DISCUSSION PAPER SERIES

No. 9819

COAL AND THE EUROPEAN INDUSTRIAL REVOLUTION

Alan Fernihough and
Kevin Hjortshøj O'Rourke

ECONOMIC HISTORY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP9819.php**

ISSN 0265-8003

COAL AND THE EUROPEAN INDUSTRIAL REVOLUTION

Alan Fernihough, Trinity College Dublin
Kevin Hjortshøj O'Rourke, Oxford University and CEPR

Discussion Paper No. 9819
February 2014

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9819

February 2014

ABSTRACT

Coal and the European Industrial Revolution*

We examine the importance of geographical proximity to coal as a factor underpinning comparative European economic development during the Industrial Revolution. Our analysis exploits geographical variation in city and coalfield locations, alongside temporal variation in the availability of coal-powered technologies, to quantify the effect of coal availability on historic city population sizes. Since we suspect that our coal measure could be endogenous, we use a geologically derived measure as an instrumental variable: proximity to rock strata from the Carboniferous era. Consistent with traditional historical accounts of the Industrial Revolution, we find that coal had a strong influence on city population size from 1800 onward. Counterfactual estimates of city population sizes indicate that our estimated coal effect explains at least 60% of the growth in European city populations from 1750 to 1900. This result is robust to a number of alternative modelling assumptions regarding missing historical population data, spatially lagged effects, and the exclusion of the United Kingdom from the estimation sample.

JEL Classification: J10, N13, N53, O13 and O14
Keywords: coal, geography, historical population and industrial revolution

Alan Fernihough
Institute for International Integration Studies
Trinity College
Dublin 2
IRELAND

Email: alan.fernihough@gmail.com

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=178007

Kevin Hjortshøj O'Rourke
Chichele Professor of Economic History
All Souls College
Oxford University
OX1 4AL

Email: kevin.orourke@all-souls.ox.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=114845

*This research has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013), ERC grant agreement no. 249546. We are grateful to Bob Allen,Kristine Asch, Michael Athanson, Nick Crafts, James Fenske, Morgan Kelly, Alexander Mirko Müller, Cormac Ó Gráda and Colin Waters for useful comments, advice and help with data. We are also grateful to seminar and conference participants at the University of Southern Denmark, Nottingham, Oxford, the 9th BETA Workshop in Historical Economics at the University of Strasbourg, the University of Warwick Summer School on Economic Growth 2013, and a conference in honour of Karl Gunnar Persson held at the University of Copenhagen. The usual disclaimer applies.

Submitted 08 January 2014

1 Introduction

There is a long tradition in economic history that places coal at the centre of the Industrial Revolution. For many economic historians trained in history departments, the Industrial Revolution *was* a switch towards coal, above all else. This paper will focus on two claims that have been made for coal, one temporal and one spatial. Focusing on variation over time, it was argued that harnessing coal explained a large share of subsequent economic growth (what we will refer to as the *growth hypothesis*). Focusing on variation across space, it was argued that the location of industry was strongly influenced by the location of coalfields (what we will refer to as the *location hypothesis*).[1]

In contrast, several economic historians trained in economics departments have been more circumspect about the importance of coal. In some instances, they have been downright sceptical, especially about the location hypothesis, but in some cases about the growth hypothesis also. Both sides have used various data to support their claims, and cliometricians have used back of the envelope calculations to try to measure the importance of coal. However, systematic econometric studies of coal's importance during the Industrial Revolution have been surprisingly scarce. Careful quantification would thus seem essential in order to distinguish between the two positions.

This paper provides an econometric test of both the location and growth hypotheses in a pan-European context. More precisely, it tests (and quantifies) the latter, based on an identification strategy that relies on (and tests) the former. Because of our interest in the location of economic activity, we need fine-grained spatial data spanning the periods both before and after the Industrial Revolution. Unfortunately, data on European regional GDP or industrial output are not available over the time scale needed, and in any event regional data might not be geographically fine-grained enough to test the influence of proximity to coalfields in a sufficiently rigorous manner. We therefore use data on city sizes, in line with several other recent studies (Nunn and Qian 2011, Dittmar 2011, and Cantoni 2013).[2] While these do not provide us with information on income per capita, they are a good indicator of the geographical distribution of economic activity at a point in time. Furthermore, both Dittmar (2011) and Cantoni (2013) use city sizes to assess the economic growth consequences of the introduction of the printing press, and

[1]Until declines in transport costs sufficiently weakened the gravitational pull of coalfields, some time in the late 19th or early 20th century (Wright 1990).

[2]We cannot explore urbanization rates, since data on the denominator (total population) are not available at the level of regional disaggregation required for our purposes.

Protestantism, respectively.

More precisely, we use a difference-in-differences (DID) strategy, to see whether cities that were located closer to coalfields grew more rapidly during and after the Industrial Revolution (but not before) than those further away. We find that being located close to a coalfield mattered for city sizes after the Industrial Revolution, but not before; we take this to be strong evidence in favour of the location hypothesis. We also find that the introduction of the coal-using technologies of the Industrial Revolution can account for up to 60% of European urban growth between 1750 and 1900, even when controlling for period fixed effects. Because we do not have data on city GDPs, some readers may not regard this as constituting a direct test of the growth hypothesis, Dittmar (2011) and Cantoni (2013) notwithstanding. However, the finding that coal mattered to this extent in determining city growth after the Industrial Revolution is an important finding in its own right, and is certainly consistent with the growth hypothesis.

The paper proceeds as follows. Section 2 summarizes the debate between those who think that coal was central to the Industrial Revolution, and those who have downplayed its significance. Section 3 provides a brief literature survey, while Section 4 discusses some historical issues relating to identification. Section 5 introduces our data. Section 6 establishes the main empirical results of the paper, and Section 7 concludes.

2 The Historical Debate

There are two distinct arguments that were traditionally made about coal.

The first is that the switch to using coal as a source of energy, above all in metallurgy and in steam engines, was the central fact permitting the take-off to modern economic growth. According to Deane (1965, p. 129), "The most important achievement of the industrial revolution was that it converted the British economy from a wood-and-water basis to a coal-and-iron basis." For Landes (1965, p. 274), three phenomena "constituted the Industrial Revolution": "the substitution of machines...for human skill and effort; the substitution of inanimate for animate sources of power, in particular, the introduction of engines for converting heat into work, thereby opening to man a new and almost unlimited supply of energy; the use of new and far more abundant raw materials, in particular, the substitution of mineral for vegetable or animal substances." "Mechanisation required large sources of power" wrote Braudel (1973, pp. 274–275), but "they were not available until after the eighteenth century...Came steam and everything was, as if by magic, speeded up."

The best-known exponent of this view over the past fifty years has been E.A. Wrigley, who regards the switch to coal as having been "a necessary condition for the industrial revolution" (Wrigley 2010, p. 23), although not a sufficient one. For Wrigley, the Industrial Revolution was above all a transition from an "organic economy" to an "energy-rich economy". In the former, photosynthesis was the major source of energy (the others being water and wind), whether this energy took the form of human or animal power, wood, or charcoal. Land was thus an indispensable input into all material production, even of metallic products (since ores were heated with wood or charcoal) (*ibid.*, p. 9).

In such circumstances, it is not hard to see how Malthusian constraints could bind tightly. "Iron, for instance, has many physical properties that make it of the greatest value to man but as long as the production of 10,000 tons of iron involved the felling of 100,000 acres of woodland, it was inevitable that it was used only where a few hundredweight or at most a few tons of iron would suffice for the task in hand" (Wrigley 1988, p. 80). The switch to coal allowed humans to tap into a vast capital reserve of energy—"stored sunlight", as Cipolla (1978, p. 62) calls it—which allowed them to break free of these constraints. In a famous calculation, Wrigley estimated that English coal production in 1800 yielded energy that would otherwise have required 11 million acres of woodland (and that British coal production was equivalent to 15 million acres of woodland). This compares with a total English land area of 32 million acres (and a British land area of 57 million acres) (Wrigley 1988, pp. 54–55; Wrigley 2010, p. 39). By the 1820s, British coal production "liberated" an area as large as the entire island (Sieferle 2001, p. 103).[3]

The second argument traditionally made about coal is that local supplies of coal were essential, or at least highly desirable, if a region was to industrialize during the 19th century. Matthias (1983, p. 11) puts the argument starkly: "The logistics of energy inputs based upon coal, translated against available transport in a pre-railway age, precluded any major industrial complex in heavy industry from developing except where coal and ore were plentiful and adjacent to one another or to water carriage." Coal was bulky, heavy and costly to transport. It was also a fuel, whose weight vanished when

[3]Landes (1965, p. 327) calculates that the UK was consuming coal in 1870 whose calorific content could have fed 850 million adult males.

it was used in the production process: there were thus substantial cost savings if coal was used close to where it was mined (Wrigley 1961, pp. 6–7).[4] Pollard (1981, p. 4) remarks that "the map of the British Industrial Revolution, it is well known, is simply the map of the coalfields", and Britain's good fortune in being well-endowed with coal has often been noted (*ibid.*, p. 40). On the European Continent, the coalfields of Belgium and northern France, and later on the Ruhr, became major centres of heavy industry, and other industrial regions "only survived if they had reasonable access by water to a supply of good coal" (*ibid.*, p. 121).

On a grander scale, Pomeranz (2000) has argued that coal was a crucial reason why the Industrial Revolution happened in Europe rather than in China. In Europe the coal was abundantly located in the most dynamic economy of the 18th and 19th centuries, Britain. By contrast, China's coal resources were in "the wrong place", in the north and northeast, far from the southern coastal regions where China's most dynamic regions were located. And in a recent cliometric contribution, Allen (2009) has argued that innovation during the British Industrial Revolution was geared towards saving expensive labour, and replacing it with cheap capital and coal. These new technologies were unprofitable where coal was too expensive relative to labour, although as their efficiency improved over time they diffused over an increasingly wide area. Directed technological progress can thus help to explain why modern industrial techniques were initially adopted close to sources of cheap coal.

There are several reasons why cliometricians have tended to downplay the role of coal.

First, the increased use of coal was a symptom of technological change, which all authors accept was the main driver of the Industrial Revolution, and which has been the focus of a series of major works by Mokyr (1990, 2002, 2009). England always had coal, but it took the Industrial Revolution for this geographical advantage to achieve its full economic potential. There is no real dispute on this point: as we have seen Wrigley does not view coal as being a sufficient condition for the Industrial Revolution, but rather as a necessary one.[5]

Second, several authors dispute the notion that coal, or more broadly the search

[4] "If the full weight of the raw material is embodied in the product there is no saving in the total cost of transport when the source of the raw material is also the point of manufacture: but if a part or the whole of the weight disappears during manufacture, the saving in transport costs which follows from manufacture at the source of the raw material may be considerable" (*ibid.*).

[5] See also Pomeranz (2000, pp. 66–68), who is also cited in Clark (2007, p. 260).

for energy efficiency, was a driver of technological change. "In the absence of coal, the ingenuity applied to using it would have been directed towards replacing it...Resource scarcities, like demand, are a steering mechanism, not a *primum movens*, of technological progress" (Mokyr 1990, pp. 160, 162; see also Mokyr 1993, p. 31; McCloskey 2011, pp. 188–189). This paper does not take a position on the issue of what drove technological change: its concern is whether, once the Industrial Revolution was in progress, proximity to coal started to matter for the location of economic activity, and whether coal-using technological change mattered a lot or a little for urban growth.

Third, cliometricians have pointed out that coal could be transported, albeit at a cost, and that coal only accounted for a fraction of the cost in several leading industries of the Industrial Revolution, notably textiles.[6] Mokyr (1976, pp. 204–208) argues that local supplies of coal in Belgium cannot explain why it industrialized, while the Netherlands did not: the Dutch could import coal by sea, and use both peat and wind. In a similar vein, he dismisses the argument that pre-Famine Ireland did not industrialize because of a lack of suitable coal deposits (Mokyr 1985, pp. 152-158). Ireland imported coal from Britain, with the result that its coal prices were between 100 and 150% higher; fuel costs in the "nonmetallurgical industries" were at most 4% of total costs. The lack of suitable local coal supplies thus increased Irish costs by at most 10% relative to British costs, and by less once substitution possibilities are taken into account. Lower Irish wages should have more than compensated for this. True, being close to coal mattered more in metallurgy than in textiles, and it mattered more in the days before widespread and efficient railway transport: even coal sceptics like Mokyr and McCloskey (2011, p. 187) recognize this. But both authors doubt whether coal was as significant a locational factor as was traditionally claimed. In a recent contribution, Clark and Jacks (2007) admit that coal may have been an important locational factor as far as iron making was concerned, but that the latter sector contributed little to Industrial Revolution productivity growth. "In a counterfactual world where the coal reserves were located in Ireland or Scotland or elsewhere in northwest Europe the history of Industrial Revolution England need not have resulted in much slower economic growth" (Clark and Jacks 2007, p. 65).

Fourth, some authors have disputed the growth hypothesis. Both McCloskey and Mokyr stress that technological progress was extremely broad-based during this period: "The industrial revolution was not the Age of Cotton or of Railways or even of Steam

[6]On the other hand, Balderston (2010) points out that between 1875 and 1884, the UK cotton industry consumed 10 pounds of coal for every one pound of raw cotton: the fact that coal accounted for such a small share of the industry's costs reflects its abundance and consequent cheapness.

entirely; it was an age of improvement" (McCloskey 1981, p. 118). In this context, an argument often made against the importance of coal, or any other single factor thought to have "caused" the Industrial Revolution, has to do with substitution: "a coal theory, or any other one-step geographical theory, has an appointment with Harberger" (McCloskey 2011, pp. 186–187). "The Industrial Revolution did not absolutely 'need' steam..., nor was steam power absolutely dependent on coal" (Mokyr 2009, pp. 101–102). Water power, peat and wood were all potential substitutes for coal, and water power was very important well into the 19th century (and was used with increasing technical efficiency) (*op. cit.*, p. 127). Scarcer and dearer coal would have implied greater fuel efficiency, and an economy producing fewer energy-intensive goods (*op. cit.*, p. 104): the net cost to the economy might have been modest. Von Tunzelman's (1978) social savings calculations for steam engines in 1800 are tiny. Clark and Jacks (2007) provide an even more heroic calculation, assuming that in the absence of any European coal Britain would have had to import the equivalent of its coal consumption in the form of Baltic timber. They estimate that this more expensive fuel would have cost the British economy no more than 4% of GDP as late as the 1860s, and while by that stage Baltic timber supplies would have come under strain, the textiles revolution "would have been well under way in the 1820s and 1830s before energy constraints became even a significant issue" (p. 68).

3 Previous Literature

Did coal matter a lot or a little for the location of economic activity? Did it matter a lot or a little for post-Industrial Revolution growth? These are empirical issues requiring econometric investigation.

There have been some country-specific studies testing the importance of coal for the location of specific industries within a Heckscher-Ohlin framework: in this model, abundant coal should matter for the location of fuel-intensive industries. The evidence is mixed: Crafts and Mulatu (2006) find strong evidence that coal abundance mattered for the location of steam-intensive industries within late 19th century Britain.[7] By contrast, Wolf (2007) finds no evidence that mineral endowments explained the location of fuel-intensive industries in interwar Poland, Klein and Crafts (2012) find little evidence

[7] Crafts and Wolf (2012) find mixed results regarding the importance of coal for the location of cotton mills in Britain in 1838.

that coal prices mattered for the location of fuel-intensive industries in the United States between 1880 and 1920, and Martinez-Galarraga (2012) finds an effect of mineral endowments on the location of mineral-intensive industries in Spain in 1913, but not in other years. In a careful recent study, Gutberlet (2012) finds that access to coal mattered for the location not only of metallurgy in late 19th century Germany, but of cotton textiles production as well.

In contrast to these studies, we are interested in the overall location of economic activity, as proxied by city size, rather than in the location of particular manufacturing sectors. Those cliometricians who deny the importance of coal to aggregate growth after the Industrial Revolution would not deny that being close to coal might have mattered for the location of particularly fuel-intensive industries. They would, on the other hand, argue that not being close to coal would have led to regions and national economies specializing in industries that were not fuel-intensive, and that therefore the aggregate impact of a lack of coal would have been small. The studies cited above do not deal with this issue. Furthermore, in contrast to these papers, in this article we adopt a pan-European rather than a national approach.

There are other papers which have used a DID strategy to study the evolution of European urbanization, or city sizes, over time. For example, Andersen *et al.* (2013) measure the benefits associated with the introduction of the heavy plough during the medieval period. The cross-section variation that they exploit comes from soil type, since some soils were more conducive to the introduction of the heavy plough than others. They find that the heavy plough had a positive impact on urbanization (measured as the number of cities per square kilometre) and population, accounting for around 10% of the growth in these variables during the high medieval period. Dittmar (2011) analyses city sizes, as we do, to explore the macroeconomic impact of the printing press. He finds that cities which adopted the printing press in the 15th century grew 60% more rapidly in the 16th century than those which did not. Cantoni (2013) also uses city sizes to explore the growth consequences of Protestantism within the Holy Roman Empire. Unlike the other papers just cited, the innovation studied in this paper—religious reform—does not appear to have had an impact on city size, or by implication on economic growth.

Nunn and Qian (2011) use a DID approach to study the impact of the potato on population and urbanization between 1700 and 1900. Their identification strategy is based on the fact that some areas are better suited than others for the cultivation of potatoes. They then ask whether, after the potato was introduced to the Old World, these more suitable areas experienced higher levels of population growth and urbanization than less suitable regions. Nunn and Qian find that the potato's introduction can explain about a quarter of Old World population growth and urbanization during the 18th and 19th centuries.

4 Identification Issues

In Nunn and Qian (2011), the "treatment" is the introduction of the potato; in Dittmar (2011) it is the introduction of the printing press. In this paper, the treatment is the introduction of the coal-using technologies of the Industrial Revolution. Similarly, Nunn and Qian achieve identification by exploiting the fact that some areas are better suited to potato cultivation than others. In our case, we identify the impact of the new coal-using technologies by exploiting the fact that some cities were located closer to coalfields than others. There is of course a possibility that coalfields might have been discovered close to cities, and we take account of this by using proximity to Carboniferous rock strata as an instrument for proximity to coal.

When did the treatment which we are interested in take place? It is important to note that coal was used in pre-Industrial Revolution Britain for a wide variety of purposes, both domestic (heating) and industrial: "brickmaking, glass, ceramics, soap-boiling, lime burning, forging, distilling, and brewing" (Mokyr 2009, p. 22). Cheap domestic heating, for example, could have facilitated higher population densities even before the Industrial Revolution (Balderston 2010, p. 574). What changed during the Industrial Revolution was the use of coal in the iron and steel industry, and the introduction of the steam engine. In 1709, Abraham Darby discovered how to smelt iron ore using coke (a purified form of coal) rather than charcoal as a fuel, and the process started becoming widespread in Britain in the second half of the century. Three years later, in 1712, Thomas Newcomen developed his famous steam engine to pump water from mines. James Watt started working on an improved design in 1763, and by 1776 his steam engines were being used commercially. Steam then started to diffuse across the economy, slowly at first, and then more rapidly, so that eventually coal was being used to fuel not just the metallurgical industries, but textiles and many other sectors as well. Steam accounted for 35,000 out of the 170,000 horsepower installed in Britain in 1800; for 165,000 out of 350,000 in 1830; for 2,060,000 out of 2,300,000 in 1870; and for 9,659,000 out of 9,842,000 in 1907 (Crafts 2004, p. 342).

To summarize, the new coal-using technologies of the Industrial Revolution were invented in stages over the course of the 18th century, and were then progressively improved and increasingly adopted during both the 18th and 19th centuries. They were first invented and used in Britain, but then diffused with a lag to the rest of Europe. By the middle of the 19th century, both coke-smelting and steam engines were being used in all the coalfields of northern France, Belgium and western Germany (Wrigley 1961, p. 4). An appropriate "treatment date" would thus be 1750 or later.

Why would being closer to a coalfield have led to bigger city sizes once these new technologies had been introduced? The argument is straightforward, and relies on a combination of the growth and location hypotheses. Adopting new coal-using technologies directly spurred economic growth, and once these technologies had been adopted increasing production required higher inputs of coal. All of this was more profitable where coal was cheaper, and coal was cheaper close to coalfields. Greater economic activity, in turn, could lead to agglomeration economies, permitting further growth.[8] Economic growth in turn stimulated population growth: indeed, the connection between industrial growth and population growth in the coalfield regions of northwest continental Europe was so tight that Wrigley (1961) used the latter as a proxy for the former.

Our identification strategy relies crucially on whether industry tended to locate closer to coalfields because of the costs of transporting coal. The strategy would break down when other forms of energy, such as electricity, became widely available, or when the costs of transporting coal became sufficiently cheap. The historical record suggests that being close to coalfields should have started mattering less by the end of the 19th century, as electricity was increasingly adopted, and an increasingly dense and efficient railway network lowered freight rates. "For example, a point was reached about 1890 when it became cheaper to carry coke to the Lorraine iron ore fields than to carry the ore to the Ruhr, because blast furnaces had grown much more economical in their use of coke than in the early days of the coke-fired furnace, and the lean ores of Lorraine were unusually costly to transport" (Wrigley 1961, p. 6). For this reason, we end our analysis in 1900.

[8]For example, Balderston (2010) argues that coal was crucial to the development of agglomeration economies in the Lancashire cotton textile industry.

5 Our Data

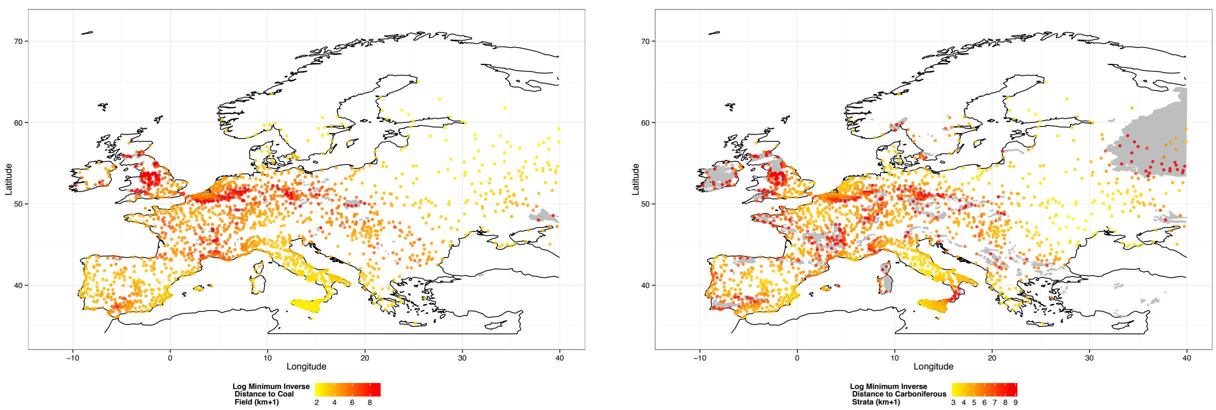
Our empirical analysis combines a number of different sources. To measure economic activity we use historic population size for a panel of European cities based in the first instance on Bairoch *et al.* (1988). The panel consists of around 2,200 cities that satisfy the criterion of having at least 5,000 inhabitants at some stage between 800 and 1800. City populations are measured at 100 year intervals between 800 and 1700, and at 50 year intervals after 1700. The panel is unbalanced, since evidence on city population levels is understandably lacking for many cities as we look further back in time.

These missing observations are potentially problematic, and we therefore take several steps to remedy this. First, as in Nunn and Qian (2011), we supplement these data with observations from De Vries (1984). Second, we begin our analysis in 1300, ignoring all data prior to this date. The vast majority of cities included in the Bairoch *et al.* dataset do not have population data prior to 1300, and we also suspect that the available population evidence prior to 1300 is less accurate than more recent figures.

Bairoch *et al.* (1988) only provide city population sizes up until 1850. Given that we want to measure economic development through the end of the 19th century, we extend the Bairoch *et al.* panel to include city populations in 1900. For this purpose we use a contemporary resource: *Lippincott's New Gazetteer* (1906). The gazetteer lists the location and population figures for the majority of the cities included in the Bairoch *et al.* panel. We restrict our sample to cities positioned west of the 40 degree line of longitude, and above the 30 degree line of latitude. This leaves a sample of 2,147 cities.

The aim of this paper is to link city populations to the availability of coal. Thus, we need to create a measure of access to coal for each city. To do this we digitize the *Les Houillères Européennes* map in Châtel and Dollfus (1931). This atlas contains the location of 124 major coalfields within Europe. We include lignite fields in our calculations since lignite (or brown coal) played an important role in the introduction of the steam engine in Prussia (Redlich 1944). We digitize this map, and then calculate the minimum distance from each city to a coalfield. For simplicity, we use great-circle distance for this calculation. Another approach would be to use a least-cost distance measure, like Özak (2012). However, dramatic changes in the form and speed of transportation methods over our eight century sample period would greatly complicate such a calculation. Coal prices, as used in Crafts and Wolf (2013), would be an economically more meaningful measure of access to coal. Unfortunately we lack city-level price data for our panel of more than 2,000 cities over 800 years. However, it is reasonable to presume that our proximity measure is correlated with the spatial variation in coal prices.

Figure 1: City's Proximity to Coalfields or Carboniferous Strata (Grey Areas) in Europe.



(a) Coalfields

(b) Carboniferous Strata

Our great-circle distance-based proximity measure could bias our empirical exercises for a number of reasons. As already noted, we are not using a least cost measure of coal proximity. Thus, our coal measure does not reflect the fact that cities at, or close to, the coast may have better access to coal than cities further away from the coast and/or located on more rugged terrain. (However, we do include proximity to coasts, ruggedness and a number of other geographical control variables in our empirical specifications.) Another issue is that that while we can measure the distance to each coalfield, we cannot measure the extent of the coalfield in terms of either abundance or quality. Both of these features will result in measurement error in a regression model estimating the impact of coal on city populations. Finally, we cannot rule out the possibility of reverse causality, that is to say the possibility that the coalfields in our dataset were developed because of their proximity to cities.

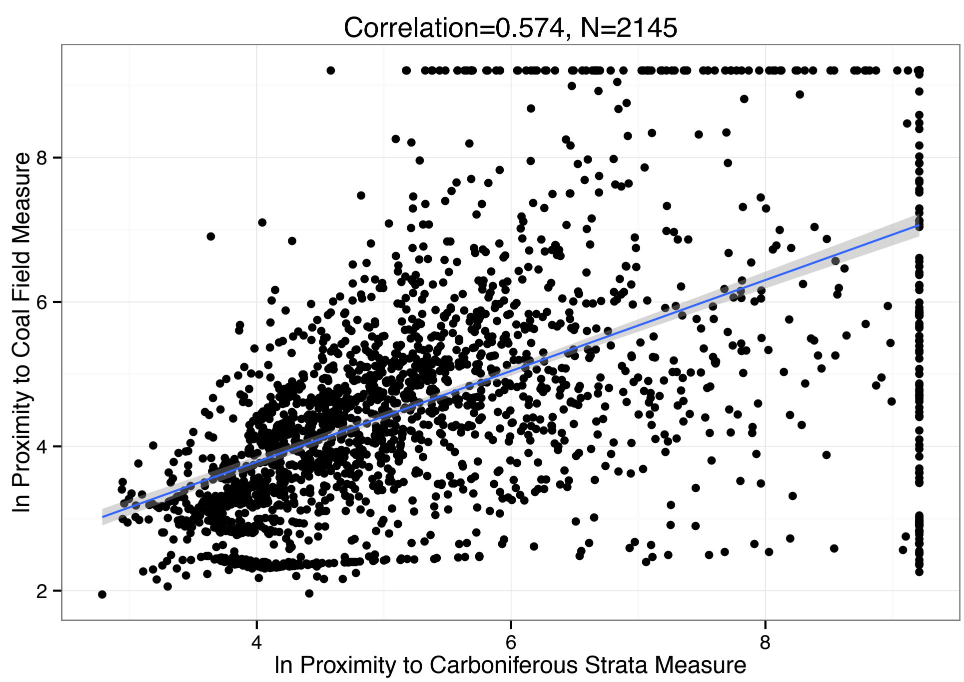
To address these measurement and endogeneity concerns, we instrument our proximity to coal measure with a variable that measures the proximity of cities to Carboniferous geological strata. Coal is often found in rock strata from the Carboniferous age, and thus the coalfield locations should be on, or very close to, rock strata from the Carboniferous epoch.[9] To construct a Carboniferous measure that corresponds to our coal proximity variable, we use data collated by the German Federal Institute for Geosciences and Natural Resources (BGR) for a project that mapped the European geological

[9]Carboniferous literally means "coal-bearing".

landscape: the 1:5 Million International Geological Map of Europe and Adjacent Areas (IGME 5000). This was a pan-European project that involved a high level of collaboration across a number of national geological offices; the result was a high resolution GIS map containing a number of geological features including age and rock type (Asch 2005). We create a proximity to Carboniferous strata variable for each city in an equivalent manner to that of the coal proximity variable.

Figure 1 contains two panels illustrating the location of our cities and their proximity to both coalfields and Carboniferous strata. Most coalfields overlap with areas whose rock strata are categorized as being of the Carboniferous epoch, although this overlap is not perfect. There are some Carboniferous areas that do not contain any coalfields, and some coalfields not located within a Carboniferous area. For example, lignite is geographically younger than black coal, typically originating in the Tertiary period. Our coal and Carboniferous measures are evidently correlated, but not perfectly. Figure 2 illustrates the strength of this relationship. A weak relationship between these variables would invalidate our IV strategy; however Figure 2 shows that the relationship between these two variables is sufficiently strong for our purposes.[10]

Figure 2: Proximity to Coalfields and Proximity to Carboniferous Strata Scatterplot.



[10]Our empirical section formally discusses the issue of weak instruments, and we provide a series of appendix tables displaying the relevant weak instrument test measures.

In addition to measurement error and simultaneity, omitted variable bias poses a threat to our identification strategy. For example, Acemoğlu *et al.* (2005) argue that the acceleration of growth in post-1500 Europe was caused by the proximity of certain economies to the Atlantic coast (used for colonial trade) and the institutions that emerged in these economies. Similarly, as we have seen Nunn and Qian (2011) find that the introduction of the potato played an important role in explaining comparative development in the Old World during the Industrial Revolution.

We therefore include a rich set of control variables in our regression analysis. These include distance to coastlines; more precisely, we create three variables that measure the distance to the Atlantic, the Mediterranean, and to all coasts. We also measure cities' distance to primary rivers. Given the prominence of state institutions in Acemoğlu *et al.*'s research, we also match each city to historical state borders. It is important to acknowledge that both state borders and institutions changed, and had differential impacts over time. We therefore digitize European state borders for a series of years between 1500 and 1913 (1500, 1618, 1699, 1748, 1804, 1848, and 1908), and match these borders as closely as possible to each year in our sample.

We also include variables measuring terrain ruggedness; the suitability of land for cultivating potatoes, wheat, and oats; altitude; and temperature. An appendix table provides more detailed information on the construction of these variables and the sources used. We also address the issue of spatial spill-overs and clustering in a number of ways. First, we include controls for absolute latitude and longitude in all model specifications. Second, we explicitly incorporate a spatially lagged variable in a number of model specifications. Given that working with spatially lagged longitudinal data requires balanced panels (Millo and Piras 2012), we interpolate missing data for city populations. The algorithm used to interpolate these data is included as an appendix.

6 Empirical Results

6.1 Empirical Methodology

Our empirical strategy follows the standard DID approach. We estimate city population as a function of the interaction between a city's proximity to coal and a post-treatment year indicator. In general, we estimate the following linear regression model:

$$\ln(Pop_{it}) = \alpha_i + \gamma_t + \beta \ln(Coal_i) I_t^{Post} + \varepsilon_{it} \tag{1}$$

where the natural log of each city's population level is a function of city (α_i) and time (γ_t) fixed effects, the interaction between a city's proximity to a coalfield ($Coal_i$) and a binary variable equal to one when the year is after the specified cut-off point (I_t^{Post}), and an idiosyncratic error term (ε_{it}), with the i and t subscripts corresponding to city and time domains respectively. The β parameter represents the causal effect of coal on city population size. Because it is more intuitive to interpret the effect of proximity to coal than the effect of distance from it (so there is a positive relationship if coal causes city population size), we define our coal variable by inverting the distance from each city to the nearest coalfield.[11]

The inclusion of the interaction effect between proximity to coal and a post-treatment year indicator implicitly assumes temporally heterogeneous coal effects. This assumption is crucial for identifying the causal parameter beta in our DID setup. If, for example, the impact of coal had been time invariant, beta would be subsumed into the city fixed-effects and we would find that coal had no impact on city sizes. What we are interested in is whether something happened during the course of the 18th or 19th centuries which made proximity to coal matter for city sizes in a way that it had not mattered before. That "something" was of course the introduction of the coal-using technologies of the Industrial Revolution. In the following subsection we implement a flexible modelling procedure that allows us to assess the plausibility of our DID strategy, and detect the "treatment" cut-off year after which the new coal-using technologies started to matter for city size.

The model in eq. 1 assumes that heterogeneous coal effects are the only observable systematic variable that causes differences in city populations, aside from common time effects and fixed city factors. Realistically, this assumption is likely to be violated, and will result in an omitted variables bias when our coal proximity measure is correlated with such omitted variables. Therefore, we include a rich set of covariates (interacted with time effects) in all of our estimated models. We amend eq. 1 to take this into account:

$$\ln(Pop_{it}) = \alpha_i + \gamma_t + \beta \ln(Coal_i) I_t^{Post} + \sum_{j=1400}^{1900} \mathbf{X}'_i \mathbf{I}_t^j \boldsymbol{\Psi}_j + \varepsilon_{it} \tag{2}$$

[11] Full details of this transformation are contained in a Data Appendix.

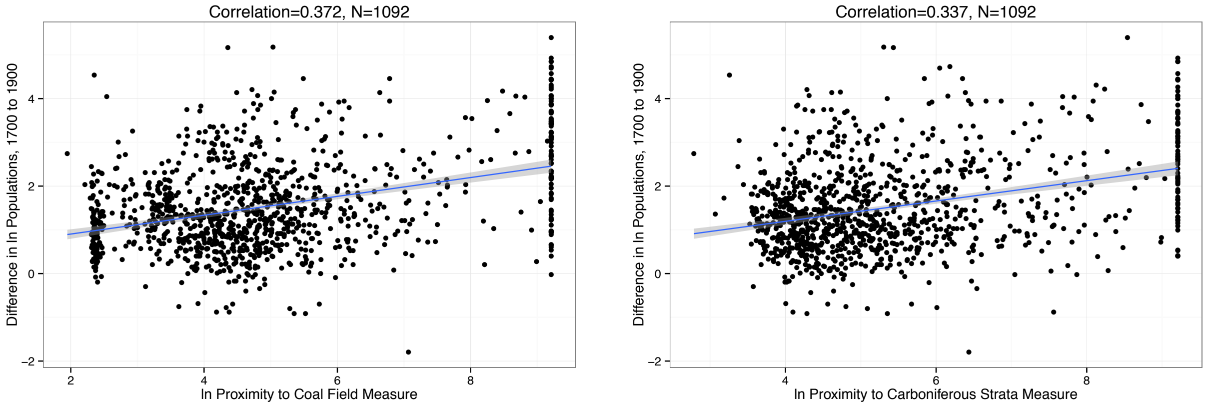
where the $\sum_{j=1400}^{1900} \mathbf{X'}_i \mathbf{I}_t^j$ term interacts a large number of geographic and historic control variables ($\mathbf{X_i'}$), described in Section 5, with time indicators ($\mathbf{I}_t^j$), with the j subscript denoting specific years. Alongside the city (α_i) and time (γ_t) fixed effects, the matrix $\mathbf{X}_i$ also contains multiple historic state border fixed effects. These multiple fixed effects allow us to control for a variety of potentially confounding factors. For example, the city fixed effects allow us to control for the possibility that administrative cities such as London were consistently larger than other cities; the time fixed effects allow us to control for the general rise in European urban populations over time; and the time-varying border fixed effects allow us to control for increases in national urban populations, relative to the European trend, due to national demographic developments, institutional environments, or other factors. Estimating a linear regression model with multiple categorical variables can be computationally burdensome, so we use the recent algorithm provided in Gaure (2013) to simplify the estimation procedure.[12]

The measurement of our coal proximity variable represents another potential threat to our empirical strategy, as does the potential for endogeneity. Mismeasurement could lead to a downward bias in our estimates of the causal beta parameter, while endogeneity could lead to an upward bias. We tackle these concerns by using an equivalent measure of proximity to Carboniferous strata as an instrumental variable (IV), and estimate eq. 2 via two-stage least-squares (2SLS). Our rationale for using this IV has already been discussed, with Figure 2 illustrating the strength of the relationship between these two variables. Figure 3 presents scatter plots illustrating the relationship between both our coal and Carboniferous measures and the change in natural logged population between 1700 and 1900 for cities where we have recorded population totals in both years. Both panel (a) and (b) in Figure 3 are consistent with the hypothesis that coal was an important factor related to city population growth during the Industrial Revolution.

While Figure 3 is indicative, these relationships may be subject to the biases discussed earlier. Furthermore, Figure 3 leaves open the possibility that coal was an equally important determinant of city growth before the Industrial Revolution, which would be inconsistent with the argument of Wrigley *et al.* We therefore proceed to more formal econometric analysis.

[12]We implement the procedure of Gaure (2013) in the R package lfe available on the Comprehensive R Achieve Network.

Figure 3: City Population Growth and Proximity to Coal Fields or Proximity to Carboniferous Strata Scatterplots.



(a) Coal Fields (b) Carboniferous Strata

6.2 Flexible Model Results

We begin our formal empirical modeling by taking a flexible approach, and estimating the following regression model:

$$\ln(Pop_{it}) = \alpha_i + \gamma_t + \sum_{j=1400}^{1900} \beta_j \ln(Coal_i) I_t^j + \sum_{j=1400}^{1900} \mathbf{X'}_i \mathbf{I}_t^j \mathbf{\Psi}_j + \varepsilon_{it} \tag{3}$$

where we include multiple interaction effects between the eight time periods and each city's proximity to coal. Thus, we obtain eight beta parameters, which indicate the influence of coal on city population size relative to our 1300 baseline (since 1300 is the excluded year). The "flexible" element in this approach is that it allows us to be relatively agnostic about when coal started to matter; that is to say, about the definition of the post-treatment period. The flexible model not only permits us to detect this cut-off, but also provides an assessment of the DID approach's plausibility. If coal was relatively unimportant for city population sizes prior to the emergence of coal-based technological innovations, we would not expect to see the earlier beta parameters (β_{1400}–β_{1600}) have a substantial impact on city population levels. In this sense, these additional β estimates form placebo tests.

Table 1 displays the estimated coal-year interaction OLS coefficients from eq. 3. Columns (1) and (2) estimate the model on the original sample (that is to say excluding interpolated estimates where city population data are missing), with and without cities

Table 1: Flexible Regression Estimates, OLS: Full Controls with Border × Year Fixed Effects.

	Dependent Variable is Log City Population			
Coal Variables	(1)	(2)	(3)	(4)
Coal × Year=1400	0.054	0.006	0.014*	0.019*
	(0.056)	(0.064)	(0.008)	(0.010)
Coal × Year=1500	0.062	0.027	0.021**	0.022*
	(0.043)	(0.045)	(0.010)	(0.011)
Coal × Year=1600	0.101**	0.066	0.035***	0.033***
	(0.043)	(0.044)	(0.011)	(0.012)
Coal × Year=1700	0.076*	0.059	0.028**	0.032**
	(0.045)	(0.044)	(0.013)	(0.015)
Coal × Year=1750	0.083*	0.044	0.043***	0.028*
	(0.046)	(0.044)	(0.014)	(0.016)
Coal × Year=1800	0.114**	0.040	0.062***	0.023
	(0.046)	(0.044)	(0.016)	(0.017)
Coal × Year=1850	0.166***	0.074*	0.115***	0.061***
	(0.048)	(0.045)	(0.020)	(0.021)
Coal × Year=1900	0.196***	0.101**	0.142***	0.090***
	(0.049)	(0.047)	(0.025)	(0.028)
Excludes UK	N	Y	N	Y
Includes Interpolated Cities	N	N	Y	Y
Num. obs.	10773	9799	19305	17613

All regressions include both year and city fixed effects. Standard errors and p-values have been corrected to account for clustering at the city level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. These regressions also include year interactions with the following variables: Latitude, Longitude, Potato, Wheat, Oat, Ruggedness, Altitude, Temperature, Atlantic, Mediterranean, Coast, Rivers, Atlantic^2, Mediterranean^2, Coast^2, and Rivers^2. The border × year fixed effects are as follows: 1500 borders × 1400, 1500 borders × 1500, 1618 borders × 1600, 1699 borders × 1700, 1748 borders × 1750, 1804 borders × 1800, 1848 borders × 1850, and 1908 borders × 1900.

from the United Kingdom (UK) respectively. We also report standard errors that are clustered on each city unit, thus taking into account any within-city serial correlation that would lead to an underestimate of the standard error. Columns (3) and (4) show the results of models equivalent to (1) and (2), but where we now include observations with interpolated city populations in the sample. Overall, the results in Table 1 are consistent with the hypothesis that the Industrial Revolution greatly increased the advantages of being located close to coal. In all of the models the coal-year interactions grow stronger over time. It appears that a city's proximity to coal matters later when the sample only includes non-UK cities, consistent with the Industrial Revolution starting in Britain and diffusing later to the Continent. Contrasting the results between the samples excluding and including cities for which we have imputed population figures we find that, as expected, the inclusion of previously omitted observations reduces the standard errors. Furthermore, the somewhat large β_{1600} coefficient in column (1) is reduced in column (3), which is reassuring. Since both the dependent variable and coal variable are in natural logarithms we can interpret these results as elasticities. However, we do not place too much emphasis on this, since we present a counterfactual framework in Section 6.4 that provides a much more intuitive way of interpreting our findings.

Table 2 provides equivalent estimates to those in Table 1, except that it uses the Carboniferous variable interacted with the year indicators to instrument for coal. We use the approach advocated in Angrist and Pischke (2009) for detecting weak instruments in the presence of multiple endogenous regressors. This approach enables one to calculate multiple partial F-test statistics (one for each endogenous regressor) analogous to the method used in cases with a single endogenous regressor. The F-test statistics associated with Table 2 are reported in an appendix alongside similar measures for the relevant IV regression model results we report later in this section; none of these tests indicate that there is an issue with weak instruments. We find support for our suspicion that measurement error may result in a downward bias in the OLS coefficients, since the estimated β coefficients for later years are substantially larger in Table 2 than those reported in Table 1. The difference between the comparable non-interpolated and interpolated model coefficients in Table 2 is small. Furthermore, the relatively large coefficient on β_{1600} in column (1) in Table 1 is absent in the equivalent specification in Table 2.

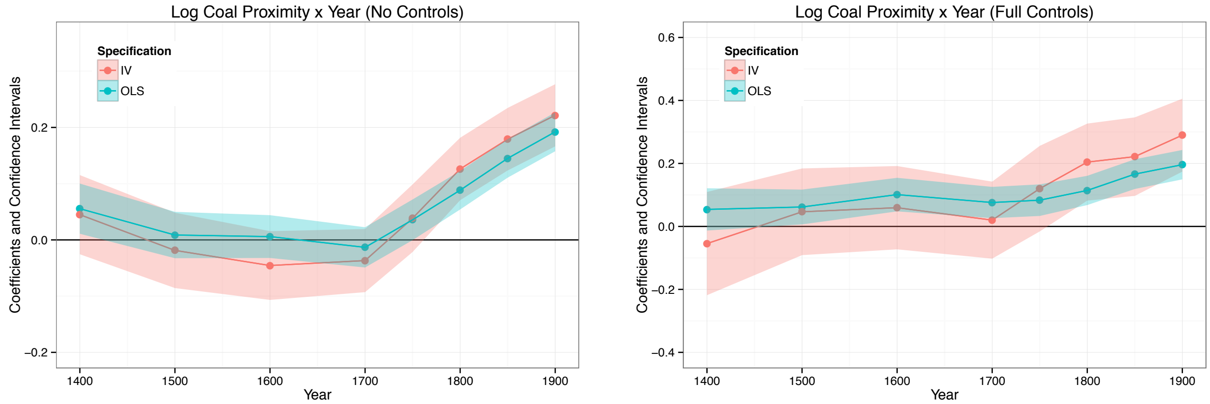
Figure 4 provides an illustration of these results. Panel (a) plots the estimated β_j values when the model in eq. 3 is estimated without any control variables (that is, the

Table 2: Flexible Regression Estimates, IV: Full Controls with Border $\times$ Year Fixed Effects.

	Dependent Variable is Log City Population			
Coal Variables	(1)	(2)	(3)	(4)
Coal $\times$ Year=1400	-0.055	-0.155	-0.011	-0.017
	(0.102)	(0.137)	(0.018)	(0.025)
Coal $\times$ Year=1500	0.046	-0.053	0.011	0.000
	(0.098)	(0.124)	(0.021)	(0.029)
Coal $\times$ Year=1600	0.059	-0.028	0.039	0.028
	(0.103)	(0.129)	(0.025)	(0.034)
Coal $\times$ Year=1700	0.020	-0.085	0.062^{**}	0.058
	(0.103)	(0.123)	(0.031)	(0.041)
Coal $\times$ Year=1750	0.120	0.015	0.126^{***}	0.090^{*}
	(0.107)	(0.126)	(0.035)	(0.047)
Coal $\times$ Year=1800	0.204^{*}	0.057	0.173^{***}	0.096
	(0.108)	(0.128)	(0.044)	(0.061)
Coal $\times$ Year=1850	0.222^{**}	0.053	0.190^{***}	0.097
	(0.109)	(0.129)	(0.050)	(0.069)
Coal $\times$ Year=1900	0.290^{***}	0.186	0.232^{***}	0.213^{***}
	(0.109)	(0.125)	(0.061)	(0.080)
Excludes UK	N	Y	N	Y
Includes Interpolated Cities	N	N	Y	Y
Num. obs.	10773	9799	19305	17613

All regressions include both year and city fixed effects. Standard errors and p-values have been corrected to account for clustering at the city level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. These regressions also include year interactions with the following variables: Latitude, Longitude, Potato, Wheat, Oat, Ruggedness, Altitude, Temperature, Atlantic, Mediterranean, Coast, Rivers, Atlantic^2, Mediterranean^2, Coast^2, and Rivers^2. The border $\times$ year fixed effects are as follows: 1500 borders $\times$ 1400, 1500 borders $\times$ 1500, 1618 borders $\times$ 1600, 1699 borders $\times$ 1700, 1748 borders $\times$ 1750, 1804 borders $\times$ 1800, 1848 borders $\times$ 1850, and 1908 borders $\times$ 1900.

Figure 4: Coal Coefficients from Flexible Models.



(a) Model without Controls.

(b) Model with Full Controls.

Shaded areas indicate +/- 2 cluster robust standard errors.

matrix $\mathbf{X}_i$ is null.[13] In other words, these results correspond to those in column (1) of Tables 1 and 2, when we omit our control variables, but not the time or city fixed effects, from the specification. Panel (b) displays the results corresponding to column (1) in Tables 1 and 2. The results in both plots are consistent with one another, and show the coal effect becoming important in the 18th century.

6.3 Rolling Growth Regression Results

The flexible model results presented in the preceding passage suggest that coal became important for aggregate economic activity during the period typically associated with the Industrial Revolution. In this section, we examine the evolution of this relationship using an alternative flexible modelling approach, estimating the following regression model:

$$\Delta \ln(Pop_{it}) = \alpha + \omega \ln(Pop_{it-1}) + \beta \ln(Coal_i) + \mathbf{X'}_i \boldsymbol{\Psi} + \varepsilon_{it} \tag{4}$$

Here, the change in the natural log of population between t and $t-1$ is a function of population in the initial period ($\ln(Pop_{it-1})$), our coal measure, and geographic and historic control variables. We estimate this cross-section model for six successive centuries, implying that all six columns are comparable. The individual city fixed effects are now omitted from the model, since we are estimating cross-section regressions. However,

[13]We have omitted these results from the text for the sake of brevity.

Table 3: Rolling Growth Regressions, Century Intervals.

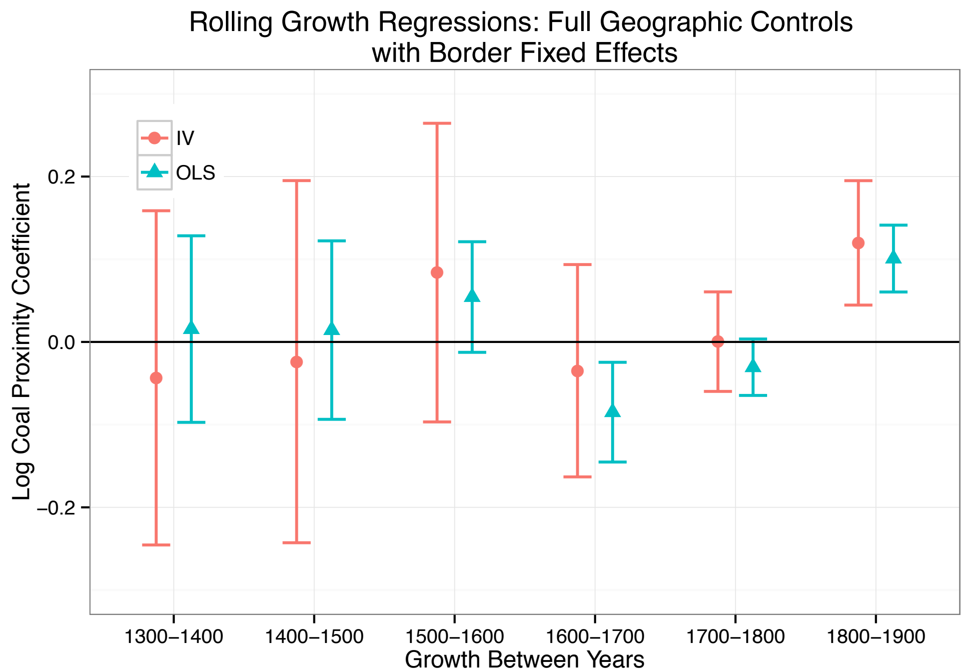
	$\Delta \ln \text{Pop}_{1400}$	$\Delta \ln \text{Pop}_{1500}$	$\Delta \ln \text{Pop}_{1600}$	$\Delta \ln \text{Pop}_{1700}$	$\Delta \ln \text{Pop}_{1800}$	$\Delta \ln \text{Pop}_{1900}$
Coal Coefficients	(1)	(2)	(3)	(4)	(5)	(6)
OLS	0.037	0.022	0.053^{*}	-0.055^{**}	-0.033^{**}	0.098^{***}
	(0.054)	(0.049)	(0.027)	(0.027)	(0.016)	(0.019)
IV	-0.045	-0.023	0.079	-0.032	0.000	0.128^{***}
	(0.105)	(0.106)	(0.085)	(0.060)	(0.030)	(0.040)
Num. obs.	278	292	495	764	1165	1808
OLS: Excl. UK	0.047	-0.022	0.056^{*}	-0.037	-0.036	0.064^{***}
	(0.081)	(0.059)	(0.033)	(0.029)	(0.022)	(0.021)
IV: Excl. UK	-0.352	-0.133	0.122	-0.082	0.001	0.179^{***}
	(0.314)	(0.194)	(0.156)	(0.101)	(0.065)	(0.058)
Num. obs.	258	270	452	699	1012	1629
OLS: Incl. Interpolated	0.002	0.007	0.017^{**}	-0.017^{*}	-0.013	0.095^{***}
	(0.005)	(0.006)	(0.007)	(0.010)	(0.009)	(0.020)
IV: Incl. Interpolated	-0.004	-0.004	0.037^{**}	0.025	0.029	0.155^{***}
	(0.012)	(0.013)	(0.018)	(0.023)	(0.023)	(0.044)
Num. obs.	2145	2145	2145	2145	2145	2145
OLS: Excl UK & Incl. Interpolated	-0.001	0.003	0.013	-0.007	-0.010	0.070^{***}
	(0.006)	(0.007)	(0.008)	(0.011)	(0.011)	(0.024)
IV: Excl UK & Incl. Interpolated	-0.028	-0.014	0.046^{*}	0.051	0.047	0.239^{***}
	(0.019)	(0.021)	(0.028)	(0.036)	(0.038)	(0.068)
Num. obs.	1957	1957	1957	1957	1957	1957

Standard errors and p-values have been corrected to account for heteroscedasticity. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. These regressions also include the following control variables: Latitude, Longitude, Potato, Wheat, Oat, Ruggedness, Altitude, Temperature, Atlantic, Mediterranean, Coast, Rivers, Atlantic^2, Mediterranean^2, Coast^2, Rivers^2, and logged population levels from the start period (for example column (4) includes the population levels in 1600). These regressions also include the following border year fixed effects: 1500, 1618, 1699, 1748, 1804, 1848, and 1908.

we include population in the initial period to capture the possibility that smaller cities might grow faster than larger ones. Each of the regressions also contains the complete set of historic state border fixed effects (i.e. for the years 1500, 1618, 1699, 1748, 1804, 1848, and 1908). These may capture the legacy effects that remain after the changing of state borders, or indeed deep-rooted common cultural factors which shape future state borders.

Table 3 provides the results when we estimate eq. 4 for century intervals from 1300–1400 to 1800–1900. Once again, we provide OLS and IV coefficients for data samples that have been stratified according to the inclusion or omission of non-UK cities, and the inclusion or exclusion of interpolated missing city population figures. The results in Table 3 are largely consistent with those in the previous subsection. In essence, columns (1) to (5) can be seen as placebo regressions, since we do not expect a city's proximity to coal to influence its population growth prior to the Industrial Revolution. We do not see any consistent pattern when we compare these columns. By contrast, proximity to coal is clearly related to city growth during the 19th century. The estimated t-statistics associated with some of the coefficients in columns (1) to (5) in Table 3 yield p-values lower than what might be used as conventional levels for statistical significance, although this is not consistent across the different samples. Furthermore, if we look at the size of the coefficients, we find that those in column (6) are always the largest across each of the columns, and are estimated with a far greater degree of precision compared to their counterparts in columns (1) to (5).

Figure 5: Coefficients from Rolling Growth Model



Bars indicate +/- 2 robust standard errors.

Figure 5 provides a coefficient plot of the results shown in the first panel of Table 3. We do not see any consistent pattern prior to the 19th century, but a clear positive effect of coal on city growth after 1800.

6.4 Fixed Treatment Effect Results

Our analysis thus far has been primarily concerned with detecting the treatment year after which coal endowments become a factor influencing comparative population sizes in our panel of European cities. The analysis indicates that the earliest date for this post-treatment cut-off is 1750. One drawback of the flexible modelling approach is that it can be difficult to assess the economic significance of these results. Therefore, we follow Nunn and Qian (2011) and estimate models with a single treatment effect—as detailed in eq. 2—and then use this singular causal parameter to create counterfactual population totals. These counterfactual totals represent the estimated population of each city in the absence of a coal effect; that is to say, in the absence of the introduction of the coal-using technologies of the Industrial Revolution which implied that proximity to coal became a factor influencing city sizes after the treatment year.

More formally, the counter-factual population for city i at the end of our estimation sample, 1900, is $\ln \widetilde{Pop_{i1900}} = \ln Pop_{i1900} - \hat{\beta} \ln Coal_i$. If we sum over all cities, we can calculate a counterfactual total urban population for 1900, and thus a counterfactual growth rate for the total urban population between the treatment year and 1900. This can then be compared with the actual growth of the total urban population over the same period, yielding an estimate of the percentage of total urban population growth explained by the introduction of the coal-using technologies of the Industrial Revolution:

$$\text{Total Effect} = 1 - \frac{\ln \sum \widetilde{Pop_{i1900}} - \ln \sum Pop_{iPost}}{\ln \sum Pop_{i1900} - \ln \sum Pop_{iPost}} \tag{5}$$

where the *Post* term refers to the treatment year. If the estimated coal effect, β, was zero, the numerator would equal the denominator in the second term and the estimated effect would be zero.

Table 4 presents the beta coefficients obtained from various estimates of the fixed treatment effect model. Once again we present results both including and excluding UK cities, and including and excluding interpolated data points. Given that one can make a case for several treatment dates, we present comparable estimates when three cut-offs are used: post-1750, post-1800, and post-1850. The top three panels in Table 4 display the OLS results, and the following three our preferred IV estimates. Once again, we omit the relevant indicators of weak instruments from the table, but we find no evidence that the Carboniferous variable interacted with the post-treatment effect indicator suffers from the weak instruments problem. We have included these tests in an appendix. In addition to the coefficients and their associated standard errors, we also include the "Total Effect" percentage calculated as above.

Consistent with our hypothesis, proximity to coal has a positive effect on city growth in each of the specifications in Table 4. The most noticeable feature of these results is that, as before, the IV coefficients are larger than their OLS counterparts, a fact that we attribute to measurement error in the coal proximity measure. Looking at the IV results with a post-1750 cut-off in column (1), we see that over 60% of the city population

Table 4: Fixed Treatment Effects.

	Dependent Variable is Log City Population			
	(1)	(2)	(3)	(4)
OLS				
Coal × Post-1750	0.115^{***}	0.036^{*}	0.114^{***}	0.043^{***}
	(0.020)	(0.020)	(0.015)	(0.016)
Counterfactual Explained (%)	34.729	10.969	36.148	13.829
Coal × Post-1800	0.105^{***}	0.038^{**}	0.123^{***}	0.051^{***}
	(0.017)	(0.017)	(0.017)	(0.018)
Counterfactual Explained (%)	38.323	13.680	44.837	18.512
Coal × Post-1850	0.090^{***}	0.036^{*}	0.122^{***}	0.061^{***}
	(0.017)	(0.019)	(0.019)	(0.020)
Counterfactual Explained (%)	51.939	18.463	70.694	31.878
IV				
Coal × Post-1750	0.208^{***}	0.144^{***}	0.187^{***}	0.099^{**}
	(0.040)	(0.053)	(0.035)	(0.045)
Counterfactual Explained (%)	60.974	42.526	57.868	31.248
Coal × Post-1800	0.153^{***}	0.102^{**}	0.173^{***}	0.092^{*}
	(0.033)	(0.042)	(0.037)	(0.048)
Counterfactual Explained (%)	54.873	36.049	62.049	32.658
Coal × Post-1850	0.131^{***}	0.126^{***}	0.157^{***}	0.107^{*}
	(0.033)	(0.043)	(0.041)	(0.055)
Counterfactual Explained (%)	74.830	63.905	89.923	55.490
Excludes UK	N	Y	N	Y
Includes Interpolated Cities	N	N	Y	Y
Num. obs.	10773	9799	19305	17613

All regressions include both year and city fixed effects. Standard errors and *p*-values have been corrected to account for clustering at the city level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. These regressions also include year interactions with the following variables: Latitude, Longitude, Potato, Wheat, Oat, Ruggedness, Altitude, Temperature, Atlantic, Mediterranean, Coast, Rivers, Atlantic^2, Mediterranean^2, Coast^2, and Rivers^2. The border × year fixed effects are as follows: 1500 borders × 1400, 1500 borders × 1500, 1618 borders × 1600, 1699 borders × 1700, 1748 borders × 1750, 1804 borders × 1800, 1848 borders × 1850, and 1908 borders × 1900.

growth experienced between 1750 and 1900 can be attributed to the introduction of the coal-using technologies of the Industrial Revolution. The "Counterfactual Explained" percentage usually increases as we move the treatment year forward, a finding that is consistent with the diffusion of coal-based technologies over time. Furthermore, comparing results from samples including and excluding cities in the United Kingdom suggests that the coal effect was stronger in the UK, but diffused spatially over time. This result is consistent with the fact that the coal-based technologies that drove the Industrial Revolution first emerged in the United Kingdom.

6.5 Fixed Treatment Effect Results with Spatially Lagged Effects

Thus far, we have not formally accounted for the possibility that cities' sizes might depend on each other, as a result of market potential effects or other spatial spillovers. A wide literature on spatial agglomeration and/or market potential effects exists, and previous empirical research has found these effects to be substantial. Bosker, Buringh and van Zanden (2013) calculate a measure of foreign urban potential for a number of European and Islamic cities and find that this spatial interaction is positively correlated with Western European city population growth from the 12th century onwards. Crafts and Wolf (2013) estimate the impact of market potential on the location of UK cotton textile manufacturers in 1838. They find that conditional on a number of key geographic variables, including coal prices, market potential was an important factor that explained variation in the location of the UK cotton industry in the 19th century.

We therefore estimate a spatial regression model in the spirit of Kelejian and Prucha (1998), introducing both a spatially lagged dependent variable and a spatially correlated error term to control for spatially interrelated cross-sections in the panel. Formally, we revise eq. 2 as follows:

$$\ln(Pop_{it}) = \lambda(\mathbf{W} \otimes \mathbf{I}_T)\ln(Pop_{it}) + \alpha_i + \gamma_t + \beta \ln(Coal_i) I^{Post} + \sum_{j=1400}^{1900} \mathbf{X'}_i \mathbf{I}_t^j \mathbf{\Psi}_j + u_{it} \quad (6)$$

$$u_{it} = \rho(\mathbf{W} \otimes \mathbf{I}_T) u_{it} + \varepsilon_{it} \quad (7)$$

where $\mathbf{I}_T$ is the identity matrix corresponding to the number of time periods T (nine in our case), $\mathbf{W}$ is a cross-section spatial weights matrix, and the error term u has been modified so that it permits spatial correlation alongside the usual idiosyncratic error: $\varepsilon_{it} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Our row-standardized spatial weights matrix $\mathbf{W}$ categorizes cities as neighbours if they lie within 50km of one another. We found that this scheme worked well, since it strikes a balance between cities having too many or too few neighbours (the median number of neighbours each city in our sample had is 5).

Complications exist when estimating the models in eq. 6 and 7 due to simultaneity, as the spatial lag term is an endogenous regressor. This simultaneity arises due to the reflection problem (Manski 1993): a city's neighbour's population influences the city's population, but the city's population itself, in turn, influences neighbouring cities. Kelejian and Prucha (1998) show how this spatial model can be estimated via a Generalized Moments (GM) framework using what they call Generalized Spatial Two-Stage Least Squares (GS2SLS)—a three step estimator analogous to the Cochrane-Orcutt procedure used in time-series econometrics. In the first step we estimate eq. 6 via 2SLS, using both first and second order spatial lags of the exogenous variables ($\sum_{j=1400}^{1900} \mathbf{W}\mathbf{X}'_i \mathbf{I}_t^j$ and $\sum_{j=1400}^{1900} \mathbf{W}^2 \mathbf{X}'_i \mathbf{I}_t^j$) as instruments for the spatially lagged dependent variable.[14] The second step estimates the spatial error coefficient via the GM procedure originally proposed in Kelejian and Prucha (1999). The final step uses the estimated ρ parameter to transform both the dependent variables and regressors, before applying the first step again on the transformed variables.

For our application we estimate a modified version of the Kelejian and Prucha technique and fit the model outlined above, but taking the spatially lagged dependent variable as an exogenous regressor. It is important to underline the fact that we have estimated our entire set of spatial models using the full Kelejian and Prucha estimator, and the results we obtain are almost identical. We have included these results in an appendix. The full Kelejian and Prucha estimator would require us to model both the spatial lag and the coal variable as endogenous regressors. However, post-regression diagnostics indicate that this approach is problematic. The problem stems from the inclusion of many irrelevant instruments in the first-stage coal equation resulting in a weaker first stage regression (we again document this issue in appendix). This is why we model the spatially lagged variable as an exogenous regressor in this section (although as we show in the appendix this is not crucial for our results). Since the spatially lagged dependent variable's coefficient might be biased, it is worthwhile considering the potential consequences of this bias. The presence of a high degree of simultaneity should lead to an over-estimate of the spatially lagged dependent variable, and consequently attenuation

[14]In our case, we omit the historical state border by year interactions from the vector of exogenous variables because we are using the Gaure (2013) technique to estimate these factor/categorical variables as fixed effects.

Table 5: Fixed Treatment Effects, OLS and IV with Spatial Lag and Error: Full Sample and Controls 1300–1900.

	Dependent Variable is Log City Population					
Post-Year	1750 OLS (1)	1800 OLS (2)	1850 OLS (3)	1750 IV (4)	1800 IV (5)	1850 IV (6)
Includes UK						
Coal × Post-Year	0.068^{***}	0.082^{***}	0.087^{***}	0.132^{***}	0.124^{***}	0.128^{***}
	(0.016)	(0.019)	(0.022)	(0.053)	(0.057)	(0.062)
λ (Spatial Lag)	0.347^{***}	0.346^{***}	0.354^{***}	0.287^{***}	0.311^{***}	0.330^{***}
	(0.015)	(0.017)	(0.019)	(0.039)	(0.041)	(0.046)
ρ (Spatial Error Lag)	-0.065	-0.065	-0.066	-0.043	-0.053	-0.057
σ_ϵ^2 (Spatial Error Variance)	0.177	0.177	0.177	0.179	0.178	0.178
Counterfactual Explained (%)	32.918	45.959	77.905	56.929	62.662	96.334
Excludes UK						
Coal × Post-Year	0.038^{**}	0.055^{***}	0.064^{***}	0.098^{*}	0.103^{*}	0.143^{**}
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
λ (Spatial Lag)	0.299^{***}	0.298^{***}	0.298^{***}	0.273^{***}	0.278^{***}	0.270^{***}
	(0.040)	(0.040)	(0.040)	(0.039)	(0.039)	(0.039)
ρ (Spatial Error Lag)	-0.056	-0.056	-0.056	-0.044	-0.048	-0.044
σ_ϵ^2 (Spatial Error Variance)	0.168	0.168	0.168	0.169	0.169	0.169
Counterfactual Explained (%)	42.316	28.934	53.729	41.885	50.255	98.678
Includes Interpolated Cities	Y	Y	Y	Y	Y	Y
Num. obs. Incl. UK	19305	19305	19305	19305	19305	19305
Num. obs. Excl. UK	17613	17613	17613	17613	17613	17613

All regressions include both year and city fixed effects. Standard errors and p-values have been corrected to account for clustering at the city level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. These regressions also include year interactions with the following variables: Latitude, Longitude, Potato, Wheat, Oat, Ruggedness, Altitude, Temperature, Atlantic, Mediterranean, Coast, Rivers, Atlantic^2, Mediterranean^2, Coast^2, and Rivers^2. The border × year fixed effects are as follows: 1500 borders × 1400, 1500 borders × 1500, 1618 borders × 1600, 1699 borders × 1700, 1748 borders × 1750, 1804 borders × 1800, 1848 borders × 1850, and 1908 borders × 1900.

bias in the other regressors. Therefore, our results for the coal coefficients represent an underestimate in the presence of an endogenous spatially lagged dependent variable (Franzese and Hays 2007). Nevertheless, a comparison of the results we present in the text and those in the appendix fails to indicate the existence of any such simultaneity bias.

Table 5 presents the results obtained when we estimate the spatial model outlined above. The use of unbalanced spatial panels can be problematic (this is equivalent to assuming that each missing observation has a value of zero) and thus we only use samples including interpolated populations. As we have already seen, including interpolated city populations does not imply results substantially different from those obtained with samples excluding these observations. The results are stratified in a number of ways. The top and bottom panels indicate results obtained when we include and exclude UK cities from the estimation sample, whereas the alternate columns present results where we first assume that the coal variable enters as an exogenous regressor with different treatment cut-off points (1750, 1800, and 1850), and we then instrument for coal with our Carboniferous variable in the final three columns. We also include a supplementary table in the appendix that includes the relevant weak instrument detection values.

The results in Table 5 are similar to those presented in Table 4. Once again, we see that the conditional relationship between coal and city population size is positive. The coefficient on the spatially lagged dependent variable is positive in all the models. This result is unsurprising as we would expect there to be positive externalities from population growth in neighbouring cities. Interestingly, the spatial error lag is quantitatively small, indicating the absence of a spatial relationship in any of the model's residuals. As before, the "Counterfactual Explained (%)" provides a measure of the economic impact of our results. In this case the figure has been adjusted to account for the relevant spatial effects, so that all of the figures in Table 5 relate to counterfactuals in a full spatial equilibrium. These counterfactual effects again indicate that coal was indeed an important element driving urban growth during the Industrial Revolution, with the effect again being weaker if we exclude the UK from the analysis, consistent with the gradual diffusion of the Industrial Revolution from Britain to the rest of Europe.

7 Conclusion

The role that coal played in shaping economic development during and after the Industrial Revolution has been the subject of considerable debate in the economic history literature. Two schools of thought exist. The first sees coal and the geographical distribution of coalfields as a crucial factor underpinning aggregate and comparative development during this period. The other sees the distribution of coal as relatively unimportant when compared with other factors, such as intellectual and cultural traditions or the quality of political institutions. This paper exploits the spatial variation in the location of Europe's coalfields, and the emergence of coal-based industrial technologies, to quantify the impact of coal on Europe's city populations between 1300 and 1900.

The results indicate that Wrigley is right, and spectacularly so. No less than 60% of urban growth in Europe between 1750 and 1900 can be attributed to the introduction of the coal-using technologies of the Industrial Revolution, even when controlling for period fixed effects and many other factors. Subject to the caveat that our data are for city populations, rather than GDP, this is an impressive vindication of the growth hypothesis. Moreover, the only reason that we have been able to identify this overall growth impact is because proximity to coal mattered so much for city sizes after the Industrial Revolution, but not before. The location hypothesis therefore emerges with flying colours when confronted with the data.

None of this is to suggest that access to coal was a sufficient cause of the Industrial Revolution, or to deny that the underlying force driving the breakthrough to modern economic change was technological progress. Indeed, all our results hinge on the fact that the new coal-using technologies of the Industrial Revolution emerged when they did. What our results do however clearly indicate is that the technological nature of the Industrial Revolution was such that, during the 19th century, access to coal became extremely important in driving economic development. The ultimate sources of growth may have been elsewhere, but we cannot ignore the role of fossil fuels in fuelling growth after the Industrial Revolution, or of geography in determining who experienced that growth during the 19th century.

References

[1] Daron Acemoğlu, Simon Johnson, and James Robinson. The rise of Europe: Atlantic trade, institutional change, and economic growth. *American Economic Review*, 95(3):546–579, 2005.

[2] Robert C. Allen. *The British Industrial Revolution in Global Perspective*. Cambridge University Press, Cambridge, UK, 2009.

[3] Thomas Barnebeck Andersen, Peter Sandholt Jensen, and Christian Stejner Skovsgaard. The heavy plough and the agricultural revolution in medieval Europe. *Department of Business and Economics, University of Southern Denmark, Discussion Papers of Business and Economics, No. 6/2013*, 2013.

[4] Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ, 2009.

[5] Kristine Asch. *The 1:5 Million International Geological Map of Europe and Adjacent Areas*. German Federal Institute for Geosciences and Natural Resources (BGR), Hannover, DE, 2005.

[6] Paul Bairoch, Jean Batou, and Pierre Chèvre. *The Population of European Cities from 800 to 1850*. Librairie Droz, Geneva, CH, 1988.

[7] Theo Balderston. The economics of abundance: Coal and cotton in Lancashire and the world. *Economic History Review*, 63(3):569–590, 2010.

[8] Maarten Bosker, Eltjo Buringh, and Jan Luiten van Zanden. From Baghdad to London: Unraveling urban development in Europe, the Middle East, and North Africa, 800–1800. *Review of Economics and Statistics*, 95(4):1418–1437, 2013.

[9] Fernand Braudel. *Capitalism and Material Life, 1400–1800*. Weidenfeld and Nicolson, London, UK, 1973.

[10] Davide Cantoni. The economic effects of the protestant reformation: Testing the Weber hypothesis in the German lands. *University of Munich, Department of Economics, Discussion Papers in Economics, No. 14811*, 2013.

[11] Châtel and Dollfus. *Atlas Châtel et Dollfus. Les Houillères Européennes*. Société de Documentation Industrielle, Paris, FR, 1931.

[12] Carlo M. Cipolla. *The Economic History of World Population*. Harvester Press, Sussex, UK, 1978.

[13] Gregory Clark. *A Farewell to Alms: A Brief Economic History of the World*. Princeton University Press, Princeton, NJ, 2007.

[14] Gregory Clark and David Jacks. Coal and the Industrial Revolution, 1700–1869. *European Review of Economic History*, 11(1):39–72, 2007.

[15] Nicholas Crafts. Steam as a general purpose technology: A growth accounting perspective. *The Economic Journal*, 114(495):338–351, 2004.

[16] Nicholas Crafts and Abay Mulatu. How did the location of industry respond to falling transport costs in Britain before World War I? *The Journal of Economic History*, 66(3):575–607, 2006.

[17] Nicholas Crafts and Nikolaus Wolf. The location of the UK cotton textiles industry in 1838: A quantitative analysis. *European Historical Economics Society (EHES), Working Papers, No. 45*, 2013.

[18] Jan De Vries. *European Urbanization, 1500–1800*. Havard University Press, Cambridge, MA, 1984.

[19] Phyllis Deane. *The First Industrial Revolution*. Cambridge University Press, Cambridge, UK, 1965.

[20] Jeremiah E. Dittmar. Information technology and economic change: The impact of the printing press. *The Quarterly Journal of Economics*, 126(3):1133–1172, 2011.

[21] FAO/IIASA. *Global Agro-ecological Zones (GAEZ v3.0)*. FAO Rome, IT and IIASA, Laxenburg, AT, 2011.

[22] Robert J. Franzese and Jude C. Hays. Spatial-econometric models of cross-sectional interdependence in political-science panel and time-series-cross-section data. *Political Analysis*, 15(2):140–164, 2007.

[23] Simen Gaure. OLS with multiple high dimensional category variables. *Computational Statistics & Data Analysis*, 66:8–18, 2013.

[24] Theresa Gutberlet. *Cheap Coal, Market Access, and Industry Location in Germany 1846 to 1882*. Mimeo, University of Arizona, 2012.

[25] Angelo Heilprin and Louis Heilprin. *Lippincott's New Gazetteer: A Complete Pronouncing Gazetteer or Geographical Dictionary of the World*. J.B. Lippincott, London, UK, 1906.

[26] Herbert Stanley Jevons. *The British Coal Trade*. Kegan Paul, French Trübner, London, UK, 1915.

[27] Harry H. Kelejian and Ingmar R. Prucha. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, 17(1):99–121, 1998.

[28] Harry H. Kelejian and Ingmar R. Prucha. A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40(2):509–533, 1999.

[29] Alexander Klein and Nicholas Crafts. Making sense of the manufacturing belt: Determinants of U.S. industrial location, 1880–1920. *Journal of Economic Geography*, 12(4):775–807, 2012.

[30] David S. Landes. *The Rise of Capitalism*. Collier-Macmillan, Macmillan, NY, 1966.

[31] Charles F. Manski. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60(3):531–542, 1993.

[32] Julio Martinez-Galarraga. The determinants of industrial location in Spain, 1856–1929. *Explorations in Economic History*, 49(2):255–275, 2012.

[33] Peter Mathias. *The First Industrial Nation: An Economic History of Britain 1700–1914*. Methuen, London, UK, 1983.

[34] Deirdre N. McCloskey. *Bourgeois Dignity: Why Economics Can't Explain the Modern World*. University of Chicago Press, Chicago, IL, 2011.

[35] Giovanni Millo and Gianfranco Piras. splm: Spatial panel data models in R. *Journal of Statistical Software*, 47(1):1–38, 2012.

[36] Joel Mokyr. *Industrialization in the Low Countries, 1795–1850*. Yale University Press, New Haven, CT, 1976.

[37] Joel Mokyr. *Why Ireland Starved: A Quantitative and Analytical History of the Irish Economy, 1800–1850*. Allen & Unwin, London, UK, 1983.

[38] Joel Mokyr. *The Lever of Riches: Technological Creativity and Economic Progress*. Oxford University Press, Oxford, UK, 1990.

[39] Joel Mokyr. Editor's introduction: The new economic history and the industrial revolution. In Joel Mokyr, editor, *The British Industrial Revolution: An Economic Perspective*, pages 1–131. Westview Press, Boulder, CO, 1993.

[40] Joel Mokyr. *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton University Press, Princeton, NJ, 2002.

[41] Joel Mokyr. *The Enlightened Economy: An Economic History of Britain, 1700–1850*. Yale University Press, New Haven, CT, 2009.

[42] Nathan Nunn and Nancy Qian. The potato's contribution to population and urbanization: Evidence from a historical experiment. *The Quarterly Journal of Economics*, 126(2):593–650, 2011.

[43] Ömer Özak. Distance to the technological frontier and economic development. *Southern Methodist University, Department of Economics, Departmental Working Papers, No. 1201*, 2012.

[44] Sidney Pollard. *Peaceful Conquest: The Industrialization of Europe 1760–1970*. Oxford University Press, Oxford, UK, 1981.

[45] Kenneth Pomeranz. *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton University Press, Princeton, NJ, 2000.

[46] Fritz Redlich. The leaders of the German steam-engine industry during the first hundred years. *The Journal of Economic History*, 4(2):121–148, 1944.

[47] Rolf P. Sieferle. *The Subterranean Forest: Energy Systems and the Industrial Revolution*. The White Horse Press, Cambridge, UK, 2001.

[48] G. N. Von Tunzelmann. *Steam power and British Industrialization to 1860*. Clarendon Press, Oxford, UK, 1978.

[49] Margaret F. J. Wilson, Brian O'Connell, Colin Brown, Janine C. Guinan, and Anthony J. Grehan. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30(1–2):3–35, 2007.

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

[51] Gavin Wright. The origins of American industrial success, 1879–1940. *American Economic Review*, 80(4):651–668, 1990.

[52] Edward Anthony Wrigley. *Industrial Growth and Population Change; A Regional Study of the Coalfield Areas of North-West Europe in the Later Nineteenth Century*. Cambridge University Press, Cambridge, UK, 1961.

[53] Edward Anthony Wrigley. *Continuity, Chance and Change: The Character of the Industrial Revolution in England*. Cambridge University Press, Cambridge, UK, 1988.

[54] Edward Anthony Wrigley. *Energy and the English Industrial Revolution*. Cambridge University Press, Cambridge, UK, 2010.

A Data Appendix

Variable Name: City Population Size

Variable Construction: Natural logarithm transformation to normalize variable.

Variable Sources: Bairoch *et al.* (1988), De Vries (1984), Heilprin and Heilprin (1906).

Data partially retrieved from: `http://scholar.harvard.edu/nunn/pages/data-0`.

Note: We use the original source for the population totals in Paris and London for 1850 as these differ from the digitised source.

Variable Name: Coal

Variable Construction: Minimum distance (km) from any of Europe's major coal fields. When a city is positioned within a coal field we assume an arbitrary distance of 1km. We transform this value into a proximity measure by dividing into one (inverse-distance measure). To help interpret the model coefficients and normalize the distribution, we use a multiplicative transformation (multiplying the inverse distance by 10,000) and then take the natural logarithm of our inverse distance measure.

Variable Source: Châtel and Dollfus (1931).

Variable Name: Carbon

Variable Construction: Minimum distance (km) from any onshore geological area classified as being Carboniferous in the IGME 5000. When a city is positioned within such an area we assume an arbitrary distance of 1km. We transform this value into a proximity measure by dividing into one (inverse-distance measure). To help interpret the model coefficients and normalize the distribution, we use a multiplicative transformation (multiplying the inverse distance by 10,000) and then take the natural logarithm of our inverse distance measure.

Variable Source: Asch (2005).

Variable Names: Latitude and Longitude.

Variable Construction: The line of absolute latitude/longitude a city is positioned on. We use these data to construct both the Coal and the Carbon variables.

Variable Source: Bairoch *et al.* (1988).

Variable Names: Potato, Wheat, and Oat

Variable Construction: Suitability of city location for growing specific crops. Data were constructed as in Nunn and Qian (2011), measuring the amount of land suitable for cultivating each particular crop within 100km of the city location. The underlying data for these variables originates from a series of raster images produced by the Food and Agriculture Organization of the United Nations (FAO) under the Global Agro-Ecological Zones (GAEZ) assessment methodology. Our data come from the data portal released in 2011. We use the crop suitability index class to define whether or not land is suitable for cultivating each crop. Like Nunn and Qian, we consider land suitability defined as very high, high, good, and medium as suitable, and land in other classes as unsuitable. We use rain-fed crop conditions with an intermediate input level, to capture historical conditions as accurately as possible. The raster image that underpins these data is in a 5 arc-minute resolution, meaning that the raster cells are typically (although this changes based on distance to the equator) 10km apart. Our suitability variable measures the area of raster cells considered suitable for growing the crop in question where the centroid of the raster cell lies within 100km radius of the city location. This yields a measure in squared km. We take the natural logarithm of this value, after adding the arbitrary value of 1 as some cities have no land suitable for a particular crop.

Variable Sources: FAO/IIASA (2011).

Data retrieved from: `http://gaez.fao.org/Main.html#`.

Variable Name: Altitude

Variable Construction: These values were extracted from a raster image produced by the FAO, and relate to the median meters above sea level measured within each raster cell. We take the natural logarithm of this value, after adding the arbitrary value of 5m as some cities are slightly below sea level.

Variable Sources: FAO/IIASA (2011).

Data retrieved from: `http://gaez.fao.org/Main.html#`.

Variable Names: Ruggedness

Variable Construction: A terrain ruggedness index. These data were constructed by measuring the average terrain ruggedness within 100km of the city location. The underlying data for this variable originates from the altitude raster image produced by the FAO. Our data come from the data portal released in 2011. We convert this altitude raster image to terrain ruggedness indices using the method proposed in Wilson *et al.* (2007), where each cell represents the mean of the absolute differences between the value of that cell and the value of its 8 surrounding cells. The raster image that underpins these data is in a 5 arc-minute resolution, meaning that the raster cells are typically (although this changes based on distance to the equator) 10km apart. We take the natural logarithm of this value.

Variable Sources: FAO/IIASA (2011).

Data retrieved from: `http://gaez.fao.org/Main.html#`.

Variable Name: Temperature

Variable Construction: The mean annual temperature for each city. The underlying data were extracted from a GIS raster image showing the mean annual temperature calculated over the period 1961–1990. We then take the natural logarithm of this value.

Variable Sources: FAO/IIASA (2011).

Data retrieved from: `http://gaez.fao.org/Main.html#`.

Variable Names: Atlantic, Mediterranean, Coast, and Rivers

Variable Construction: We use 1:10m Physical Vector data on coastlines and rivers and measure the minimum distance from each city to each of these individual features. The minimum distance to the coast relates to all coasts worldwide, and we separate the Atlantic from the Baltic and Mediterranean seas at the most northerly point of Denmark and most southerly point of Spain respectively. We then transform these measures into proximity values by inverting, before taking the natural logarithm to normalize the distribution.

Data retrieved from: `http://www.naturalearthdata.com/downloads/10m-physical-vectors/`.

Variable Names: Historic Borders

Variable Construction: We digitize a series of maps that chart the evolution of European borders over the period 1500–2008. We geo-reference and make shapefiles for the following border years (roughly) corresponding to the population data: 1500, 1618, 1699, 1748, 1804, 1848, and 1908.

Data retrieved from: `http://www.iegmaps.de/map2-4.htm` and `http://www.iegmaps.de/map2-1.htm`.

B Data Interpolation Procedure

The use of a spatial panel methodology necessitates the use of balanced panels. However, our source for city populations (Bairoch *et al.*, 1988) does not contain population values for every year in each city. To reconcile our data with the structure required for spatial panel analysis we interpolate these missing values.

Our imputation algorithm begins by examining all cities that have a population value for the year 1300. For each of these cities we iteratively interpolate the city populations by regressing logged population on year and, for cities containing 5 or more data points, year squared. The missing values thus correspond to the predictions from these models.

At the next step we perform an equivalent calculation for cities that have an observation in 1400, but not in 1300. We do not predict missing values in 1300 for this group in this manner; our imputation method for these observations is described later. After calculating predictions for the missing values in the sample of cities with a population value in 1400 but missing data thereafter, we replicate our procedure for 1500, 1600, 1700, 1750, and 1800.

The cities are then divided into 10 geographic clusters based on their contemporary isocodes. These clusters were chosen in order to balance considerations of geographical proximity (we want the cities close to each other) and sample size (we want a sufficient number). The remaining missing data points were calculated by once again iteratively examining each individual city, finding the nearest available data point, and then interpolating using the median growth rate for each particular cluster.

Cluster 1: Albania, Greece, Croatia, Macedonia, Slovenia, Yugoslavia, and Bosnia and Herzegovina.

Cluster 2: Austria, Bulgaria, Czech Republic, Hungary, Moldova, Romania, Slovakia, and Poland.

Cluster 3: Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Lithuania, Latvia, Russia, and Ukraine.

Cluster 4: Belgium, Netherlands, and Luxembourg.

Cluster 5: Switzerland, and Germany.

Cluster 6: Denmark, Norway, Sweden, and Finland.

Cluster 7: Spain, Portugal, and Gibraltar.

Cluster 8: France.

Cluster 9: Ireland and Great Britain.

Cluster 10: Italy and Malta.

C Weak Instrument Statistics

Table 6: Angrist-Pischke Multivariate F-test Statistics for Flexible Regression Models.

	F-Test Statistic			
Endogenous Eqn.	(1)	(2)	(3)	(4)
Coal × Year=1400	37.415	24.195	189.299	105.159
Coal × Year=1500	63.769	41.046	189.299	105.159
Coal × Year=1600	97.438	58.889	182.022	98.515
Coal × Year=1700	143.316	55.200	166.792	85.975
Coal × Year=1750	89.860	52.630	163.473	83.351
Coal × Year=1800	157.752	77.812	150.804	72.446
Coal × Year=1850	159.081	82.242	156.610	77.276
Coal × Year=1900	214.498	129.149	184.674	105.930
Excludes UK	N	Y	N	Y
Includes Interpolated Cities	N	N	Y	Y
Num. obs.	10773	9799	19305	17613

The number in each cell displays the Angrist-Pischke multivariate F-test statistic for the corresponding endogenous variable in the regression estimates reported in Table 2. Each F-test statistic has been corrected to account for clustering at the city level.

Table 7: First-Stage F-Statistics, Rolling Growth Regressions.

	F-Test Statistic					
	(1)	(2)	(3)	(4)	(5)	(6)
IV	22.554	25.723	30.719	68.907	171.233	202.852
IV: Excl. UK	11.598	11.109	13.996	29.419	49.605	88.272
IV: Incl. Interpolated	177.212	177.306	177.365	178.070	178.250	177.723
IV: Excl UK & Incl. Interpolated	80.149	80.223	80.253	80.212	80.032	80.039

The number in each cell displays the partial first-stage F-test statistic for a corresponding growth regression model reported in Table 3 wherein the Coal variable is modelled as an endogenous regressor. Each F-test statistic has been corrected to account for heteroscedasticity.

Table 8: First-Stage F-Statistics, Fixed Treatment Effects.

	F-Test Statistic			
Endogenous Eqn.	(1)	(2)	(3)	(4)
Coal $\times$ Post-1750	207.579	99.082	211.403	110.968
Coal $\times$ Post-1800	259.606	134.614	211.403	110.968
Coal $\times$ Post-1850	259.606	134.614	211.403	110.968
Excludes UK	N	Y	N	Y
Includes Interpolated Cities	N	N	Y	Y
Num. obs.	10773	9799	19305	17613

The number in each cell displays the partial first-stage F-test statistic for the corresponding regression model reported in Table 4 wherein the Coal variable is modelled as an endogenous regressor. Each F-test statistic has been corrected to account for clustering at the city level.

Table 9: First-Stage F-Statistics, Fixed Treatment Effects, OLS and IV with Spatial Lag and Error: Full Sample and Controls 1300–1900.

	F-Test Statistic		
Post-Year	1750 (1)	1800 (2)	1850 (3)
Includes UK			
F-Statistic	180.787	180.651	184.431
Excludes UK			
F-Statistic	96.029	97.761	104.191
Includes Interpolated Cities	Y	Y	Y
Num. obs. Incl. UK	19305	19305	19305
Num. obs. Excl. UK	17613	17613	17613

The number in each cell displays the partial first-stage F-test statistic for a corresponding spatial regression model reported in Table 5 wherein the Coal variable is modelled as an endogenous regressor. Each F-test statistic has been corrected to account for clustering at the city level.

D Spatial Estimators with Endogenous Lagged Dependent Variable

Table 10: Fixed Treatment Effects, OLS and IV with Spatial Lag and Error: Full Sample and Controls 1300–1900. Spatial Lag Modeled as Endogenous Regressor.

	Dependent Variable is Log City Population					
Post-Year	1750 OLS (1)	1800 OLS (2)	1850 OLS (3)	1750 IV (4)	1800 IV (5)	1850 IV (6)
Includes UK						
Coal $\times$ Post-Year	0.061^{***}	0.074^{***}	0.078^{***}	0.112^{***}	0.122^{***}	0.132^{***}
	(0.017)	(0.021)	(0.024)	(0.043)	(0.048)	(0.052)
λ (Spatial Lag)	0.296^{***}	0.296^{***}	0.297^{***}	0.262^{***}	0.262^{***}	0.271^{***}
	(0.018)	(0.020)	(0.023)	(0.039)	(0.043)	(0.047)
ρ (Spatial Error Lag)	0.110	0.106	0.117	0.128	0.126	0.129
σ^2_ϵ (Spatial Error Variance)	0.179	0.179	0.179	0.179	0.179	0.179
Counterfactual Explained (%)	27.735	38.782	64.551	47.219	57.553	91.585
Excludes UK						
Coal $\times$ Post-Year	0.031^{*}	0.046^{**}	0.051^{**}	0.068	0.081^{*}	0.100^{*}
	(0.078)	(0.078)	(0.081)	(0.089)	(0.089)	(0.089)
λ (Spatial Lag)	0.284^{***}	0.279^{***}	0.280^{***}	0.280^{***}	0.276^{***}	0.276^{***}
	(0.082)	(0.081)	(0.082)	(0.085)	(0.085)	(0.085)
ρ (Spatial Error Lag)	0.091	0.089	0.090	0.094	0.091	0.091
σ^2_ϵ (Spatial Error Variance)	0.170	0.169	0.169	0.170	0.170	0.170
Counterfactual Explained (%)	30.334	23.675	42.067	29.909	39.508	71.084
Includes Interpolated Cities	Y	Y	Y	Y	Y	Y
Num. obs. Incl. UK	19305	19305	19305	19305	19305	19305
Num. obs. Excl. UK	17613	17613	17613	17613	17613	17613

All regressions include both year and city fixed effects. Standard errors and p-values have been corrected to account for clustering at the city level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. These regressions also include year interactions with the following variables: Latitude, Longitude, Potato, Wheat, Oat, Ruggedness, Altitude, Temperature, Atlantic, Mediterranean, Coast, Rivers, Atlantic^2, Mediterranean^2, Coast^2, and Rivers^2. The border $\times$ year fixed effects are as follows: 1500 borders $\times$ 1400, 1500 borders $\times$ 1500, 1618 borders $\times$ 1600, 1699 borders $\times$ 1700, 1748 borders $\times$ 1750, 1804 borders $\times$ 1800, 1848 borders $\times$ 1850, and 1908 borders $\times$ 1900.

Table 11: Kleibergen-Paap rk Wald F statistic, Fixed Treatment Effects, OLS and IV with Spatial Lag and Error: Full Sample and Controls 1300–1900.

	F-Test Statistic		
Post-Year	1750	1800	1850
	(1)	(2)	(3)
Includes UK			
Kleibergen-Paap rk Wald F-statistic	4.034	3.814	3.274
Excludes UK			
Kleibergen-Paap rk Wald F-statistic	3.438	3.261	3.003
Includes Interpolated Cities	Y	Y	Y
Num. obs. Incl. UK	19305	19305	19305
Num. obs. Excl. UK	17613	17613	17613

The number in each cell displays the Kleibergen-Paap rk Wald F statistic for the corresponding spatial regression models reported in Table 10, in which the Coal variable and the spatial lag are modelled as endogenous regressors. Each F-test statistic has been corrected to account for clustering at the city level.