically significant characteristics of Manchester and Preston.

We consider first the implications of counterfactuals in which observed characteristics are replaced one at a time by the sample average observation for the variable. The results for both Manchester and Preston show that bringing coal price and market access to their average values would have reduced employment by about 50 percent. In contrast, ruggedness has relatively weak effects and, for Manchester, the direction of change is the opposite. In terms of water power, as historians of the Lancashire cotton industry are well aware, Preston was a favored location but Manchester was not.

The counterfactual prediction is that with average access to water power employment in Manchester would have increased significantly but in Preston it would have decreased appreciably. Reducing textile inventiveness (Patent Citations) to the sample average has a big effect on Manchester's cotton employment but a more muted one for Preston (a smaller place).

Turning to the counterfactuals where a 10 percent increase is imposed on the independent variables, the most striking result is the strength of the impact of changes in market access compared with the other characteristics. For both Manchester and Preston, in terms of total cotton employment, a 10 percent increase in market access has a bigger effect than a simultaneous 10 percent change for each of the other four characteristics. Since market access effects were growing with the size of the industry while most other characteristics were stable, it is likely that cotton textiles was becoming increasingly 'locked in' to its existing heartlands. Nevertheless, good market access could be offset by other disadvantages. For example, for Preston if each of water power, coal price and textile inventiveness had been only average the model predicts cotton employment would have been close to zero and there were places even within Lancashire itself with better market access than Preston but no cotton mills.

Comparing the impacts of 10 percent changes in different variables reveals something about their relative strength and the tradeoffs that choice of location entailed. For example, Table 6 shows that a 10 percent rise in coal price could be offset by a 44 percent increase in water flow, a 15 percent rise in textile inventiveness, being 28 percent more rugged, or by a 1.6 percent improvement in market access. These results support the emphasis Rodgers (1960) gave to coal prices as a major influence in the location of the cotton industry without endorsing his claim that was the only factor. It might also be noted that, as the industry became more spatially concentrated over time, some locations lost their mills. Based on the data in Peter M. Solar and John S. Lyons (2011), we can identify observations in our data set for which there was no mill in 1838 but where there is evidence for cotton textiles at an earlier date. Exit was positively correlated with high coal prices and negatively with textile inventiveness.

Finally, there is an important caveat as to what this methodology cannot do. The counterfactuals show, at best, what would be the implications in 1838 of taking away Lancashire's advantages. They do not provide estimates of the total impact that the 'original advantages' had on the location of the industry. Indirect, legacy effects working through accrued agglomeration benefits in the form of sunk costs or otherwise obviously mattered but are not accounted for in Table 6.

Conclusion

We have shown that UK cotton textiles factories in 1838 preferred those locations with good availability of water power, rugged terrain, a history of textile invention, close to ports, and with good market access. We have found that lower coal prices did not have a significant effect on the probability that there was a cotton mill in a given location but did increase employment in the cotton industry and the size of mills. We have not found evidence to support the hypotheses that humidity and access to soft water mattered to these issues.

These results vindicate most of Farnie's claims (1979). In particular it is clear that Lancashire's dominance of cotton textiles was based both on 'original' and 'acquired' advantages. In the former category, our evidence supports the roles played by 'poverty', 'water supply' and 'textile tradition' and, in the latter category, by agglomeration benefits, cheap coal, and access to foreign markets. We

do not, however, agree with Farnie's emphasis on the value of a humid climate. Our results are consistent with the view of Rodgers (1960) that locations further away from coalfields were at a significant disadvantage in the age of steam power but not that coal prices dominated everything else.

Our results also highlight the close connection between the original geography of the industry and its subsequent location. In particular we found a strong correlation between access to water power and the use of steam engines. This seems to be explained by the impact of the sunk costs of earlier investment in equipment such as waterwheels. We have also found strong evidence for the agglomeration benefits of locating near to other producers which eventually acted to 'lock in' the industry to its heartlands. Here is a channel through which original advantages could have legacy effects in much later periods even though their initial relevance had evaporated.

This paper is based on one cross-section – the Factory Inspectors' report for 1838. While the results provide considerable insight into the determinants of the location of the cotton industry, obviously, there are some important limitations on what can be inferred especially in the presence of endogenous variables. A valuable next step using panel data will be to analyze the evolution of the industry over the remainder of the nineteenth century as it became even more spatially concentrated and reliant on steam power. It will also be important to investigate comparisons with the location patterns of cotton textiles in other countries during industrialization where persistent spatial concentration in one location was not the case.[15]

Inter alia, this will allow investigation of issues relating to the nature of technical change during and after the Industrial Revolution. The Lancashire agglomeration was central to the long-run success of British cotton textiles which were increasingly spatially concentrated in an area characterized by high wages and cheap coal. This emerging geography intensified the rationale identified by Allen (2009) for technological progress whose factor-saving bias was energy-using and labor-saving. So, an analysis of 'directed technical change' in the cotton industry over the course of the nineteenth century would seem to offer an excellent opportunity to explore Allen's hypothesis in greater depth.

---

[15] For example, in India where the industry spread out and Bombay and Bengal provinces became less important (Roy 2011) and the United States where the industry moved from New England to the South (Galenson 1985).

Technical Appendix: Second-Nature Geography and the Effect of Sunk Costs

The early 19th century was a time of dramatic technological change in the cotton industry. Especially steam power was rapidly replacing water power, which should have triggered a relocation of the industry towards places with cheap coal. In this appendix we want to show how in the presence of technological change, sunk costs can delay of even prevent the relocation of production. With their assumption regarding fixed costs, Head and Mayer (2004) do not take into account that a considerable part of the fixed costs are typically sunk costs in the sense that they have been incurred in the past (such as costs for a building), and they are location-specific so that they cannot be recovered upon exit (Motta and Thisse 1994). A simple reformulation of equation (5) in the text (section 3) can capture this idea:

$$Max(Profit_j) = Max[cm_j^{-(\sigma-1)}MP_j - F_j S_{ij}], j \in \{1, \dots, J\}, \tag{A1}$$

where S_{ij} is an indicator variable equal to some positive value bounded by 1 whenever the location of an existing mill is evaluated against a new location, and zero otherwise. In this formulation, some fraction of fixed costs will now matter for a firm that considers relocating, even if fixed costs are perfectly equal across locations. The existence of sunk costs will *dampen* the impact of current or prospective changes in the differences in variable costs and market access between locations and will *increase* the impact of past differences. To see this, consider a situation with only two possible locations (j= 1,2) at two different points in time (t=0,1) and treating market access as exogenous for simplicity (i.e. each firm is small relative to the local market). Initially (t=0), location 1 is more attractive than location 2, due to lower variable costs ($m_1 < m_2$) ceteris paribus, so that:

$$[cm_1^{-(\sigma-1)}MP_1 - F_1] > [cm_2^{-(\sigma-1)}MP_2 - F_2] \quad \text{ or simply}$$

$$Profit_1(m_1) > Profit_2(m_2). \tag{A2}$$

This will attract a larger share of firms to location 1. Now consider a reversal in variable costs in the next period (t=1), for example due to biased technological change that affects relative input prices and variable costs. With this, we have $m_1 > m_2$ or $m_1 = (1+\Delta) m_2$. A mill owner who has chosen location 1 in t=0 will have to evaluate the benefit of moving to the lower cost location 2 against the opportunity cost of moving, which will now depend on the level of sunk costs. Let us define the operating profit as profit net of fixed and sunk cost. Then we can see that the critical value of sunk costs FS*, at which a mill owner is indifferent between relocating and staying, is increasing in the difference in operating costs due to the change in variable costs:

$$OpProfit_1(m_1) = OpProfit_2(m_2) - F_2 S_{12}{}^* \quad \text{ or}$$

$$OpProfit_1((1+\Delta)m_2) = OpProfit_2(m_2) - F_2 S_{12}{}^* \quad \text{ or}$$

$$F_2 S_{12}{}^* = OpProfit_2(m_2) - OpProfit_1((1+\Delta)m_2) \tag{A3}$$

Put differently, in the presence of positive sunk costs, a higher differential in (current and prospective) operating profits is necessary as an incentive for relocation (see Redding et al 2011). By the same token, past differences in operating profits can have a more persistent effect because more firms will stay at a location even after earlier cost advantages have disappeared.

References

Acemoglu, Daron. "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *Quarterly Journal of Economics* 113, no. 4 (1998): 1055-1089.

Allen, Robert. *The British Industrial Revolution in Global Perspective*. Cambridge: Cambridge University Press, 2009.

Ashmore, Owen. *The Industrial Archaeology of Lancashire*. Newton Abbot: David and Charles, 1969.

Atwood, Rollin Salisbury. "Localization of the Cotton Industry in Lancashire, England." *Economic Geography* 4, no. 1 (1928): 187-195.

Balderston, Theo. "The Economics of Abundance: Coal and Cotton in Lancashire and the World." *Economic History Review* 63, no. 3 (2010): 569-590.

Bartholomew, John George. *The Survey Gazetteer of the British Isles.* London: George Newnes, 1904.

Bowley, Arthur Lyon. *Wages in the United Kingdom in the Nineteenth Century*. Cambridge: Cambridge University Press, 1990.

Broadberry, Stephen, and Andrew Marrison. "External Economies of Scale in the Lancashire Cotton Industry, 1900-1950." *Economic History Review* 55, no.1 (2002): 51-77.

Carlton, Dennis. "Why New Firms Locate Where They Do: an econometric Model." In *Interregional Movement and Regional Growth* edited by William Wheaton, 13-50. Washington, DC.: Urban Institute, 1979.

Combes, Pierre-Philippe, Thierry Mayer, and Jacques-François Thisse. *Economic Geography*. Princeton: Princeton University Press, 2008.

Crafts, N. F. R.. *British Economic Growth during the Industrial Revolution*. Oxford: Clarendon Press, 1985.

Deane, Phyllis and William Alan Cole. *British Economic Growth, 1688-1959*. Cambridge: Cambridge University Press, 1962.

Ellison, Glenn and Glaeser, Edward. L.. "Geographic Concentration in U.S. Manufacturing Industries: Dartboard Approach." *Journal of Political Economy* 105, no. 5 (1997): 889-927.

Farnie, Douglas. *The English Cotton Industry and the World Market, 1815-1896*. Oxford: Clarendon Press, 1979.

Farnie, Douglas and Jeremy, David. "The Role of Cotton as a World Power, 1780-1990." In *The Fibre that Changed the World: the Cotton Industry in International Perspective, 1600-1990s* edited by Douglas Farnie and David Jeremy, . Oxford: Oxford University Press, 2004.

Galenson, Alice. *The Migration of the Cotton Textile Industry from New England to the South, 1880-1930*. New York: Garland Publishing, 1985.

Gatrell, V.A.G.. "Labour, Power and the Size of Firms in Lancashire Cotton in the Second Quarter of the Nineteenth Century." *Economic History Review* 30, no. 1 (1977): 95-139.

Ginter, Donald. E.. *A Measure of Wealth: the English Land Tax in Historical Analysis.* Montreal: McGill- Queen's University Press, 1992

Guimaraes, Paulo, Octávio Figueirdo, and Douglas Woodward. "A Tractable Approach to the Firm Location Decision." *Review of Economics and Statistics* 85, no. 1 (2003): 201-204.

Hanlon, W. Walker. "Industry Concentration and the Geographic Location of Economic Activity." mimeo, 2013.

Head, Keith and Thierry Mayer. "Market Potential and the Location of Japanese Investments in the European Union." *Review of Economics and Statistics* 86, no. 4 (2004): 959-972.

———————————————. "Regional Wage and Employment Responses to Market Potential in the EU." *Regional Science and Urban Economics* 36, vol. 5 (2006): 573-594.

Holden, Roger N.. "Water Supplies to Steam-Powered Mills." *Industrial Archaeology Review* 21, no. 1 (1999): 41-51.

Kane, Robert. *The Industrial Resources of Ireland*. Dublin: Hodges and Smith, 1845.

Krugman, Paul. "Increasing Returns and Economic Geography." *Journal of Political Economy* 99, no. 3 (1991): 483-499.

Lander, Cecil H.. *Ventilation and Humidity in Textile Mills and Factories*. London: Longmans, Green and Co, 1914.

Langton, John. *Geographical Change and Industrial Revolution*. Cambridge: Cambridge University Press, 1979.

Lawrence, Mark G.. "The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: a Simple Conversion and Applications." *Bulletin of the American Meteorological Society* 86, no. 2. (2005): 225-233.

Leunig, Timothy. "A British Industrial Success: Productivity in the Lancashire and New England Cotton Spinning Industries a Century Ago." *Economic History Review* 56, no. 1 (2003): 90-117.

Marshall, Alfred. *Industry and Trade*. London: Macmillan, 1919.

Maw, Peter, Terry Wyke, and Alan Kidd. "Canals, Rivers, and the Industrial City: Manchester's Industrial Waterfront, 1790-1850." *Economic History Review* 65, no. 4 (2012): 1495-1523.

McFadden, Daniel. "Modeling the Choice of Residential Location." In *Spatial Interaction Theory and Residential Location* edited by A. Karlquist et al., 75-96. Amsterdam: North-Holland, 1978.

Melitz, Marc J.. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71, no. 6 (2003): 1695-1725.

Melitz, Marc J. and Redding, Stephen. J.. "Heterogeneous Firms and Trade." NBER Working Paper No. 18652, Cambridge, MA, December 2012.

Motta, Massimo, and Jacques-François Thisse. "Does Environmental Dumping lead to De-Location?" *European Economic Review* 38, no. 3-4 (1994): 563-576.

Nuvolari, Alessandro, and Valentina Tratari. "Bennet Woodcroft and the Value of English Patents, 1617-1841." *Explorations in Economic History* 48, no. 1 (2011): 97-115.

Ogden, H. W. (1927), "The Geographical Basis of the Lancashire Cotton Industry", *Journal of the Manchester Geographical Society* 43, no.1 (1927): 8-30.

Redding, Stephen, Daniel M. Sturm, and Nikolaus Wolf. "History and Industry Location: Evidence from German Airports." *The Review of Economics and Statistics* 93 no. 3 (2011): 814-831.

Rodgers, H. B.. "The Lancashire Cotton Industry in 1840." *Transactions of the Institute of British Geographers* 28, no. 1 (1960): 135-153.

Rose, Mary B.. "Introduction: the Rise of the Cotton Industry in Lancashire to 1830." In *The Lancashire Cotton Industry: a History since 1700* edited by Mary B. Rose, 1-28. Preston: Lancashire County Books,1996.

Rosés, Joan R.. "Why Isn't the Whole of Spain Industrialized? New Economic Geography and Early Industrialization, 1791-1910." *Journal of Economic History* 63, no. 4 (2003): 995-1022.

Roy, Tirthankar. *The Economic History of India, 1857-1947*. Oxford: Oxford University Press, 2003

Schmidheiny, Kurt, and Marius Brulhart. "On the Equivalence of Location Choice Models: Conditional Logit, Nested Logit and Poisson", *Journal of Urban Economics* 69, no. 2 (2011): 214-222.

Solar, Peter. M., and John S. Lyons. "The English Cotton Spinning Industry, 1780-1840, as Revealed in the Columns of the London Gazette." *Business History* 53, no. 3 (2011): 302-323.

Taylor, A. J.. "Concentration and Specialization in the Lancashire Cotton Industry, 1825-1850." *Economic History Review* 1, no. 2-3 (1949): 114-122.

University of East Anglia Climatic Research Unit (2008), CRU Time Series High Resolution Gridded Datasets. Available from http://badc.nerc.ac.uk/view/badc.nerc.ac.uk ATOM datent 1256223773328276

Von Tunzelmann, Nick. "Coal and Steam Power." In *Atlas of Industrializing Britain* edited by in John Langton and Robert J. Morris, 72-79. London: Methuen, 1986.

Williams, M.. *Cotton Mills in Greater Manchester*. Preston: Carnegie Publishing, 1992.

Woodcroft, Bennet. *Titles of Patents of Invention Chronologically Arranged, 1617-1852*. London: Eyre and Spottiswoode, 1854.

Woodcroft, Bennet. *Reference Index of English Patents of Invention, 1617-1852*. London: Eyre and Spottiswoode, 1862.

Official Publications

BPP (1839). Return of All the Mills and Factories.

BPP (1843). Account of the Prices of the Prices of Articles of Consumption at the Poor Law Unions throughout England and Wales.

BPP (1844a). Report of the Poor Law Commissioners for Scotland.

BPP (1844b). Return Showing the Annual Value of Real Property in each Parish of Each County.

BPP (1844c). Occupational Abstract: 1841 Population Census.

BPP (1850). Return of the Number of Cotton, Woollen, Worsted, Flax and Silk Factories subject to the Factories Acts in the United Kingdom.

BPP (1854). Vessels and Tonnage at Each of the 12 Principal Ports of the United Kingdom.

BPP (1862). Poor Law Unions: a Return of the Area and Population.

BPP (1871). Report of the Commissioners Appointed to Inquire into the Several Matters relating to Coal in the United Kingdom.

BPP (1903). Cotton Factories.

UK (2008). Hydrometric Register.

Table 1: The Location of the Cotton Industry in 1838

		Number of Cotton Mills	Share in UK in %	Employment	Share in UK in %
England	North-West	1533	84.1	211454	81.1
	…of which Lancashire	1125	61.7	152145	58.3
	North, North-East and Central	45	2.4	5806	2.0
	South-West	1	0.1	29	<0.1
	South-East	19	1.0	942	0.4
Wales		5	0.3	1010	0.4
Scotland		196	10.8	36918	14.2
Ireland		24	1.3	4622	1.8
United Kingdom		1823	100	260781	100
Herfindahl-Index of Concentration Lancashire, 31 Poor Law Unions [Normalized]		0.106 [0.076]		0.117 [0.088]	
Herfindahl-Index of Concentration England, 41 Counties [Normalized]		0.519 [0.507]		0.520 [0.508]	
Herfindahl-Index of Concentration UK, 118 Counties [Normalized]		0.404 [0.398]		0.371 [0.366]	

Source: Factory Inspectors' report for 1838 (BPP, 1839).

North West is Cheshire, Derbyshire, Lancashire, and Yorkshire WR. North is Northumberland, Cumberland, Durham, Westmoreland, North East is Lincolnshire, Nottinghamshire, Yorkshire ER and Yorkshire NR. Central is Warwickshire, Leicestershire, Northamptonshire, Huntingdonshire, Rutland, Cambridgeshire, Bedfordshire, Buckinghamshire, Oxfordshire, Hertfordshire, Middlesex, Berkshire. South East is Norfolk, Suffolk, Surrey, Essex, Kent. South West is Cornwall, Devon, Dorset, Gloucestershire, Hampshire, Herefordshire, Somerset Staffordshire, Worcestershire, and Wiltshire.

Table 2: Characteristics of Different Locations, c.1838

	Manchester	Preston	Full Sample Average	Lowest	Highest
Market Access 1838	10215	7143	3541	782 (Kerry)	20110 (Chorlton, Lancs.)
Market Access 1801	112	89	73	29 (Caithness)	125 (Chorlton, Lancs.)
Distance to major ports (km)	185	196	310	182 (Warrington,)	630 (Caithness)
Coal Price 1842/43 (d/ton)	64	89.5	173.5	64.0 (Leigh, Wigan, Manchester,)	369.6 (Hertfordshire)
Coal Deposits 1871	13.1	0	120608	0	5140589 (Carmarthen-shire)
1690 Land Tax (d/acre)	126.65	4.69	9.81	0.04 (Sutherland)	252.55 (Middlesex)
Ruggedness (Index)	24.8	46.4	42.2	1.6 (Norfolk)	148.7 (Stirlingshire)
Water power: Q95 (m^3/sec)	0.5	4.4	1.11	0.02 (Oxfordshire, Anglesey)	5.82 (Montgomery-shire)
Relative Humidity (%)	88.26	88.25	89.081	83.486 (Banff)	96.394 (Fermanagh)
Water Hardness (mg/litre)	15.212	15.212	36.903	1.79 (Cromarty)	132.41 (Huntingdon-shire)
Textile Inventions (Patent Citation Index)	33.6	6.55	4.8	0 (21 locations)	239.8 (Middlesex)
Agricultural Wages ca 1770 (d/day)	13	13	11	5 (Kings County)	20 (Kent)
Agricultural Wages ca 1830 (d/day)	25.3	25.3	18	6 (Longford)	27 (Kent)

Sources: see text

Table 3: Poisson Model of Cotton Industry Location, Britain 1838

Dep. Var.:	Number of cotton mills		Number of cotton mill employees		
	(1)	(2)	(3)	(4)	(5)
Ln(Area)	0.145 (0.186)	0.198 (0.118)*	0.193 (0.175)	0.238 (0.118)	0.088 (0.084)
Ln(Ruggedness)	1.249 (0.462)***	0.898 (0.193)***	0.580 (0.375)	0.328 (0.168)*	0.133 (0.138)
Ln(Relative Humidity)	-34.917 (24.551)	-	-27.197 (17.963)	-	-
Ln(Waterpower)	0.314 (0.035)***	0.282 (0.035)***	0.255 (0.043)***	0.207 (0.053)***	0.311 (0.154)**
Ln(Totalhardness)	-0.031 (0.13)	0.035 (0.257)	0.023 (0.043)	0.070 (0.426)	-0.199 (0.252)
Ln(Patent Citations)	0.788 (0.032)***	0.717 (0.069)***	0.804 (0.030)***	0.753 (0.062)***	0.038 (0.054)
Ln(Coalprice)	-0.426 (0.745)	-0.411 (0.673)	-1.081 (0.651)*	-1.046 (0.537)*	-1.551 (0.272)***
Ln(Weighted Dist)	-3.823 (0.962)***	-1.778 (1.420)	-3.292 (1.302)**	-1.472 (1.821)	-5.219 (1.357)***
Ln(MP1801)	4.405 (1.073)***	4.603 (1.109)***	4.561 (0.845)***	4.773 (1.039)***	-2.733 (1.207)**
Constant and Region dummies	yes	yes	yes	Yes	Yes
# of Obs.	148	148	148	148	148
Pseudo *R2*	0.76	0.74	0.76	0.74	0.30
Log Pseudolikelihood	-869.20	-944.93	-123258.55	-131045.49	-7722.47

Note: robust standard errors in parentheses, * indicates significance at the 10 percent level; ** significance at the 5 percent level; *** significance at the 1 percent level.

Sources: see text.

Table 4a: Poisson Model of Cotton Industry Location, 118 Counties of Great Britain

Dep. Var.:	# of cotton mills	# of employed persons in cotton mills	# of employed persons per mill
Ln(Area)	-0.131 (0.186)	0.065 (0.163)	0.084 (0.264)
Ln(Ruggedness)	-0.573 (0.443)	-0.979 (0.324)**	-0.059 (0.214)
Ln(Waterpower)	0.769 (0.403)*	0.910 (0.279)***	0.477 (0.206)**
Ln(Totalhardness)	-0.374 (0.342)	-0.258 (0.341)	-0.226 (0.200)
Ln(Patent Citations)	0.684 (0.301)**	0.258 (0.178)*	-0.022 (0.201)
Ln(Coalprice)	-2.752 (0.842)***	-3.278 (0.396)***	-1.903 (0.362)***
Ln(Weighted Dist)	-1.703 (2.125)	-0.506 (2.674)	-5.832 (1.456)***
Ln(MP1801)	4.143 (0.783)***	5.384 (0.906)***	-3.390 (1.401)***
Constant	11.032 (14.929)	5.959 (19.876)	63.181 (17.277)***
# of Obs.	118	118	118
Pseudo *R2*	0.950	0.947	0.326
Log Pseudolikelihood	-296.965	-43595.905	-6328.562

Note: robust standard errors in parentheses, * indicates significance at the 10 percent level; ** significance at the 5 percent level; *** significance at the 1 percent level.

Sources: see text.

Table 4b: Poisson Model of Cotton Industry Location, 31 Poor Law Unions of Lancashire 1838

Dep. Var.:	# of cotton mills	# of employed persons in cotton mills	# of employed persons per mill
Ln(Area)	-0.295 (0.319)	-0.085 (0.291)	0.315 (0.192)*
Ln(Ruggedness)	3.011 (0.933)***	2.095 (0.722)***	0.082 (0.281)
Ln(Waterpower)	0.201 (0.167)	0.021 (0.129)	-0.017 (0.101)
Ln(Totalhardness)	2.793 (1.030)**	1.556 (0.971)*	0.254 (1.101)
Ln(Patent Citations)	0.476 (0.266)*	0.754 (0.276)***	0.391 (0.197)**
Ln(Coalprice)	-1.111 (0.732)	-1.412 (0.756)*	-1.462 (0.612)**
Ln(Weighted Dist)	-13.239 (13.225)	-0.618 (10.618)	6.662 (5.846)
Ln(MP1801)	2.517 (3.225)	3.769 (3.140)	2.100 (2.621)
Constant	48.613 (80.332)	-11.657 (67.220)	-36.340 (40.123)
# of Obs.	31	31	31
Pseudo *R2*	0.712	0.772	0.286
Log Pseudolikelihood	-328.263	-34345.758	-970.676

Note: robust standard errors in parentheses, * indicates significance at the 10 percent level; ** significance at the 5 percent level; *** significance at the 1 percent level.

Sources: see text.

Table 5: The persistent effect of water power: sunk costs and market access

Dep. Var.	# of steam engines in cotton industry		
	(1)	(2)	(3)
Ln(Area)	0.249 (0.138)*	0.189 (0.057)***	-0.182 (0.123)
Ln(Ruggedness)	0.254 (0.248)	-	-
Ln(Waterpower)	0.255 (0.044)***	0.181 (0.041)***	0.032 (0.029)
Ln(Totalhardness)	0.186 (0.407)	-	-
Ln(PatentCitations)	0.761 (0.072)***	-	-
Ln(Coalprice)	-1.062 (0.557)*	-	-
Ln(Weighted Dist)	-0.792 (1.379)	-	-
Ln(MP1801)	5.707 (0.949)***	-	-
Ln(MPCotton38)	-	0.585 (0.233)**	0.655 (0.584)
Ln(Waterwheels)	-	-	0.474 (0.210)***
Constant and Region dummies	Yes	yes	Yes
# of Obs.	148	148	148
Pseudo *R2*	0.744	0.473	0.738
Log Pseudolikelihood	-24995.51	-51619.61	-25671.53

Sources: see text.

Table 6: Counterfactuals

|		Manchester			Preston		
		Orig. Variable	Change in Variable	Cotton Employment	Orig. Variable	Change in Variable	Cotton Employment
Actual data 1838		-	-	39363	-	-	8061
Benchmark Prediction (Table 2, column4)		-	-	30645	-	-	10021
Counterfactuals: predicted employment effect with sample average	Ruggedness	24.852	42.181	36450	46.416	42.181	9711
	WaterPower	0.5	1.11	36199	4.4	1.11	7532
	Patent Citations	33.583	4.847	8040	6.555	4.847	8263
	Coalprice	64	173.530	10790	89.5	173.530	5012
	Market Access 1801	111.758	73.230	4073	89.286	73.230	3890
Counterfactuals: predicted employment effect of a 10% increase of…	Ruggedness	24.852	27.337	31618	46.416	51.057	10339
	Water Power	0.5	0.55	31258	4.4	4.84	10221
	Patent Citations	33.583	36.941	32859	6.555	7.210	10669
	Coalprice	64	70.4	27735	89.5	98.45	9069
	Market Access 1801	111.758	122.934	48299	89.286	98.2156	15794
Counterfactuals: predicted employment effect of a 10% decrease of…	Ruggedness	24.852	22.368	29604	46.416	41.774	9681
	Water Power	0.5	0.45	29981	4.4	3.96	9804
	Patent Citations	33.583	30.224	27688	6.555	5.899	8318
	Coalprice	64	57.6	*34218*	89.5	80.55	11191
	Market Access 1801	111.758	100.582	18532	89.286	80.358	6060

Sources: see text.

Figure 1: Scatterplot of Market Potential (population 1801) and Market Potential (cotton textiles employment 1838)
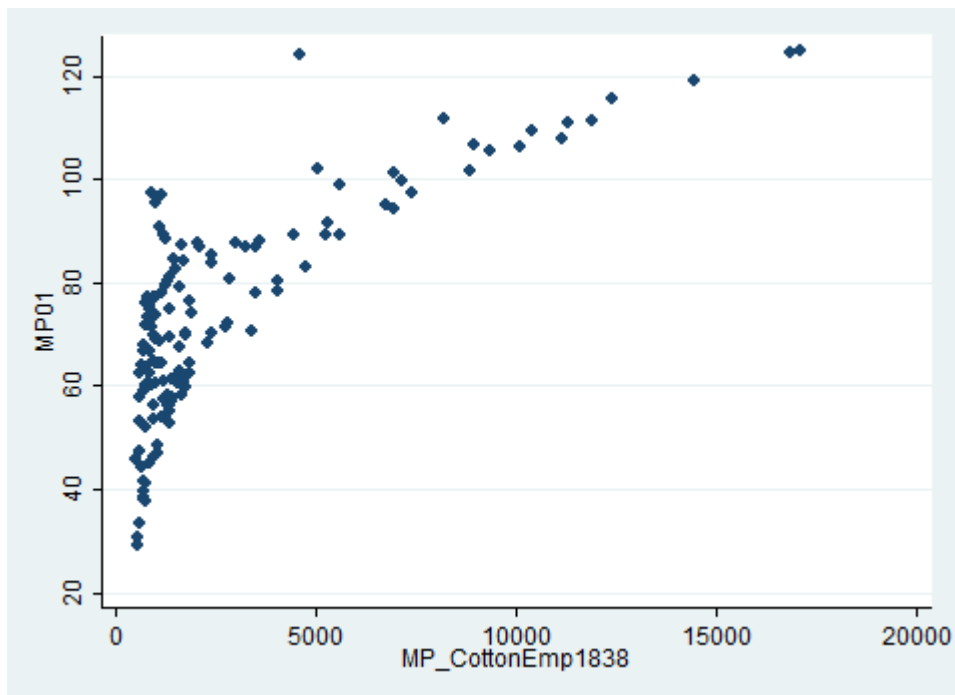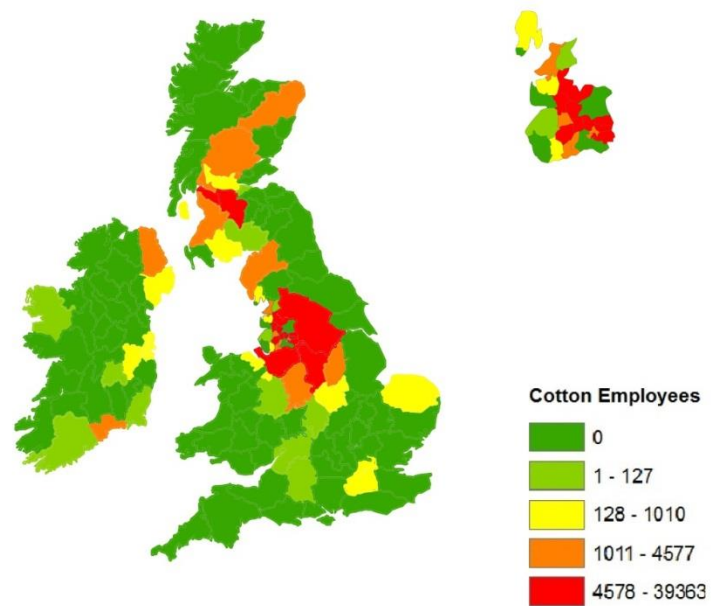

Source: see text.

Map 1: the location of cotton mills in Britain 1838
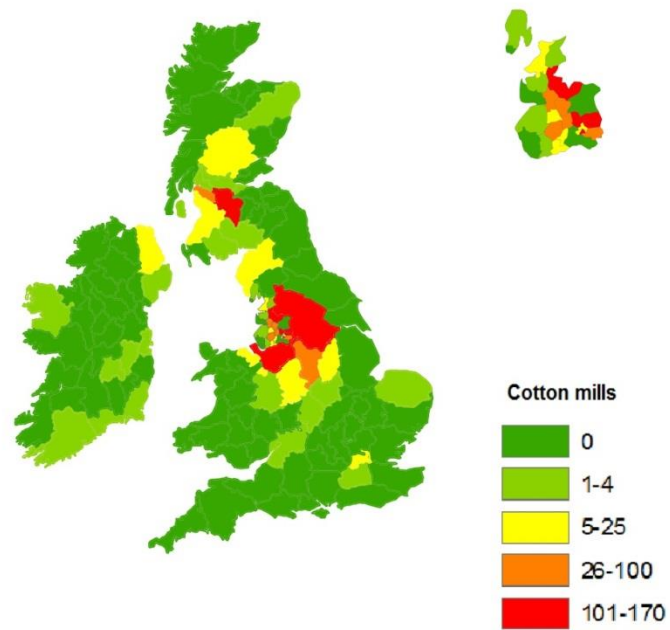
Cotton mills
- 0
- 1-4
- 5-25
- 26-100
- 101-170

Note: the inlay in the right upper corner shows Lancashire and its 31 Poor Law Unions.

Source: own map based on Factory Inspectors' report for 1838 (BPP 1839).
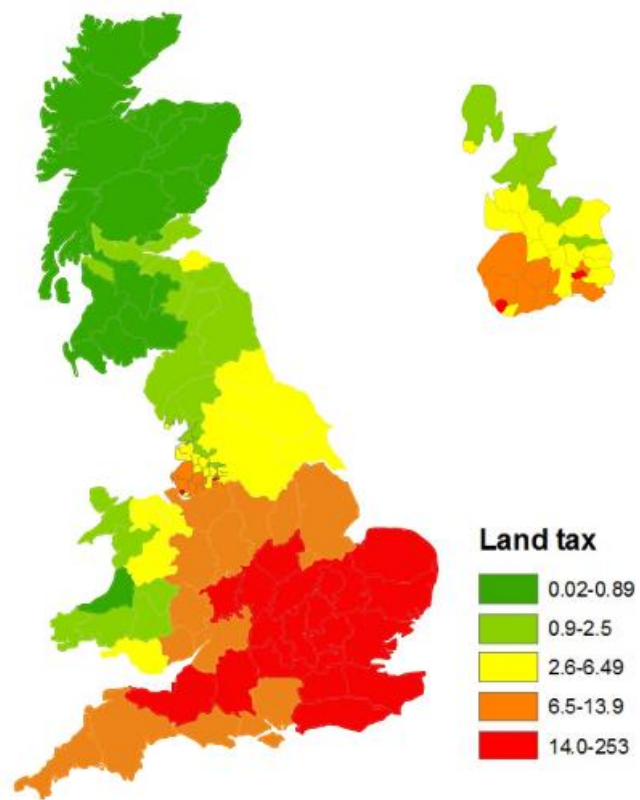
Map 2: the location of employment in the cotton industry in Britain 1838

Note: the inlay in the right upper corner shows Lancashire and its 31 Poor Law Unions.
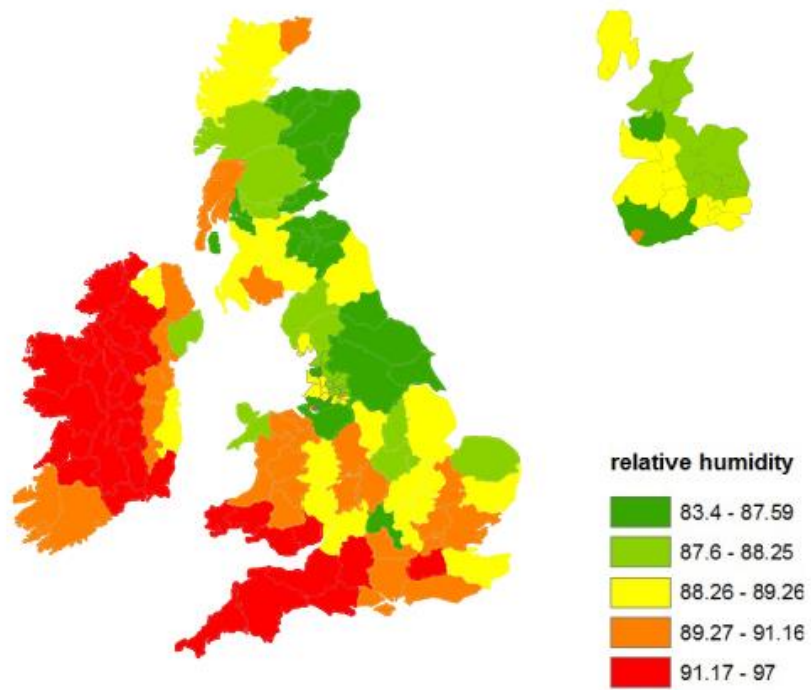
Source: own map based on Factory Inspectors' report for 1838 (BPP 1839).

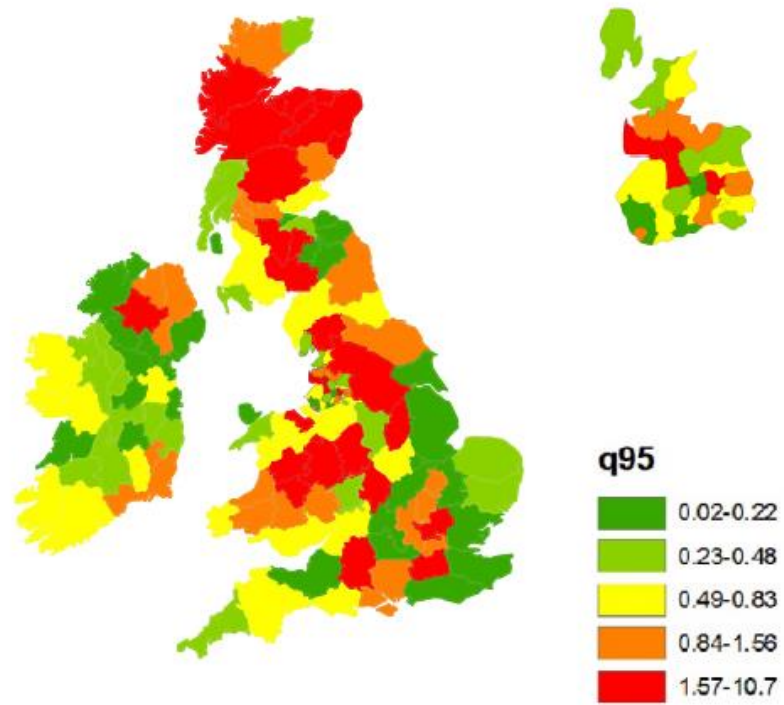**Map 3: land taxes in England, Wales and Scotland (shilling per acre, ca 1700)**

Sources: see text.

Map 4: average relative humidity in %

relative humidity

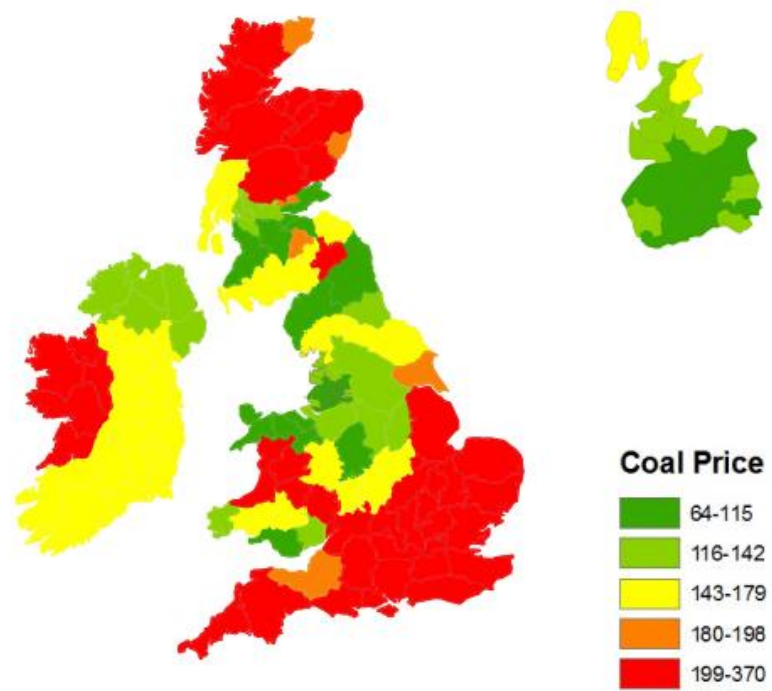|	
![green]	83.4 - 87.59
![light green]	87.6 - 88.25
![yellow]	88.26 - 89.26
![orange]	89.27 - 91.16
![red]	91.17 - 97

Sources: see text.

Map 5: Water flow experienced at least 95 percent of the time at gauging stations in the UK and Ireland (m³/sec)
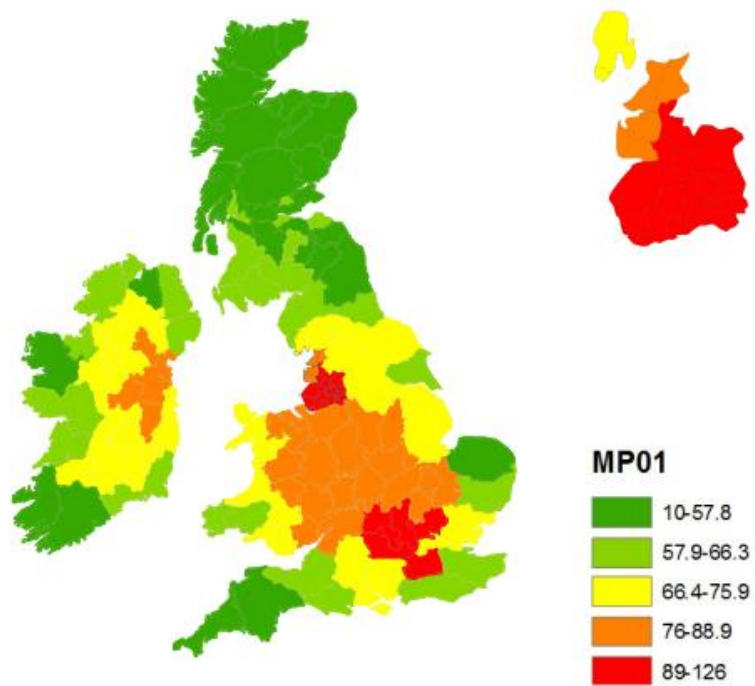

Source: see text.

Map 6: British coal prices (shillings per ton, around 1842)

Coal Price

- 64-115
- 116-142
- 143-179
- 180-198
- 199-370

Note: the inlay in the right upper corner shows Lancashire and its 31 Poor Law Unions.

Source: see text.

Map 7: Market-potential (access to population in 1801, weighted by distance)

Sources: see text.